Composition de Mathématiques

Le 27 janvier 2021 – De 13 heures à 17 heures

Si, au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il le signale sur sa copie et poursuit sa composition en expliquant les raisons des initiatives qu'il est amené à prendre.

La présentation et la rédaction comptent pour une part importante dans l'appréciation de la copie.

Les calculatrices sont interdites.
Les téléphones portables doivent être éteints et rangés.

Partie commune

❖ I – Problème ❖

Soit X, une variable aléatoire discrète définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$. On suppose que cette variable aléatoire est bornée : il existe donc un réel M tel que

$$\forall\, \omega \in \Omega, \quad \big|X(\omega)\big| \leqslant M.$$

1. La variable aléatoire X est-elle d'espérance finie ? Admet-elle un moment d'ordre deux ?

2. Dans un premier temps, un entier $n \geqslant 2$ est fixé et on suppose qu'il existe n variables aléatoires discrètes, indépendantes et de même loi, définies sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, telles que la somme $S_n = X_1 + \cdots + X_n$ ait même loi que X.

2. a. Démontrer que

$$\bigcap_{1\leqslant i\leqslant n} \left[X_i > \frac{M}{n}\right]$$

est un événement, puis qu'il est négligeable.

2. b. En déduire que

$$\forall\, 1 \leqslant i \leqslant n, \qquad \mathbf{P}\left(X_i > \frac{M}{n}\right) = 0.$$

2. c. Démontrer que

$$\forall\, 1 \leqslant i \leqslant n, \qquad \mathbf{P}\left(|X_i| \leqslant \frac{M}{n}\right) = 1.$$

2. d. Démontrer que $\mathbf{E}(X_1^2) \leqslant \frac{M^2}{n^2}$ puis que $\mathbf{V}(X) \leqslant \frac{M^2}{n}$.

3. On suppose maintenant que la variable aléatoire X est **indéfiniment divisible** au sens où, pour tout entier $n \geqslant 2$, il existe des variables aléatoires discrètes $X_{n,1}, \ldots, X_{n,n}$ définies sur $(\Omega, \mathcal{A}, \mathbf{P})$, indépendantes, de même loi et telles que la somme

$$X_{n,1} + \cdots + X_{n,n}$$

ait même loi que X.

3. a. Appliquer l'inégalité de Bienaymé-Tchebychev à X.

3. b. En déduire que

$$\mathbf{P}\big(X = \mathbf{E}(X)\big) = 1$$

en considérant les événements $A_n = [|X - \mathbf{E}(X)| \geqslant 1/n]$.

Que signifie cette propriété ?

❖ II – Problème ❖

On considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ sur lequel sont définies trois variables aléatoires X, Y_1 et Y_2 à valeurs dans $E = \{0; 1\}$.

On suppose que X suit la loi de Bernoulli $\mathscr{B}(p)$ avec $0 < p < 1$. On suppose de plus que les variables aléatoires Y_1 et Y_2 sont conditionnellement indépendantes sachant X au sens où :

$$\forall\, \varepsilon_0, \varepsilon_1, \varepsilon_2 \in E, \qquad \mathbf{P}_{[X=\varepsilon_0]}(Y_1 = \varepsilon_1, Y_2 = \varepsilon_2) = \mathbf{P}_{[X=\varepsilon_0]}(Y_1 = \varepsilon_1) \cdot \mathbf{P}_{[X=\varepsilon_0]}(Y_2 = \varepsilon_2). \tag{1}$$

On suppose enfin que

$$\mathbf{P}_{[X=1]}(Y_1 = 1) = \mathbf{P}_{[X=1]}(Y_2 = 1) = 1/2, \tag{2}$$
$$\mathbf{P}_{[X=0]}(Y_1 = 1) = \mathbf{P}_{[X=0]}(Y_2 = 1) = 1/6. \tag{3}$$

1. Pourquoi les relations (2) et (3) ont-elles un sens ?

2. Quelle est la loi de Y_1 ? celle de Y_2 ?

3. Les variables aléatoires X et Y_1 sont-elles indépendantes ?

4. Les variables aléatoires Y_1 et Y_2 sont-elles indépendantes ? (Considérer les événements $[Y_1 = 1]$ et $[Y_2 = 1]$.)

5. Calculer la probabilité $\mathbf{P}(X = 1)$, puis les probabilités conditionnelles $\mathbf{P}_{[Y_1=1]}(X = 1)$ et $\mathbf{P}_{[Y_1=1, Y_2=1]}(X = 1)$.

6. On dispose d'un sac contenant 100 dés cubiques, dont 25 sont truqués. On admet que la probabilité d'obtenir 6 en lançant un dé truqué est égale à $1/2$.

On choisit un dé au hasard, en se demandant s'il est truqué. On lance ce dé et on obtient 6. On lance à nouveau ce dé et on obtient encore 6.

Que penser de cela ?

Piste rouge (Centrale PC)

❖ III – Problème ❖

Dans ce problème, on considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ et une variable aléatoire discrète

$$X : \Omega \to E$$

définie sur $(\Omega, \mathcal{A}, \mathbf{P})$, à valeurs dans une partie dénombrable $E \subset \mathbb{R}$. On supposera connue une énumération de cette partie :

$$E = \{x_n,\ n \in \mathbb{N}\}$$

et on notera

$$\forall\, n \in \mathbb{N}, \quad a_n = \mathbf{P}(X = x_n).$$

L'objectif de ce problème est d'étudier la fonction

$$\varphi_X : \mathbb{R} \to \mathbb{C}$$

définie par

$$\forall\, t \in \mathbb{R}, \quad \varphi_X(t) = \mathbf{E}(e^{itX}).$$

Partie A. Premières propriétés

1. Démontrer que la série $\sum a_n$ converge et que sa somme est égale à 1.

2. Démontrer que la série $\sum a_n e^{itx_n}$ est absolument convergente pour tout $t \in \mathbb{R}$.

3. En déduire que φ_X est bien définie sur $\mathbb{R}$, que

$$\forall\, t \in \mathbb{R}, \quad \varphi_X(t) = \sum_{n=0}^{+\infty} a_n e^{itx_n}$$

et que φ_X est continue sur $\mathbb{R}$.

4. Exprimer $\varphi_X(-t)$ en fonction de $\varphi_X(t)$. Que peut-on en déduire si la fonction φ_X est paire ?

Partie B. Image de φ_X

Pour $(a, b) \in \mathbb{R}^2$, on note

$$a + b\mathbb{Z} = \{a + bk,\ k \in \mathbb{Z}\}.$$

5. Démontrer que

$$\forall\, t \in \mathbb{R}, \quad |\varphi_X(t)| \leqslant 1.$$

6. Dans cette question, on suppose qu'il existe $a \in \mathbb{R}$ et $t_0 \in \mathbb{R}^*$ tels que

$$E \subset a + \frac{2\pi}{t_0}\mathbb{Z}.$$

Démontrer que $|\varphi_X(t_0)| = 1$.

7. Réciproquement, nous supposons maintenant qu'il existe $t_0 \in \mathbb{R}^*$ tel que

$$|\varphi_X(t_0)| = 1.$$

7. a. Démontrer qu'il existe $a \in \mathbb{R}$ tel que

$$\sum_{n=0}^{+\infty} a_n \exp\big[i(t_0 x_n - t_0 a)\big] = 1.$$

7. b. En déduire que

$$\sum_{n=0}^{+\infty} a_n \big[1 - \cos(t_0 x_n - t_0 a)\big] = 0.$$

7. c. Démontrer pour tout $n \in \mathbb{N}$ l'alternative suivante :
- ou bien $a_n = 0$,
- ou bien $x_n \in a + \frac{2\pi}{t_0}\mathbb{Z}$.

7. d. En déduire que

$$\mathbf{P}\Big(X \in a + \frac{2\pi}{t_0}\mathbb{Z}\Big) = 1.$$

Partie C. Fonction caractéristique et loi de X

L'objectif de cette partie est de montrer que, comme son nom l'indique, la fonction caractéristique de X *caractérise* la loi de X.

On note sinc, la fonction *sinus cardinal*, définie par

$$\operatorname{sinc} 0 = 1 \quad \text{et} \quad \forall\, x \in \mathbb{R}^*, \quad \operatorname{sinc} x = \frac{\sin x}{x}.$$

On **admet** que sinc est continue sur $\mathbb{R}$ et que

$$\forall\, x \in \mathbb{R}, \quad |\operatorname{sinc} x| \leqslant 1.$$

On fixe un réel m. Pour tout $T > 0$, on pose

$$V_m(T) = \frac{1}{2T}\int_{-T}^{T} \varphi_X(t) e^{-imt}\, dt$$

ainsi que, pour tout $n \in \mathbb{N}$,

$$g_n(T) = \operatorname{sinc}\big((x_n - m)T\big) \cdot \mathbf{P}(X = x_n).$$

8. Démontrer que

$$\forall\, T > 0, \quad V_m(T) = \sum_{n=0}^{+\infty} g_n(T).$$

9. Démontrer que la fonction V_m est continue sur $\mathbb{R}_+^*$.

10. Démontrer que chaque fonction g_n tend vers une limite finie au voisinage de $+\infty$.

11. Démontrer que

$$\lim_{T \to +\infty} V_m(T) = \mathbf{P}(X = m).$$

12. On suppose que $X : \Omega \to \mathbb{R}$ et $Y : \Omega \to \mathbb{R}$ sont deux variables aléatoires discrètes telles que

$$\forall\, t \in \mathbb{R}, \quad \varphi_X(t) = \varphi_Y(t).$$

Démontrer que

$$\forall\, m \in \mathbb{R}, \quad \mathbf{P}(X = m) = \mathbf{P}(Y = m).$$

Comment interpréter ce résultat ?

Piste noire (Centrale MP)

❖ IV – Problème ❖

On considère ici une variable aléatoire X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, à valeurs dans $\mathbb{R}$ et telle que
$$\mathbf{P}(X = 0) > 0.$$
Sa fonction génératrice est notée G_X.

1. Démontrer qu'il existe une unique suite réelle $(\lambda_i)_{i\geqslant 1}$ telle que
$$\forall\, k \in \mathbb{N}^*, \quad k\,\mathbf{P}(X = k) = \sum_{j=1}^{k} j\lambda_j\,\mathbf{P}(X = k - j).$$

2. Pour $k \in \mathbb{N}^*$, démontrer que
$$\left|\lambda_k\right|\,\mathbf{P}(X = 0) \leqslant \mathbf{P}(X = k) + \sum_{j=1}^{k-1} |\lambda_j|\,\mathbf{P}(X = k - j)$$
puis que
$$\left|\lambda_k\right|\,\mathbf{P}(X = 0) \leqslant \left[1 - \mathbf{P}(X = 0)\right] \cdot \left(1 + \sum_{j=1}^{k-1} |\lambda_j|\right).$$

3. Démontrer que
$$\forall\, k \in \mathbb{N}^*, \quad 1 + \sum_{j=1}^{k} |\lambda_j| \leqslant \frac{1}{\mathbf{P}(X = 0)^k}.$$

4. On note $\rho(X)$, le rayon de convergence de la série entière $\sum \lambda_k t^k$. Démontrer que
$$\rho(X) \geqslant \mathbf{P}(X = 0).$$

5. Pour tout réel $t \in \left]-\rho(X), \rho(X)\right[$, on pose
$$H_X(t) = \ell n\,\mathbf{P}(X = 0) + \sum_{k=1}^{+\infty} \lambda_k t^k.$$

5. a. Démontrer que
$$\forall\, t \in \left]-\rho(X), \rho(X)\right[, \quad G_X'(t) = H_X'(t) G_X(t).$$

5. b. En déduire l'expression de $H_X(t)$ en fonction de la fonction génératrice $G_X(t)$.

❖ V – Problème ❖

Partie A.

Soient X et Y, deux variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ à valeurs dans $\mathbb{N}$. Les fonctions génératrices de X et Y sont notées G_X et G_Y respectivement.

1. Démontrer que $[X \neq Y]$ est un événement.

2. Soient A et B, deux événements.

Démontrer que les probabilités $\mathbf{P}(A \cap B^c)$ et $\mathbf{P}(A^c \cap B)$ sont bien définies, puis que
$$\left|\mathbf{P}(A) - \mathbf{P}(B)\right| \leqslant \mathbf{P}(A \cap B^c) + \mathbf{P}(A^c \cap B).$$

3. En déduire que
$$\forall\, t \in [-1, 1], \quad \left|G_X(t) - G_Y(t)\right| \leqslant 2\,\mathbf{P}(X \neq Y).$$

Partie B.

Soit $(U_k)_{k\geqslant 1}$, une suite de variables aléatoires définies sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, indépendantes, à valeurs dans $\mathbb{N}$ et telles que la série $\sum \mathbf{P}(U_k \neq 0)$ soit convergente.

On note $\mathbb{N}' = \mathbb{N} \cup \{+\infty\}$.

4. Démontrer que l'ensemble $\mathbb{N}'$ est dénombrable. Est-il contenu dans $\mathbb{R}$?

5. Pour tout entier $n \in \mathbb{N}^*$, on pose
$$Z_n = \left\{\omega \in \Omega \;:\; \exists\, k \geqslant n, \quad U_k(\omega) \neq 0\right\}.$$
Démontrer que $(Z_n)_{n\geqslant 1}$ est une suite décroissante d'événements et que
$$\lim_{n\to+\infty} \mathbf{P}(Z_n) = 0.$$

6. On considère l'application $W \;:\; \Omega \to \mathbb{N}'$ définie par
$$\forall\, \omega \in \Omega, \quad W(\omega) = \#\left\{k \in \mathbb{N}^* \;:\; U_k(\omega) \neq 0\right\}.$$
Démontrer que W est une variable aléatoire discrète et que
$$\mathbf{P}(W = +\infty) = 0.$$

7. Pour tout $n \in \mathbb{N}^*$, on pose $S_n = U_1 + \cdots + U_n$ et, pour tout $\omega \in \Omega$, on pose
$$S(\omega) = \sum_{k=1}^{+\infty} U_k(\omega)$$
si la série $\sum U_k(\omega)$ converge et $S(\omega) = +\infty$ dans le cas contraire.

On pose enfin
$$\forall\, t \in [0, 1], \quad G(t) = \sum_{i=0}^{+\infty} \mathbf{P}(S = i) t^i.$$

7. a. Démontrer que S_n est une variable aléatoire à valeurs dans $\mathbb{N}$ pour tout $n \in \mathbb{N}^*$.

7. b. Démontrer que S est une variable aléatoire à valeurs dans $\mathbb{N}'$ et que $[S \in \mathbb{N}]$ est un événement presque sûr.

7. c. Démontrer que G est continue sur le segment $[0, 1]$ et de classe $\mathscr{C}^\infty$ sur l'intervalle $[0, 1[$.

7. d. Pour tout $n \in \mathbb{N}^*$, on note G_n, la fonction de répartition de S_n. Démontrer que la suite $(G_n)_{n\geqslant 1}$ converge uniformément sur $[0, 1]$ vers G.

Solution I ❀ Variables bornées indéfiniment divisibles

1. Une variable aléatoire bornée admet des moments de tout ordre, en particulier d'ordre deux, mais aussi d'ordre un : c'est donc une variable aléatoire d'espérance finie et sa variance est bien définie.

2. a. Comme X_i est une variable aléatoire sur $(\Omega, \mathcal{A})$ à valeurs réelles, alors

$$\forall\, \alpha \in \mathbb{R}, \quad [X_i > \alpha] \in \mathcal{A}.$$

Comme $\mathcal{A}$ est une tribu, elle est stable par intersection finie ou dénombrable, donc

$$B \overset{\text{déf.}}{=} \bigcap_{1 \leqslant i \leqslant n} \left[X_i > \frac{M}{n}\right] \in \mathcal{A}.$$

❧ Soit $\omega \in B$. On a donc

$$\forall\, 1 \leqslant i \leqslant n, \quad X_i(\omega) > \frac{M}{n}$$

et donc, en sommant ces inégalités,

$$S_n(\omega) = X_1(\omega) + \cdots + X_n(\omega) > M.$$

On a ainsi démontré que

$$B \subset [S_n > M]$$

et donc que

$$0 \leqslant \mathbf{P}(B) \leqslant \mathbf{P}(S_n > M)$$

par croissance de $\mathbf{P}$.

Or S_n et X ont même loi, donc

$$\mathbf{P}(S_n > M) = \mathbf{P}(X > M) = 0$$

puisque

$$[X > M] \subset [|X| \leqslant M]^c = \Omega^c = \varnothing.$$

L'événement B est donc négligeable.

Remarque.— On ne peut pas démontrer que B est impossible, seulement que B est négligeable : on sait en effet que S_n et X ont même loi, mais rien ne nous assure que les fonctions S_n et X soient égales. Il se pourrait donc que S_n prît des valeurs strictement supérieures à M (mais avec une probabilité nulle, bien entendu).

2. b. Comme les variables aléatoires X_i sont indépendantes et de même loi,

$$\forall\, 1 \leqslant k \leqslant n, \quad \mathbf{P}\left(\bigcap_{1 \leqslant i \leqslant n} \left[X_i > \frac{M}{n}\right]\right) = \left[\mathbf{P}\left(X_k > \frac{M}{n}\right)\right]^n$$

et donc, puisque cette probabilité est nulle,

$$\forall\, 1 \leqslant k \leqslant n, \quad \mathbf{P}\left(X_k > \frac{M}{n}\right) = 0.$$

❧ Par passage au complémentaire, on en déduit que

$$\forall\, 1 \leqslant i \leqslant n, \qquad \mathbf{P}\left(X_i \leqslant \frac{M}{n}\right) = 1.$$

2. c. On raisonne de même sur l'événement

$$C = \bigcap_{1 \leqslant i \leqslant n} \left[X_i < \frac{-M}{n}\right]$$

et on en déduit pareillement que

$$\forall\, 1 \leqslant i \leqslant n, \qquad \mathbf{P}\left(X_i \geqslant \frac{-M}{n}\right) = 1.$$

Une intersection finie (voire dénombrable) d'événements presque sûrs est encore un événement presque sûr, donc l'événement

$$\left[X_i \leqslant \frac{M}{n}\right] \cap \left[X_i \geqslant \frac{-M}{n}\right] = \left[|X_i| \leqslant \frac{M}{n}\right]$$

est presque sûr (pour tout $1 \leqslant i \leqslant n$).

2. d. Comme l'application $\left[x \mapsto x^2\right]$ est strictement croissante sur $\mathbb{R}_+$,

$$\left[|X_1| \leqslant \frac{M}{n}\right] = \left[X_1^2 \leqslant \frac{M^2}{n^2}\right]$$

et d'après la question précédente,

$$\mathbf{P}\left(X_1^2 \leqslant \frac{M^2}{n^2}\right) = 1.$$

Les inégalités presque sûres sont conservées par l'espérance, donc

$$\forall\, 1 \leqslant i \leqslant n, \quad \mathbf{E}(X_1^2) \leqslant \frac{M^2}{n^2}.$$

❧ Comme X et S_n ont même loi, alors $\mathbf{V}(X) = \mathbf{V}(S_n)$. Comme S_n est la somme de variables aléatoires indépendantes et de même loi,

$$\mathbf{V}(S_n) = \sum_{i=1}^{n} \mathbf{V}(X_i) = n\,\mathbf{V}(X_1)$$

et d'après Koenig-Huyghens,

$$\mathbf{V}(X_1) = \mathbf{E}(X_1^2) - [\mathbf{E}(X_1)]^2 \leqslant \mathbf{E}(X_1^2).$$

Finalement

$$\mathbf{V}(X) = \mathbf{V}(S_n) = n\,\mathbf{V}(X_1) \leqslant \frac{M^2}{n}.$$

3. a. Comme X admet un moment d'ordre deux,

$$\forall\, \alpha > 0, \quad \mathbf{P}\big(|X - \mathbf{E}(X)| > \alpha\big) \leqslant \frac{\mathbf{V}(X)}{\alpha^2}.$$

3. b. Les hypothèses de la question **2.** sont ici vérifiées pour tout entier $n \geqslant 2$.

D'après **2.d.**,

$$\forall\, n \geqslant 2, \qquad \mathbf{V}(X) \leqslant \frac{M^2}{n}$$

et d'après l'inégalité de Bienaymé-Tchebychev rappelée à la question précédente,

$$\forall\, \alpha > 0,\ \forall\, n \geqslant 2, \quad \mathbf{P}\big(|X - \mathbf{E}(X)| > \alpha\big) \leqslant \frac{M^2}{n\alpha^2}.$$

Cette inégalité large étant vérifiée pour tout $n \geqslant 2$, on peut faire tendre n vers $+\infty$ pour obtenir :

$$\forall\, \alpha > 0, \quad \mathbf{P}\big(\big|X - \mathbf{E}(X)\big| > \alpha\big) = 0$$

ce qui prouve que tous les événements A_n sont négligeables.

❧ Une union dénombrable d'événements négligeables est encore un événement négligeable, donc l'événement

$$B = \Big(\bigcup_{n\geqslant 2} A_n\Big)^c = \bigcap_{n\geqslant 2} A_n^c$$

est presque sûr (en tant que complémentaire d'un événement négligeable). Or

$$\begin{aligned}
\omega \in B &\iff \forall\, n \geqslant 2, \quad \omega \notin A_n\\
&\iff \forall\, n \geqslant 2, \quad \big|X(\omega) - \mathbf{E}(X)\big| < \frac{1}{n}\\
&\iff \big|X(\omega) - \mathbf{E}(X)\big| = 0\\
&\iff X(\omega) = \mathbf{E}(X).
\end{aligned}$$

On a donc bien démontré que la variable aléatoire X était presque sûrement constante :

$$\mathbf{P}\big(X = \mathbf{E}(X)\big) = 1.$$

Solution II ❁ Lois conditionnelles

1. Par hypothèse, $\mathbf{P}(X = 1) = p$ et, par conséquent,

$$\mathbf{P}(X = 0) = 1 - \mathbf{P}(X = 1) = 1 - p.$$

Comme $0 < p < 1$, ni l'événement $[X = 1]$, ni l'événement $[X = 0]$ n'est négligeable et les probabilités conditionnelles (2) et (3) ont bien un sens.

2. Comme Y_1 est une variable aléatoire à valeurs dans $\{0; 1\}$, sa loi est une loi de Bernoulli, qui est donc caractérisée par son paramètre $\mathbf{P}(Y_1 = 1)$.

Comme X est une variable aléatoire de Bernoulli, le couple $\big([X = 0], [X = 1]\big)$ est un système complet d'événements. D'après la Formule des probabilités totales,

$$\mathbf{P}(Y_1 = 1) = \mathbf{P}(Y_1 = 1 \mid X = 1)\,\mathbf{P}(X = 1) + \mathbf{P}(Y_1 = 1 \mid X = 0)\,\mathbf{P}(X = 0)$$

et donc

$$\mathbf{P}(Y_1 = 1) = \frac{p}{2} + \frac{1-p}{6} = \frac{2p+1}{6}.$$

❧ D'après (2) et (3), la loi de Y_2 est la même que la loi de Y_1.

3. Par hypothèse, $\mathbf{P}(X = 1) = p \in \,]0, 1[$ et

$$\mathbf{P}(X = 1, Y_1 = 1) = \frac{1}{2} \times p.$$

D'après la question précédente,

$$\mathbf{P}(X = 1)\,\mathbf{P}(Y_1 = 1) = p \times \frac{2p+1}{6}.$$

Si X et Y_1 étaient indépendantes, alors on aurait

$$\frac{p}{2} = \frac{p}{2} \times \frac{2p+1}{3}$$

et donc $p = 0$ (exclu a priori) ou $2p + 1 = 3$, c'est-à-dire $p = 1$ (exclu a priori).

Les variables aléatoires X et Y_1 ne sont donc pas indépendantes.

4. Comme $([X = 1], [X = 0])$ est un système complet d'événements (bis),

$$\begin{aligned}
\mathbf{P}(Y_1 = 1, Y_2 = 1) &= \mathbf{P}(X = 1, Y_1 = 1, Y_2 = 1) + \mathbf{P}(X = 0, Y_1 = 1, Y_2 = 1)\\
&= \mathbf{P}(Y_1 = 1, Y_2 = 1 \mid X = 1)\,\mathbf{P}(X = 1) + \mathbf{P}(Y_1 = 1, Y_2 = 1 \mid X = 0)\,\mathbf{P}(X = 0)\\
&= \mathbf{P}_{[X=1]}(Y_1 = 1)\,\mathbf{P}_{[X=1]}(Y_2 = 1)\,\mathbf{P}(X = 1) + \mathbf{P}_{[X=0]}(Y_1 = 1)\,\mathbf{P}_{[X=0]}(Y_2 = 1)\,\mathbf{P}(X = 0) \qquad \text{d'après (1)}\\
&= \frac{p}{4} + \frac{1-p}{36} = \frac{1+8p}{36}.
\end{aligned}$$

D'après **2.**,

$$\mathbf{P}(Y_1 = 1)\,\mathbf{P}(Y_2 = 1) = \frac{(2p+1)^2}{36}.$$

Ainsi, si Y_1 et Y_2 étaient indépendantes, il faudrait que $(2p+1)^2 = 8p+1$, c'est-à-dire $4p(p-1) = 0$, ce qui est impossible puisque $0 < p < 1$ par hypothèse.

Les variables aléatoires Y_1 et Y_2 ne sont donc pas indépendantes.

REMARQUE.— L'indépendance des variables aléatoires est une notion relative : par hypothèse, les variables aléatoires Y_1 et Y_2 sont indépendantes pour les mesures de probabilité $\mathbf{P}_{[X=1]}$ et $\mathbf{P}_{[X=0]}$ mais, comme on vient de le vérifier, elles ne sont pas indépendantes pour la mesure $\mathbf{P}$.

Il est aisé d'être prudent dans un cadre théorique (comme c'est le cas ici), c'est beaucoup moins simple dans un cadre trop concret (comme c'est le cas au **6.**, où l'hypothèse d'indépendance des variables aléatoires Y_1 et Y_2 n'est pas évidente à traduire).

5. Par définition du paramètre d'une loi de Bernoulli, $\mathbf{P}(X = 1) = p$.

D'après la définition des probabilités conditionnelles, d'après **2.** et **3.**,

$$\mathbf{P}_{[Y_1=1]}(X = 1) = \frac{\mathbf{P}(X = 1, Y_1 = 1)}{\mathbf{P}(Y_1 = 1)} = \frac{p/2}{(2p + 1)/6} = \frac{3p}{2p + 1}.$$

De même, d'après **4.**,

$$\mathbf{P}_{[Y_1=1,Y_2=1]}(X = 1) = \frac{\mathbf{P}(X = 1, Y_1 = 1, Y_2 = 1)}{\mathbf{P}(Y_1 = 1, Y_2 = 1)} = \frac{p/4}{(1 + 8p)/36} = \frac{9p}{8p + 1}.$$

6. On modélise le choix du dé par la variable aléatoire X en convenant que $X = 1$ si on choisit un dé truqué. Il est alors raisonnable de poser $p = 1/4$.

Une fois que le dé est choisi, on le lance deux fois. Il est raisonnable de penser que ces deux lancers peuvent être modélisés par des variables aléatoires Y_1 et Y_2, indépendantes et de même loi : loi $\mathscr{B}(1/2)$ pour un dé truqué, loi $\mathscr{B}(1/6)$ pour un dé non truqué.

On précise que l'hypothèse d'indépendance est postérieure au choix du dé (cf remarque au **4.**).

D'après **5.**,
- la probabilité de choisir un dé truqué est égale à $p = 1/4$;
- en lançant le dé une première fois et en obtenant 6, on reçoit une information supplémentaire, grâce à laquelle la probabilité d'avoir tiré un dé truqué monte à $1/2$;
- en lançant le dé une seconde fois et en obtenant à nouveau 6, on reçoit une nouvelle information qui d'une certaine façon confirme l'information précédente et nous dit que la probabilité d'avoir tiré un dé truqué s'élève maintenant à $3/4$.

❧ Dans cette optique, calculer la probabilité conditionnelle de B sachant A revient à calculer la probabilité de l'événement B *en prenant en compte l'information supplémentaire* apportée par la réalisation de l'événement A.

Si A et B sont indépendants, alors $\mathbf{P}_A(B) = \mathbf{P}(B)$ et la réalisation de A n'apporte aucune information supplémentaire au sujet de B.

Dans le cas qui nous intéresse, en choisissant le dé, on imagine avoir une chance sur quatre d'avoir un dé truqué. En le lançant et en obtenant 6 deux fois de suite, nous sommes incités à penser que le dé est plus probablement truqué (trois chances sur quatre d'après nos calculs).

Comme disait ma grand'mère, c'est aux fruits qu'on reconnaît l'arbre.

(Exercice tiré de l'énoncé 105 de la banque CCINP.)

Solution III ❀ Fonction caractéristique

Partie A. Premières propriétés

1. Comme $\{x_n,\ n \in \mathbb{N}\}$ est une énumération de E, la famille

$$\big([X = x_n]\big)_{n\in\mathbb{N}}$$

est un système complet d'événements. Par σ-additivité de $\mathbf{P}$, on en déduit d'une part que la série $\sum \mathbf{P}(X = x_n)$ est convergente et d'autre part que

$$\sum_{n=0}^{+\infty} \mathbf{P}(X = x_n) = \mathbf{P}(\Omega) = 1.$$

2. Pour tout $t \in \mathbb{R}$, on a

$$\left|a_n e^{itx_n}\right| = a_n = \mathbf{P}(X = x_n),$$

et on déduit de la question précédente que la série $\sum a_n e^{itx_n}$ converge absolument.

3. Comme X est une variable aléatoire discrète, il en va de même pour e^{itX} (qui est une fonction de X).

D'après la Formule de transfert, la variable aléatoire e^{itX} est une variable d'espérance finie si, et seulement si, la famille

$$\big(e^{itx_n}\, \mathbf{P}(X = x_n)\big)_{n\in\mathbb{N}}$$

est sommable, c'est-à-dire si la série $\sum a_n e^{itx_n}$ est absolument convergente : c'est bien le cas d'après **2.**

La fonction φ_X est donc bien définie sur $\mathbb{R}$ et

$$\forall\, t \in \mathbb{R}, \quad \varphi_X(t) = \sum_{x\in E} e^{itx}\, \mathbf{P}(X = x) = \sum_{n=0}^{+\infty} a_n e^{itx_n}.$$

❧ Pour tout $n \in \mathbb{N}$, la fonction $\left[t \mapsto a_n e^{itx_n}\right]$ est évidemment continue sur $\mathbb{R}$.

On a vu à la question précédente que

$$\forall\, n \in \mathbb{N},\ \forall\, t \in \mathbb{R}, \quad |a_n e^{itx_n}| \leqslant a_n$$

où a_n est indépendant de t et $\sum a_n$ est une série convergente.

La série *de fonctions* continues $\sum a_n e^{itx_n}$ converge donc normalement sur $\mathbb{R}$, ce qui prouve que sa somme φ_X est continue sur $\mathbb{R}$.

4. Pour tout $n \in \mathbb{N}$ et tout $t \in \mathbb{R}$,

$$a_n e^{i(-t)x_n} = \overline{a_n e^{itx_n}}$$

puisque les a_n et les x_n sont <u>réels</u>.

Comme la conjugaison est $\mathbb{R}$-linéaire,

$$\forall\, N \in \mathbb{N}, \quad \sum_{n=0}^{N} a_n e^{i(-t)x_n} = \overline{\sum_{n=0}^{N} a_n e^{itx_n}}$$

et qu'elle est 1-lipschitzienne (c'est en fait une isométrie),

$$\forall\, t \in \mathbb{R}, \quad \varphi_X(-t) = \overline{\varphi_X(t)}.$$

☙ Si la fonction φ_X est paire, on en déduit que

$$\forall\, t \in \mathbb{R}, \quad \varphi_X(t) = \varphi_X(-t) = \overline{\varphi_X(t)}$$

et donc que φ_X est une fonction *à valeurs réelles*.

Partie B. Image de φ_X

5. Comme on l'a déjà remarqué au **2.** et au **3.**,

$$\forall\, n \in \mathbb{N}, \quad \left|a_n e^{itx_n}\right| \leqslant a_n.$$

Par inégalité triangulaire et **1.**,

$$\left|\varphi_X(t)\right| \leqslant \sum_{n=0}^{+\infty} |a_n e^{itx_n}| = \sum_{n=0}^{+\infty} \mathbf{P}(X = x_n) = 1$$

pour tout $t \in \mathbb{R}$.

6. Pour tout $n \in \mathbb{N}$, il existe $k_n \in \mathbb{Z}$ tel que

$$x_n = a + \frac{2k_n\pi}{t_0}.$$

Par conséquent,

$$e^{it_0x_n} = e^{it_0a} \cdot e^{i2k_n\pi} = e^{it_0a}$$

et donc (encore **1.** !)

$$\varphi_X(t_0) = e^{it_0a} \sum_{n=0}^{+\infty} \mathbf{P}(X = x_n) = e^{it_0a},$$

ce qui prouve que $\left|\varphi_X(t_0)\right| = 1$.

7. a. Comme $\left|\varphi_X(t_0)\right| = 1$, alors il existe un réel θ tel que

$$\sum_{n=0}^{+\infty} \mathbf{P}(X = x_n) \cdot e^{it_0x_n} = \varphi_X(t_0) = e^{i\theta}.$$

Or $t_0 \neq 0$, donc il existe $a = \theta/t_0 \in \mathbb{R}$ tel que $e^{i\theta} = e^{it_0a}$ et donc tel que

$$\sum_{n=0}^{+\infty} \mathbf{P}(X = x_n) \cdot e^{it_0x_n} e^{-it_0a} = 1.$$

7. b. D'après **1.** et l'égalité précédente,

$$\sum_{n=0}^{+\infty} a_n \exp\left[i(t_0x_n - t_0a)\right] = \sum_{n=0}^{+\infty} a_n$$

et donc, par différence :

$$\sum_{n=0}^{+\infty} a_n\left(1 - e^{it_0(x_n-a)}\right) = 0.$$

L'application $\mathfrak{Re}$ est $\mathbb{R}$-linéaire et lipschitzienne et les a_n sont tous *réels*, donc

$$\begin{aligned} 0 = \mathfrak{Re}(0) &= \sum_{n=0}^{+\infty} a_n\, \mathfrak{Re}\left(1 - e^{it_0(x_n-a)}\right) \\ &= \sum_{n=0}^{+\infty} a_n\left(1 - \cos[t_0(x_n - a)]\right). \end{aligned}$$

REMARQUE.— Ici comme au **5.**, la linéarité seule ne permet pas de conclure : on ne peut invoquer la linéarité que sur une somme d'un nombre FINI de termes. Pour justifier

$$\varphi\left(\sum_{n=0}^{+\infty} u_n\right) = \sum_{n=0}^{+\infty} \varphi(u_n),$$

il faut s'assurer que l'application linéaire φ est continue (c'est-à-dire lipschitzienne).

7. c. On reconnaît ici une somme de *réels positifs* et cette somme est nulle, donc tous les termes sont nuls.

Or un produit de réels est nul si, et seulement si, l'un des deux facteurs est nul. Par conséquent, pour tout $n \in \mathbb{N}$,

- ou bien $a_n = 0$,
- ou bien $\cos[t_0(x_n - a)] = 1$ et ce cas équivaut au fait que $t_0(x_n - a)$ soit un multiple entier de 2π, c'est-à-dire

$$\frac{t_0(x_n - a)}{2\pi} \in \mathbb{Z}.$$

On conclut en rappelant que $t_0 \neq 0$ par hypothèse.

7. d. On utilise le système complet d'événements associé à X pour décomposer cet événement :

$$\left[X \notin a + \frac{2\pi}{t_0}\mathbb{Z}\right] = \bigsqcup_{\substack{n \,\in\, \mathbb{N} \text{ tels que} \\ x_n \notin a + \frac{2\pi}{t_0}\mathbb{Z}}} [X = x_n].$$

Par σ-additivité de $\mathbf{P}$ (on a une famille finie ou dénombrable d'événements deux à deux disjoints), on en déduit que

$$\mathbf{P}\left(X \notin a + \frac{2\pi}{t_0}\mathbb{Z}\right) = \sum_{\substack{n \,\in\, \mathbb{N} \text{ tels que} \\ x_n \notin a + \frac{2\pi}{t_0}\mathbb{Z}}} a_n = 0$$

(puisque, d'après l'alternative précédente, tous les a_n qui figurent dans cette somme sont nuls).

Par passage au complémentaire, on en déduit que

$$\mathbf{P}\left(X \in a + \frac{2\pi}{t_0}\mathbb{Z}\right) = 1.$$

Partie C. Fonction caractéristique et loi de X

8. On intègre sur le *segment* $[-T, T]$ l'expression

$$\varphi_X(t)e^{-imt} = \sum_{n=0}^{+\infty} a_n e^{i(x_n-m)t}.$$

Comme au **3.**, il s'agit de la somme d'une série de fonctions continues sur $\mathbb{R}$ qui converge normalement sur $\mathbb{R}$. On peut donc intégrer terme à terme sur le segment $[-T, T]$.

Pour calculer chaque terme, il faut distinguer deux cas :

– ou bien $x_n = m$ et dans ce cas,

$$\frac{1}{2T}\int_{-T}^{T} e^{i(x_n-m)t}\,dt = \frac{1}{2T}\int_{-T}^{T} 1\,dt = 1$$

– ou bien $x_n \neq m$ et dans ce cas,

$$\frac{1}{2T}\int_{-T}^{T} e^{i(x_n-m)t}\,dt = \frac{1}{2T}\cdot\left[\frac{e^{i(x_n-m)t}}{i(x_n-m)}\right]_{-T}^{T}.$$

Dans les deux cas,

$$\frac{1}{2T}\int_{-T}^{T} e^{i(x_n-m)t}\,dt = \operatorname{sinc}\big[(x_n-m)T\big]$$

et par conséquent, pour tout $T > 0$,

$$\begin{aligned} V_m(T) &= \sum_{n=0}^{+\infty} a_n \frac{1}{2T}\int_{-T}^{T} e^{i(x_n-m)t}\,dt \\ &= \sum_{n=0}^{+\infty} a_n \operatorname{sinc}\big[(x_n-m)T\big] = \sum_{n=0}^{+\infty} g_n(T). \end{aligned}$$

9. Comme sinc est continue sur $\mathbb{R}$, chaque fonction g_n est continue sur $\mathbb{R}$ (par composition).

De plus,

$$\forall\, n \in \mathbb{N},\ \forall\, T > 0,\quad \big|a_n \operatorname{sinc}\big[(x_n-m)T\big]\big| \leqslant a_n$$

et la série $\sum a_n$ converge (propriété déjà utilisée en **2.**, **3.** et **5.**...), donc la série de fonctions continues $\sum g_n$ converge normalement sur $\mathbb{R}_+^*$ et sa somme V_m est par conséquent continue sur $\mathbb{R}_+^*$.

10. On reprend la discussion menée au **8.**

– Si $m = x_n$, alors $g_n(T) = a_n$ pour tout $T > 0$, donc

$$\lim_{T\to+\infty} g_n(T) = a_n.$$

– Si $m \neq x_n$, alors

$$\lim_{T\to+\infty} g_n(T) = 0$$

par composition de limites (le produit $(x_n - m)T$ tend vers $\pm\infty$ et la fonction sinc tend vers 0 au voisinage de $\pm\infty$).

11. Chaque fonction g_n tend vers une limite finie au voisinage de $+\infty$; la série de fonctions g_n converge normalement (et donc uniformément) sur un voisinage de $+\infty$ (sur $\mathbb{R}_+^*$ en fait).

D'après le Théorème de la double limite, la somme V_m de cette série de fonctions tend vers une limite finie au voisinage de $+\infty$ et

$$\lim_{T\to+\infty} V_m(T) = \sum_{n=0}^{+\infty} \lim_{T\to+\infty} g_n(T).$$

D'après la question précédente, la seule limite qui n'est pas nécessairement nulle au second membre est égale à $a_n = \mathbf{P}(X = x_n)$ et cela, seulement si $m = x_n$. Donc

$$\forall\, m \in \mathbb{R},\qquad \sum_{n=0}^{+\infty} \lim_{T\to+\infty} g_n(T) = \mathbf{P}(X = m).$$

12. Si les fonctions caractéristiques φ_X et φ_Y sont égales, alors pour tout $m \in \mathbb{R}$, les fonctions V_m calculées avec φ_X ou avec φ_Y sont les mêmes et en particulier, elles ont même limite au voisinage de $+\infty$.

On déduit de la question précédente que

$$\forall\, m \in \mathbb{R},\quad \mathbf{P}(X = m) = \mathbf{P}(Y = m)$$

c'est-à-dire que les variables aléatoires X et Y suivent la même loi.

Solution IV ❁ Série génératrice

1. On procède par récurrence.

❧ Pour $k = 1$, l'équation devient

$$\mathbf{P}(X = 1) = \lambda_1\,\mathbf{P}(X = 0)$$

et comme $\mathbf{P}(X = 0) > 0$ par hypothèse, on en déduit que la seule solution est

$$\lambda_1 = \frac{\mathbf{P}(X = 1)}{\mathbf{P}(X = 0)}.$$

❧ HR : Supposons que, pour un rang $n \geqslant 1$, on ait démontré qu'il existe une unique famille $(\lambda_i)_{1\leqslant i\leqslant n}$ de réels telle que

$$\forall\, 1 \leqslant k \leqslant n,\quad k\,\mathbf{P}(X = k) = \sum_{j=1}^{k} j\lambda_j\,\mathbf{P}(X = k - j).$$

Alors au rang $(n + 1)$, l'équation est

$$\begin{aligned} (n+1)\,\mathbf{P}(X = n+1) = \sum_{j=1}^{n} &\underbrace{j\lambda_j\,\mathbf{P}(X = k - j)}_{\text{connu}} \\ &+ (n+1)\lambda_{n+1}\,\mathbf{P}(X = 0). \end{aligned}$$

Comme $\mathbf{P}(X = 0) > 0$, il s'agit d'une équation du premier degré en λ_{n+1} (seule inconnue de cette équation !) et elle admet par conséquent une unique solution.

❧ L'hypothèse de récurrence est donc établie pour tout $n \geqslant 1$ et par conséquent il existe une, et une seule, famille réelle $(\lambda_i)_{i\geqslant 1}$ telle que

$$\forall\, k \in \mathbb{N}^*,\quad k\,\mathbf{P}(X = k) = \sum_{j=1}^{k} j\lambda_j\,\mathbf{P}(X = k - j).$$

2. Isolons le terme en $j = k$:

$$k\,\mathbf{P}(X = k) = k\lambda_k\,\mathbf{P}(X = 0) + \sum_{j=1}^{k-1} j\lambda_j\,\mathbf{P}(X = k - j)$$

puis divisons par $k \geqslant 1$:

$$\lambda_k \, \mathbf{P}(X=0) = \mathbf{P}(X=k) - \sum_{j=1}^{k-1} \frac{j}{k} \lambda_j \, \mathbf{P}(X=k-j)$$

et prenons la valeur absolue avant d'appliquer l'inégalité triangulaire :

$$\begin{aligned} |\lambda_k| \, \mathbf{P}(X=0) &\leqslant \underbrace{\mathbf{P}(X=k)}_{\geqslant 0} + \sum_{j=1}^{k-1} \underbrace{\frac{j}{k}}_{\in[0,1]} \underbrace{|\lambda_j| \, \mathbf{P}(X=k-j)}_{\geqslant 0} \\ &\leqslant \mathbf{P}(X=k) + \sum_{j=1}^{k-1} |\lambda_j| \, \mathbf{P}(X=k-j) \end{aligned}$$

❧ Pour $0 \leqslant j < k$, on a $k-j \geqslant 1$ et par conséquent

$$[X = k-j] \subset [X=0]^c.$$

Par croissance de la mesure $\mathbf{P}$, on en déduit que

$$\mathbf{P}(X=k-j) \leqslant \mathbf{P}([X=0]^c) = 1 - \mathbf{P}(X=0).$$

Comme les $|\lambda_j|$ sont positifs, on en déduit que

$$\begin{aligned} |\lambda_k| \, \mathbf{P}(X=0) &\leqslant \mathbf{P}(X=k) + \big[1 - \mathbf{P}(X=0)\big] \left(\sum_{j=1}^{k-1} |\lambda_j| \right) \\ &\leqslant \big[1 - \mathbf{P}(X=0)\big] \left(1 + \sum_{j=1}^{k-1} |\lambda_j| \right) \end{aligned}$$

pour tout $k \geqslant 1$.

3. On procède encore par récurrence.

❧ Pour $k = 1$, la relation précédente se réduit à

$$|\lambda_1| \; \mathbf{P}(X=0) \leqslant 1 - \mathbf{P}(X=0)$$

donc

$$1 + |\lambda_1| \leqslant 1 + \frac{1}{\mathbf{P}(X=0)} - 1 = \frac{1}{\mathbf{P}(X=0)}.$$

❧ HR : on suppose qu'il existe $k \in \mathbb{N}^*$ tel que

$$1 + \sum_{j=1}^{k} |\lambda_j| \leqslant \frac{1}{\mathbf{P}(X=0)^k}.$$

D'après **2.** et HR,

$$\begin{aligned} |\lambda_{k+1}| &\leqslant \frac{1 - \mathbf{P}(X=0)}{\mathbf{P}(X=0)} \cdot \frac{1}{\mathbf{P}(X=0)^k} \\ &\leqslant \frac{1}{\mathbf{P}(X=0)^{k+1}} - \frac{1}{\mathbf{P}(X=0)^k} \end{aligned}$$

et on déduit alors de HR que

$$\begin{aligned} 1 + \sum_{j=1}^{k+1} |\lambda_j| &= \left(1 + \sum_{j=1}^{k} |\lambda_j| \right) + |\lambda_{k+1}| \\ &\leqslant \frac{1}{\mathbf{P}(X=0)^k} + \frac{1}{\mathbf{P}(X=0)^{k+1}} - \frac{1}{\mathbf{P}(X=0)^k} \\ &\leqslant \frac{1}{\mathbf{P}(X=0)^{k+1}} \end{aligned}$$

et le résultat est démontré par récurrence.

4. Par **2.** et **3.**,

$$|\lambda_k t^k| \underset{k \to +\infty}{=} \mathcal{O}\left(\frac{t^k}{\mathbf{P}(X=0)^k} \right)$$

et pour $|t| < \mathbf{P}(X=0)$, la série géométrique de raison $\frac{t}{\mathbf{P}(X=0)}$ est absolument convergente.

Par comparaison, la série $\sum \lambda_k t^k$ est absolument convergente au moins pour

$$|t| < \mathbf{P}(X=0),$$

ce qui prouve que le rayon de convergence de la série entière $\sum \lambda_k t^k$ est au moins égal à $\mathbf{P}(X=0)$ et en particulier qu'il est strictement positif.

5. a. Le rayon de convergence de la série génératrice de X est au moins égal à 1 ; le rayon de convergence de la série $\sum \lambda^k t^k$ est au moins égal à $\mathbf{P}(X=0) \leqslant 1$ d'après la question précédente. Par conséquent, l'intervalle

$$I_0 = \left]-\rho(X), \rho(X)\right[$$

est contenu dans l'intervalle ouvert de convergence de ces deux séries entières.

❧ Comme les deux rayons de convergence sont *strictement* positifs, les sommes G_X et H_X sont de classe $\mathscr{C}^\infty$ sur I_0 et on peut dériver terme à terme pour exprimer les dérivées.

D'une part,

$$\forall\, t \in I_0, \quad G'_X(t) = \sum_{k=0}^{+\infty} (k+1) \, \mathbf{P}(X=k+1) t^k$$

et d'autre part,

$$\forall\, t \in I_0, \quad H'_X(t) = \sum_{j=0}^{+\infty} (j+1) \lambda_{j+1} t^j.$$

❧ Le rayon de convergence de la série dérivée est le même que celui de la série entière initiale. Par conséquent, les séries $\sum j\lambda_j t^{j-1}$ et $\sum \mathbf{P}(X=k)t^k$ sont *absolument* convergentes sur l'intervalle ouvert I_0 et on peut appliquer le théorème sur le produit de Cauchy.

Ainsi, pour tout $t \in I_0$ (au moins),

$$H'_X(t) G_X(t) = \sum_{k=0}^{+\infty} w_k(t)$$

où on a posé, pour tout $k \in \mathbb{N}$,

$$\begin{aligned} w_k(t) &= \sum_{j=0}^{k} \big[(j+1)\lambda_{j+1} t^j\big] \cdot \big[\mathbf{P}(X=k-j) t^{k-j}\big] \\ &= \left[\sum_{j=0}^{k} (j+1) \lambda_{j+1} \, \mathbf{P}(X=k-j) \right] t^k \\ &= \left[\sum_{j=1}^{k+1} j\lambda_j \, \mathbf{P}(X=[k+1]-j) \right] t^k \\ &= (k+1) \, \mathbf{P}(X=k+1) \, t^k \end{aligned}$$

par définition des λ_i. On a bien démontré que

$$\forall\, t \in I_0, \quad G'_X(t) = H'_X(t) G_X(t).$$

5. b. On déduit de la question précédente que H'_X est la dérivée logarithmique de G_X et donc que $H_X - \ell n\, G_X$ est constante sur l'intervalle I_0. Or, par définition,

$$H_X(0) = \ell n\, \mathbf{P}(X = 0) = \ell n\, G_X(0)$$

donc

$$\forall\, t \in I_0, \quad H_X(t) = \ell n\, G_X(t).$$

REMARQUE.— Ceux qui n'aiment pas la dérivée logarithmique ont bien tort !

Solution V ❀ Séries de variables aléatoires à valeurs entières

Partie A.

1. Comme X et Y sont des variables aléatoires à valeurs dans $\mathbb{N}$, la famille

$$\big([X = m, Y = n]\big)_{(m,n)\in\mathbb{N}^2}$$

est un système complet d'événements et, quels que soient les entiers m et n,

$$\begin{cases} X(\omega) \neq Y(\omega) \\ X(\omega) = m \\ Y(\omega) = n \end{cases} \iff \begin{cases} m \neq n \\ X(\omega) = m \\ Y(\omega) = n \end{cases}$$

donc

$$[X \neq Y] \cap [X = m, Y = n] = \left| \begin{array}{ll} \varnothing & \text{si } m = n, \\ [X = m, Y = n] & \text{sinon} \end{array} \right.$$

et par conséquent

$$[X \neq Y] = \bigsqcup_{\substack{(m,n)\in\mathbb{N}^2 \\ \text{tels que } m \neq n}} [X = m, Y = n].$$

Comme X et Y sont des variables aléatoires, alors

$$\forall\, (m, n) \in \mathbb{N}^2, \quad [X = m, Y = n] \in \mathcal{A}.$$

D'autre part, $\mathbb{N}^2$ est dénombrable et toute partie d'un ensemble dénombrable est elle-même dénombrable. Donc $[X \neq Y]$ est une union dénombrable d'événements et par conséquent

$$[X \neq Y] \in \mathcal{A}.$$

REMARQUE.— On aurait aussi bien pu raisonner sur le complémentaire et justifier que

$$[X = Y] = \bigsqcup_{n\in\mathbb{N}} [X = n, Y = n] \in \mathcal{A}.$$

La seule différence est une affaire de goût.

2. Une tribu est stable par passage au complémentaire et par intersection (finie ou dénombrable). Par hypothèse, A et B appartiennent à la tribu $\mathcal{A}$. Par conséquent,

$$A \cap B^c \in \mathcal{A} \quad \text{et} \quad A^c \cap B \in \mathcal{A}$$

et comme la mesure de probabilité $\mathbf{P}$ est définie sur $\mathcal{A}$, les deux probabilités $\mathbf{P}(A\cap B^c)$ et $\mathbf{P}(A^c\cap B)$ sont bien définies.

❧ En décomposant A sur le système complet d'événements (B, B^c), on obtient

$$A = (A \cap B) \sqcup (A \cap B^c) \subset B \sqcup (A \cap B^c).$$

Par croissance et additivité de $\mathbf{P}$,

$$\mathbf{P}(A) \leqslant \mathbf{P}\big(B \sqcup (A \cap B^c)\big) = \mathbf{P}(B) + \mathbf{P}(A \cap B^c)$$

et a fortiori

$$\mathbf{P}(A) \leqslant \mathbf{P}(B) + \mathbf{P}(A \cap B^c) + \mathbf{P}(A^c \cap B)$$

puisque $\mathbf{P}(A^c \cap B) \geqslant 0$.

Symétriquement, en décomposant B sur (A, A^c), on obtient

$$\begin{aligned} \mathbf{P}(B) &\leqslant \mathbf{P}(A) + \mathbf{P}(A^c \cap B) \\ &\leqslant \mathbf{P}(A) + \mathbf{P}(A^c \cap B) + \mathbf{P}(A \cap B^c) \end{aligned}$$

et ces deux inégalités nous donnent bien

$$\big|\mathbf{P}(A) - \mathbf{P}(B)\big| \leqslant \mathbf{P}(A \cap B^c) + \mathbf{P}(A^c \cap B).$$

REMARQUE.— On doit toujours se rappeler que la relation $|x| \leqslant y$ n'est pas une inégalité, mais un <u>encadrement</u> :

$$|x| \leqslant y \iff -y \leqslant x \leqslant y.$$

3. On sait que les séries

$$\sum \mathbf{P}(X = n)t^n \quad \text{et} \quad \sum \mathbf{P}(Y = n)t^n$$

convergent absolument pour tout $t \in [-1, 1]$. Par inégalité triangulaire,

$$\begin{aligned} \big|G_X(t) - G_Y(t)\big| &\leqslant \sum_{n=0}^{+\infty} \big|\mathbf{P}(X = n) - \mathbf{P}(Y = n)\big|\, t^n \\ &\leqslant \sum_{n=0}^{+\infty} \big|\mathbf{P}(X = n) - \mathbf{P}(Y = n)\big| \end{aligned}$$

(puisque le terme général pour t quelconque est majoré par le terme pour $t = 1$).

D'après **2.**,

$$\big|\mathbf{P}(X = n) - \mathbf{P}(Y = n)\big| \leqslant \mathbf{P}(X = n, Y \neq n) + \mathbf{P}(X \neq n, Y = n).$$

Or (décomposition sur le système complet d'événements associé à X)

$$[X \neq Y] = \bigsqcup_{n\in\mathbb{N}} [X = n] \cap [X \neq Y] = \bigsqcup_{n\in\mathbb{N}} [X = n, Y \neq n]$$

et de même (par symétrie)

$$[X \neq Y] = \bigsqcup_{n\in\mathbb{N}} [X \neq n, Y = n].$$

On a ici deux unions dénombrables d'événements deux à deux disjoints donc, par σ-additivité de $\mathbf{P}$, les deux séries

$$\sum \mathbf{P}(X = n, Y \neq n) \quad \text{et} \quad \sum \mathbf{P}(X \neq n, Y = n)$$

sont convergentes et

$$\sum_{n=0}^{+\infty} \mathbf{P}(X = n, Y \neq n) = \sum_{n=0}^{+\infty} \mathbf{P}(X \neq n, Y = n) = \mathbf{P}(X \neq Y).$$

Par conséquent,

$$\forall\, t \in [-1, 1], \quad \big|G_X(t) - G_Y(t)\big| \leqslant 2\,\mathbf{P}(X \neq Y).$$

REMARQUE.— On notera que le majorant trouvé est *indépendant* de $t \in [-1, 1]$.

Partie B.

4. Par définition, $\mathbb{N}'$ est l'union d'un ensemble dénombrable ($\mathbb{N}$) et d'un ensemble fini (le singleton $\{+\infty\}$), donc $\mathbb{N}'$ est dénombrable.

On ne sait pas trop ce que désigne $+\infty$ mais en tout cas, ce n'est pas un nombre réel. Donc $\mathbb{N}'$ n'est pas une partie de $\mathbb{R}$.

5. Par définition,

$$\forall\, n \geqslant 1, \quad Z_n = \bigcup_{k \geqslant n} [U_k \neq 0].$$

Comme les U_k sont des variables aléatoires, alors

$$[U_k = 0] \in \mathcal{A},$$

donc

$$[U_k \neq 0] = [U_k = 0]^c \in \mathcal{A}$$

et donc $Z_n \in \mathcal{A}$ en tant qu'union dénombrable d'événements. D'autre part,

$$Z_n = [U_n \neq 0] \cup \bigcup_{k \geqslant n+1} [U_k \neq 0] = [U_n \neq 0] \cup Z_{n+1}$$

et par conséquent, $(Z_n)_{n \geqslant 1}$ est une suite décroissante d'événements :

$$\forall\, n \geqslant 1, \quad Z_{n+1} \subset Z_n.$$

La suite de terme général $\mathbf{P}(Z_n)$ est donc convergente (décroissante et positive) Par sous-σ-additivité (puisque, pour une fois, les événements ne sont pas deux à deux disjoints),

$$\forall\, n \geqslant 1, \quad 0 \leqslant \mathbf{P}(Z_n) \leqslant \sum_{k=n}^{+\infty} \mathbf{P}(U_k \neq 0).$$

Le majorant est, par hypothèse, le reste d'une série convergente, donc il tend vers 0 et par encadrement

$$\lim_{n \to +\infty} \mathbf{P}(Z_n) = 0.$$

6. La valeur de $W(\omega)$ est entière lorsque l'ensemble

$$\big\{k \in \mathbb{N}^* \;:\; U_k(\omega) \neq 0\big\} \subset \mathbb{N}$$

est fini et égale à $+\infty$ dans le cas contraire. Plus précisément, $W(\omega) = n$ si, et seulement si, il existe n entiers

$$1 \leqslant k_1 < k_2 < \cdots < k_n$$

tels que

$$U_{k_1}(\omega) \neq 0, \ldots, U_{k_n}(\omega) \neq 0$$

et tels que

$$\forall\, k \notin \{k_1, \ldots, k_n\}, \quad U_k(\omega) = 0.$$

Par conséquent, $[W = n]$ est égal à

$$\bigsqcup_{1 \leqslant k_1 < k_2 < \cdots < k_n} \left(\bigcap_{i=1}^{n} [U_{k_i} \neq 0] \cap \bigcap_{\substack{k \in \mathbb{N}^* \\ k \notin \{k_1, \ldots, k_n\}}} [U_k = 0] \right)$$

On a déjà justifié que $[U_k = 0] \in \mathcal{A}$ et $[U_k \neq 0] \in \mathcal{A}$ pour tout entier $k \geqslant 1$. La tribu $\mathcal{A}$ est stable par intersection finie ou dénombrable (comme toutes les tribus), donc l'ensemble entre parenthèses est bien un événement. Enfin, l'union disjointe est indexée par une partie de $\mathbb{N}^n$, qui est dénombrable en tant que produit cartésien d'un nombre fini d'ensembles dénombrables, donc

$$\forall\, n \in \mathbb{N}, \quad [W = n] \in \mathcal{A}.$$

❧ Comme W est à valeurs dans $\mathbb{N}'$, on en déduit que

$$[W = +\infty] = \Big(\bigsqcup_{n \in \mathbb{N}} [W = n] \Big)^c \in \mathcal{A}$$

et on a maintenant démontré que

$$\forall\, x \in \mathbb{N}', \qquad [W = x] \in \mathcal{A},$$

c'est-à-dire que W est bien une variable aléatoire sur $(\Omega, \mathcal{A})$ à valeurs dans $\mathbb{N}'$.

Cependant, cette expression ne permet pas de calculer simplement $\mathbf{P}(W = +\infty)$...

❧ Une partie de $\mathbb{N}$ est finie si, et seulement si, elle est majorée, donc

$$W(\omega) \neq +\infty \iff \exists\, n \geqslant 1, \forall\, k \geqslant n, \quad U_k(\omega) = 0$$

c'est-à-dire (en écrivant la négation de ce qui précède)

$$W(\omega) = +\infty \iff \forall\, n \geqslant 1, \exists\, k \geqslant n, \quad U_k(\omega) \neq 0$$

et par conséquent

$$[W = +\infty] = \bigcap_{n \geqslant 1} Z_n \in \mathcal{A}.$$

D'après **5.** et la continuité décroissante de $\mathbf{P}$ (puisque la suite des événements Z_n est décroissante pour l'inclusion),

$$\mathbf{P}\Big(\bigcap_{n \geqslant 1} Z_n \Big) = \lim_{n \to +\infty} \mathbf{P}(Z_n) = 0$$

donc l'événement $[W = +\infty]$ est bien négligeable.

7. a. En tant que somme d'un nombre fini de variables aléatoires à valeurs dans $\mathbb{N}$, la fonction S_n est bien une variable aléatoire à valeurs dans $\mathbb{N}$.

7. b. Pour tout $\omega \in \Omega$, la série $\sum U_k(\omega)$ est une série d'*entiers*.

Pour qu'une série converge, il faut que son terme général tende vers 0. Comme le terme général est un *entier*, il faut donc que le terme général soit nul à partir d'un certain rang.

Réciproquement, si le terme général est nul à partir d'un certain rang, alors la série $\sum U_k(\omega)$ est évidemment convergente.

❧ On a donc

$$[S \in \mathbb{N}] = [W \in \mathbb{N}] = \bigcup_{n\geqslant 1} Z_n^c$$

et, par passage au complémentaire,

$$[S = +\infty] = [W = +\infty] \in \mathcal{A}.$$

On déduit de **6.** que $\mathbf{P}(S = +\infty) = 0$ et donc que

$$\mathbf{P}(S \in \mathbb{N}) = 1.$$

❧ Pour tout $s \in \mathbb{N}$, on a donc

$$[S = s] = [S = s] \cap [S \in \mathbb{N}] = \bigcup_{n\geqslant 1} [S = s] \cap Z_n^c.$$

Or, pour tout $n \geqslant 2$,

$$\begin{cases} S(\omega) = s \\ \omega \notin Z_n \end{cases} \iff \begin{cases} S(\omega) = s \\ \forall\, k \geqslant n, \quad U_k(\omega) = 0 \end{cases} \iff \begin{cases} S(\omega) = s \\ \forall\, k \geqslant n, \quad U_k(\omega) = 0 \\ S(\omega) = S_{n-1}(\omega) \end{cases}$$

(puisque le terme général de la série est positif ou nul). On en déduit que

$$[S = s] = \big([S = s] \cap Z_1^c\big) \cup \left(\bigcup_{n\geqslant 2} [S_{n-1} = s] \cap Z_n^c\right).$$

Si $\omega \notin Z_1$, alors $U_k(\omega) = 0$ pour tout $k \geqslant 1$ et par conséquent $S(\omega) = 0$.

On a ainsi démontré que

$$[S = 0] = Z_1^c \cup \left(\bigcup_{n\geqslant 2} [S_{n-1} = 0] \cap Z_n^c\right)$$

et que, pour tout $s \in \mathbb{N}^*$,

$$[S = s] = \bigcup_{n\geqslant 2} [S_{n-1} = s] \cap Z_n^c.$$

❧ On sait par **5.** que les Z_n sont des événements ; on a justifié plus haut que $[S_m = s]$ est un événement pour tout $m \in \mathbb{N}^*$ et tout $s \in \mathbb{N}$; la tribu $\mathcal{A}$ est stable par passage au complémentaire, par intersection et par union dénombrable (c'est son métier de tribu qui veut ça...), donc

$$\forall\, s \in \mathbb{N}, \quad [S = s] \in \mathcal{A}.$$

❧ Finalement,

$$\forall\, x \in \mathbb{N}', \qquad [S = x] \in \mathcal{A}$$

et S est bien une variable aléatoire discrète définie sur $(\Omega, \mathcal{A})$ à valeurs dans $\mathbb{N}'$.

7. c. Bien que S ne soit pas une variable aléatoire à valeurs dans $\mathbb{N}$, on a quand même

$$\sum_{i=0}^{+\infty} \mathbf{P}(S = i) = \mathbf{P}(S \in \mathbb{N}) = 1$$

et on peut donc appliquer les résultats du cours sur les séries génératrices : la fonction G est continue sur le segment $[0, 1]$ (comme somme d'une série entière qui converge normalement sur cet intervalle) et de classe $\mathscr{C}^\infty$ sur $[0, +\infty[$ (comme somme d'une série entière dont le rayon de convergence est au moins égal à 1).

7. d. Nous allons bien sûr appliquer l'estimation établie en **3.** :

$$\forall\, t \in [0, 1], \quad |G(t) - G_n(t)| \leqslant 2\,\mathbf{P}(S \neq S_n).$$

Or

$$S(\omega) = S_n(\omega) + \sum_{k=n+1}^{+\infty} \underbrace{U_k(\omega)}_{\geqslant 0}$$

donc

$$S(\omega) \neq S_n(\omega) \iff \exists\, k \geqslant n + 1, \quad U_k(\omega) \neq 0$$

c'est-à-dire

$$[S \neq S_n] = \bigcup_{k\geqslant n+1} [U_k \neq 0] = Z_{n+1}.$$

Comme $\mathbf{P}(Z_{n+1})$ tend vers 0 d'après **5.**, on a bien trouvé un majorant de

$$\big|G(t) - G_n(t)\big|$$

qui est indépendant de $t \in [0, 1]$ et qui tend vers 0 : la suite $(G_n)_{n\geqslant 1}$ converge donc uniformément sur $[0, 1]$ vers G.