Centrale 2017 - MP2
Un corrigé

Le sujet est très long et il est parfois difficile de savoir ce qui correspond au cours et ce qui doit être démontré. On a ainsi implicitement utilisé à plusieurs reprises le fait que si X est une variable positive d'espérance nulle alors X est presque surement constante. Ceci permet, entre autres, de dire que si $\mathbb{V}(X) = 0$ alors X est presque surement constante.
De façon similaire, certain résultats proches du cours ont parfois été admis (par exemple, l'expression de la fonction génératrice pour les lois usuelles) ou justifiés un peu lestement (par exemple la recurrence permettant de calculer la fonction génératrice d'une somme de fonctions indépendantes).

1 Variables aléatoires entières décomposables

1.1 Premiers exemples

1. Si $X \sim X'$ alors (on sait dans cette question que $X(\Omega) = X'(\Omega') \subset \mathbb{N}$) $\forall k \in \mathbb{N}$, $\mathbb{P}(X = x) = \mathbb{P}'(X' = k)$ et ainsi $G_X = G_{X'}$.
Réciproquement, si $G_X = G_{X'}$ alors, on a deux séries entières de rayon de convergence au moins égal à 1 dont les sommes sont égales. Par unicité des DSE, $\forall k \in \mathbb{N}$, $\mathbb{P}(X = k) = \mathbb{P}'(X' = k)$. Ainsi, $X \sim X'$ (puisque les variables sont supposées à valeurs dans $\mathbb{N}$).
$$\boxed{X \sim X' \iff G_X = G_{X'}}$$
2. On a, quand cela existe, $G_X(t) = \mathbb{E}(t^X)$ et comme $X \sim Y + Z$, $\mathbb{E}(t^X) = \mathbb{E}(t^{Y+Z}) = \mathbb{E}(t^Y t^Z)$. Comme Y et Z sont indépendantes, il en est de même de t^Z et t^Z et $\mathbb{E}(t^Y t^Z) = \mathbb{E}(t^Y)\mathbb{E}(t^Z)$. Finalement,
$$\boxed{X \sim Y + Z \Rightarrow G_X = G_Y G_Z}$$
3. Le cours indique que $G_x(t) = (pt + (1-p))^n$.
Supposons que $X \subset Y + Z$ avec Y, Z à valeurs dans $\mathbb{N}$. On a $(Y + Z)(\Omega') = X(\Omega) = [|0, n|]$. Comme Y, Z sont à valeurs dans $\mathbb{N}$, on en déduit que $Y(\Omega') \subset [|0, n|]$ et $Z(\Omega') \subset [|0, n|]$. L'égalité $G_X = G_Y G_Z$ est une identité entre polynômes. Comme Y, Z ne sont pas presque sûrement constantes, $\deg(G_Y)$ et $\deg(G_Z)$ sont ≥ 1. Ainsi $\deg(G_X) = n \geq 2$.
Réciproquement, supposons $n \geq 2$. On peut trouver des variables $X_1, \dots, X_n$ définies sur un même espace probabilisé, indépendantes et de loi de Bernoulli $\mathcal{B}(p)$. $S = X_1 + \cdots + X_n$ vérifie $G_S = G_{X_1} \dots G_{X_n} = G_X$ et donc $X \sim S$. En posant $Y = X_1 + \cdots + X_{n-1}$ et $Z = X_n$, on a $X \sim Y + Z$ avec Y, Z indépendantes et non presque sûrement constante. X est donc décomposable.
$$\boxed{X \hookrightarrow \mathcal{B}(n, p) \text{ est décomposable ssi } n \geq 2}$$
4. (a) Ue étude de fonctions montre que A est décroissante sur $]-\infty, -1/\sqrt[3]{2}]$ puis croissante sur $[-1/\sqrt[3]{2}, +\infty[$. Comme $A(-1) = 0$ et $A(0) > 0$, A possède donc exactement deux racines réelles qui sont -1 et $t_0 \in]-1/\sqrt[3]{2}, 0[$. On a donc
$$A(T) = (T+1)(T^3 - T^2 + T + 1) = (T+1)(T - t_0)(T^2 + (t_0 - 1)T - \frac{1}{t_0})$$
Notons que le terme de degré 2 n'admet pas de racine réelle et est irreductible dans $\mathbb{R}$.
Supposons que $A(T) = U(T)V(T)$ avec U et V à coefficients réels positifs et unitaires (hypothèse non réductrice puisque l'on peut multiplier le polynôme U par un scalaire pour le rendre unitaire et V par l'inverse de ce scalaire, V devenant alors unitaire). Si U et V ne sont pas constants et si $\deg(U) \leq \deg(V)$ (ils jouent des rôles symétriques) on a deux possibilités :

- $\deg(U) = 1$ et alors $U = (T+1)$ et $V = (T^3 - T^2 + T + 1)$ ou $U = (T - t_0)$ et $V = T^3 + t_0 T^2 + (t_0 - 1 - \frac{1}{t_0})T - \frac{1}{t_0}$. Dans les deux cas, V a un coefficient négatif.
- $\deg(U) = \deg(V) = 2$ et les deux facteurs sont $(T+1)(T-t_0)$ et $(T^2 + (t_0 - 1)T - \frac{1}{t_0})$ et comme $t_0 - 1 < 0$, on a encore un coefficient négatif.

$$\boxed{A + UV \Rightarrow U \text{ ou } V \text{ est constant}}$$

(b) Comme $1/4 + 1/2 + 1/4 = 1$, il existe une variable aléatoire X telle que $X(\Omega) = \{0, 1, 2\}$ et $\mathbb{P}(X = 0) = 1/4$, $\mathbb{P}(X = 1) = 1/2$ et $\mathbb{P}(X = 2) = 1/4$. On a $G_X(t) = \frac{1}{4}(1+t)^2$. On a en fait $X \hookrightarrow \mathcal{B}(2, 1/2)$ et on a vu en question 3 que X est décomposable.
$C = X^2$ est telle que $C(\Omega) = \{0, 1, 4\}$ et $G_C(t) = \frac{1}{4}A(t)$. Si C était décomposable, alors comme en question 3 on pourrait écrire G_C comme produit de deux polynômes non constants à coefficients positifs (puisque ces polynômes seraient des fonctions génératrices) et on vient de voir que c'est impossible.

$$\boxed{\text{On peut avoir } X \text{ décomposable et } X^2 \text{ non décomposable.}}$$

1.2 Variables uniformes

1. (a) Supposons, par conditions nécessaires, disposer de Q et R. Pour tout $\omega \in \Omega$, on a

$$X(\omega) = aQ(\omega) + R(\omega)$$

Comme $Q(\omega)$ et $R(\omega)$ sont des entiers avec $0 \leq R(\omega) \leq a - 1$, $Q(\omega)$ et $R(\omega)$ sont nécessairement le quotient et le reste dans la division euclidienne de $X(\omega)$ par a.
Réciproquement, notons Q et R les applications définies sur Ω comme ci-dessus. On a immédiatement $X = aQ + R$. Il reste à voir si Q et R sont des variables alétoires. Or,

$$R(\Omega) \subset [|0, a-1|] \text{ et } \forall k \in [|0, a-1|],\ (R = k) = \bigcup_{q \in \mathbb{N}} (X = k + qa)$$

$$Q(\Omega) \subset \mathbb{N} \text{ et } \forall q \in \mathbb{N},\ (Q = q) = \bigcup_{k=0}^{a-1} (X = qa + k)$$

ce qui montre que $(R = k)$ et $(Q = q)$ sont des éléments de $\mathcal{A}$ (stable par réunion dénombrable ou fini) et donc que Q, R sont des variables alétaoires.

$$\boxed{\exists!(Q, R), \text{ variable entières sur } \Omega, \text{ telles que } X = aQ + R \text{ et } R(\Omega) \subset [|0, a-1|]}$$

(b) On a $(Q, R)(\Omega) \subset \mathbb{N} \times [|0, a-1|]$ et

$$\forall (q, k) \in \mathbb{N} \times [|0, a-1|],\ \mathbb{P}((Q, R) = (q, k)) = \mathbb{P}(X = aq + k) = \begin{cases} \frac{1}{n} & \text{si } aq + k \leq n - 1 \\ 0 & \text{sinon} \end{cases}$$

En reprenant les égalités de la question précédente et comme les réunions mises en jeu sont disjointes,

$$\forall k \in [|0, a-1|],\ \mathbb{P}(R = k) = \sum_{q \in \mathbb{N}} \mathbb{P}(X = k + qa)$$

$$\forall q \in \mathbb{N},\ \mathbb{P}(Q = q) = \sum_{k=0}^{a-1} \mathbb{P}(X = k + qa)$$

Il s'agit de compter à k (resp q) fixé combien il y a de q (resp k) tels que $0 \leq qa + k \leq n - 1 = ab - 1$.

A k fixé, les q convenables sont $0, 1, \ldots, b-1$ et il y en a b.
A q fixé dans $[|0, b-1|]$, il y a a valeurs pour k et si $q \geq b$, il n'y en a aucune.

$$\forall k \in [|0, a-1|],\ \mathbb{P}(R = k) = \frac{b}{n} = \frac{1}{a}$$

$$\forall q \in [|0, b-1|],\ \mathbb{P}(Q = q) = \frac{a}{n} = \frac{1}{b}$$

$$\boxed{R \hookrightarrow \mathcal{U}([|0, a-1|]) \text{ et } Q \hookrightarrow \mathcal{U}([|0, b-1|])}$$

(c) On vient de voir que $X = aQ + R$ et comme $a, b \geq 2$, aQ et R ne sont pas des variables presque sûrement constantes. Ces variables sont indépendantes. En effet, pour $0 \leq q \leq b-1$ et $0 \leq k \leq a-1$,

$$\mathbb{P}(Q = q \cap R = k) = \mathbb{P}((Q, R) = (q, k)) = \frac{1}{n} = \mathbb{P}(Q = q)\mathbb{P}(R = k)$$

X est donc décomposable et $G_X = G_{aQ+R} = G_{aQ}G_R$ et ainsi

$$\boxed{G_X(t) = \frac{1}{n}\left(\sum_{k=0}^{a-1} t^k\right)\left(\sum_{q=0}^{b-1} t^{aq}\right)}$$

2. (a) Supposons que X soit décomposable et qu'on ait donc $X \sim Y + Z$. On a alors $G_X = G_Y G_Z$ et, comme en question **1.1.3**, G_Y et G_Z sont des polynômes non constants (et à coefiicients dans $\mathbb{R}^+$). Or,

$$G_X(t) = \frac{1}{n}\sum_{k=0}^{n-1} t^k$$

Si X est décomposable, il existe donc une décomposition de $1+T+\cdots+T^{n-1}$ comme produit de deux polynômes non constants à coefficients dans $\mathbb{R}^+$. En contraposant, on obtient le résultat demandé (se ramener au cas de polynôme unitaire ne pose pas de problème car quitte à multiplier l'un par un scalaire positif - et l'autre par l'inverse - on peut en supposer un unitaire et l'autre l'est alors aussi).

(b) Les racines de UV sont les racines n-ièmes de 1 différentes de 1. Comme n est premier, il est égal à 2 ou impair. Si $n = 2$ le cas est simples ($U(T) = 1 + T$ et $V(T) = 1$ ou l'inverse). Si n est impair et -1 n'est pas racine n-ième de 1. Les racines se regoupent donc par paires de conjuguées qui sont aussi des inverses ($e^{-i\theta} = \frac{1}{e^{i\theta}}$).
Comme U est à coefficients réels, deux racines conjuguées se trouvent groupées dans U et U est de la forme

$$U(T) = \prod_{j=1}^{k}(T - e^{i\theta_k})(T - e^{-i\theta_k})$$

On vérifie alors que

$$U(T) = T^{2k}\prod_{j=1}^{k}(1 - e^{i\theta_k}/T)(1 - e^{-i\theta_k}/T) = T^{2k}\prod_{j=1}^{k}(e^{-i\theta_k} - 1/T)(e^{i\theta_k} - 1/T) = T^{2k}U(1/T)$$

On procède de même pour V.

$$\boxed{U(T) = T^r U\left(\frac{1}{T}\right) \text{ et } V(T) = T^s V\left(\frac{1}{T}\right)}$$

(c) Comme $U(T) = T^r U(\frac{1}{T})$, on a l'égalité polynomiale

$$u_0 + u_1 T + \cdots + u_{r-1}T^{r-1} + T^r = u_0 T^r + u_1 T^{r-1} + \cdots + u_{r-1}T + 1$$

Ceci montre que $u_0 = 1$ (ce que l'énoncé donne sans le justifier) et plus généralement $u_{r-i} = u_i$.
En regardant le coefficient de T^r dans le polynôme $UV = 1 + \cdots + T^{n-1}$ on obtient

$$\sum_{i=0}^{r} u_i v_{r-i} = 1$$

Comme $u_r = v_0 = 1$, on obtient

$$\sum_{i=0}^{r-1} u_i v_{r-i} = 0$$

Les u_k, v_k étant positifs, les termes dans la somme sont positifs et comme la somme est nulle, ils sont tous nuls. Ainsi

$$\forall i \in [|0, r-1|],\ u_i v_{r-i} = 0$$

Mais on a vu que $u_i = u_{r-i}$ et comme $r - i$ varie dans $[|1, r|]$ quand i varie dans $[|0, r-1|]$, on conclut que

$$\boxed{\forall k \in [|1, r|],\ u_k v_k = 0}$$

(d) Montrons par récurrence le résultat demandé.
- En regardant le coefficient de degré 1 de UV on obtient $u_1 + v_1 = 1$. Comme u_1 ou v_1 est nul, on a donc u_1 et v_1 qui sont dans $\{0, 1\}$.
- Supposons le résultat vrai jusqu'à un rang $k - 1 \leq r - 1$. En regardant le coefficient de degré k de UV on obtient

$$v_k + \sum_{i=1}^{k-1} u_i v_{k-i} + u_k = 1$$

Chaque élément de la somme vaut 0 ou 1 par hypothèse de récurrence.
S'ils sont tous nuls, $u_k + v_k = 1$. Comme u_k ou v_k est nul, on a donc l'un des deux qui est null et l'autre qui vaut 1.
Sinon, la somme vaut au moins 1 et $u_k + v_k \leq 0$ ce qui entraîne $u_k = v_k = 0$ (car $u_k, v_k \geq 0$).
Dans les deux cas, on a u_k et v_k dans $\{0, 1\}$.

$$\boxed{\forall k \in [|1, r|],\ u_k \in \{0, 1\} \text{ et } v_k \in \{0, 1\}}$$

(e) Regardons le coefficient de T^{r+1} dans UV. On obtient

$$1 = v_1 + v_2 u_{r-1} + \cdots + v_r u_1 + v_{r+1}$$

Les termes $v_1, v_2 u_{r-1}, \ldots, v_r u_1$ valent tous 0 ou 1. S'ils sont tous nuls, $v_{r+1} = 1$. Sinon $v_{r+1} \leq 0$ et comme c'est un terme positif, $v_{r+1} = 1$. Ainsi $v_{r+1} \in \{0, 1\}$. On peut poursuivre et montrer par récurrence comme en question précédente que

$$v_{r+1}, \ldots, v_{s-1} \in \{0, 1\}$$

$$\boxed{\text{Si } X \hookrightarrow \mathcal{U}([|0, n-1|]) \text{ et } n \text{ premier alors } X \text{ n'est pas décomposable}}$$

2 Variables infiniment divisibles : exemples

2.1 Variables bornées

1. On suppose X constante égale à a. Pour tout $m \in \mathbb{N}^*$, on peut trouver m variables mutuellement indépendantes $X_{m,i}$ et toutes constantes à $\frac{a}{m}$. La variable $X_{m,1} + \cdots + X_{m,m}$ est alors constante égale à a et a donc même loi que X.

$$\boxed{\text{Une variable constante est infiniment divisible}}$$

2. (a) On a évidemment

$$\bigcap_{i=1}^{n} \left(X_i > \frac{M}{n} \right) \subset (X_1 + \cdots + X_n > M)$$

En passant aux probabilités, et par indépendance des X_i,

$$\prod_{i=1}^{n} \mathbb{P}' \left(X_i > \frac{M}{n} \right) \leq \mathbb{P}'(X_1 + \cdots + X_n > M)$$

Comme $X \sim X_1 + \cdots + X_n$, le membre de droite vaut $\mathbb{P}(X > M)$ et est donc nul. Comme le membre de gauche est positif, il est en fait nul et

$$\exists\, i / \; \mathbb{P}' \left(X_i > \frac{M}{n} \right) = 0$$

Comme les X_i, on même loi, on en déduit que

$$\forall i, \; \mathbb{P}' \left(X_i > \frac{M}{n} \right) = 0$$

En passant à l'événement contraire,

$$\forall i, \; \mathbb{P}' \left(X_i \leq \frac{M}{n} \right) = 1$$

On prouve de même que

$$\forall i, \; \mathbb{P}' \left(X_i < \frac{-M}{n} \right) = 0$$

Ainsi, $\mathbb{P}'(|X_i| > M/n) = \mathbb{P}'(X_i > M/n \cup X_i < -M/n) = \mathbb{P}'(X_i > M/n) + \mathbb{P}'(X_i < -M/n)$ (événements incompatibles) et cette quantité est nulle. Donc

$$\forall i, \; \mathbb{P}' \left(|X_i| > \frac{M}{n} \right) = 0$$

et finalement

$$\boxed{\forall i, \; \mathbb{P}' \left(|X_i| \leq \tfrac{M}{n} \right) = 1}$$

(b) X_i est presque surement à valeurs dans $[-M/n, M/n]$ et donc $|\mathbb{E}(X_i)| \leq \frac{M}{n}$ (croissance de l'espérance).
$|X_i - \mathbb{E}(X_i)|$ est à valeurs dans $[0, M/n]$ et donc $(X_i - \mathbb{E}(X_i))^2$ est à valeurs dans $[0, (M/n)^2]$ et donc $\mathbb{E}((X_i - \mathbb{E}(X_i))^2) \in [0, (M/n)^2]$ (toujours par croissance de l'espérance). On en déduit que $0 \leq \mathbb{V}(X_i) \leq \frac{M^2}{n^2}$.
Les X_i étant indépendantes, $\mathbb{V}(X_1 + \cdots + X_n) = \mathbb{V}(X_1) + \cdots + \mathbb{V}(X_n) \leq \frac{M^2}{n}$. Comme X et $X_1 + \cdots + X_n$ ont même loi,

$$\boxed{\mathbb{V}(X) \leq \tfrac{M^2}{n}}$$

3. Ceci a lieu pour tout n et on peut faire tendre n vers l'infini pour obtenir $\mathbb{V}(X) = 0$ et ainsi

$$\boxed{X \text{ est presque sûrement constante}}$$

2.2 Etude du caractère infiniment divisible de quelques variables entières

1. Une variable binomiale est bornée (puisque finie). La partie précédente montre qu'elle n'est infiniment que si elle est constante. Ainsi

$$\boxed{\text{Si } X \hookrightarrow \mathcal{B}(n,p) \text{ avec } p \in]0,1[,\ X \text{ n'est pas infiniment divisible}}$$

2. Le cours nous indique que si $X \hookrightarrow \mathcal{P}(\lambda)$ alors $G_x(t) = e^{\lambda(t-1)}$. Si $X_1, \dots, X_n$ sont indépendantes, on a de plus $G_{X_1+\dots+X_n} = G_{X_1} \dots G_{X_n}$ (avec la question **1.1.2** et à l'aide d'une récurrence utilisant le lemme des coalitions).
Dans le cadre de la question $G_{X_1+\dots+X_n}(t) = e^{(\lambda_1+\dots+\lambda_n)(t-1)}$. On reconnaît la fonction génératrice d'une loi de Poisson de paramètre $\lambda_1 + \dots + \lambda_n$ et avec la question **1.1.1**

$$\boxed{X_1 + \dots + X_n \hookrightarrow \mathcal{P}(\lambda_1 + \dots + \lambda_n)}$$

3. On suppose $X \hookrightarrow \mathcal{P}(\lambda)$. Soit $m \in \mathbb{N}^*$. On peut trouver m variables indépendantes $X_{m,i}$ suivant la loi de Poisson de paramètre λ/m. On a alors $X \sim X_{m,1} + \dots + X_{m,m}$.

$$\boxed{\text{Si } X \hookrightarrow \mathcal{P}(\lambda) \text{ alors } X \text{ est infiniment divisible}}$$

4. Chacune des variables X_i est infiniment divisible et il en est de même des iX_i (il suffit de multiplier par i les variables de la décomposition). Pour conclure, il suffit donc de montrer qu'une somme de variables indépendantes et infiniment divisibles est infiniment divisible. Et quitte à conclure par une récurrence simple, il suffit de traiter le cas d'une somme de deux variables.
Soient donc X et Y infiniment divisibles et indépendantes. Soit $m \in \mathbb{N}^*$. On a alors une décomposition $X \sim X_1 + \dots + X_m$ où les X_i sont indépendantes de même loi et $Y \sim Y_1 + \dots + Y_m$ où les Y_i sont indépendantes de même loi. Notons F la fonctions caractéristique de X_1 et H celle de Y_1 en sorte que $G_X = F^m$ et $G_Y = H^m$.
$FH = A$ est une fonction DSE de rayon de convergence au moins 1. La suite (a_n) de ses coefficients est le produit de Cauchy des suites (f_n) et (h_n) des coefficients de F et H. Comme les f_i et h_i sont positifs, les a_n le sont aussi. De plus, l'égalité $FH = A$ est valable sur $[-1,1]$ (car il y a convergence normale sur $[-1,1]$ donc absolue en tout point de $[-1,1]$). Ainsi $A(1) = F(1)H(1)$.
Finalement, A peut s'interpréter comme lafonction génératrice d'une variable aléatoire Z. On peut alors trouver m variables indépendantes Z_i de même loi que Z. $Z_1 + \dots + Z_m$ a pour fonction génératrice $A^m = F^m H^m = G_X G_Y = G_{X+Y}$ (car X et Y sont indépendantes). On en déduit que $X + Y \sim Z_1 + \dots + Z_m$ et on conclut.

$$\boxed{\sum_{i=1}^{n} iX_i \text{ est infiniment divisible}}$$

2.3 Séries de variables aléatoires à valeurs entières

1. (a) On remarque que $A \cup B$ est la réunion disjointe de $A \cap \overline{B}$, $\overline{A} \cap B$ et $A \cap B$. On a donc

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cap \overline{B}) + \mathbb{P}(\overline{A} \cap B) + \mathbb{P}(A \cap B)$$

Comme $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$, on en déduit que

$$\mathbb{P}(A) + \mathbb{P}(B) - 2\mathbb{P}(A \cap B) = \mathbb{P}(\overline{A} \cap B) + \mathbb{P}(A \cap B)$$

Comme $\mathbb{P}(A \cap B)$ est plus petit que $\mathbb{P}(A)$ et que $\mathbb{P}(B)$ on a donc $\mathbb{P}(B) - \mathbb{P}(A)$ et son opposés plus petits que $\mathbb{P}(\overline{A} \cap B) + \mathbb{P}(A \cap B)$.

$$\boxed{|\mathbb{P}(A) - \mathbb{P}(B)| \leqslant \mathbb{P}(A \cap \overline{B}) + \mathbb{P}(\overline{A} \cap B)}$$

(b) En particulier, pour tout n on a $|\mathbb{P}(X=n)-\mathbb{P}(Y=n)| \leq \mathbb{P}(X=n\cap Y\neq n)+\mathbb{P}(X\neq n\cap Y=n)$. On somme les inégalités sous reserve d'existence :

$$\sum_{n=0}^{\infty} |\mathbb{P}(X=n)-\mathbb{P}(Y=n)| \leq \sum_{n=0}^{\infty} (\mathbb{P}(X=n\cap Y\neq n)+\mathbb{P}(X\neq n\cap Y=n))$$

Les événements $(X=n)\cap(Y\neq n)$ sont incompatibles et tous inclus dans $(X\neq Y)$. Il en est de même pour les événements $(X\neq n)\cap(Y\neq n)$. Le majorant est donc plus petit que $2\mathbb{P}(X\neq Y)$. Ceci justifie les existences (car on somme des quantités positives) et donne

$$\sum_{n=0}^{\infty} |\mathbb{P}(X=n)-\mathbb{P}(Y=n)| \leq 2\mathbb{P}(X\neq Y)$$

Pour $t\in[-1,1]$, on a $|\mathbb{P}(X=n)t^n-\mathbb{P}(Y=n)t^n| \leq |\mathbb{P}(X=n)-\mathbb{P}(Y=n)\|$ qui est le terme général d'une série convergente avec ce qui précède. On a donc $\sum((\mathbb{P}(X=n)t^n-\mathbb{P}(Y=n))t^n)_{n\in\mathbb{N}}$ qui converge et est de somme en module inférieure à $2\mathbb{P}(X\neq Y)$.

$$\boxed{\forall t\in[-1,1],\ \ |G_X(t)-G_Y(t)| \leqslant 2\mathbb{P}(X\neq Y)}$$

2. (a) Si on peut trouver $i\geq n+1$ tel que $U_i(\omega)\neq 0$, on peut a fortiori trouver un $i\geq n$. Ainsi $Z_{n+1}\subset Z_n$. On remarque que

$$Z_n=\bigcup_{i=n}^{+\infty}(U_i\neq 0)$$

Comme $\sum(\mathbb{P}(U_i\neq 0))$ converge, on en déduit (reste de série convergente) que

$$0\leq \mathbb{P}(Z_n)\leq \sum_{i=n}^{+\infty}\mathbb{P}(U_i\neq 0)\to 0$$

$$\boxed{(Z_n) \text{ décroît et } \mathbb{P}(Z_n)\to 0}$$

(b) Par continuité décroissante, on a aussi

$$0=\lim_{n\to+\infty}\mathbb{P}(Z_n)=\lim_{n\to+\infty}\mathbb{P}\left(\bigcap_{k=1}^{n} Z_k\right)=\mathbb{P}\left(\bigcap_{k=1}^{+\infty} Z_k\right)$$

Il est ainsi quasiment impossible qu'une infinité des événement Z_k soient satisfaits simultanément, et donc l'ensemble des ω pour lesquels il existe une infinité de i tels que $U_i(\omega)\neq 0$ est de probabilité nulle. Autrement dit

$$\boxed{\{i\in\mathbb{N}^*/\ U_i\neq 0\} \text{ est presque sûrement fini}}$$

(c) Notons C l'ensemble des $\omega\in\Omega$ tels qu'il n'existe qu'un nombre fini de i pour lesquels $U_i(\omega)\neq 0$. Si $\omega\in C$ alors $S(\omega)$ existe (comme somme finie). S est définie au moins sur C qui est de probabilité 1 (question précédente) et donc

$$\boxed{S \text{ est donc définie presque sûrement}}$$

Avec la question **2.3.1b**, on a

$$\forall t\in[-1,1],\ \ |G_{S_n}(t)-G_S(t)|\leq 2\mathbb{P}(S_n\neq S)$$

On en déduit que

$$\|G_{S_n}-G_S\|_{\infty,[-1,1]}\leq 2\mathbb{P}(S_n\neq S)$$

La suite d'événements $((S_n\neq S))_{n\in\mathbb{N}^*}$ est décroissante car les U_i sont à valeurs positives. Comme plus haut, la continuité décroissante donne

$$\lim_{n\to+\infty}\mathbb{P}(S_n\neq S)=\mathbb{P}\left(\bigcap_{n=1}^{\infty}(S_n\neq S)\right)=\mathbb{P}(\overline{C})=0$$

$$\boxed{(G_{S_n}) \text{ converge uniformément vers } G_S \text{ sur } [-1,1]}$$

3. (a) On a $\mathbb{P}(X_i \neq 0) = 1 - \mathbb{P}(X_i = 0) = 1 - e^{-\lambda_i} \sim \lambda_i$ car $\lambda_i \to 0$ (ou $0 \leq \mathbb{P}(X_i \neq 0) \leq \lambda_i$ par inégalité de convexité si on est gêné par le fait que λ_i peu être nul). Par comparaison des séries positives,

$$\boxed{\sum \mathbb{P}(X_i \neq 0) \text{ converge}}$$

(b) On est dans le cadre de la question **2.3.2** et $X = \sum_{i=1}^{\infty} X_i$ est définie presque sûrement. De plus, la fonction génératrice G_X est la limite uniforme sur $[-1,1]$ de la suite de fonctions de terme général

$$G_{X_1+\cdots+X_n} = G_{X_1} \dots G_{X_n} \ : \ t \mapsto \prod_{i=1}^{n} e^{\lambda_i(t-1)} = \exp\left((t-1)\sum_{i=1}^{n} \lambda_i\right)$$

La limite uniforme de cette suite de fonctions est aussi sa limite simple et ainsi

$$\forall t \in [-1,1],\ G_X(t) = \exp\left((t-1)\sum_{i=1}^{\infty} \lambda_i\right)$$

On reconnaît la fonction génératrice d'une loi de Poisson.

$$\boxed{\sum_{i=1}^{\infty} X_i \hookrightarrow \mathcal{P}\left(\sum_{i=1}^{\infty} \lambda_i\right)}$$

(c) De même, $\mathbb{P}(iX_i \neq 0) = 1 - \mathbb{P}(iX_i = 0) = 1 = e^{-\lambda_i} \sim \lambda_i$ est le terme général d'une série convergente. La variable aléatoire $X = \sum_{i=1}^{\infty} iX_i$ est donc définie presque sûrement. Sa fonction génératrice est la limite uniforme de la suite (G_n) où G_n est la fonction génératrice de $S_n = \sum_{i=1}^{n} iX_i$.
Soit $m \in \mathbb{N}^*$. Comme S_n est infiniment divisible, il existe des variables $Y_{n,i}$ indépendantes de même loi telles que $S_n \sim Y_{n,1} + \cdots + Y_{n,m}$. On a alors $\forall t \in [-1,1],\ G_n(t) = F_n(t)^m$ où F_n est la fonction génératrice de $Y_{n,1}$. $F_n(t) = G_n(t)^{1/m} \to G_X(t)^{1/m}$. En posant $F(t) = G_X(t)^{1/m}$, on a alors $G_X(t) = F(t)^m$. On pourra conclure comme en **2.2.4** si on montre que $F(t)$ est la fonction génératrice d'une variable aléatoire.
Remarquons alors que

$$G_{iX_i}(t) = \sum_{k=0}^{\infty} \mathbb{P}(X_i = k) t^{ik} = G_{X_i}(t^i)$$

en sorte que

$$G_n = \prod_{i=1}^{n} G_{iX_i} \ : \ t \mapsto \prod_{i=1}^{n} e^{\lambda_i(t^i-1)} = \exp\left(-\sum_{i=1}^{n} \lambda_i\right) \times \exp\left(\sum_{i=1}^{n} \lambda_i t^i\right)$$

et donc que

$$\forall t \in [-1,1],\ G_X(t) = \exp\left(-\sum_{i=1}^{\infty} \lambda_i\right) \times \exp\left(\sum_{i=1}^{\infty} \lambda_i t^i\right)$$

ou encore que

$$\forall t \in [-1,1],\ F(t) = \exp\left(-\sum_{i=1}^{\infty} \frac{\lambda_i}{m}\right) \times \exp\left(\sum_{i=1}^{\infty} \frac{\lambda_i}{m} t^i\right)$$

C'est bien la fonction génératrice d'une variable aléatoire : il suffit de reprendre le calcul de la question en travaillant avec $\mu_i = \lambda_i/m$ au lieu de λ.

$$\boxed{\sum_{i=1}^{\infty} iX_i \text{ est infiniment divisible}}$$

3 Variables entières infiniment divisibles : étude générale

3.1 Série entière auxiliaire

1. Si la suite (λ_i) convient alors, on a nécessairement

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

$$\forall k \geq 2,\ \lambda_k = \frac{1}{k\mathbb{P}(X=0)}\left(k\mathbb{P}(X=k) - \sum_{j=1}^{k-1} j\lambda_j \mathbb{P}(X=k-j)\right)$$

   Réciproquement, ceci définit bien une suite de manière récurrente (pour définir λ_k, on n'utilise que les valeurs précédemment définies) et elle vérifie les relations demandées.
   Remarque : la formule générale pour $k \geq 2$ est encore valable si $k = 1$ en convenant que $\sum_{j=1}^{0} y_j = 0$.

2. On multiplie la relation précédente par $\mathbb{P}(X = 0)$ et on passe au module puis on utilise l'inégalité triangulaire :

$$\forall k \geq 1,\ |\lambda_k|\mathbb{P}(X=0) \leq \mathbb{P}(X=k) + \sum_{j=1}^{k-1} \frac{j}{k}|\lambda_j|\mathbb{P}(X=k-j)$$

   Les quantités $\frac{j}{k}$ étant ≤ 1, on en déduit que

$$\forall k \geq 1,\ |\lambda_k|\mathbb{P}(X=0) \leq \mathbb{P}(X=k) + \sum_{j=1}^{k-1} |\lambda_j|\mathbb{P}(X=k-j)$$

   En remarquant que $\mathbb{P}(X = k - j) \leq 1 - \mathbb{P}(X = 0)$ quand $k - j \neq 0$, on en déduit que

$$\forall k \geq 1,\ |\lambda_k|\mathbb{P}(X=0) \leq (1-\mathbb{P}(X=0))\left(1 + \sum_{j=1}^{k-1} |\lambda_j|\right)$$

3. On prouve le résultat par récurrence sur k.
   - On a $1 + |\lambda_1| = 1 + \frac{\mathbb{P}(X=1)}{\mathbb{P}(X=0)} \leq \frac{\mathbb{P}(X=1)}{\mathbb{P}(X=0)} \leq \frac{1}{\mathbb{P}(X=0)}$. Le résultat est donc vrai au rang 1.
   - Supposons le résultat vrai jusqu'au rang $k - 1$. On a donc

$$1 + \sum_{j=1}^{k-1} |\lambda_j| \leq \frac{1}{\mathbb{P}(X=0)^{k-1}} \qquad (*)$$

     La question précédente donne alors

$$|\lambda_k|\mathbb{P}(X=0) \leq \frac{1-\mathbb{P}(X=0)}{\mathbb{P}(X=0)^{k-1}} \qquad (**)$$

     En sommant $(*)$ et $(**)$ divisée par $\mathbb{P}(X = 0)$, on en déduit que

$$1 + \sum_{j=1}^{k} |\lambda_j| \leq \frac{1}{\mathbb{P}(X=0)^{k-1}} + \frac{1-\mathbb{P}(X=0)}{\mathbb{P}(X=0)^k} = \frac{1}{\mathbb{P}(X=0)^k}$$

     ce qui donne le résultat au rang k.

4. En particulier, $|\lambda_k|\mathbb{P}(X=0)^k \leq \mathbb{P}(X=0)^k \sum_{j=1}^k |\lambda_j| \leq 1$. La suite $(\lambda_k \mathbb{P}(X=0)^k)$ est donc bornée. Par définition du rayon de convergence, $\sum \lambda_k t^k$ est de rayon de convergence

$$\rho(X) \geq \mathbb{P}(X=0)$$

5. Je ne vois pas pourquoi $\rho(X) \geq 1$ et je ne crois pas que cela importe pour la suite. Je pose donc $R = \min(\rho(X), 1) > 0$. Comme on peut dériver terme à terme les séries entières sur l'intervalle ouvert de convergence, on a

$$\forall t \in]-R, R[,\ G'_X(t) = \sum_{n=0}^{\infty} (n+1)\mathbb{P}(X=n+1)t^k$$

$$\forall t \in]-R, R[,\ H'_X(t) = \sum_{k=0}^{\infty} (k+1)\lambda_{k+1} t^k$$

Pour $t \in]-R, R[$, les séries définissant $H'_X(t)$ et $G_X(t)$ sont absolument convergentes. Le produit des quantités est alors la somme d'un produit de Cauchy :

$$H'_X(t)G_X(t) = \sum_{n=0}^{\infty} \mu_n t^n \quad \text{avec} \quad \mu_n = \sum_{k=0}^{n} (k+1)\lambda_{k+1}\mathbb{P}(X=n-k)$$

Le changement d'indice $j = k+1$ montre que

$$\mu_n = \sum_{j=1}^{n+1} j\lambda_j \mathbb{P}(X=n+1-j) = (n+1)\mathbb{P}(X=n+1)$$

On a finalement montré que

$$\forall t \in]-R, R[,\ H'_X(t)G_X(t) = G'_X(t)$$

On sait résoudre cette équation différentielle. Il existe une constante c telle que

$$\forall t \in]-R, R[,\ G_X(t) = c\exp(H_X(t))$$

et comme $G_X(0) = \mathbb{P}(X=0)$ et $H_X(0) = \ln(\mathbb{P}(X=0))$, $c = 1$.

$$\boxed{\forall t \in]-R, R[,\ G_X(t) = \exp(H_X(t))}$$

6. Pour $|t| < \min(\rho(X), \rho(Y), 1)$, on a

$$H_{X+Y}(t) = \ln(G_{X+Y}(t)) = \ln(G_X(t)G_Y(t)) = \ln(G_X(t)) + \ln(G_Y(t)) = H_X(t) + H_Y(t)$$

L'égalité étant valable sur un intervalle $]-r, r[$ avec $r > 0$, les coefficients des séries entières sont égaux et l'égalité est en fait vraie pour $|t| < \min(\rho(X), \rho(Y))$.

3.2 Variables aléatoires entières λ-positives

1. Si les λ_j sont positifs, la relation obtenue en **3.1.1** donne immédiatement

$$\forall k \geq 1,\ 0 \leq \lambda_k \leq \frac{\mathbb{P}(X=k)}{\mathbb{P}(X=0)}$$

Le majorant est le terme général d'une série positive convergente (sommes partielles majorées par $\frac{1-\mathbb{P}(X=0)}{\mathbb{P}(X=0)}$). Par théorème de comparaison des séries positives,

$$\boxed{\sum(\lambda_k) \text{ converge}}$$

2. En particulier, le rayon de convergence $\rho(X)$ de $\sum(\lambda_k t^k)$ est plus grand que 1. Mieux, la série entière définissant H_X est normalement convergente sur $[-1,1]$. Il en est de même pour celle définissant G_X. La partie **3.1** donne une identité valable sur $]-1,1[$ et qui le reste sur $[-1,1]$ (par passage à la limite aux bornes). Ainsi

$$\boxed{\forall t \in [-1,1],\ G_X(t) = \exp(H_X(t))}$$

La relation pour $t = 1$ donne

$$\ln(\mathbb{P}(X=0)) = H_X(1) - \sum_{k=1}^{\infty} \lambda_k = \ln(G_X(1)) - \sum_{k=1}^{\infty} \lambda_k$$

Comme $G_X(1) = \sum_{k=0}^{\infty} \mathbb{P}(X=k) = 1$, on conclut que

$$\boxed{\sum_{k=1}^{\infty} \lambda_k = -\ln\left(\mathbb{P}(X=0)\right)}$$

3. Remarquons tout d'abord que

$$G_{iX_i}(t) = \sum_{k=0}^{\infty} \mathbb{P}(X_i = k) t^{ik} = G_{X_i}(t^i)$$

Posons $S_n = \sum_{i=1}^{n} iX_i$. On a

$$G_{S_n} = \prod_{i=1}^{n} G_{iX_i} \ :\ t \mapsto \prod_{i=1}^{n} e^{\lambda_i(t^i - 1)} = \exp\left(-\sum_{i=1}^{n} \lambda_i\right) \times \exp\left(\sum_{i=1}^{n} \lambda_i t^i\right)$$

On obtient la fonction génératrice de $S = \sum_{i=1}^{\infty} iX_i$ en passant à la limite (d'après la partie **2**) et ainsi

$$\forall t \in [-1,1],\ G_S(t) = \exp\left(-\sum_{i=1}^{\infty} \lambda_i\right) \times \exp\left(\sum_{i=1}^{\infty} \lambda_i t^i\right) = \exp(H_X(t)) = G_X(t)$$

On conclut donc

$$\boxed{X \sim S = \sum_{i=1}^{\infty} iX_i}$$

3.3 Caractérisation des variables entières infiniment divisibles

Dans les questions suivantes, je note $S_n = X_{n,1} + \cdots + X_{n,n}$ qui a la même loi que X. Les variables S et X ne sont pas forcément définies sur le même espace probabilisé mais pour la simplicité de la rédaction, on utilisera toujours le symbole $\mathbb{P}$ et on écrira ainsi $\mathbb{P}(S_n = k)\mathbb{P}(X = k)$ (au lieu de $\mathbb{P}'(S_n = k) = \mathbb{P}(X = k)$)

1. (a) On a

$$\bigcap_{k=1}^{n} (X_{n,k} < 0) \subset (S_n < 0)$$

On en déduit que

$$0 \le \mathbb{P}\left(\bigcap_{k=1}^{n} (X_{n,k} < 0)\right) \le \mathbb{P}(S_n < 0) = 0$$

Par indépendance des $X_{n,k}$,

$$\prod_{k=1}^{n} \mathbb{P}(X_{n,k} < 0) = 0$$

il existe donc un k tel que $\mathbb{P}(X_{n,k} < 0) = 0$. Comme les $X_{n,j}$ ont tous la même loi, $\mathbb{P}(X_{n,1} < 0)$ et $X_{n,1}$ est presque sûrement positive ou nulle.

(b) On a
$$\mathbb{P}(S_n=0)=\mathbb{P}(S_n=0|X_{n,1},\dots,X_{n,n}\geq 0)\mathbb{P}(X_{n,1},\dots,X_{n,n}\geq 0)$$
Mais par indépendance,
$$\mathbb{P}(X_{n,1},\dots,X_{n,n}\geq 0)=\prod_{k=1}^{n}\mathbb{P}(X_{n,k}\geq 0)=1$$
et par ailleurs
$$\mathbb{P}(S_n=0|X_{n,1},\dots,X_{n,n}\geq 0)=\mathbb{P}(X_{n,1}=0\cap\dots\cap X_{n,n}=0)=\prod_{k=1}^{n}\mathbb{P}(X_{n,k}=0)$$
En conclusion,
$$\mathbb{P}(S_n=0)=\prod_{k=1}^{n}\mathbb{P}(X_{n,k}=0)$$
Cette quantité étant non nulle, aucun des facteurs ne l'est et
$$\mathbb{P}(X_{n,1}=0)>0$$

(c) Soit $k\in\mathbb{N}$. L'événement $(X_{n,1}\in]i,i+1[)\cap(X_{n,2}=0)\cap\dots\cap(X_{n,n}=0)$ est impossible puisque S_n est à valeurs dans $\mathbb{N}$. Par indépendance des variables,
$$\mathbb{P}(X_{n,1}\in]i,i+1[)\mathbb{P}(X_{n,2}=0)\dots\mathbb{P}(X_{n,n}=0)=0$$
Les $\mathbb{P}(X_{n,k}=0)$ étant non nuls, $\mathbb{P}(X_{n,1}\in]i,i+1[)=0$. Ceci étant vrai pour tout i, on conclut que $X_{n,1}$ (et donc tous les $X_{n,k}$) est presque surement à valeurs entières.

2. (a) On sait que $\mathbb{P}(X_{n,1}=0)^n=\mathbb{P}(X=0)$. D'autre part, la suite $(\mathbb{P}(X_{n,1}=0))$ est bornée et pour prouver qu'elle converge, il suffitde prouver qu'elle admet une unique valeur d'adhérence. Soit ℓ une telle valeur. $\ell\in[0,1]$ et si $\ell\neq 1$, l'identité $\mathbb{P}(X_{n,1}=0)^n=\mathbb{P}(X=0)$ montre que $\mathbb{P}(X=0)=0$ ce qui est exclus. 1 est ainsi l'unique valeur d'adhérence.
$$\boxed{\lim_{n\to\infty}\mathbb{P}\left(X_{n,1}=0\right)=1}$$

(b) On a $0\leq\mathbb{P}(X_{n,1}\geq 1)\leq 1-\mathbb{P}(X_{n,1})\to 0$ quand $n\to+\infty$. Comme $0\leq\mathbb{P}(X_{n,1}=i)\leq\mathbb{P}(X_{n,1}\geq 1)$ pour tout $i\geq 1$, on en déduit que
$$\forall i\in\mathbb{N}^*,\ \lim_{n\to\infty}\mathbb{P}\left(X_{n,1}=i\right)=0$$

3. Comme $0\leq\mathbb{P}(X=0)\leq\mathbb{P}(X_{n,1}=0)$, les séries entières H_n et H_X ont un rayon de convergence convergence au moins égal à $\mathbb{P}(X=0)$ (question **3.1.4**). Ci-dessous, les égalités entre séries entières ont donc au moins lieu sur $I=]-\mathbb{P}(X=0),\mathbb{P}(X=0)[$ et comme $\mathbb{P}(X=0)>0$, cela suffit pour une éventuelle identification des coefficients des séries entières. Dans la suite, les égalités entre fonctions sont à comprendre comme des égalités sur I.

(a) On a $G_X=\prod_{i=1}^{n}G_{X_{i,n}}=G_{X_{1,n}}^n$. En passant au logarithme, on en déduit que
$$nH_n=H_X$$

(b) On peut alors identifier les coefficients des séries entières comme indiqué plus haut pour conclure que, pour $k,n\in\mathbb{N}*$,
$$n\lambda_k(X_{1,n})=\lambda_k(X)$$
où on note $\lambda_k(Y)$ le coefficients λ_k associé à la variable Y. Ainsi
$$\sum_{j=1}^{k}j\lambda_j(X)\mathbb{P}\left(X_{n,1}=k-j\right)=n\sum_{j=1}^{k}j\lambda_j(X_{1,n})\mathbb{P}\left(X_{n,1}=k-j\right)=nk\mathbb{P}\left(X_{n,1}=k\right)$$

4. On fixe $k \in \mathbb{N}^*$. Pour $j \in [|1, k-1|]$, $\mathbb{P}(X_{n,1} = k - j) \to 0$ quand $n \to +\infty$ alors que $\mathbb{P}(X_{n,1} = 0) \to 1$. Le passage à la limite $n \to +\infty$ dans la relation de la question précédente donne donc
$$\forall k \in \mathbb{N}^*,\ \lim_{n\to+\infty} n\mathbb{P}(X_{n,1} = k) = \lambda_k$$
Les λ_k sont donc tous positifs comme limite de quantités positives et X est λ-positive.

5. (a) On vient de voir que si X est infiniment divisible, elle est λ-positive. $(ii) \Rightarrow (iii)$ provient de **3.2.3**. La dernière implication provient de **2.3.3c** (si $X \sim Y$ et Y est infiniment divisible, alors X l'est aussi).
   (b) Dans le cas d'une variable à valeurs dans $\mathbb{N}^*$, on a $\mathbb{P}(X = 0) = 0$. Si $\mathbb{P}(X = 1) \neq 0$, il suffit de travailler avec la variable $Y = X - 1$ qui est à valeur dans $\mathbb{N}$. X est infiniment divisible si et seulement si Y l'est et on peut appliquer ce qui précède à Y.
   (c) $Y = X - 1$ suit la loi $\mathbb{P}(Y = k) = p(1-p)^k$, la suite λ associée est la suite des coefficients du DSE de
$$H_Y(t) = \ln(G_Y(t)) = \ln(p) - \ln(1 - (1-p)t)$$
   On a donc
$$\forall j \geq 1,\ \lambda_j = \frac{(1-p)^j}{j} \geq 0$$
   Y est donc infiniment divisible et X l'est donc aussi.