Mathématiques 2

Présentation du sujet

Le problème porte sur les conditions sous lesquelles une variable aléatoire à valeurs dans $\mathbb{N}$ peut être écrite comme somme de deux variables entières non constantes et indépendantes (variable décomposable), ou bien comme somme de variables discrètes indépendantes et de même loi (variable divisible).

La première partie s'intéresse à quelques exemples de variables décomposables ; on y établit notamment les conditions sous lesquelles une variable binomiale, ou une variable uniforme, sont décomposables.

Dans la deuxième partie, on introduit la notion de variables infiniment divisibles et on étudie des exemples : variables constantes, variables bornées, variables de Poisson et somme de multiples entiers de telles variables. On termine cette partie en établissant qu'une telle somme pondérée de variables de Poisson converge, sous certaines hypothèses, vers une variable infiniment divisible.

La troisième partie propose une caractérisation des variables à valeurs dans $\mathbb{N}$ infiniment divisibles ; une étude très guidée, faisant intervenir le logarithme de la fonction génératrice, permettait de voir que ces variables sont exactement celles qui ont été construites en fin de deuxième partie.

Analyse globale des résultats

Une très grande partie des candidats confond les notions les plus importantes en probabilités, notamment les variables aléatoires et leurs lois. Dans ce contexte, la propriété d'indépendance est très mal comprise ; et son importance capitale dans les définitions proposées par le sujet n'a pas toujours été perçue.

Peu de candidats font l'effort nécessaire pour s'approprier les notions du sujet, en vérifiant scrupuleusement les définitions à l'aide des premiers exemples traités. Ceci va de pair avec une maitrise insuffisante de la langue : il est souvent bien difficile, par exemple, de distinguer une hypothèse supplémentaire d'une étape (censée avoir été démontrée) dans le raisonnement.

Les questions d'analyse (séries entières) et d'algèbre (polynômes) ont été décevantes ; le jury note un manque de rigueur général sur ces sujets, sans parler de la tendance à inventer de nouveaux « théorèmes » en cas de difficulté dans une question.

Les meilleurs candidats sont ceux qui ont abordé le problème en citant précisément les hypothèses du sujet et les théorèmes du cours qui permettaient de résoudre les questions posées. Les questions plus délicates leur ont permis d'exprimer des idées pertinentes, et même si leur mise en œuvre n'était pas parfaite, elles ont été valorisées.

Commentaires sur les réponses apportées et conseils aux futurs candidats

Afin d'aider la préparation des futurs candidats et de préciser les attentes du jury, un corrigé commenté est disponible en annexe de ce rapport.

Partie I

Pour la première question, les candidats oublient presque toujours de préciser que la rayon de convergence de G_X est strictement positif ; il n'est pas possible d'utiliser l'unicité des coefficients sans cette hypothèse.

Trop souvent, on lit que les fonctions génératrices intervenant dans cette partie sont des polynômes, sans qu'une explication en soit donnée.

Dans **I.A.3** et suivantes, et également dans **II.B.3**, il faut rappeler au candidats que l'indépendance est une notion qui s'applique aux variables aléatoires, pas aux lois ; on ne peut pas déduire du choix des lois de deux variables qu'elles sont indépendantes. Et l'indépendance est essentielle dans la définition de décomposition, il faut absolument en faire état, pour montrer qu'on a compris la notion.

Trop souvent, dans cette question et dans **I.B.2a**, le caractère indécomposable d'une variable est confondu avec la simple irréductibilité de sa fonction génératrice, lorsque c'est un polynôme.

Dans **I.A.4a**, l'énoncé invite explicitement à considérer tous les cas possibles pour les degrés de U et V ; traiter le cas $(1,3)$ et expliquer que l'autre cas $(2,2)$ est similaire est alors un gain de temps, mais une perte de points. Deux polynômes qui factorisent un polynôme unitaires ne sont pas forcément unitaires aussi. Dans le cas où l'on voit que $A(-1) = 0$, ce n'est pas parce qu'un polynôme unitaire de degré 1 divise A qu'il s'agit forcément de $T + 1$; il y a une autre racine réelle... Pour résoudre la question en raisonnant avec les deux racines réelles, donc ici avec les trois facteurs irréductibles dans $\mathbb{R}[X]$, il faut être très clair, rigoureux, et surtout exhaustif.

Dans **I.A.4b**, indiquer qu'on choisit X suivant la loi $\mathcal{B}(2, 1/2)$ permet d'éviter un verbiage inutile.

La question **I.B.1a** n'a pas été comprise ; le théorème de la division euclidienne permettait de définir les applications Q et R (ce point précis a rarement été fait explicitement) et il fallait vérifier que ces applications étaient les seules qui convenaient. Moins d'un candidat sur 100 s'est intéressé ensuite à la démonstration du fait que Q et R sont des variables aléatoires ; c'est pourtant explicitement demandé.

La question **I.B.2a** n'a pas toujours été comprise ; un candidat a éclairé les correcteurs, en affirmant que la question n'avait aucun sens, puisqu'on ne demandait pas de prouver un résultat, mais de prouver qu'il suffisait de prouver un autre résultat ! Le lien entre U, V, G_Y et G_Z n'a pas toujours été bien explicité et très peu de candidats se sont intéressés aux hypothèses imposées à U et V (positivité des coefficients, polynômes unitaires).

Plus de la moitié des candidats traitent la question (difficile) **I.B.2b**, en inventant un théorème d'identification : si on a deux factorisations $A = UV = U_1V_1$ avec égalité des degrés (même pas d'égalité des coefficients dominants), alors $U = U_1$ et $V = V_1$. Il est clair qu'une telle affirmation est lourdement pénalisante pour la suite de la copie.

Partie II

Beaucoup de candidats ne traitent pas correctement la première question, en omettant d'évoquer l'indépendance, ou bien en oubliant de choisir la même loi pour les variables qui interviennent.

La question **II.A.2a** a montré une grande confusion dans l'esprit des candidats : plus de 80% de ceux qui traitent la question pensent que $X_1 + \cdots + X_n$ a même loi que nX_1, puisque les X_i sont indépendantes et de même loi. Aucun d'entre eux ne semble réagir en se disant que dans ces conditions, la notion de divisibilité ne présenterait aucun intérêt.

L'existence de la variance dans **II.A.2b** n'a que rarement été mentionnée ; on y a souvent vu par contre l'enchainement absurde $\mathbb{V}(X_1) \leqslant \mathbb{V}(M/n) = M^2/n^2$.

En **II.A.2c**, il fallait démontrer explicitement que la nullité de la variance entrainait que la variable est presque sûrement constante. L'inégalité de Bienaymé-Tchebitchev, même avec $\mathbb{V}(X) = 0$, ne suffisait pas, seule, pour conclure.

La question **II.B.1** est souvent résolue en faisant référence aux résultats de la partie I, sans s'apercevoir que le contexte n'est plus le même (les variables ne sont plus à valeurs entières). Les cas particuliers $p = 0$ et $p = 1$ sont oubliés.

La question **II.B.2** est une réussite, mais les candidats doivent démontrer la formule donnant la fonction caractéristique d'une variable de Poisson. En procédant par récurrence, les étudiants oublient souvent d'établir l'indépendance des variables à l'aide du lemme des coalitions.

La question difficile **II.B.4** a permis aux candidats ayant le plus de recul de proposer une solution, fortement valorisée, bien que la démonstration de l'indépendance des variables soit souvent incomplète.

La formule des probabilités totales n'est qu'exceptionnellement citée pour le calcul de la probabilité $\mathbb{P}(X \neq Y)$ dans la question **II.C.1b**.

Dans la question **II.C.2a**, il est souvent fait mention du théorème de continuité décroissante, alors qu'il ne permet pas de conclure ; c'est l'inégalité de Boole qu'il fallait utiliser ici et le théorème de continuité à la question suivante.

Dans la question **II.C.2c**, les candidats croient le plus souvent que $\lim S_n = S$ entraine automatiquement que $\lim \mathbb{P}(S_n \neq S) = 0$.

Partie III

Cette partie a été peu abordée. La première question a souvent donné lieu a des rédactions très longues, alors qu'il suffisait d'expliciter (λ_k) comme suite récurrente pour conclure en quelques lignes. Dans la question **III.A.6**, les candidats prennent le logarithme des fonctions génératrices à l'intérieur du disque de convergence sans s'inquiéter de la positivité des fonctions ; d'autres croient pouvoir conclure sans faire référence à l'indépendance des variables, ni en fait à un lien quelconque entre G_{X+Y}, G_X et G_Y.

Conclusion

Faire des probabilités, c'est avant tout faire des mathématiques ; l'intuition y joue certes un rôle accru, mais elle ne remplace pas une vraie démonstration explicitant les hypothèses et les théorèmes utilisés.

Il convient de s'approprier rapidement et surtout précisément les notions du sujet, afin de bien mettre en valeur les points les plus importants des démonstrations. Ici par exemple, l'indépendance des variables était centrale, et il convenait de la mentionner explicitement à chaque fois que c'était nécessaire.

Rappelons pour finir quelques évidences. Une copie rédigée correctement permet de suivre sans effort la démarche proposée par le candidat, et ceci est toujours apprécié et valorisé. À l'inverse, une copie difficilement lisible, écrite dans un français approximatif, présentant de nombreuses ratures ou fautes d'orthographe, ne mettant pas en valeur les résultats démontrés, est forcément sanctionnée, plus ou moins consciemment.

Corrigé commenté

I Variables aléatoires entières décomposables

I.A – Premiers exemples

I.A.1) Si $X \sim X'$, alors $\mathbb{P}_X = \mathbb{P}_{X'}$ et donc par définition des fonctions génératrices : $G_X = G_{X'}$.

Réciproquement, si $G_X = G_{X'}$, comme ces séries entières ont un rayon de convergence au moins égal à 1, l'unicité des coefficients montre que $\mathbb{P}(X = n) = \mathbb{P}(X' = n)$ pour tout $n \in \mathbb{N}$, c'est-à-dire que $X \sim X'$.

I.A.2) Si $X \sim Y + Z$, alors $G_X = G_{Y+Z}$ et si Y et Z sont indépendantes, on a de plus $G_{Y+Z} = G_Y G_Z$; ainsi $G_X = G_Y G_Z$.

I.A.3) Si $n \geqslant 2$, choisissons Y et Z indépendantes telles que $Y \hookrightarrow \mathcal{B}(1,p)$ et $Z \hookrightarrow \mathcal{B}(n-1,p)$. Comme $p \in]0,1[$, Y et Z ne sont pas constantes et on a bien $X \sim Y + Z$ par stabilité de la loi binomiale.

> On peut utiliser le résultat fourni par l'énoncé pour justifier l'existence de telles variables Y et Z, en prenant des variables X_1, ..., X_n mutuellement indépendantes suivant la même loi de Bernoulli de paramètre p et en utilisant ensuite le lemme des coalitions.
>
> Pour la réciproque, on peut procéder de différentes façons.

Première méthode. Supposons $X \sim Y + Z$, où Y et Z ne sont pas constantes presque sûrement ; alors $\mathbb{P}(Y = 0) \neq 1$, donc il existe $k \geqslant 1$ tel que $\mathbb{P}(Y = k) > 0$; et de même il existe $\ell \geqslant 1$ tel que $\mathbb{P}(Z = \ell) > 0$. Mais alors, comme $[Y = k] \cap [Z = \ell] \subset [X + Y = k + \ell]$, on a par indépendance :

$$\mathbb{P}(X \geqslant 2) \geqslant \mathbb{P}(X = k + \ell) = \mathbb{P}(Y + Z = k + \ell) \geqslant \mathbb{P}(Y = k,\ Z = \ell) = \mathbb{P}(Y = k)\,\mathbb{P}(Z = \ell) > 0$$

et ainsi $2 \in X(\Omega) = [\![0, n]\!]$, donc $n \geqslant 2$.

> Pour les méthodes suivantes, on procède par contraposée et on démontre que si $X \hookrightarrow \mathcal{B}(1,p)$, alors X n'est pas décomposable.

Supposons que $X \sim Y + Z$, on commence par remarquer que Y et Z sont bornées presque sûrement. En effet, comme $Y(\Omega) \subset \mathbb{N}$ et $Z(\Omega) \subset \mathbb{N}$, on a :

$$0 \leqslant \mathbb{P}(Y > 1) = \mathbb{P}(Y + Z > 1 + Z) \leqslant \mathbb{P}(Y + Z > 1) = \mathbb{P}(X > 1) = 0$$

et de même $\mathbb{P}(Z > 1) = 0$. Ainsi $Y \in \{0,1\}$ et $Z \in \{0,1\}$ presque sûrement.

Deuxième méthode. On écrit

$$0 = \mathbb{P}(X = 2) = \mathbb{P}(Y + Z = 2) = \mathbb{P}(Y = 1, Z = 1) = \mathbb{P}(Y = 1)\,\mathbb{P}(Z = 1)$$

par indépendance ; donc l'une des probabilités de droite est nulle.

Si par exemple $\mathbb{P}(Y = 1) = 0$, c'est que $\mathbb{P}(Y = 0) = 1$ et donc Y est constante presque sûrement. Il n'existe pas de décomposition $X \sim Y + Z$ où X et Y ne sont pas constantes presque sûrement : c'est que X n'est pas décomposable.

Troisième méthode. D'après la question précédente, on doit avoir $G_X = G_Y G_Z$; mais ici $G_X(t) = 1 - p + pt$ est un polynôme de degré 1, tandis que G_Y et G_Z sont également des polynômes (puisque Y et Z sont bornées presque sûrement). Nécessairement l'un des polynômes G_Y ou G_Z est constant, et donc Y ou Z est une variable aléatoire constante presque sûrement. On conclut comme plus haut.

Remarquons que, si on ne démontre pas que Y et Z sont bornées, rien n'oblige à priori dans l'égalité de séries entières $G_X = G_Y G_Z$ à avoir G_Y et G_Z polynomiales ; c'est la positivité des coefficients qui permet de démontrer ce résultat vrai ici. On peut méditer avec intérêt l'égalité $1 = (1-t)\sum t^k$, ou encore $1 + t = (\sqrt{1+t})^2$...

I.A.4a) Raisonnons par l'absurde et supposons qu'il existe deux polynômes U et V non constants, à coefficients positifs, tels que $U(T)V(T) = A(T)$. Quitte à les diviser par leur coefficient dominant, on peut supposer qu'ils sont unitaires (puisque A l'est) et, quitte à les permuter, on peut supposer que $1 \leqslant \deg U \leqslant \deg V$.

– Si $\deg U = 1$, alors $\deg V = 3$, soient $(a,b,c,d) \in (\mathbb{R}_+)^4$ tels que $U(T) = T + a$ et $V(T) = T^3 + bT^2 + cT + d$. Alors

$$U(T)V(T) = T^4 + (a+b)T^3 + (ab+c)T^2 + (ac+d)T + ad = T^4 + 2T + 1$$

Par identification, on a alors $a+b = 0$ et donc par positivité, $a = b = 0$; mais alors $ad = 0 \neq 1$, contradiction.

– Si $\deg U = \deg V = 2$, soient $(a,b,c,d) \in (\mathbb{R}_+)^4$ tels que $U(T) = T^2 + aT + b$ et $V(T) = T^2 + cT + d$. Alors

$$U(T)V(T) = T^4 + (a+c)T^3 + (ac+b+d)T^2 + (ad+bc)T + bd = T^4 + 2T + 1$$

Ici encore, $a + c = 0$ conduit à $a = c = 0$, puis à $ad + bc = 0 \neq 2$, contradiction.

Il n'existe donc pas de polynômes à coefficients positifs, tous deux non constants, dont le produit soit A.

On aurait également pu trouver (numériquement, à l'aide de la calculatrice) la décomposition en facteurs irréductibles de A dans $\mathbb{R}[X]$ et examiner à partir de là toutes les possibilités non triviales de factorisation, pour constater qu'elles conduisent toutes à des facteurs dont les coefficients ne sont pas tous positifs. Mais ce n'est pas plus simple (il faut être exhaustif et bien expliquer la méthode) et cela ne correspond pas à ce que suggérait l'indication de l'énoncé.

I.A.4b) Pour appliquer ce qui précède et vu l'indication, le polynôme $A/4$ doit être la fonction génératrice de la variable X^2. Ceci conduit à prendre pour X une variable suivant la loi $\mathcal{B}(2, 1/2)$; d'après **I.A.3**, X est décomposable, et on vérifie que

$$\begin{aligned} G_{X^2}(t) &= \mathbb{P}(X^2 = 0) + \mathbb{P}(X^2 = 1)t + \mathbb{P}(X^2 = 4)t^4 \\ &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1)t + \mathbb{P}(X = 2)t^4 \\ &= \frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^4 \\ &= \frac{1}{4}A(t) \end{aligned}$$

Supposons que $X^2 \sim Y + Z$ où Y et Z sont à valeurs dans $\mathbb{N}$ et indépendantes. Comme plus haut, puisque $\mathbb{P}(X^2 > 4) = 0$ et que Y et Z sont positives, on voit que

$$\mathbb{P}(Y > 4) = \mathbb{P}(Z > 4) = 0$$

Ainsi $\mathbb{P}(Y \in [\![0,4]\!]) = 1$ et donc G_Y est un polynôme (de degré au plus 4) et de même G_Z est un polynôme.

Comme $X^2 \sim Y + Z$, on a par indépendance $G_{X^2} = G_Y G_Z$, donc $A(T) = 4G_Y(T)G_Y(T)$. Ceci est une égalité entre polynômes à coefficients positifs, puisque les coefficients de G_Y et G_Z le sont, s'agissant de probabilités. D'après la question précédente, l'un des polynômes G_Y ou G_Z est constant et donc l'une des variables Y ou Z est constante presque sûrement.

La variable X^2 n'est donc pas décomposable.

I.B – Variables uniformes

I.B.1) Variables uniformes décomposables

I.B.1a) S'agissant d'un problème d'existence et d'unicité, procédons par analyse et synthèse.

Supposons que Q et R soient des applications bien définies.

Alors, pour $\omega \in \Omega$, on a $X(\omega) = aQ(\omega) + R(\omega)$ avec $R(\omega) \in [\![0, a-1]\!]$, donc nécessairement $Q(\omega)$ et $R(\omega)$ sont le quotient et le reste de la division euclidienne de $X(\omega)$ par a, puisque ceux-ci sont uniques.

Définissons l'application $Q : \Omega \to \mathbb{N}$ de la façon suivante : à tout $\omega \in \Omega$, on associe le quotient $Q(\omega)$ de la division euclidienne de $X(\omega)$ par a. On définit de même $R : \Omega \to [\![0, a-1]\!]$, de sorte que $R(\omega)$ soit le reste dans cette même division euclidienne. L'existence et l'unicité (encore) du quotient et du reste dans la division euclidienne font que ces applications sont bien définies et on a bien $X = aQ + R$.

Il reste à vérifier que Q et R sont bien des variables aléatoires.

Première méthode. En prévision de la question suivante, on peut constater que $Q(\Omega) \subset \mathbb{N}$ et écrire, pour tout $q \in \mathbb{N}$

$$[Q = q] = \bigcup_{r \in [\![0, a-1]\!]} [X = aq + r]$$

ce qui permet de voir $[Q = q]$ comme une réunion finie d'événements et donc comme un événement lui-même. Ceci étant vrai pour tout $q \in \mathbb{N}$, Q est une variable aléatoire.

De même, $R(\Omega) \subset [\![0, a-1]\!]$ et si $r \in [\![0, a-1]\!]$, alors

$$[R = r] = \bigcup_{q \in \mathbb{N}} [X = aq + r]$$

est une réunion dénombrable d'événements, donc un événement aussi, et ainsi R est également une variable aléatoire.

> On peut s'éviter ce travail avec R puisque $R = X - aQ$ est différence de deux variables aléatoires, donc également une variable aléatoire.

Deuxième méthode. On peut également constater que Q et R sont en fait des fonctions de X ; concrètement, si $\phi : \mathbb{N} \to \mathbb{N}$ est l'application qui à un entier k associe le quotient de la division euclidienne de k par a, alors $Q = \phi \circ X$. La variable X étant discrète et la fonction ϕ étant déterministe, on en déduit que Q est une variable aléatoire. De même, avec l'application « reste », ou bien par différence, on voit que R est également une variable aléatoire.

Deuxième méthode (bis). On explicite les choses : $Q = \lfloor X/a \rfloor$ et $R = X - a \lfloor X/a \rfloor$ sont des fonctions (déterministes) de la variable aléatoire discrète X, donc Q et R sont elles-mêmes des variables aléatoires discrètes.

I.B.1b) Par double inclusion et compte tenu de la définition de Q et R, on a l'égalité d'évènements

$$\forall (q,r) \in \mathbb{N} \times [\![0, a-1]\!], \qquad [(Q,R) = (q,r)] = [X = aq + r]$$

Comme $X \hookrightarrow \mathcal{U}([\![0, n-1]\!])$ et que

$$aq + r \in [\![0, n-1]\!] \iff (q,r) \in [\![0, b-1]\!] \times [\![0, a-1]\!]$$

vu que $n = ab$, on en déduit

$$\mathbb{P}((Q,R) = (q,r)) = \begin{cases} 1/n & \text{si } (q,r) \in [\![0, b-1]\!] \times [\![0, a-1]\!] \\ 0 & \text{sinon} \end{cases}$$

On en déduit alors les lois de Q et R, en tant que lois marginales de ce couple

$$\mathbb{P}(Q = q) = \sum_{r \in R(\Omega)} \mathbb{P}((Q,R) = (q,r)) = \sum_{r=0}^{a-1} \mathbb{P}((Q,R) = (q,r))$$

$$\mathbb{P}(Q = q) = \begin{cases} a/n = 1/b & \text{si } q \in [\![0, b-1]\!] \\ 0 & \text{sinon} \end{cases}$$

Et aussi

$$\mathbb{P}(R = r) = \sum_{q=0}^{b-1} \mathbb{P}((Q,R) = (q,r)) = \begin{cases} b/n = 1/a & \text{si } q \in [\![0, a-1]\!] \\ 0 & \text{sinon} \end{cases}$$

On reconnait $Q \hookrightarrow \mathcal{U}([\![0, b-1]\!])$, $R \hookrightarrow \mathcal{U}([\![0, a-1]\!])$.

I.B.1c) Comme $X = aQ + R$, on a bien évidemment $X \sim aQ + R$. Il reste à vérifier que aQ et R sont bien indépendantes et non constantes.

Aucune n'est constante, puisqu'elles suivent des lois uniformes sur des ensembles à plus d'un élément, car $a \geqslant 2$, $b \geqslant 2$.

Et elles sont bien indépendantes, puisque Q et R le sont car

$$\forall (q,r) \in [\![0, b-1]\!] \times [\![0, a-1 \div]\!], \quad \mathbb{P}((Q,R) = (q,r)) = \frac{1}{n} = \frac{1}{b}\frac{1}{a} = \mathbb{P}(Q = q)\,\mathbb{P}(R = r)$$

On en déduit que $G_X = G_{aQ} G_R$. Or, comme Q ne prend que des valeurs entières,

$$G_{aQ}(t) = \sum_{k=0}^{+\infty} \mathbb{P}(aQ = k) t^k = \sum_{i=0}^{+\infty} \mathbb{P}(aQ = ai) t^{ai} = \frac{1}{b} \sum_{i=0}^{b-1} t^{ai} = \frac{1}{b} \frac{1 - t^{ab}}{1 - t^a} = \frac{1}{b} \frac{1 - t^n}{1 - t^a}$$

On calcule aussi

$$G_R(t) = \sum_{k=0}^{a-1} \mathbb{P}(R = k) t^k = \frac{1}{a} \frac{1 - t^a}{1 - t}$$

Finalement, le calcul de G_X est semblable et on obtient

$$G_X(t) = \frac{1}{n} \frac{1 - t^n}{1 - t} = \left(\frac{1}{b} \frac{1 - t^n}{1 - t^a}\right) \left(\frac{1}{a} \frac{1 - t^a}{1 - t}\right)$$

ce qui peut se réécrire sous la forme

$$1 + t + \cdots + t^{n-1} = (1 + t^a + t^{2a} + \cdots + t^{n-a})(1 + t + t^2 + \cdots + t^{a-1})$$

et ce n'est pas un scoop.

I.B.2) Variables uniformes non décomposables

I.B.2a) Supposons que le résultat proposé par l'énoncé est déjà démontré et prenons $X \sim Y + Z$, où $X \hookrightarrow \mathcal{U}(\llbracket 0, n-1 \rrbracket)$, Y et Z à valeurs dans $\mathbb{N}$ et indépendantes.

De la même façon que démontré plus haut, Y et Z sont presque sûrement bornées, puisque $\mathbb{P}(Y \geqslant n) \leqslant \mathbb{P}(Y + Z \geqslant n) = 0$. On en déduit que les fonctions génératrices G_Y et G_Z sont des polynômes, à coefficients positifs puisque ce sont des probabilités.

On a donc $G_X(t) = \frac{1}{n}(1 + t + \cdots + t^{n-1}) = G_Y(t)G_Z(t)$, donc

$$1 + T + \cdots + T^{n-1} = (nG_Y(T))G_Z(T)$$

En divisant ces polynômes par leur coefficient dominant (strictement positif), on en déduit une égalité du type $1+T+\cdots+T^{n-1} = U(T)\,V(T)$, où U et V sont des polynômes unitaires à coefficients positifs.

Si le résultat proposé par l'énoncé est vrai, on en déduit alors que U ou V est un polynôme constant ; et alors, en revenant aux fonctions génératrices, que G_Y ou G_Z est constante, c'est-à-dire que Y ou Z est constante presque sûrement. Et donc que X n'est pas décomposable.

I.B.2b) Le polynôme $1 + T + \cdots + T^{n-1}$ est scindé dans $\mathbb{C}$ et ses racines, simples, sont les racines n-ièmes de l'unité autres que 1, puisque

$$1 + T + \cdots + T^{n-1} = \frac{1 - T^n}{1 - T}$$

On en déduit que U et V sont également scindés à racines simples.

Soit z une racine de U ; comme U est un polynôme à coefficients réels, $\overline{z}$ est également racine de U, mais ici $|z| = 1$ donc $\overline{z} = 1/z$. On en déduit que U et $T^r U(1/T)$ (qui est bien un polynôme) ont les mêmes racines, forcément simples, et ces polynômes sont donc proportionnels. Mais ils prennent la même valeur en 1 et cette valeur est non nulle puisque 1 n'est pas racine de U. On en déduit donc qu'ils sont égaux : $U(T) = T^r U(1/T)$. On a de même $V(T) = T^s V(1/T)$.

I.B.2c) Avec les notations de l'énoncé, ces égalités signifient : $u_{r-i} = u_i$ pour tout $i \in \llbracket 1, r-1 \rrbracket$, et $v_{s-j} = v_j$ pour tout $j \in \llbracket 1, s-1 \rrbracket$.

Considérons le terme de degré r du produit $U(T)V(T)$; il vaut

$$\sum_{k=0}^{r} u_{r-k}v_k = 1 + \sum_{k=1}^{r} u_k v_k = 1, \qquad \text{donc} \qquad \sum_{k=1}^{r} u_k v_k = 0$$

Comme $u_k v_k \geqslant 0$, on en déduit que $u_k v_k = 0$ pour tout $k \in \llbracket 1, r \rrbracket$.

I.B.2d) On procède par récurrence forte.

- Le coefficient de degré 1 du produit UV est $u_1 + v_1 = 1$. Or $u_1 v_1 = 0$, donc l'un des facteurs est nul, et l'autre vaut alors forcément 1. On a bien $\{u_1, v_1\} \subset \{0, 1\}$.
- Supposons, pour un entier $k \in \llbracket 1, r-1 \rrbracket$, que $\{u_1, u_2, ..., u_k, v_1, v_2, ..., v_k\} \subset \{0, 1\}$. Considérons alors le coefficient de degré $k+1$ dans le produit UV. Il s'écrit

$$\sum_{i=0}^{k+1} u_{k+1-i} i v_i = u_{k+1} + v_{k+1} + \sum_{i=1}^{k} u_{k+1-i} v_i = 1$$

  Par hypothèse de récurrence, chacun des termes de la somme vaut 0 ou 1, il s'agit donc d'un entier positif. Comme u_{k+1} et v_{k+1} sont eux mêmes des réels positifs, on en déduit que $u_{k+1} +$

$v_{k+1} \in \{0,1\}$, et sachant que $u_{k+1}v_{k+1} = 0$, on en déduit encore que l'un est nul, et que l'autre vaut 0 ou 1. Ainsi $\{u_{k+1}, v_{k+1}\} \subset \{0,1\}$, la transmission est assurée.

I.B.2e) En posant $u_{r+1} = u_{r+2} = \cdots = u_s = 0$, la même transmission peut continuer à se faire jusqu'au rang s, en exploitant successivement les coefficients de degré $r+1, r+2, ..., s$ du produit UV; on en déduit que tous les coefficients des polynômes U et V sont égaux à 0 ou à 1.

On conclut en évaluant l'égalité $1 + T + \cdots + T^{n-1} = U(T)V(T)$ en 1 : on obtient $n = pq$, où $p = U(1) \geqslant 1 + 1^r = 2$ et $q = V(1) \geqslant 2$ sont des entiers, ce qui contredit la primalité de n.

L'hypothèse $r \geqslant 1$ et $s \geqslant 1$ conduit ainsi à une contradiction.

On en déduit que, si $1 + T + \cdots + T^{n-1} = UV$ où U et V sont des polynômes unitaires, à coefficients positifs, alors l'un des deux a un degré nul, c'est-à-dire qu'il est constant.