

# Variables aléatoires décomposables

*J'ai mis un sujet de DS des années précédentes, c'est un sujet de concours complet avec une partie que j'ai considérée "en trop" pour le format DS.*

Ce problème a pour objet la représentation de la loi d'une variable aléatoire comme loi d'une somme de variables aléatoires indépendantes.

On s'intéresse d'abord au cas d'une somme de deux variables à valeurs entières, puis au cas de variables aléatoires dont la loi est celle de la somme d'un nombre quelconque de variables indépendantes de même loi.

## Notations

Toutes les variables aléatoires considérées dans ce problème sont discrètes. On note  $\mathbb{P}_X$  la loi d'une variable aléatoire  $X$ .

Si  $X$  et  $X'$  sont deux variables aléatoires définies sur les espaces probabilisés respectifs  $(\Omega, \mathcal{A}, \mathbb{P})$  et  $(\Omega', \mathcal{A}', \mathbb{P}')$ , la notation  $X \sim X'$  signifie que  $X$  et  $X'$  ont même loi, c'est à dire que  $\mathbb{P}_X = \mathbb{P}_{X'}$ .

Pour toute variable aléatoire  $X$  à valeurs dans  $\mathbb{N}$  on note  $G_X$  sa fonction génératrice, définie, pour  $t \in \mathbb{R}$ , par

$$G_X(t) = \sum_{n=0}^{\infty} \mathbb{P}(X = n) t^n$$

lorsque la série converge.

On pourra si nécessaire utiliser librement le résultat suivant.

Si  $m \in \mathbb{N}^*$  et si  $\mathcal{L}$  est une loi de probabilité sur un espace probabilisé  $\Omega$ , alors il existe des variables aléatoires  $X_1, \dots, X_m$  définies sur un espace probabilisé  $\Omega_m$ , mutuellement indépendantes et de loi  $\mathcal{L}$ .

Si  $a$  et  $b$  sont deux entiers tels que  $a \leq b$ , on désigne par  $\llbracket a, b \rrbracket$  l'ensemble des entiers  $k$  tels que  $a \leq k \leq b$ .

## I Variables aléatoires entières décomposables

Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{N}$ . On appelle *décomposition* de  $X$  toute relation de la forme  $X \sim Y + Z$  où  $Y$  et  $Z$  sont deux variables aléatoires indépendantes à valeurs dans  $\mathbb{N}$ , définies sur un espace probabilisé pouvant être distinct de celui sur lequel  $X$  est définie.

On dit que  $X$  est *décomposable* si  $X$  admet une décomposition où  $Y$  et  $Z$  ne sont pas constantes presque sûrement.

### I.A - Premiers exemples

1. Soit  $X$  et  $X'$  deux variables aléatoires à valeurs dans  $\mathbb{N}$ . Justifier que  $X \sim X'$  si et seulement si  $G_X = G_{X'}$ .
2. Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{N}$  admettant une décomposition  $X \sim Y + Z$ , où  $Y$  et  $Z$  sont des variables aléatoires indépendantes à valeurs dans  $\mathbb{N}$ . Quelle relation lie  $G_X$ ,  $G_Y$  et  $G_Z$  ?
3. Soit  $X$  une variable aléatoire suivant la loi binomiale  $\mathcal{B}(n, p)$  où  $n \geq 1$  et  $p \in ]0, 1[$ . Montrer que  $X$  est décomposable si et seulement si  $n \geq 2$ .
4. Soit  $A(T) \in \mathbb{R}[T]$  le polynôme :  $A(T) = T^4 + 2T + 1$ .
  - (a) Soit  $U(T)$  et  $V(T)$  deux polynômes à coefficients réels positifs ou nuls tels que  $U(T)V(T) = A(T)$ . montrer que l'un des polynômes  $U(T)$  ou  $V(T)$  est constante.  
*On pourra distinguer les cas selon les valeurs des degrés de  $U(T)$  et  $V(T)$ .*

(b) En déduire qu'il existe une variable aléatoire décomposable  $X$  telle que  $X^2$  ne soit pas décomposable.  
*On pourra considérer le polynôme  $\frac{1}{4}A(T)$ .*

## I.B - Variables uniformes

Dans cette sous-partie,  $n$  est un entier naturel supérieur ou égal à 2 et  $X$  est une variable aléatoire à valeurs dans  $\mathbb{N}$ , définie sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$  et suivant la loi uniforme sur  $\llbracket 0, n-1 \rrbracket$  :

$$\mathbb{P}(X = k) = \frac{1}{n} \text{ si } k \in \llbracket 0, n-1 \rrbracket \text{ et } \mathbb{P}(X = k) = 0 \text{ sinon}$$

### 1. Variables uniformes décomposables

On suppose dans cette question que  $n$  n'est pas premier : il existe des entiers  $a$  et  $b$ , supérieurs ou égaux à 2, tels que  $n = ab$ .

(a) Montrer qu'il existe un unique couple de variables aléatoires entières  $(Q, R)$  définies sur  $\Omega$  telles que

$$X = aQ + R \text{ et } \forall \omega \in \Omega, R(\omega) \in \llbracket 0, a-1 \rrbracket$$

*On pourra considérer une division euclidienne.*

(b) Préciser la loi de  $(Q, R)$ , puis les lois de  $Q$  et de  $R$ .  
(c) Montrer que  $X$  est décomposable. En déduire une expression de  $G_X$  comme produit de deux polynômes non constants que l'on précisera.

Section hors DS (je l'ai quand même mise dans le corrigé)

### 2. Variables uniformes non décomposables

On suppose dans cette question que  $n$  est un nombre premier et on établit que  $X$  n'est pas décomposable.

(a) Montrer qu'il suffit de prouver le résultat suivant : si  $U$  et  $V$  sont des polynômes de  $\mathbb{R}[T]$  unitaires et à coefficients dans  $\mathbb{R}^+$  tels que  $U(T)V(T) = 1 + T + \dots + T^{n-1}$ , alors l'un des deux polynômes  $U$  ou  $V$  est constant.

Dans ce qui suit, on fixe des polynômes  $U$  et  $V$  de  $\mathbb{R}[T]$  unitaires à coefficients dans  $\mathbb{R}^+$  tels que

$$U(T)V(T) = 1 + T + \dots + T^{n-1}$$

On pose  $r = \deg(U)$  et  $s = \deg(V)$  et on suppose par l'absurde que  $r$  et  $s$  sont non nuls.

(b) Montrer que  $U(T) = T^r U\left(\frac{1}{T}\right)$  et  $V(T) = T^s V\left(\frac{1}{T}\right)$ .

On note alors  $U(T) = 1 + u_1 T + \dots + u_{r-1} T^{r-1} + T^r$  et  $V(T) = 1 + v_1 T + \dots + v_{s-1} T^{s-1} + T^s$  avec  $r \leq s$  (quitte à échanger les rôles de  $U$  et  $V$ ).

(c) Montrer que  $\forall k \in \llbracket 1, r \rrbracket$ ,  $u_k v_k = 0$ .  
(d) En déduire que  $\forall k \in \llbracket 1, r \rrbracket$ ,  $u_k \in \{0, 1\}$  et  $v_k \in \{0, 1\}$ .  
(e) Conclure.

*On pourra d'abord montrer que tous les coefficients de  $V$  sont à valeurs dans  $\{0, 1\}$ .*

Fin de la section hors DS

## II Variables infiniment divisibles : exemples

Soit  $X$  une variable aléatoire discrète à valeurs dans  $\mathbb{R}$ . On dit que  $X$  est *infiniment divisible* si, pour tout  $m \in \mathbb{N}^*$ , il existe des variables aléatoires réelles discrètes  $X_{m,1}, \dots, X_{m,m}$  mutuellement indépendantes, de même loi, et vérifiant  $X \sim X_{m,1} + \dots + X_{m,m}$ . Dans cette définition, l'espace probabilisé  $\Omega_m$  sur lequel sont définies les  $X_{m,i}$  peut dépendre de  $m$ .

## II.A - Variables bornées

1. On suppose que  $X$  est constante égale à  $a \in \mathbb{R}$ . Montrer que  $X$  est infiniment divisible.

L'objectif de cette sous-partie est de montrer que toute variable aléatoire bornée infiniment divisible est presque sûrement constante.

Soit  $X$  une variable aléatoire bornée infiniment divisible définie sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .

On note  $M = \sup_{\Omega} |X|$ , de sorte que  $|X(\omega)| \leq M$  pour tout  $\omega \in \Omega$ .

2. Soit  $n \in \mathbb{N}^*$  et soient  $X_1, \dots, X_n$  des variables aléatoires indépendantes et de même loi, et telles que  $X_1 + \dots + X_n$  ait même loi que  $X$ .

(a) Pour tout  $i \in \llbracket 1, n \rrbracket$ , montrer que  $X_i \leq \frac{M}{n}$  presque sûrement, puis  $|X_i| \leq \frac{M}{n}$  presque sûrement.

(b) En déduire que  $\mathbb{V}[X] \leq \frac{M^2}{n}$  où  $\mathbb{V}[X]$  désigne la variance de  $X$ .

3. Conclure que  $X$  est presque sûrement constante.

## II.B - Étude du caractère infiniment divisible de quelques variables entières

1. Une variable binomiale est-elle infiniment divisible ?
2. Soit  $n$  un entier naturel non nul et soit  $X_1, \dots, X_n$  des variables aléatoires mutuellement indépendantes suivant des lois de Poisson de paramètres respectifs  $\lambda_1, \dots, \lambda_n$ .

Montrer que  $X_1 + \dots + X_n$  suit une loi de Poisson de paramètre  $\lambda_1 + \dots + \lambda_n$ .

3. Soit  $X$  une variable aléatoire de Poisson. Montrer que  $X$  est infiniment divisible.
4. Soit  $r$  un entier naturel non nul et soit  $X_1, \dots, X_r$  des variables aléatoires de Poisson mutuellement indépendantes.

Montrer que  $\sum_{i=1}^r iX_i$  est une variable aléatoire infiniment divisible.

## II.C - Séries de variables aléatoires à valeurs entières

1. Soit  $X$  et  $Y$  deux variables aléatoires définies sur  $(\Omega, \mathcal{A}, \mathbb{P})$  et à valeurs dans  $\mathbb{N}$ .

(a) Montrer que si  $A$  et  $B$  sont des événements de  $\mathcal{A}$ , et si  $\overline{A}$  et  $\overline{B}$  sont leurs événements contraires respectifs, alors

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

(b) En déduire que, pour tout  $t \in [-1, 1]$ ,  $|G_X(t) - G_Y(t)| \leq 2\mathbb{P}(X \neq Y)$ .

2. Soit  $(U_i)_{i \in \mathbb{N}^*}$  une suite de variables aléatoires mutuellement indépendantes à valeurs dans  $\mathbb{N}$  telle que la série des  $\mathbb{P}(U_i \neq 0)$  soit convergente.

(a) Soit  $Z_n = \{\omega \in \Omega, \exists i \geq n, U_i(\omega) \neq 0\}$ . Montrer que  $(Z_n)$  est une suite décroissante d'événements et que  $\lim_{n \rightarrow +\infty} \mathbb{P}(Z_n) = 0$ .

(b) En déduire que l'ensemble  $\{i \in \mathbb{N}^*, U_i \neq 0\}$  est presque sûrement fini.

(c) On pose  $S_n = \sum_{i=1}^n U_i$  et  $S = \sum_{i=1}^{\infty} U_i$ . Justifier que  $S$  est définie presque sûrement. Montrer que  $G_{S_n}$  converge uniformément vers  $G_S$  sur  $[-1, 1]$ .

3. Soit  $(\lambda_i)_{i \in \mathbb{N}^*}$  une suite de réels positifs ou nuls. On suppose que la série  $\sum \lambda_i$  est convergente, et on note  $\lambda = \sum_{i=1}^{\infty} \lambda_i$ .

Soit  $(X_i)_{i \in \mathbb{N}^*}$  une suite de variables aléatoires indépendantes telles que, pour tout  $i$ ,  $X_i$  suive une loi de Poisson de paramètre  $\lambda_i$ . On convient que, si  $\lambda_i = 0$ ,  $X_i$  est la variable aléatoire nulle.

(a) Montrer que la série  $\sum \mathbb{P}(X_i \neq 0)$  est convergente.

(b) Montrer que la série  $\sum_{i \geq 1} X_i$  est presque sûrement convergente et que sa somme (définie presque sûrement) suit une loi de Poisson de paramètre  $\lambda$ .

(c) Montrer que la série  $\sum_{i \geq 1} iX_i$  est presque sûrement convergente et que sa somme  $X = \sum_{i=1}^{\infty} iX_i$  définit une variable aléatoire infiniment divisible.

## III V.a. entières infiniment divisibles : étude générale

### III.A - Série entière auxiliaire

Dans cette sous-partie,  $X$  est une variable aléatoire à valeurs dans  $\mathbb{N}$  telle que  $\mathbb{P}(X = 0) > 0$ .

- Montrer qu'il existe une unique suite réelle  $(\lambda_i)_{i \in \mathbb{N}^*}$  telle que, pour tout  $k \in \mathbb{N}^*$

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

- Pour tout  $k \in \mathbb{N}^*$ , montrer

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

- Pour tout  $k \in \mathbb{N}^*$ , montrer :  $1 + \sum_{j=1}^k |\lambda_j| \leq \frac{1}{\mathbb{P}(X = 0)^k}$ .

- Montrer que la série entière  $\sum \lambda_k t^k$  a un rayon de convergence  $\rho(X)$  supérieur ou égal à  $\mathbb{P}(X = 0)$ .

Pour tout réel  $t$  de  $]-\rho(X), \rho(X)[$ , on pose

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

À toute variable aléatoire  $X$  à valeurs dans  $\mathbb{N}$  et telle que  $\mathbb{P}(X = 0) > 0$ , on associe ainsi une série entière  $H_X$ .

Dans la suite du problème,  $H_X$  sera appelée *série entière auxiliaire* de  $X$ .

- Pour  $t \in ]-\rho(X), \rho(X)[ \cap [-1, 1]$ , montrer  $G'_X(t) = H'_X(t)G_X(t)$ , puis  $G_X(t) = \exp(H_X(t))$ .
- Soit  $X$  et  $Y$  deux variables aléatoires indépendantes, définies sur l'espace  $\Omega$  et à valeurs dans  $\mathbb{N}$ , et soit  $H_X$  et  $H_Y$  leurs séries entières auxiliaires.

Montrer  $H_{X+Y}(t) = H_X(t) + H_Y(t)$  pour tout réel  $t$  vérifiant  $|t| < \min(\rho(X), \rho(Y))$ .

### III.B - Variables aléatoires entières $\lambda$ -positives

Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{N}$  telle que  $\mathbb{P}(X = 0) > 0$ , et soit  $H_X$  sa série entière auxiliaire :

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

On dira que  $X$  est  $\lambda$ -positive si  $\lambda_k \geq 0$  pour tout  $k \geq 1$ .

On suppose dans cette sous-partie que  $X$  est  $\lambda$ -positive.

- Pour tout  $k \in \mathbb{N}^*$ , montrer que  $\lambda_k \leq \frac{\mathbb{P}(X = k)}{\mathbb{P}(X = 0)}$ . En déduire que la série  $\sum \lambda_k$  converge.
- Montrer que, pour tout  $t \in [-1, 1]$ ,  $G_X(t) = \exp(H_X(t))$  et que  $\sum_{k=1}^{\infty} \lambda_k = -\ln(\mathbb{P}(X = 0))$ .
- Soit  $(X_i)$  la suite de variables aléatoires définie au II.C.3. Montrer que  $X \sim \sum_{i=1}^{\infty} iX_i$ .

### III.C - Caractérisation des variables entières infiniment divisibles

Soit  $X$  une variable aléatoire infiniment divisible à valeurs dans  $\mathbb{N}$  et telle que  $\mathbb{P}(X = 0) > 0$ .

Le but de cette sous-partie est de montrer que les trois assertions suivantes sont équivalentes.

(i)  $X$  est infiniment divisible;

(ii)  $X$  est  $\lambda$ -positive;

(iii) il existe une suite  $(X_i)_{i \geq 1}$  de variables de Poisson indépendantes, comme au II.C.3, telle que  $X \sim \sum_{i=1}^{\infty} iX_i$ .

Dans les questions III.C.1 à III.C.4, on suppose que  $X$  est une variable aléatoire infiniment divisible à valeurs dans  $\mathbb{N}$  et telle que  $\mathbb{P}(X = 0) > 0$ . Pour tout  $n \in \mathbb{N}^*$ , il existe donc  $n$  variables aléatoires indépendantes  $X_{n,1}, \dots, X_{n,n}$  de même loi telles que la variable aléatoire  $X_{n,1} + \dots + X_{n,n}$  suive la loi de  $X$ .

1. (a) Pour tout  $n \in \mathbb{N}^*$ , montrer que  $X_{n,1}$  est presque sûrement positive ou nulle.  
 (b) Pour tout  $n \in \mathbb{N}^*$ , montrer que  $\mathbb{P}(X_{n,1} = 0) > 0$ .  
 (c) Montrer que les variables aléatoires  $X_{n,i}$  sont presque sûrement à valeurs dans  $\mathbb{N}$ .
2. (a) Montrer  $\lim_{n \rightarrow \infty} \mathbb{P}(X_{n,1} = 0) = 1$ .  
 (b) En déduire que, pour tout  $i \in \mathbb{N}^*$ ,  $\lim_{n \rightarrow \infty} \mathbb{P}(X_{n,1} = i) = 0$ .
3. Soit  $H_X$  la série entière auxiliaire de  $X$ , comme elle est définie à la question III.A.4, et soit  $\rho(X)$  son rayon de convergence.

Pour tout  $n \in \mathbb{N}^*$ , soit  $H_n$  la série entière auxiliaire de  $X_{n,1}$ .

- (a) Pour tout  $n \in \mathbb{N}^*$ , montrer  $nH_n = H_X$ .  
 (b) En déduire, pour tous  $n$  et  $k$  dans  $\mathbb{N}^*$

$$kn\mathbb{P}(X_{n,1} = k) = \sum_{j=1}^k j\lambda_j \mathbb{P}(X_{n,1} = k-j)$$

4. Pour tout  $k \in \mathbb{N}^*$ , montrer que la suite  $(n\mathbb{P}(X_{n,1} = k))_{n \in \mathbb{N}^*}$  converge vers  $\lambda_k$ . En déduire que  $X$  est  $\lambda$ -positive.

## 5. Conclusion

- (a) Montrer le résultat annoncé au début de cette sous-partie III.C.  
 (b) Comment adapter ce résultat aux variables aléatoires à valeurs dans  $\mathbb{N}^*$  ?  
 (c) Soit  $X$  une variable aléatoire suivant la loi géométrique  $\mathcal{G}(p)$ , où  $p \in ]0, 1[$  :

$$\forall k \in \mathbb{N}^* \quad \mathbb{P}(X = k) = (1 - p)^{k-1} p$$

La variable aléatoire  $X$  est-elle infiniment divisible ?