cours 19, le jeudi 1er avril 2010

Exemple : calcul de l'espérance d'une variable aléatoire T de loi exponentielle de paramètre $\lambda > 0$. La loi $\mathrm{P_T}$ de T est donnée par $\mathrm{dP_T}(x) = \mathbf{1}_{x\ge 0}\lambda\,\mathrm{e}^{-\lambda x}\,\mathrm{d}x$; d'après la « formule de transfert », appliquée à $f(x) = x$, on a $f(\mathrm{T}) = \mathrm{T}$ et

$$\mathrm{E\,T} = \mathrm{E}(f\circ \mathrm{T}) = \int_\Omega (f\circ\mathrm{T})\,\mathrm{dP} = \int_{\mathbb{R}} f(x)\,\mathrm{dP_T}(x)$$
$$= \int_0^{+\infty} x\lambda\,\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda}\int_0^{+\infty} u\,\mathrm{e}^{-u}\,\mathrm{d}u = \frac{1}{\lambda}\cdot$$

Moments

Soit X une variable aléatoire (en abrégé : v.a.) sur un espace de probabilité $(\Omega, \mathcal{F}, \mathrm{P})$; pour tout entier $n \ge 1$, le *moment d'ordre* n de X est l'espérance $\mathrm{E\,X}^n$ de la puissance X^n (sous-réserve d'existence, garantie par la finitude du « moment absolu », $\mathrm{E}\,|\mathrm{X}|^n < +\infty$).

On peut considérer pour tout réel $p > 0$ la quantité

$$\mathrm{E}\,|\mathrm{X}|^p$$

finie ou égale à $+\infty$.

Lemme. *La fonction*

$$p \;\rightarrow\; \bigl(\mathrm{E}\,|\mathrm{X}|^p\bigr)^{1/p}$$

est croissante sur $]0, +\infty[$.

Preuve. — On commence par comparer les valeurs en 1 et $r > 1$; on introduit l'exposant conjugué s défini par $1/s = 1 - 1/r$ et on trouve d'après l'inégalité de Hölder, pour toute variable aléatoire Y positive

$$\mathrm{E\,Y} = \int_\Omega \mathrm{Y}(\omega).1\,\mathrm{dP}(\omega) \le \Bigl(\int_\Omega \mathrm{Y}(\omega)^r\,\mathrm{dP}(\omega)\Bigr)^{1/r}\Bigl(\int_\Omega 1^s\,\mathrm{dP}(\omega)\Bigr)^{1/s} =$$
$$= \Bigl(\mathrm{E\,Y}^r\Bigr)^{1/r}\bigl(\mathrm{P}(\Omega)\bigr)^{1/s} = \Bigl(\mathrm{E\,Y}^r\Bigr)^{1/r};$$

on traite le cas général $0 < p < q < +\infty$ en introduisant $r = q/p > 1$ et $\mathrm{Y} = |\mathrm{X}|^p$; alors $\mathrm{Y}^r = |\mathrm{X}|^q$ et l'inégalité précédente donne

$$\mathrm{E}\,|\mathrm{X}|^p \le \bigl(\mathrm{E}\,|\mathrm{X}|^q\bigr)^{p/q},$$

d'où le résultat.

Ainsi, pour une probabilité P, on a les inclusions

$$\mathrm{L}^\infty(\Omega, \mathcal{F}, \mathrm{P}) \subset \mathrm{L}^2(\Omega, \mathcal{F}, \mathrm{P}) \subset \mathrm{L}^1(\Omega, \mathcal{F}, \mathrm{P}).$$

Exercice. Si X est une variable aléatoire presque sûrement bornée, montrer qu'on a l'égalité $\|\mathrm{X}\|_\infty = \lim_{p\to+\infty} \|\mathrm{X}\|_p$.

Variance

Si $X \in L^2(\Omega, \mathcal{F}, P)$, on définit la *variance de* X, notée Var(X), par

$$\mathrm{Var}(X) = E(X - E\,X)^2 = E\,X^2 - (E\,X)^2.$$

La variance donne une idée de la concentration d'une v.a. autour de sa valeur moyenne $E\,X$; la variance de X est nulle si et seulement si X est presque sûrement constante, égale à sa moyenne. Un contrôle quantitatif est donnée par l'inégalité de Tchebychev plus bas.

La fonction constante sur Ω qui est égale à la moyenne $E\,X$ est la projection orthogonale de X sur le sous-espace de L^2 formé des fonctions constantes : si c est une fonction constante, la différence $X - E\,X$ est orthogonale à c dans $L^2(\Omega, \mathcal{F}, P)$,

$$\langle X - E\,X, c\rangle = \int_\Omega (X - E\,X)\,\overline{c}\,dP = E\big((X - E\,X)\,\overline{c}\big) = E(X\overline{c}) - (E\,X)\,\overline{c} = 0\,;$$

la variance est donc le carré de la distance L^2 de X aux constantes.

Exemple. Si G est une variable gaussienne centrée réduite, c'est-à-dire que sa loi est $(2\pi)^{-1/2}\,e^{-x^2/2}\,dx$, on a vu que $E\,G = 0$ et $E\,G^2 = 1$; on a donc

$$\mathrm{Var}(G) = E(G - E\,G)^2 = E\,G^2 = 1.$$

Inégalité de Tchebychev

Proposition. *Pour toute variable aléatoire* X *de carré intégrable et tout* $a > 0$, *on a*

$$P(|X - E\,X| > a) \le \frac{1}{a^2}\,\mathrm{Var}(X).$$

Preuve. — C'est une conséquence immédiate de l'inégalité de Markov, appliquée à l'intégrale de $Y = (X - E\,X)^2$ et à la valeur a^2,

$$P(|X - E\,X| > a) = P(|X - E\,X|^2 > a^2) = P(Y > a^2) \le \frac{1}{a^2}\int_\Omega Y\,dP = \frac{1}{a^2}\,E(X - E\,X)^2.$$

Loi jointe

Si X et Y sont deux v.a. définies sur le même espace $(\Omega, \mathcal{F}, P)$, la loi jointe $P_{(X,Y)}$ est la loi image de P par l'application couple $\omega \to (X(\omega), Y(\omega))$ à valeurs dans $\mathbb{R}^2$. La « formule de transfert » pour le couple dit que

$$E\,f(X, Y) = \int_{\mathbb{R}^2} f(x, y)\,dP_{(X,Y)}(x, y)$$

pour toute fonction borélienne $f : \mathbb{R}^2 \to [0, +\infty]$.

Exemple : calcul de l'espérance de $\inf(U_1, U_2)$, variables uniformes sur $[0, 1]$, indépendantes. Comme U_1 et U_2 sont indépendantes, la loi jointe est le produit tensoriel de deux copies de la loi uniforme $\mathbf{1}_{[0,1]}(x)\,dx$, donc

$$E\inf(U_1, U_2) = \int_{\mathbb{R}^2} \inf(u_1, u_2)\,\mathbf{1}_{[0,1]}(u_1)\,du_1\,\mathbf{1}_{[0,1]}(u_2)\,du_2.$$

On peut appliquer Fubini positif,

$$\mathrm{E}\inf(\mathrm{U}_1,\mathrm{U}_2)=\int_{\mathbb{R}}\Big(\int_{\mathbb{R}}\inf(u_1,u_2)\,\mathbf{1}_{[0,1]}(u_1)\,\mathbf{1}_{[0,1]}(u_2)\,\mathrm{d}u_1\Big)\,\mathrm{d}u_2$$

$$=\int_{\mathbb{R}}\Big(\int_0^{u_2}u_1\,\mathrm{d}u_1+u_2\int_{u_2}^1\mathrm{d}u_1\Big)\,\mathbf{1}_{[0,1]}(u_2)\,\mathrm{d}u_2=\int_0^1\big(u_2^2/2+u_2(1-u_2)\big)\,\mathrm{d}u_2$$

$$=\int_0^1(u_2-u_2^2/2)\,\mathrm{d}u_2=1/2-1/6=1/3.$$

Marginales de la loi jointe

Si Q est une loi de probabilité sur $(\mathbb{R}^2,\mathcal{B}_{\mathbb{R}^2})$, on appelle *lois marginales* de Q les deux images Q_1 et Q_2 de la mesure Q par les deux projections de $\mathbb{R}^2$ sur $\mathbb{R}$ données par les deux applications coordonnées,

$$\pi_1:(x_1,x_2)\in\mathbb{R}^2\to x_1,\quad \pi_2:(x_1,x_2)\in\mathbb{R}^2\to x_2.$$

Les deux mesures $\mathrm{Q}_j=\pi_j(\mathrm{Q})$, $j=1,2$ sont des probabilités sur $(\mathbb{R},\mathcal{B})$.

Si on compose l'application couple (X,Y) avec la première projection π_1 de $\mathbb{R}^2$ sur $\mathbb{R}$, on retrouve X (et Y avec la deuxième projection). Il en résulte que l'image de $\mathrm{P}_{(\mathrm{X},\mathrm{Y})}$ par la première projection est la loi P_X de X. De même l'image de $\mathrm{P}_{(\mathrm{X},\mathrm{Y})}$ par la deuxième projection π_2 est la loi P_Y de Y. La connaissance de la loi du couple est (c'est pas étonnant) une donnée plus précise que les simples lois de chacune des v.a.

Exemple : on peut obtenir de très nombreuses lois sur le carré $[0,1]^2$ dont les deux marginales sont égales à la mesure de Lebesgue sur $[0,1]$, par exemple la loi du « couple » $(\mathrm{U}_1,\mathrm{U}_1)$, portée par la diagonale $u_1=u_2$ du carré, ou bien la mesure de Lebesgue du carré, si le couple $(\mathrm{U}_1,\mathrm{U}_2)$ est formé de deux uniformes indépendantes.

VI.2.2. *Indépendance de familles de tribus ou de v.a.*

Définition. Soit $(\Omega,\mathcal{F},\mathrm{P})$ un espace de probabilité ; on dit que les sous-tribus $\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_n$ de $\mathcal{F}$ sont indépendantes si pour tous $\mathrm{A}_1,\ldots,\mathrm{A}_n$ tels que $\mathrm{A}_j\in\mathcal{A}_j$, $j=1,\ldots,n$, on a

$$\mathrm{P}\Big(\bigcap_{j=1}^n\mathrm{A}_j\Big)=\prod_{j=1}^n\mathrm{P}(\mathrm{A}_j).$$

On dit que les sous-tribus $(\mathcal{A}_n)_{n\geqslant 0}$ de $\mathcal{F}$ sont indépendantes si toutes les sous-familles finies sont indépendantes.

Définition. Soit $(\Omega,\mathcal{F},\mathrm{P})$ un espace de probabilité ; on dit que les v.a. $\mathrm{X}_1,\mathrm{X}_2,\ldots,\mathrm{X}_n$ sont indépendantes si les tribus $\sigma(\mathrm{X}_1),\sigma(\mathrm{X}_2),\ldots,\sigma(\mathrm{X}_n)$ sont indépendantes ; cela revient à dire que pour tous boréliens $\mathrm{B}_1,\ldots,\mathrm{B}_n$ de $\mathbb{R}$, on a

$$\mathrm{P}\Big(\bigcap_{j=1}^n\{\mathrm{X}_j\in\mathrm{B}_j\}\Big)=\prod_{j=1}^n\mathrm{P}(\mathrm{X}_j\in\mathrm{B}_j).$$

On dit que les v.a. $(\mathrm{X}_n)_{n\geqslant 0}$ de $\mathcal{F}$ sont indépendantes si toutes les sous-familles finies sont indépendantes.

On dira aussi qu'une variable aléatoire X est indépendante d'une sous-tribu $\mathcal{A}$ si les tribus $\sigma(\mathrm{X})$ et $\mathcal{A}$ sont indépendantes.

Lemme. *Soit* $(\Omega, \mathcal{F}, \mathrm{P})$ *un espace de probabilité ; si un événement* $\mathrm{A} \in \mathcal{F}$ *est indépendant de tous les événements* $\mathrm{C} \in \mathcal{C}$ *d'une classe* $\mathcal{C}$ *stable par intersection finie, alors* A *est indépendant de* $\sigma(\mathcal{C})$.

Preuve. — Considérons les deux mesures μ et ν sur $(\Omega, \mathcal{F})$ définies par

$$\mathrm{d}\mu(\omega) = \mathbf{1}_{\mathrm{A}}(\omega)\,\mathrm{dP}(\omega), \quad \mathrm{d}\nu(\omega) = \mathrm{P(A)}\,\mathrm{dP}(\omega).$$

Pour tout ensemble $\mathrm{B} \in \mathcal{F}$,

$$\mu(\mathrm{B}) = \int_\Omega \mathbf{1}_{\mathrm{B}}\mathbf{1}_{\mathrm{A}}\,\mathrm{dP} = \int_\Omega \mathbf{1}_{\mathrm{B}\cap\mathrm{A}}\,\mathrm{dP} = \mathrm{P}(\mathrm{B}\cap\mathrm{A})\,; \quad \nu(\mathrm{B}) = \mathrm{P(A)}\,\mathrm{P(B)}.$$

Ces deux mesures ont la même masse totale,

$$\mu(\Omega) = \mathrm{P}(\Omega\cap\mathrm{A}) = \mathrm{P(A)}, \quad \nu(\Omega) = \mathrm{P(A)}\,\mathrm{P}(\Omega) = \mathrm{P(A)},$$

et elles coïncident sur la classe $\mathcal{C}$, d'après l'hypothèse d'indépendance,

$$\forall \mathrm{C} \in \mathcal{C}, \quad \mu(\mathrm{C}) = \mathrm{P}(\mathrm{C}\cap\mathrm{A}) = \mathrm{P(C)}\,\mathrm{P(A)} = \nu(\mathrm{C}).$$

D'après un lemme d'unicité de mesure (qui résulte du théorème de classe monotone), on déduit que $\mu = \nu$ sur la tribu $\sigma(\mathcal{C})$, ce qui donne le résultat voulu.

Corollaire. *Pour que les variables* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *soient indépendantes, il (faut et il) suffit que pour tout sous-ensemble* J *de* $\{1, \ldots, n\}$ *et tout choix de réels* $(c_j)_{j\in\mathrm{J}}$, *on ait*

$$(*) \qquad \mathrm{P}\Big(\bigcap_{j\in\mathrm{J}} \{\mathrm{X}_j > c_j\}\Big) = \prod_{j\in\mathrm{J}} \mathrm{P}(\mathrm{X}_j > c_j).$$

Preuve. — Il est évident que la condition $(*)$ est nécessaire pour que les v.a. soient indépendantes. Inversement, supposons que la propriété $(*)$ est vraie. On va montrer par récurrence sur $k = 0, 1, \ldots, n$ la propriété suivante : si J_0 et J_1 sont deux sous-ensembles de $\{1, \ldots, n\}$, J_0 et J_1 disjoints et J_1 de cardinal $\le k$, si $(c_j)_{j\in\mathrm{J}_0}$ est une famille de réels et si $(\mathrm{B}_j)_{j\in\mathrm{J}_1}$ est une famille de boréliens de $\mathbb{R}$, alors

$$\mathrm{P}\Big(\Big(\bigcap_{j\in\mathrm{J}_0} (\mathrm{X}_j > c_j)\Big) \,\&\, \Big(\bigcap_{j\in\mathrm{J}_1} (\mathrm{X}_j \in \mathrm{B}_j)\Big)\Big) = \Big(\prod_{j\in\mathrm{J}_0} \mathrm{P}(\mathrm{X}_j > c_j)\Big) \Big(\prod_{j\in\mathrm{J}_1} \mathrm{P}(\mathrm{X}_j \in \mathrm{B}_j)\Big).$$

Arrivés à $k = n$, le corollaire sera prouvé.

Le cas où $k = 0$ est donné par l'hypothèse $(*)$. Supposons la propriété vraie pour $k < n$, et considérons $\mathrm{J}_0, \mathrm{J}_2$ disjoints avec J_2 de cardinal $k+1$; si on choisit $i \in \mathrm{J}_2$, on peut écrire $\mathrm{J}_2 = \mathrm{J}_1 \cup \{i\}$ avec J_1 de cardinal $\le k$; posons

$$\mathrm{A}_0 = \bigcap_{j\in\mathrm{J}_0} (\mathrm{X}_j > c_j), \quad \mathrm{A}_1 = \bigcap_{j\in\mathrm{J}_1} (\mathrm{X}_j \in \mathrm{B}_j)$$

et $\mathrm{A} = \mathrm{A}_1 \cap \mathrm{A}_2$. D'après l'hypothèse de récurrence, on a

$$\mathrm{P(A)} = \Big(\prod_{j\in\mathrm{J}_0} \mathrm{P}(\mathrm{X}_j > c_j)\Big) \Big(\prod_{j\in\mathrm{J}_1} \mathrm{P}(\mathrm{X}_j \in \mathrm{B}_j)\Big).$$

D'après l'hypothèse de récurrence appliquée à $J_0 \cup \{i\}$ et J_1 (de cardinal k), l'ensemble A est indépendant de tous les ensembles $C = \{X_i > c_i\}$, où c_i varie dans $\mathbb{R}$. D'après le lemme, A est indépendant des ensembles de la tribu $\mathcal{T}$ engendrée par les $\{X_i > c_i\}$, et $\mathcal{T}$ contient (en fait, est égale à la classe de tous) les ensembles $C = \{X_i \in B_i\}$. On a donc

$$P\Big(\Big(\bigcap_{j \in J_0} (X_j > c_j)\Big) \,\&\, \Big(\bigcap_{j \in J_1 \cup \{i\}} (X_j \in B_j)\Big)\Big) = P(A \cap C)$$

$$= P(A)P(C) = \Big(\prod_{j \in J_0} P(X_j > c_j)\Big)\Big(\prod_{j \in J_1} P(X_j \in B_j)\Big) P(X_i \in B_i),$$

ce qui termine la récurrence.

Développement dyadique

On sait que tout nombre réel x tel que $0 \leq x < 1$ peut se représenter sous la forme

$$x = \sum_{n=1}^{+\infty} \frac{\varepsilon_n}{2^n},$$

où $\varepsilon_n = 0, 1$ est obtenu en développant x en base 2. On va montrer que les « décimales » dyadiques $(\varepsilon_n)_{n \geqslant 1}$ peuvent être considérées comme une suite de variables aléatoires indépendantes.

Définissons une fonction φ sur $\mathbb{R}$, qui ne prend que les valeurs 0 et 1, en posant $\varphi(x) = 0$ sur $[0, 1[$ et $\varphi(x) = 1$ sur $[1, 2[$, et en disant que φ est périodique de période 2 sur $\mathbb{R}$; on a donc

$$\varphi = \sum_{j \in \mathbb{Z}} \mathbf{1}_{[2j-1, 2j[},$$

et on a aussi

$$\{x \in \mathbb{R} : \varphi(x) = 0\} = \bigcup_{j \in \mathbb{Z}} [2j, 2j+1[,$$

c'est-à-dire que $\varphi(x) = 0$ si et seulement si la partie entière de x est paire. On considère l'espace $\Omega = [0, 1[$, muni des boréliens et de la mesure de Lebesgue (qui est une probabilité sur cet espace Ω). Pour tout entier $n \geq 1$ et tout $\omega \in \Omega$, on pose

$$X_n(\omega) = \varphi(2^n \omega)\,;$$

on va montrer que ces v.a. forment une suite (infinie) $(X_n)_{n \geqslant 1}$ de variables aléatoires indépendantes.

Identifions la loi de X_n : la v.a. X_n ne prend que les valeurs $0, 1$, et elle les prend avec probabilité $1/2$; en effet, l'ensemble $(X_n = 0)$ est l'ensemble des ω réels tels que $0 \leq \omega < 1$ et $\varphi(2^n\omega) = 0$, donc c'est l'ensemble des $\omega \in \Omega$ pour lesquels il existe $j \in \mathbb{Z}$ tel que $2j \leq 2^n\omega < 2j + 1$; on a $0 \leq 2^n\omega < 2j + 1$ qui entraîne $j \geq 0$ et $2j \leq 2^n\omega < 2^n$ qui entraîne $j < 2^{n-1}$, donc

$$2^n\omega \in \bigcup_{0 \leq j < 2^{n-1}} [2j, 2j+1[$$

ce qui montre que

$$\{\mathrm{X}_n = 0\} = \bigcup_{0 \le j < 2^{n-1}} [2j2^{-n}, (2j+1)2^{-n}[\,;$$

l'ensemble $\{\mathrm{X}_n = 0\}$ est formé de 2^{n-1} intervalles disjoints, chacun de longueur 2^{-n}, par conséquent

$$\mathrm{P}(\mathrm{X}_n = 0) = 2^{n-1}2^{-n} = 1/2.$$

Comme X_n ne prend que les valeurs 0 et 1, on a aussi $\mathrm{P}(\mathrm{X}_n = 1) = 1/2$.

On va montrer par récurrence sur $n \ge 1$ les deux affirmations suivantes : *la tribu $\mathcal{F}_n$ engendrée par $\mathrm{X}_1, \ldots, \mathrm{X}_n$ est engendrée par la partition*

$$\Pi_n = \big([j2^{-n}, (j+1)2^{-n}[\big)_{0 \le j < 2^n}$$

de Ω en 2^n atomes, et X_n est indépendante de $\mathcal{F}_{n-1}$.

Quand $n = 1$, la tribu $\mathcal{F}_1$ est engendrée par X_1, qui prend la valeur 0 sur $[0, 1/2[$ et la valeur 1 sur $[1/2, 1[$; la partition Π_1 de $[0, 1[$ est formée de $[0, 1/2[$ et $[1/2, 1[$; la tribu $\mathcal{F}_1$ engendrée par X_1 contient les quatre ensembles engendrés par cette partition, à savoir vide, plein et les deux intervalles précédents : c'est la tribu engendrée par la partition Π_1. Il est naturel de poser $\mathcal{F}_0 = \{\emptyset, \Omega\}$, et n'importe quelle variable aléatoire, en particulier X_1, est indépendante de cette tribu « grossière » $\mathcal{F}_0$. Le cas $n = 1$ est réglé.

Supposons la propriété vraie à l'ordre n : tout ensemble A de la tribu $\mathcal{F}_n$ est de la forme

$$\mathrm{A} = \bigcup_{j \in \mathrm{J}} \mathrm{A}_j, \quad \text{où } \ \mathrm{A}_j = [j2^{-n}, (j+1)2^{-n}[,$$

$0 \le j < 2^n$ et où J est un sous-ensemble de $\{0, \ldots, 2^n - 1\}$. La tribu $\mathcal{F}_{n+1}$ est la plus petite tribu qui contienne tous les ensembles de la forme $(\mathrm{X}_k \in \mathrm{B}_k)$ où B_k est un borélien de $\mathbb{R}$ et $k = 1, \ldots, n+1$. La tribu $\mathcal{F}_n$ contient déjà tous ces ensembles pour $k \le n$, on va voir comment il faut la raffiner pour traiter X_{n+1} aussi. Comme X_{n+1} ne prend que les valeurs 0 et 1, il suffit de garantir que l'ensemble $(\mathrm{X}_{n+1} = 0)$ soit dans la tribu $\mathcal{F}_{n+1}$: on obtiendra $(\mathrm{X}_{n+1} = 1)$ par complémentaire ; les autres ensembles de la forme $(\mathrm{X}_{n+1} \in \mathrm{B}_{n+1})$ sont $\emptyset$ et Ω, qui sont déjà dans $\mathcal{F}_n$, donc dans $\mathcal{F}_{n+1}$.

L'ensemble $(\mathrm{X}_{n+1} = 0)$ est la réunion pour $0 \le j < 2^n$ des ensembles

$$\mathrm{A}_j \cap (\mathrm{X}_{n+1} = 0) = \{\omega : (j2^{-n} \le \omega < (j+1)2^{-n}) \,\&\, (\mathrm{X}_{n+1}(\omega) = 0)\}.$$

Si $j2^{-n} \le \omega < (j+1)2^{-n}$ et $\mathrm{X}_{n+1}(\omega) = \varphi(2^{n+1}\omega) = 0$, on sait que $2j \le 2^{n+1}\omega < 2j+2$ et que $2^{n+1}\omega$ appartient à un intervalle pair $[2i, 2i+1[$ qui ne peut être que $[2j, 2j+1[$. On conclut que

$$\mathrm{A}_j \cap (\mathrm{X}_{n+1} = 0) = [2j2^{-n-1}, (2j+1)2^{-n-1}[.$$

On voit donc que l'ensemble $(\mathrm{X}_{n+1} = 0)$ est réunion de certains des intervalles de la partition Π_{n+1}. Inversement, la tribu $\mathcal{F}_{n+1}$ doit contenir $\mathcal{F}_n$, donc contenir tous les ensembles A_j, $0 \le j < 2^n$, donc tous les

$$\mathrm{A}_j \cap (\mathrm{X}_{n+1} = 0) = [2j2^{-n-1}, (2j+1)2^{-n-1}[$$

et tous les

$$\mathrm{A}_j \cap (\mathrm{X}_{n+1} = 1) = [(2j+1)2^{-n-1}, (2j+2)2^{-n-1}[.$$

On voit donc que $\mathcal{F}_{n+1}$ contient tous les ensembles de la partition Π_{n+1}, et coïncide par conséquent avec la tribu engendrée par Π_{n+1} ; cette partition a 2^{n+1} atomes et la tribu (finie) $\mathcal{F}_{n+1}$ contient $2^{2^{n+1}}$ ensembles.

Pour montrer que X_{n+1} est indépendante de $\mathcal{F}_n$, on doit montrer que

$$\mathrm{P}\big(\mathrm{A} \ \& \ (\mathrm{X}_{n+1} = 0)\big) = \mathrm{P}(\mathrm{A})\,\mathrm{P}(\mathrm{X}_{n+1} = 0) = \mathrm{P}(\mathrm{A})/2$$

pour tout $\mathrm{A} \in \mathcal{F}_n$; comme tout $\mathrm{A} \in \mathcal{F}_n$ est de la forme $\mathrm{A} = \bigcup_{j\in \mathrm{J}} \mathrm{A}_j$, réunion disjointe, il suffit de le faire sur chaque bout A_j, $j \in \mathrm{J}$. Mais on vient de voir que

$$\mathrm{A}_j \cap (\mathrm{X}_{n+1} = 0) = [2j2^{-n-1}, (2j+1)2^{-n-1}[$$

qui est de longueur

$$2^{-n-1} = 2^{-n} \times \frac{1}{2} = \mathrm{P}(\mathrm{A}_j)\,\mathrm{P}(\mathrm{X}_{n+1} = 0).$$

On a montré que X_{n+1} est indépendante de $\mathcal{F}_n$ et terminé la récurrence.

Terminons la preuve de l'indépendance des X_n : si N entier quelconque est donné, ainsi que des boréliens (B_k), $k = 1, \ldots, \mathrm{N}$, on note que par la preuve précédente, l'ensemble $(\mathrm{X}_\mathrm{N} \in \mathrm{B}_\mathrm{N})$ est indépendant de $\bigcap_{1\le k<\mathrm{N}} (\mathrm{X}_k \in \mathrm{B}_k)$ qui est dans la tribu $\mathcal{F}_{\mathrm{N}-1}$, donc

$$\mathrm{P}\Big(\bigcap_{1\le k\le \mathrm{N}} (\mathrm{X}_k \in \mathrm{B}_k)\Big) = \mathrm{P}\Big(\Big(\bigcap_{1\le k<\mathrm{N}} (\mathrm{X}_k \in \mathrm{B}_k)\Big) \,\&\, (\mathrm{X}_\mathrm{N} \in \mathrm{B}_\mathrm{N})\Big)$$

$$= \mathrm{P}\Big(\bigcap_{1\le k<\mathrm{N}} (\mathrm{X}_k \in \mathrm{B}_k)\Big)\,\mathrm{P}(\mathrm{X}_\mathrm{N} \in \mathrm{B}_\mathrm{N}),$$

et en recommençant avec $\mathrm{N}-1, \mathrm{N}-2, \ldots, 2$ on termine avec

$$\mathrm{P}\Big(\bigcap_{1\le k\le \mathrm{N}} (\mathrm{X}_k \in \mathrm{B}_k)\Big) = \prod_{k=1}^{\mathrm{N}} \mathrm{P}(\mathrm{X}_k \in \mathrm{B}_k).$$

Ainsi, on a prouvé que la suite $(\mathrm{X}_n)_{n\geqslant 1}$ est une suite de v.a. indépendantes.

Remarque. En algèbre linéaire, on peut traduire l'indépendance linéaire d'une suite de vecteurs $(v_n)_{n\geqslant 1}$ en disant que pour tout $n \ge 1$,

$$v_n \notin \mathrm{Vect}(v_1, \ldots, v_{n-1})$$

(où Vect() est interprété comme $\{0\}$ quand $n-1=0$). De façon analogue, on traduit l'indépendance d'une suite de variables aléatoires $(\mathrm{X}_n)_{n\geqslant 1}$ en disant que pour tout $n \ge 1$, la v.a. X_n est indépendante de la tribu engendrée par $\mathrm{X}_1, \ldots, \mathrm{X}_{n-1}$.

Remarque (un peu prématurée). Par construction les $\mathrm{X}_n(\omega)$ sont les « décimales » de la numération binaire, donc pour tout ω de $[0,1[$ on a l'égalité

$$\omega = \sum_{n=1}^{+\infty} 2^{-n}\mathrm{X}_n(\omega).$$

On a ainsi construit la variable aléatoire $\omega \in \Omega \to \omega \in \mathbb{R}$, de loi uniforme sur $[0,1]$, à partir d'une suite de v.a. de Bernoulli $p = 1/2$ indépendantes. On pourrait se convaincre que la loi d'une série de Bernoulli ne dépend pas de la réalisation particulière (on peut calculer toutes les probabilités des valeurs des sommes partielles de la série grâce à l'indépendance). Ainsi, pour toute sous-suite des (X_n) précédentes,

$$\sum_{k=1}^{+\infty} 2^{-k}\mathrm{X}_{n_k}$$

est une variable uniforme sur $[0,1]$.

À partir d'une suite de Bernoulli indépendantes, on peut donc construire une suite $(\mathrm{U}_m)_{m\geqslant 0}$ d'uniformes indépendantes (utiliser le fait que $\mathbb{N} \sim \mathbb{N}\times\mathbb{N}$, et le fait que des variables construites à partir de paquets disjoints pris dans une suite indépendante restent indépendantes), donc des tas d'autres suites indépendantes, par exemple des Bernoulli de paramètre $p \neq 1/2$ en considérant $\mathrm{Y}_m = \mathbf{1}_{[0,p]}(\mathrm{U}_m)$.