Variables aléatoires, Espérance, Indépendance

De l'analyse aux probabilités. Ici on considère un espace de probabilité Ω qui contient les événements w. Les ensembles mesurables seront donnés par la tribu $\mathcal{F}$ et on va travailler avec une mesure de masse 1 appelée probabilité et notée $\mathbb{P}$. Une fonction mesurable X sur $(\Omega, \mathcal{F})$ sera appelée une variable aléatoire et son intégrale contre $\mathbb{P}$ sera son espérance et notée $\mathbb{E}(X)$.
On dit qu'un évenement mesurable A a lieu $\mathbb{P}$ p.s. (pour 'presque surement') quand $\mathbb{P}(A) = 1$.

1 Variables aléatoires

Dans la suite, $(\Omega, \mathcal{F}, \mathbb{P})$ désigne un espace probabilisé.

1.1 Définitions

Définition 1.1. *Une application X de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ qui est mesurable, i.e.*

$$\forall A \in \mathcal{B}(\mathbb{R}^d),\ X^{-1}(A) = \{\omega \in \Omega;\ X(\omega) \in A\} \in \mathcal{F}$$

est appelée variable aléatoire.

Remarque 1.1. *Comme $\mathcal{B}(\mathbb{R}^d)$ est engendrée par $\{\prod_{i=1}^d (-\infty; \alpha_i);\ \alpha_i \in \mathbb{R}\}$, pour montrer que $X = (X_1, \dots, X_d)$ est une variable aléatoire à valeurs dans $\mathbb{R}^d$, il suffit de vérifier :*

$$\forall \alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d,\ \ \{\omega \in \Omega; \forall i \in \mathbb{N}_d,\ X_i(\omega) < \alpha_i\} \in \mathcal{F}$$

En particulier, ceci montre que X est une variable aléatoire à valeurs dans $\mathbb{R}^d$ ssi pour tout $i \in \mathbb{N}_d$, X_i est une variable aléatoire réelle (la raison profonde étant en fait que la tribu borélienne sur $\mathbb{R}^d$ est la tribu produit de d copies de $\mathbb{R}$).

Proposition 1.1. *L'ensemble des variables aléatoires sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une algèbre stable par* $\inf$, $\sup$, $\liminf_n$ *et* $\limsup_n$*.*
De plus, $\{\omega \in \Omega\ ;\ \lim_{n\to\infty} X_n(\omega)$ existe et appartient à $\mathbb{R}\} \in \mathcal{F}$.

Pour le voir, écrire par exemple $\{w : inf_{n\in\mathbb{N}} X_n(w) < a\} = \cup_{n\in\mathbb{N}}\{X_n(w) < a\}$, et

$$\begin{aligned}
\left\{\omega \in \Omega\ ;\ \lim_{n\to\infty} X_n(\omega) \text{ existe et appartient à} \mathbb{R}\right\} &= \{X_n \text{ suite de Cauchy}\} \\
&= \{\forall \epsilon > 0, \exists n_0 \in \mathbb{N} : \forall p \in \mathbb{N} |X_{n_0+p} - X_{n_0}| \leq \epsilon\} \\
&= \cap_{\epsilon>0} \cup_{n_0\in\mathbb{N}} \cap_{p\in\mathbb{N}} \{|X_{n_0+p} - X_{n_0}| \leq \epsilon\} \\
&= \cap_{n\in\mathbb{N}} \cup_{n_0\in\mathbb{N}} \cap_{p\in\mathbb{N}} \{|X_{n_0+p} - X_{n_0}| \leq 1/n\}
\end{aligned}$$

A partir de maintenant, on utilisera une forme abrégée pour écrire des événements faisant intervenir des variables aléatoires. Ainsi, pour noter par exemple l'événement $\{\omega \in \Omega \,;\; X(\omega) \leq C\}$, on écrira simplement $\{X \leq C\}$.

1.2 Tribu engendrée par une collection de variables aléatoires

Définition 1.2. *On se donne une famille $(Y_\gamma)_{\gamma \in \mathcal{C}}$ de variables aléatoires sur $(\Omega, \mathcal{F}, \mathbb{P})$ indexée par un ensemble $\mathcal{C}$ quelconque. La tribu engendrée par la famille $(Y_\gamma)_{\gamma \in \mathcal{C}}$, notée $\Gamma = \sigma(Y_\gamma \,;\, \gamma \in \mathcal{C})$, est définie comme la plus petite tribu sur Ω rendant mesurable chaque Y_γ pour $\gamma \in \mathcal{C}$. Autrement dit,*

$$\Gamma = \sigma\left(Y_\gamma^{-1}(B) : \; \gamma \in \mathcal{C}, \; B \in \mathcal{B}(\mathbb{R}^d)\right).$$

Proposition 1.2. *Soit X une variable aléatoire et $f : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ borélienne.*
Alors $f(X)$ est $\sigma(X)$-mesurable.
Réciproquement si une fonction $Y : \Omega \to \mathbb{R}$ est $\sigma(X)$-mesurable alors il existe une fonction borélienne f telle que $Y = f(X)$.

Cette propriété est une conséquence du théorème de la classe monotone. En effet c'est un espace vectoriel qui contient les indicatrices des ensembles $X^{-1}(A)$ puisque $1(X^{-1}(A))(w) = 1(A)(X(w))$ et $Y_n = f_n(X)$ croissante vers Y entraine $Y = \limsup f_n(X)$.
Elle est très importante dans la mesure où elle fait clairement comprendre ce que signifie "être $\sigma(X)$-mesurable".

1.3 Distributions et fonctions de répartition

Définition 1.3. *Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. On appelle mesure image de $\mathbb{P}$ par X ou encore loi de X la mesure de probabilité $\mathbb{P}_X$ sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ définie par :*

$$\forall A \in \mathcal{B}(\mathbb{R}^d), \; \mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A))$$

La loi de X est déterminée par la fonction suivante, appelée fonction de répartition.

Proposition 1.3. *Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$. On définit la fonction de répartition de X sur $\mathbb{R}^d$ par*

$$\forall x \in \mathbb{R}^d, \; F_X(x) = \mathbb{P}(X \leq x).$$

Deux variables aléatoires ayant même fonction de répartition ont même loi.

Démonstration. On sait que $\left\{\prod_{i=1}^d (-\infty; c_i] \,;\; c_i \in \mathbb{R}\right\}$ est un π-système qui génère $\mathcal{B}(\mathbb{R}^d)$. Par le théorème d'unicité des probabilités coïncidant sur un π-système. □

Les fonctions de répartitions jouissent de propriétés très caractéristiques, du moins dans le cadre de la dimension 1. Dans le reste de cette section, on se limite donc à la dimension 1.

Proposition 1.4. *Supposons que F soit la fonction de répartition d'une variable aléatoire réelle X. Alors :*
a) $F : \mathbb{R} \to [0, 1]$ est croissante.
b) $\lim_{x \to +\infty} F(x) = 1$ et $\lim_{x \to -\infty} F(x) = 0$.

c) F est continue à droite (et admet des limites à gauche).

Soit F une fonction ayant les propriétés a,b,c. Il existe alors une variable aléatoire X sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ telle que la fonction de répartition de X soit F.

Démonstration.
a) trivial car $\{X \leq x\}$ est croissant en x.
b) $\lim_{x\to+\infty} F(x) = \lim_{x\to+\infty} \mathbb{P}(\{\omega \in \Omega; X(\omega) \leq x\})$ et cette limite existe car F est croissante. Or $(A_n)_n = (\{X \leq n\})_n$ est une suite croissante d'événements tendant vers Ω. Donc par additivité dénombrable de $\mathbb{P}$, $\lim_{n\to} \mathbb{P}(A_n) = 1$. Par suite, $\lim_{x\to+\infty} F(x) = 1$. Un raisonnement analogue montre que $\lim_{x\to-\infty} F(x) = 0$.
c) Soit $(x_n)_n$ une suite de points tendant vers x avec pour tout $n \in \mathbb{N}$, $x_n \geq x$. On a $F(x_n) = \mathbb{P}(X \leq x_n)$ et $\{X \leq x_n\} \downarrow \{X \leq x\}$ donc par additivité dénombrable, $F(x_n) \downarrow F(x)$ et F est continue à droite. Le fait qu'elle admette des limites à gauches vient de sa monotonie.

Idée de la réciproque : la variable X cherchée va être distribuée comme $F^{-1}(U)$ où U v.a. uniforme sur $[0,1]$ puisque $\mathbb{P}(F^{-1}(U) \leq c) = \mathbb{P}(U \leq F(c)) = F(c)$.
Plus rigoureusement, on considère l'espace $(\Omega, \mathcal{F}, \mathbb{P}) = (]0,1[, \mathcal{B}(]0,1[), \lambda)$ où λ est la mesure de Lebesgue. On pose pour $\omega \in]0,1[$,

$$X(\omega) = \inf\{z \in \mathbb{R}\ ;\ F(z) \geq \omega\} = \sup\{y \in \mathbb{R}\ ;\ F(y) < \omega\}$$

Il est facile de voir d'après la définition de X que

$$\omega \leq F(c) \Longleftrightarrow X(\omega) \leq c$$

Par conséquent, $\lambda(w : X(w) \leq c) = \lambda(\omega \leq F(c)) = F(c)$. □

1.4 Variables aléatoires discrètes et variables aléatoires à densités

Définition 1.4. *Une variable aléatoire X est dite discrète s'il existe un ensemble dénombrable $\mathcal{E}$ tel que $\mathbb{P}(X \in \mathcal{E}) = 1$. La loi de X est alors caractérisée par une suite $(p_e)_{e\in\mathcal{E}} \in [0,1]^{\mathcal{E}}$ telle que $\sum_{e\in\mathcal{E}} p_e = 1$ et $p_e = \mathbb{P}(X = e)$.*

Exemples fondamentaux :
1) X suit une loi de Bernoulli de paramètre $p \in [0,1]$ ssi :

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

C'est le résultat d'un tirage (eventuellement biaisé) à pile ou face.
2) X suite une loi binomiale de paramètres (n,p) ssi :

$$\forall 0 \leq k \leq n,\ \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

C'est le nombre de piles obtenus après n lancers.
3) X suit une loi géométrique de paramètre $p \in [0,1]$ ssi :

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

C'est le nombre de fois où vous devez lancer la pièce pour obtenir un pile.
C'est aussi la (seule) loi sans mémoire discrète, au sens où

$$\mathbb{P}(X > n + m | X > n) = \mathbb{P}(X > m) \qquad (n, m \geq 0).$$

4) X suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$ ssi :

$$\forall k \in \mathbb{N}, \; \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Elle modélise le nombre de clients qui se présentent au bureau de poste pendant un intervalle de temps donné. En effet si chaqun des n clients décide avec probabilité p_n de se rendre à la poste, indépendamment les uns des autres, alors le nombre de personne qui se rendent à la poste est donné par une binomiale de paramètre (n, p_n) et quand n temps vers l'infini et p_n tend vers zéro avec $p_n \sim \lambda/n$, on retrouve la loi de poisson

Définition 1.5. *On dira qu'une variable aléatoire X est à densité f (par rapport à la mesure de Lebesgue) ssi*

$$\forall A \in \mathcal{B}(\mathbb{R}^d), \quad \mathbb{P}_X(A) = \int_A f(x)dx$$

Exemples fondamentaux :
0) X est une loi uniforme sur $[a, b]$ ssi X admet pour densité

$$f(x) = \mathbb{1}_{[a,b]}(x) \frac{1}{b-a}.$$

1) X est une gaussienne de moyenne $m \in \mathbb{R}^d$ et de matrice de covariance $\Gamma \in S_d^+$ (ensemble des matrices symétriques définies positives) ssi X admet la densité sur $\mathbb{R}^d$ donnée par

$$f(x) = \frac{1}{(2\pi)^{d/2} |det\Gamma|^{1/2}} \exp\left(-\frac{1}{2} < \Gamma^{-1}(x - m), (x - m) >\right)$$

où $< \cdot, \cdot >$ désigne le produit scalaire euclidien usuel sur $\mathbb{R}^d$.
La gaussienne a quelque chose d'universelle puisqu'elle donne la distribution de la somme d'une grande quantité de variable iid centrées (voir TLC). Elle joue ainsi un role important pour quantifier l'erreur dans un sondage. Pour la voir apparaitre physiquement, se rendre à la cité des sciences ou des billes tombent sur une grille de clous et percutent les clous successifs pour passer (de manière équiprobable et indépendante) à gauche ou à droite de chaque clou. Le tas de caillou formé à la forme d'une gaussienne. La gaussienne de densité sur $\mathbb{R}$:

$$f(x) = \frac{e^{-(x-m)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

2) X est une exponentielle de paramètre λ ssi X admet la densité sur $\mathbb{R}$ donnée par

$$f(x) = 1_{\mathbb{R}^+}(x) \lambda e^{-\lambda x}$$

C'est la loi sans mémoire, c'est à dire la seule à verifier que

$$\mathbb{P}(X > t + s \mid X > t) = \mathbb{P}(X > s)$$

Elle permet donc de modéliser des variables telles que le prochain moment où on gagne à un jeu, ou un accident se produit mais aussi certaines durées de vie comme le homard (qui ne vieillit pas) ou des composants qui ne s'usent pas mais peuvent casser.

2 Espérance

Définition 2.1. *L'espérance d'une variable aléatoire est donnée par* $\mathbb{E}(X) = \int_\Omega X(\omega)\mathbb{P}(d\omega)$. *Une v.a. est dite intégrable si* $\mathbb{E}(|X|) < \infty$.

Remarque 2.1. *i) Notons que si* $\mathbb{P}(X \in \{x_i : i \in \mathbb{N}\}) = 1$ *(v.a. discrète) alors*

$$\mathbb{E}(X) = \sum_{i\in\mathbb{N}} \mathbb{P}(X = x_i)x_i.$$

(ii) Si $\mathbb{P}(X \in dx) = f(x)dx$ *(v.a. à densité* f*) alors*

$$\mathbb{E}(X) = \int_{\mathbb{R}^d} xf(x)dx.$$

(iii) Si X *est intégrable, alors* $\mathbb{P}(X = \infty) = 0$.
Si X *et* Y *sont intégrables et* $X = Y$ *p.s. alors* $\mathbb{E}(X) = \mathbb{E}(Y)$.
Si $(Z_k)_k$ *est une suite de variables aléatoires positives alors*

$$\mathbb{E}\left(\sum_k Z_k\right) = \sum_k \mathbb{E}(Z_k)$$

Si $(Z_k)_k$ *est une suite de variables aléatoires positives telles que* $\sum_k \mathbb{E}(Z_k) < +\infty$ *alors presque sûrement, on a :*

$$\sum_k Z_k < +\infty \quad et \quad Z_k \to 0$$

(iv) Si X, Y *sont deux variables aléatoires de carrés intégrables, on appelle covariance de* X *et* Y *le réel défini par*

$$Cov(X,Y) = \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

La variance de X *est* $Var(X) = \sigma_X^2 = Cov(X,X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \geq 0$.

De plus si $X = \mathbb{1}_A$, $\mathbb{E}(X) = \mathbb{P}(A)$. On en déduit facilement que la loi de X est identifiée par les quantités $\mathbb{E}(f(X))$:

Proposition 2.1. *Soient* X *et* Y *deux v.a. Si pour toutes fonctions continues bornées* f *(ou continues à support compact, ou* $\mathcal{C}^\infty$ *à support compact, ou de la forme* $1(A)$ *pour* A *dans un* π*-système engendrant* $\mathcal{F}$*.) ,*

$$\mathbb{E}(f(X)) = \mathbb{E}(f(Y))$$

alors X *et* Y *ont même loi.*

C'est une manière pratique de montrer qu'une v.a. X a une loi donnée (par exemple pour obtenir sa densité par changement de variable). Pour des fonctions $f(X)$ où f est inversible et la loi de X connue ou bien pour $\max(X_i)$ où X_i sont indépendantes, on préférera souvent utiliser les fonctions de répartition pour trouver les lois.

Enfin en munissant $\Omega \times \mathbb{R}$ de $\mathbb{P} \otimes dx$ et en notant $A = \{(\omega, x) \ / \ 0 \leq x < X(\omega)\}$, A est $\mathcal{F} \otimes \mathcal{B}(\mathbb{R})$-mesurable et l'on a par Fubini :

$$\mu(A) = \int_{[0;+\infty)} \mathbb{P}(X > x)dx = \int_{\Omega} X(\omega)\mathbb{P}(d\omega) = \mathbb{E}(X)$$

et donc

$$\mathbb{E}(X) = \int_{\mathbb{R}^+} (1 - F(x))dx$$

Théorème 2.1. *Inégalité de Jensen Soit $\phi : G \to \mathbb{R}$ une fonction convexe et X une variable aléatoire $\mathbb{P}$ p.s. à valeurs dans G où G est un intervalle de $\mathbb{R}$ ouvert.*
On suppose que :

$$\mathbb{E}(|X|) < +\infty \ , \ \mathbb{E}(|\phi(X)|) < +\infty$$

Alors :

$$\phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X))$$

Démonstration. Pour $u < v < w$, on pose $\Delta_{u,v} = \dfrac{\phi(v) - \phi(u)}{v - u}$. Puisque ϕ est convexe,

$$\Delta_{u,v} \leq \Delta_{v,w}$$

$(D_-\phi)(v) = \lim_{u\uparrow v} \Delta_{u,v}$ et $(D_+\phi)(v) = \lim_{w\downarrow v} \Delta_{v,w}$ sont les dérivées à gauche et à droite de ϕ. On a donc

$$(D_-\phi)(v) \leq (D_+\phi)(v)$$

D'autre part, par un argument du même genre en prenant quatre points, $D_-\phi$, $D_+\phi$ sont croissantes.
Si $x < w < v$, $\Delta_{x,w} \leq \Delta_{w,v}$ et par passage à la limite quand w tend en croissant vers v, on obtient $\Delta_{x,v} \leq (D_-\phi)(v)$ car ϕ est continue sur G.
On a donc $\phi(v) - \phi(x) \leq (v - x)(D_-\phi)(v) \leq m(v - x)$ pour $m \in [(D_-\phi)(v); (D_+\phi)(v)]$.
En raisonnant de même si $v < x$, on a pour tout $v, x \in G$, $\phi(x) \geq m(x - v) + \phi(v)$ pour $m \in [(D_-\phi)(v); (D_+\phi)(v)]$.
Donc si $v = \mu = \mathbb{E}(X) \in G$ ($a < X < b \Rightarrow a < \mathbb{E}(X) < b$), et $x = X$, on a :

$$\phi(X) \geq m(X - \mu) + \phi(\mu)$$

et en intégrant on obtient

$$\mathbb{E}(\phi(X)) \geq \phi(\mu) = \phi(\mathbb{E}(X))$$

□

2.1 Espaces $\mathcal{L}^p$

On rappelle que $\mathcal{L}^p$ est l'espace vectoriel des variables aléatoires X telle que

$$\|X\|_p := \mathbb{E}(|X|^p)^{1/p}$$

est finie. On dit alors que X a un moment d'ordre p et on remarque que $\|.\|_p$ est une seminorme pour $\mathcal{L}^p$.
Une propriété particulière aux espaces de probabilités (par rapport aux espaces associés à une mesure quelconque) est la suivante

Proposition 2.2. *Si* $1 \leq p < r \leq +\infty$ *et si* $X \in \mathcal{L}^r$*, alors* $X \in \mathcal{L}^p$ *et on a*

$$\|X\|_p \leq \|X\|_r$$

Autrement dit, plus une variable aléatoire a des moments d'ordre élevé, mieux c'est. La démonstration facile repose sur l'inégalité de Jensen appliquée, après troncature, à la fonction convexe $\phi(x) = x^{r/p}$. Ou appliquer Holder à $1.X$ et $r/p > 1$.

On rappelle ces deux inégalités d'un usage constant et qui dérivent de l'inégalité de Jensen

Théorème 2.2. *Inégalité de Hölder*
Soit $1 < p, q < +\infty$*, tels que* $p^{-1} + q^{-1} = 1$ *et* $X \in \mathcal{L}^p, Y \in \mathcal{L}^q$*. On a alors l'inégalité de Hölder*

$$\mathbb{E}(|XY|) \leq \|X\|_p \|Y\|_q$$

Démonstration. On peut bien sûr se restreindre au cas $X, Y \geq 0$ et $\mathbb{E}(X^p) > 0$. D'autre part, quitte à considérer $X \wedge n$ et $Y \wedge n$, puis à passer à la limite quand n tend vers $+\infty$ par le théorème de convergence monotone, on peut supposer X, Y bornées. On définit une probabilité $\mathbb{Q}$ par

$$\mathbb{Q}(dw) = \frac{X^p \mathbb{P}(dw)}{\mathbb{E}(X^p)}, \qquad i.e. \mathbb{Q}(A) = \frac{\int_A X^p(w) \mathbb{P}(dw)}{\mathbb{E}(X^p)}$$

et une variable aléatoire Z par

$$Z = \mathbb{1}_{X>0} \frac{Y}{X^{p-1}}$$

On conclue avec Jensen :

$$\mathbb{E}_{\mathbb{Q}}(Z)^q = \left(\frac{\mathbb{E}(XY)}{\mathbb{E}(X^p)} \right)^q \leq \mathbb{E}_{\mathbb{Q}}(Z^q) = \frac{\mathbb{E}(Y^q)}{\mathbb{E}(X^p)}$$

□

Théorème 2.3. *Inégalité de Minkowski*
Soit $X, Y \in \mathcal{L}^p$*, on a l'inégalité de Minkowski*

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$$

Démonstration. L'inégalité de Hölder donne

$$\begin{aligned} \mathbb{E}(|X+Y|^p) &\leq \mathbb{E}(|X||X+Y|^{p-1}) + \mathbb{E}(|Y||X+Y|^{p-1}) \\ &\leq A\|X\|_p + A\|Y\|_p \end{aligned}$$

où $A = \||X+Y|^{p-1}\|_q = \mathbb{E}(|X+Y|)^{1/q}$. Un majoration grossière montre que A est finie et le résultat tombe. □

Inégalités de Markov Si $X \in \mathcal{L}^p$, $\mathbb{P}(|X| \geq c) \leq c^{-p}\mathbb{E}(|X|^p)$.
Si X est une variable aléatoire positive, pour tout réel c et tout $\lambda > 0$, on a

$$\mathbb{P}(X \geq c) \leq e^{-\lambda c}\mathbb{E}(e^{\lambda Y})$$

2.2 L'espace normé $\mathbb{L}^p$

$\mathbb{L}^p$ est le quotient de $\mathcal{L}^p$ pour la relation d'équivalence $X \sim Y \Longleftrightarrow X = Y$, $\mathbb{P}p.s.$.

Théorème 2.4. *$\mathbb{L}^p$ est un espace vectoriel normé complet.*

On rappelle que pour démontrer ce théorème, on utilise ce que l'on appelle parfois “la réciproque du théorème de convergence dominée”, à savoir :

Proposition 2.3. *Si $(X_n)_n$ est une suite de Cauchy dans $\mathcal{L}^p$ alors il existe une sous-suite $(Y_n)_n$ de $(X_n)_n$ et une variable aléatoire X ayant un moment d'ordre p telle que $(Y_n)_n$ converge en norme $\mathcal{L}^p$ vers X.*

Le cas $p = 2$

$\mathbb{L}^2$ jouit d'un place particulière au sein des espaces $\mathbb{L}^p$ puisque c'est un espace de Hilbert pour le produit scalaire $< X, Y >= \mathbb{E}(XY)$.

En anticipant sur la suite, si X et Y sont indépendantes alors $Cov(X, Y) = 0$ et la réciproque est fausse. Si $Cov(X, Y) = 0$ alors les variances s'ajoutent : $Var(X + Y) = Var(X) + Var(Y)$ (théorème de Pythagore).

2.3 Exercices

Exercice 2.1. Lois usuelles Calculer l'espérance et la variance de :

1) La loi binomiale $\mathcal{B}(n, p)$ $(p \in [0, 1],\ n \in \mathbb{N}^*)$.

2) La loi de Poisson $\mathcal{P}(\lambda)$ $(\lambda > 0)$.

3) La loi géométrique $\mathcal{G}(p)$ $(p \in]0, 1])$.

4) La gaussienne de densité sur $\mathbb{R}$:

$$f(x) = \frac{e^{-(x-m)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

Solutions : 1) L'espérance vaut np et la variance $np(1 - p)$, ce qui s'obtient en utilisant $k\binom{n}{k} = n\binom{n-1}{k-1}$, ou en dérivant $\sum_{k=0}^{n} \binom{n}{k} x^k (1 - p)^{n-k}$, ou en écrivant la loi binomiale comme somme de n Bernoulli indépendantes (cf section suivante).

2) La moyenne et la variance sont égales à λ.

3) La moyenne vaut $1/p$ et la variance $[1 - p]/p^2$.

4) La moyenne est donnée par m et la variance σ^2 (obtenue par intégration par partie).

Exercice 2.2. (Lois normales dans $\mathbb{R}^2$, Ouvrard, Tome II p32) Soit $X = (X_1, X_2)$ une v.a. à valeurs dans $\mathbb{R}^2$ de loi normale $\mathcal{N}(0, I_2)$, c.a.d. sa densité f_X est donnée par

$$f_X(x) = \frac{1}{2\pi} e^{-\|x\|^2/2},$$

où $\| x \| = \| (x_1, x_2) \| = \sqrt{x_1^2 + x_2^2}$ est la norme euclidienne usuelle.

1) Déterminer la loi de la v.a. $\| X \|^2$

2) Soit $D = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \neq x_2\}$. Démontrer que la variable aléatoire T définie par

$$T := \left(\frac{X_1 + X_2}{X_1 - X_2}\right)^2 \text{ si } (X_1, X_2) \in D; \qquad T := 0 \text{ si } (X_1, X_2) \notin D,$$

admet une densité égale à (loi de Hotelling)

$$\mathbb{1}_{\mathbb{R}_+^*}(x) \frac{1}{\pi(x+1)\sqrt{x}}$$

Solution :
1) En utilisant un changement de variable polaire puis le changement $u = \rho^2$, on obtient pour tout fonction positive à support compact

$$\mathbb{E}(f(\| X \|^2) = \int_{\mathbb{R}_+^*} f(\rho^2) e^{-\rho^2/2} \rho d\rho = \int_{\mathbb{R}_+^*} f(u) \frac{1}{2} e^{-u/2} du.$$

Ceci prouve que $\| X \|^2$ ext une variable exponentielle de paramètre $1/2$.

2) On effectue le changement de variable

$$u = \frac{x_1 + x_2}{x_1 - x_2}, \ v = x_1 + x_2 \quad \Leftrightarrow \quad x_1 = \frac{1}{2}\left(v + \frac{v}{u}\right), \ x_2 = \frac{1}{2}\left(v - \frac{v}{u}\right)$$

de jacobien $v/2u^2$ et on utlise le théorème de Fubini

$$\begin{aligned}
\mathbb{E}(f(T)) &= \frac{1}{4\pi} \int_{\mathbb{R}^*} \frac{f(u^2)}{u^2} \left[\int_{\mathbb{R}} |v| \exp(-v^2(1 + 1/u^2)/4) dv\right] du \\
&= \frac{1}{\pi} \int_{\mathbb{R}^*} \frac{f(u^2)}{u^2} \frac{u^2}{1+u^2} du \\
&= \int_{\mathbb{R}} f(x) \mathbb{1}_{\mathbb{R}_+^*} \frac{1}{\pi(x+1)\sqrt{x}} dx
\end{aligned}$$

3 Indépendance

3.1 Indépendance d'événements et de variables aléatoires

Définition 3.1. *a) Les événements A et B sont dits indépendants si et seulement si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*
b) Deux familles d'événements $\mathcal{A}_1$ et $\mathcal{A}_2$ sont dites indépendantes si et seulement si tout élément de $\mathcal{A}_1$ est indépendant de tout élément de $\mathcal{A}_2$.

Théorème 3.1. *Soient $\mathcal{C}_1$ et $\mathcal{C}_2$ deux π-systèmes contenus dans $\mathcal{F}$. On note $\mathcal{F}_i$ la tribu engendrée par $\mathcal{C}_i$. Alors $\mathcal{F}_1$ est indépendant de $\mathcal{F}_2$ si et seulement si $\mathcal{C}_1$ est indépendant de $\mathcal{C}_2$.*

Démonstration. La condition nécessaire étant triviale, passons à la condition suffisante. C'est une conséquence de la proposition d'unicité du prolongement des mesures, qui est une conséquence directe du théorème de Dynkin. En effet, pour un $C_1 \in \mathcal{C}_1$ fixé, les deux applications définies sur $\mathcal{F}$ par :

$$I \in \mathcal{F} \to \mathbb{P}(I)\mathbb{P}(C_1) \ \ , \ \ I \in \mathcal{F} \to \mathbb{P}(I \cap C_1)$$

Ce sont deux mesures finies de même masse et coïncidant sur un π-système, à savoir $\mathcal{C}_2$. Elles coïncident donc sur la tribu engendrée par $\mathcal{C}_2$:

$$\forall C_1 \in \mathcal{C}_1, \forall F_2 \in \mathcal{F}_2, \mathbb{P}(F_2)\mathbb{P}(C_1) = \mathbb{P}(F_2 \cap C_1).$$

Soit maintenant $F_2 \in \mathcal{F}_2$ fixé. Considérons les deux applications

$$I \in \mathcal{F} \to \mathbb{P}(I)\mathbb{P}(F_2) \ \ , \ \ I \in \mathcal{F} \to \mathbb{P}(I \cap F_2)$$

sont deux mesures de masses égales finies coïncidant sur le π-système $\mathcal{C}_1$: elles coïncident sur la tribu engendrée $\mathcal{F}_1$, c.q.f.d. □

On verra dans la section 3 une généralisation de ce théorème (théorème des coalitions).

Définition 3.2. *Soient X_1 et X_2 deux variables aléatoires. On dit que X_1 et X_2 sont indépendantes si et seulement si les tribus engendrées $\sigma(X_1)$ et $\sigma(X_2)$ le sont.*

Remarque 3.1. *Si X_1 et X_2 sont deux variables aléatoires indépendantes et si f_1 et f_2 sont des fonctions boréliennes alors $f_1(X_1)$ et $f_2(X_2)$ sont indépendantes (car $\sigma(f_i(X_i)) \subset \sigma(X_i)$). Par exemple, si X_1 et X_2 sont deux vecteurs aléatoires alors toute composante (ou marginale) de X_1 est indépendante de toute composante (marginale) de X_2.*

Théorème 3.2. *Soient X_1 et X_2 deux variables aléatoires à valeurs dans $\mathbb{R}^n$ et $\mathbb{R}^m$ respectivement. L'indépendance de X_1 et X_2 équivaut à chacune des propositions suivantes*

(i) $\forall A_1 \in \mathcal{B}(\mathbb{R}^n),\ \forall A_2 \in \mathcal{B}(\mathbb{R}^m)$ *:* $\mathbb{P}(X_1 \in A_1, X_2 \in A_2) = \mathbb{P}(X_1 \in A_1)\mathbb{P}(X_2 \in A_2)$ *(i.e.* $\mathbb{P}_{(X_1,X_2)} = \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2}$*).*

(i′) $\forall A_1 \in \pi_1,\ \forall A_2 \in \pi_2$ *:* $\mathbb{P}(X_1 \in A_1, X_2 \in A_2) = \mathbb{P}(X_1 \in A_1)\mathbb{P}(X_2 \in A_2)$ *avec π_1 un π-système qui engendre $\mathcal{B}(\mathbb{R}^n)$ et π_2 un π-système qui engendre $\mathcal{B}(\mathbb{R}^m)$.*

(i‴) En terme de conction de répartition

$$\mathbb{P}(X_1 \leq x_1,\ X_2 \leq x_2) = \mathbb{P}(X_1 \leq x_1)\mathbb{P}(X \leq x_2) \quad (x_1 \in \mathbb{R}^n, x_2 \in \mathbb{R}^m)$$

(ii) Pour toutes fonctions boréliennes positives f_1 et f_2, on a

$$\mathbb{E}\left[f_1(X_1)f_2(X_2)\right] = \mathbb{E}\left[f_1(X_1)\right]\mathbb{E}\left[f_2(X_2)\right]$$

(ii′) Pour toutes fonctions boréliennes bornées f_1 et f_2, on a

$$\mathbb{E}\left[f_1(X_1)f_2(X_2)\right] = \mathbb{E}\left[f_1(X_1)\right]\mathbb{E}\left[f_2(X_2)\right].$$

Il suffit en fait de le vérifier pour des fonctions continues (ou même C^∞) bornées (ou même à support compact).

Démonstration. (i) est la définition, (i') vient de l'unicité du prolongement d'une mesure définie sur un π système ou du Théorème 3.1, tout comme (i''').

Passer de (i) à (ii) ou (ii') se fait par la machine standard, et pour $(i) \Rightarrow (ii')$ on peut aussi invoquer le théorème de la classe monotone. Les réciproques sont évidentes.

□

Nous passons maintenant à une propriété vérifiée par les variables aléatoires indépendantes.

Proposition 3.1. *Soit X_1 et X_2 deux variables aléatoires réelles indépendantes.*
1) Si X_1 et X_2 sont dans $\mathcal{L}^1$ alors X_1X_2 est dans $\mathcal{L}^1$ et $\mathbb{E}(X_1X_2) = \mathbb{E}(X_1)\mathbb{E}(X_2)$.
2) Si X_1 et X_2 sont dans $\mathcal{L}^2$ alors $cov(X_1, X_2) = 0$ et $Var(X_1 + X_2) = Var(X_1) + Var(X_2)$.

Rappel : La covariance de deux variables aléatoires réelles X et Y est définie par $cov(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ et $\sigma_X^2 = Cov(X, X)$.

Démonstration. 1) Pour montrer que X_1X_2 est intégrable, on applique le (ii) du théorème précédent avec $f_1 = f_2 = |\cdot|$. Pour établir l'égalité entre les espérances, on écrit X_1 et X_2 comme différence de leurs parties positives et négatives respectives. Par linéarité de l'espérance, on se ramène ainsi au cas de variables aléatoires positives. On applique alors à nouveau la proposition (ii) avec $f_1 = f_2 = id$.
2) Evident.
□ Bien sur la réciproque de cette propriété est en général fausse. Vous pouvez construire un contrexemple où $X_1 \in \{-1, 0, 1\}$ et $X_2 \in \{0, 1\}$, en imposant même $X_1X_2 = 0$ et $\mathbb{E}(X_1) = 0$. Néanmoins, si $X_1 \in \{a, b\}$ et $X_2 \in \{c, d\}$ vérifient $Cov(X_1, X_2) = 0$, alors X_1 et X_2 sont indépendants. Exo ! (pour lequel il est conseillé de se ramener à $a = c = 0$, $b = d = 0$ par un changement affine.

3.2 En particulier (et en pratique) : les critères d'indépendance

En pratique on utilise un des points suivants, ce qui revient à regarder un bon π système dans les cas 1 et 3 ou à utiliser la mesure produit dans le cas 2.

1) Pour des variables aléatoires discrètes
Si les variables aléatoires X_1 et X_2 sont discrètes, il en est de même pour (X_1, X_2). Pour que X_1 et X_2 soient indépendantes, il faut et il suffit que

$$\forall x_1 \in X_1(\Omega), \forall x_2 \in X_2(\Omega),\ \mathbb{P}(X_1 = x_1, X_2 = x_2) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2)$$

La démonstration de cette propriété est laissée en exercice. Le point important est de remarquer que pour une variable discrète X à valeurs dans un ensemble discret E les ensembles atomiques $\{X = e\}$, $e \in E$ forment un π-système qui engendre $\sigma(X)$.

2) En termes de densités On a équivalence entre
i) X_1 et X_2 sont deux v.a. dans $\mathbb{R}^{d_1}$ et $\mathbb{R}^{d_2}$ indépendantes de densités f_1 et f_2.

ii) (X_1, X_2) est une v.a. $\mathbb{R}^{d_1+d_2}$ de densité $x = (x_1, x_2) \to f_1(x_1)f_2(x_2)$ avec $\int f_1 = \int f_2 = 1$.

Pour l'implication, on utilise à nouveau le π-système $\mathcal{C} = \left\{A_1 \times A_2;\ A_1 \in \mathcal{B}(\mathbb{R}^{d_1}), A_2 \in \mathcal{B}(\mathbb{R}^{d_2})\right\}$ qui engendre $\mathcal{B}(\mathbb{R}^n)$ et l'unicité du prolongement des mesures finies qui coïncident sur un π-système appliqué à $\mathbb{P}((X,Y) \in A)$ et $\int_A f_1(x_1)f_2(x_2)dx_1dx_2$. Ou bien Fubini.
Pour la réciproque, on se donne des boréliens A_1 et A_2, on a

$$\mathbb{P}((X_1, X_2) \in A_1 \times A_2) = \left(\int_{A_1} f_1(x_1)dx_1\right)\left(\int_{A_2} f_2(x_2)dx_2\right)$$

En prenant successivement $A_1 = \mathbb{R}^{d_1}$ puis $A_2 = \mathbb{R}^{d_2}$, on obtient que la loi de X_1 est donnée par la densité f_1 et celle de X_2 par la densité f_2. On déduit de l'égalité précédente l'indépendance de X_1 et X_2.

3.3 Indépendance généralisée

Soit I un ensemble quelconque, dénombrable ou pas.

Définition 3.3. (Indépendance de familles d'évènements et de variables aléatoires)

- *Les événements $(A_i)_{i\in I}$ sont dits indépendants ssi :*

$$\forall J \subset I, \mathbb{P}\left(\cap_{j\in J} A_j\right) = \prod_{j\in J} \mathbb{P}(A_j)$$

- *Des familles d'événements $(\mathcal{A}_i)_{i\in I}$ sont indépendantes ssi pour tout choix d'un A_i dans $\mathcal{A}_i$, la famille d'événements $(A_i)_{i\in I}$ est constituée d'événements indépendants.*
- *On dira qu'une famille de variables aléatoires $(X_i)_{i\in I}$ est une famille de variables aléatoires indépendantes ssi les $(\sigma(X_i))_{i\in I}$ le sont.*

On prouvera en exercice que trois variables aléatoires peuvent être indépendantes deux à deux mais pas dans leur ensemble. Pensez aussi par exemple aux tirages consécutifs de deux dés et aux trois variables suivantes : le premier résultat, le deuxième et la parité de la somme des deux.

Le théorème suivant qui est une généralisation du théorème vu dans la section 1 signifie simplement que l'indépendance se conserve par "regroupements disjoints".

Théorème 3.3. *("des coalitions")*
Soit une famille $(\mathcal{C}_i)_{i\in I}$ de π-systèmes indépendants contenus dans une tribu $\mathcal{F}$. On note $\mathcal{F}_i = \sigma(\mathcal{C}_i)$ la tribu engendrée par $\mathcal{C}_i$. Soit $(I_j)_{j\in J}$ une partition quelconque de I et pour tout $j \in J$, $\mathcal{G}_j = \sigma\left(\cup_{i\in I_j} \mathcal{C}_i\right)$.
Alors les tribus $(\mathcal{G}_j)_{j\in J}$ sont indépendantes. En particulier, les tribus $(\mathcal{F}_i)_{i\in I}$ sont indépendantes.

Démonstration. La démonstration dans le cas fini est faite dans le Durrett et dans l'Ouvrard. Durett donne pratiquement tous les éléments pour traiter le cas infini. Dans le cas infini, on pourra trouver une démonstration bien faite dans le Buchwalter.

On commence par montrer le théorème pour $J = I$ et $I_j = \{j\}$. Dans ce cas, il est clair que l'on peut se ramener au cas I fini. On suppose pour simplifier les notations que $I = \{1, \ldots, n\}$ et on rappelle que le cas $n = 2$ a été vu dans la section 1.
On fixe $A_2 \in \mathcal{C}_2, \ldots, A_n \in \mathcal{C}_n$ et on considère les deux applications suivantes sur $\mathcal{F}_1$

$$B_1 \in \mathcal{F}_1 \to \mathbb{P}(B_1 \cap A_2 \cap \ldots \cap A_n)$$

$$B_1 \in \mathcal{F}_1 \to \mathbb{P}(B_1)\mathbb{P}(A_2) \ldots \mathbb{P}(A_n)$$

Ce sont deux mesures de même masse finie (faire $B_1 = \Omega_1$) qui coïncident sur le π-système $\mathcal{C}_1$ donc sur $\mathcal{F}_1$.
On fixe ensuite $B_1 \in \mathcal{F}_1, A_3 \in \mathcal{C}_3, \ldots, A_n \in \mathcal{C}_n$ et on considère les deux applications

$$B_2 \in \mathcal{F}_2 \to \mathbb{P}(B_1 \cap B_2 \cap A_3 \ldots \cap A_n)$$

$$B_2 \in \mathcal{F}_2 \to \mathbb{P}(B_1)\mathbb{P}(B_2)\mathbb{P}(A_3) \ldots \mathbb{P}(A_n)$$

qui sont des mesures de même masse finie (appliquer ce qui précède en faisant $B_2 = \Omega_2$) coïncidant sur le π-système $\mathcal{C}_2$ donc sur $\mathcal{F}_2$. Et on continue ainsi pour obtenir le résultat annoncé.

On a donc obtenu le théorème pour le cas de la partition triviale. On définit

$$\mathcal{L}_j = \left\{\cap_{i \in I_j} A_i;\ A_i \in \mathcal{C}_i \text{ pour un nombre fini d'indices et } \quad A_i = \Omega \text{ autrement}\right\}$$

C'est un π-système contenant tous les $\mathcal{C}_i$ pour $i \in I_j$ donc $\sigma(\mathcal{L}_j)$ contient $\mathcal{G}_j$. Or les $\mathcal{C}_i$ sont indépendants donc les $\mathcal{L}_j$ le sont : il suffit de l'écrire. D'après ce qui précède, les $\sigma(\mathcal{L}_j)_j$ sont des tribus indépendantes et il en est donc de même des $(\mathcal{G}_j)_j$ puisque $\mathcal{G}_j \subset \sigma(\mathcal{L}_j)$. □

Remarque 3.2. *Voici une application discrète du précédent théorème. Soient $A_1, \ldots, A_n$ des évènements indépendants. On considère les évènements $\tilde{A}_1, \ldots, \tilde{A}_n$ où $A_i = A_i$ ou A_i^c. Montrer que les $\tilde{A}_i$ sont indépendants. Celà peut aussi se vérifier à la main, mais ce n'est pas si facile (faire une récurrence sur le nombre de complémentaires A_i^c dans la séquence $\tilde{A}_i$).*

3.4 Indépendance et événements asymptotiques

Le type de question que nous appréhendons ici est la suivante. Si je tire à pile ou face avec probabilité p (ou p_n au $n$ième lancer) de faire pile, est ce qu'à un moment je cesserai de voir des piles ?

Dans cette section, nous énoncons trois théorèmes concernant la tribu asymptotique : loi du 0-1, lemme de Borel-Cantelli 1 et lemme de Borel-Cantelli 2. Le lemme de Borel-Cantelli 1 n'a rien à voir avec l'indépendance mais son énoncé étant une sorte de réciproque du 2, il est d'usage de le placer ici.

Définition 3.4. *Soit $(\mathcal{F}_n)_n$ une suite de sous-tribus de $\mathcal{F}$. On appelle tribu asymptotique des $(\mathcal{F}_n)_n$ la tribu*

$$\mathcal{F}_\infty = \cap_{n \in \mathbb{N}} \sigma \quad (\cup_{k \geq n} \mathcal{F}_k)$$

Les éléments de $\mathcal{F}_\infty$ sont appelés les événements asymptotiques et une variable aléatoire $\mathcal{F}_\infty$-mesurable est dite asymptotique.
Si $(X_n)_n$ est une suite de variables aléatoires, la tribu asymptotique associée aux X_n est la tribu asymptotique associée aux $\mathcal{F}_n = \sigma(X_n)$.

Par exemple, l'événement $\{\omega \in \Omega;\ X_n(\omega) \text{ converge dans } \mathbb{R}\}$ est un événement asymptotique (par rapport aux X_n) et $\limsup X_n, \liminf X_n$ sont des variables aléatoires asymptotiques. Traitons un exemple :

$$\liminf_n X_n \geq a \Leftrightarrow (\forall l > 0, \exists n,\ \forall k \geq n, X_n \geq a - \frac{1}{l}) \Leftrightarrow \omega \in \cap_n \cap_l \cup_{p \geq n} \cap_{k \geq p} \{X_k \geq a + \frac{1}{l}].$$

Théorème 3.4. *(Loi du "Tout ou Rien', ou du "0-1")*
Si la suite $(\mathcal{F}_n)_n$ est constituée de tribus indépendantes alors

$$\forall A \in \mathcal{F}_\infty,\ \mathbb{P}(A) = 0 \ \text{ou}\ \mathbb{P}(A) = 1$$

La loi du tout ou rien dit ainsi que soit p.s. je cesserai de voir des piles, soit p.s. je verrai une infinité de piles. le fait d'etre dans un cas ou dans l'autre dépend de la suite p_n, et les lemmes de Borel Cantelli répondent à la question.

Démonstration. On démontre que $\mathcal{F}_\infty$ est indépendante d'elle-même ce qui implique trivialement le résultat. D'après le théorème des coalitions, $\sigma\left(\cup_{k\geq n}\mathcal{F}_k\right)$ est indépendant de $\mathcal{F}_p$ pour $p < n$. Or $\mathcal{F}_\infty$ est l'intersection des $\sigma\left(\cup_{k\geq n}\mathcal{F}_k\right)$ donc $\mathcal{F}_\infty$ est indépendant de $\mathcal{F}_p$ pour tout p. Il en résulte que $\mathcal{F}_\infty$ est indépendant de $\sigma\left(\cup_{p\geq 1}\mathcal{F}_p\right)$ et $\mathcal{F}_\infty \subset \sigma\left(\cup_{p\geq 1}\mathcal{F}_p\right)$ donc $\mathcal{F}_\infty$ est indépendante d'elle-même. □

On rappelle que si l'on a une suite d'événements $(A_n)_n$, on définit

$$\limsup A_n = \cap_n \cup_{p\geq n} A_p = \{\omega;\ \omega \text{ appartient à une infinité de } A_n\}$$

c'est à dire A_n se produit une infinité de fois (en anglais *infinitely often, i.o.*)

Lemme 3.1. *(Borel-Cantelli 1)*
Si $\sum_n \mathbb{P}(A_n) < +\infty$ alors $\mathbb{P}(\limsup A_n) = 0$. (C'est à dire : A_n ne se produit qu'un nombre fini de fois).

Démonstration. Pour tout n, $\limsup A_k \subset \cup_{k\geq n}A_k$ donc

$$\mathbb{P}(\limsup A_k) \leq \mathbb{P}(\cup_{k\geq n}A_k) \leq \sum_{k\geq n}\mathbb{P}(A_k)$$

et le dernier terme de droite tend vers 0 quand n tend vers l'infini par hypothèse d'où le résultat.
□

Lemme 3.2. *(Borel-Cantelli 2)*
Si les $(A_n)_n$ sont des événements indépendants et si $\sum_n \mathbb{P}(A_n) = +\infty$ alors $\mathbb{P}(\limsup A_n) = 1$. (C'est à dire : A_n se produit un nombre infini de fois).

Démonstration. Il est aisé de voir que

$$\{\limsup A_k\}^c = \cup_n \cap_{k\geq n} A_k^c$$

Posons $p_k = \mathbb{P}(A_k)$. On a $\mathbb{P}\left(\cap_{k\geq n}A_k^c\right) = \prod_{k\geq n}(1 - p_k)$ car par indépendance,

$$\mathbb{P}\left(\cap_{m\geq k\geq n}A_k^c\right) = \prod_{m\geq k\geq n}(1 - p_k)$$

et on peut faire tendre m vers l'infini grâce au théorème de convergence monotone.
Pour tout $x \geq 0$, $1 - x \leq \exp(-x)$ donc

$$\prod_{k \geq n} (1 - p_k) \leq \exp\left(- \sum_{k \geq n} p_k \right) = 0$$

Donc $\mathbb{P}((\limsup A_k)^c) = 0$. □

3.5 Exercices : Indépendance

Exercice 3.1. Soient T_1 et T_2 deux variables aléatoires indépendantes de loi géométrique de paramétre $1 - p_1$ et $1 - p_2$.

1) Calculer et reconnaitre la loi de $min(T_1, T_2)$.

2) Calculer la loi jointe de $min(T_1, T_2)$ et $T_1 - T_2$.

3) En déduire que $min(T_1, T_2)$ est indépendant de $\mathbb{1}_{T1 \leq T2}$. Quelle est la loi de $\mathbb{1}_{T1 \leq T2}$?

4) Déduire également de la question 2) que $R = max(T_1, T_2) - min(T_1, T_2)$ est indépendant de $min(T_1, T_2)$.

5) Calculer la loi de R conditionnellement à $\{R \neq 0\}$. Reconnaitre cette loi quand $p_1 = p_2$.

Solutions :
1) Utiliser que $\mathbb{P}(T_1 \geq n) = p_1^n$ *pour en déduire que* $\mathbb{P}(min(T_1, T_2) \geq n) = \mathbb{P}(T_1 \geq n)\mathbb{P}(T_2 \geq n) = (p_1 p_2)^n$*, c'est à dire que* $min(T_1, T_2)$ *est une v.a. géométrique de paramètre* $p_1 p_2$*.*
2)Pour tout $j \geq 0$, $\mathbb{P}(min(T_1, T_2) = i,\ T_1 - T_2 = j) = \mathbb{P}(T_2 = i)\mathbb{P}(T_1 = i + j) = (1 - p_1)(1 - p_2) p_2^i p_1^{i+j}$ *et* $\mathbb{P}(min(T_1, T_2) = i,\ T_1 - T_2 = -j) = (1 - p_1)(1 - p_2) p_1^i p_2^{i+j}$.
3) Il suffit de cacluler

$$\begin{aligned}\mathbb{P}(min(T_1, T_2) = i,\ T_1 \geq T_2) &= \sum_{j \geq 0} \mathbb{P}(min(T_1, T_2) = i,\ T_1 - T_2 = j) = \sum_{j \geq 0} (1 - p_1)(1 - p_2) p_2^i p_1^{i+j} \\ &= (p_1 p_2)^i (1 - p_1) = \mathbb{P}(min(T_1, T_2) = i)\Pi(T_1 \geq T_2),\end{aligned}$$

ce qui prouve l'indépendance de $min(T_1, T_2)$ *et* $T_1 - T_2$*.*
4) Pour tout $i, j \geq 1$, $\mathbb{P}(min(T_1, T_2) = i,\ R = j)$ *vaut*

$$\mathbb{P}(min(T_1, T_2) = i,\ T_1 - T_2 = j) + \mathbb{P}(min(T_1, T_2) = i,\ T_1 - T_2 = -j) = (p_1 p_2)^i (1 - p_1)(1 - p_2)(p_1^j + p_2^j),$$

ce qui prouve l'indépendance des deux variables (en traitant à part le cas $j = 0$*).*

Exercice 3.2. Un modèle pour le jeu de pile ou face (Ouvrard, Tome II, 9.3., p. 53-59), Lecon suite de v.a. de Bernoulli indépendantes
Soit $x \in [0, 1)$. On définit par récurrence les suites de terme général

$$R_0(x) = x$$

et pour tout $n > 0$ $D_n(x) = [2R_{n-1}(x)]$ et $R_n(x) = 2R_{n-1}(x) - D_n(x)$.
On a pour tout $n > 0$,

$$x = \sum_{j=1}^{n} \frac{D_j(x)}{2^j} + \frac{1}{2^n} R_n(x).$$

Le but de l'exercice est de montrer la proposition suivante :

Proposition 3.2. *Sur* $[0,1)$ *muni de la tribu des Boréliens et la mesure de Lebesgue, la suite* $(D_n)_{n\in\mathbb{N}^*}$ *est une suite de variables aléatoires indépendantes de même loi de Bernoulli de paramètre* $1/2$*. La variable aléatoire* R_n *est de loi uniforme sur* $[0,1)$ *et la variable* R_n *et les variables* $(D_1, ..., D_n)$ *sont indépendantes.*

Pour cela on utilisera le théorème des coalitions : Soit une famille $(\mathcal{C}_i)_{i\in I}$ de π-systèmes contenus dans $\mathcal{A}$ et indépendants ; on note $\mathcal{F}_i$ la tribu engendrée par $\mathcal{C}_i$. Soit $(I_j)j \in J$ une partition de I et $\mathcal{A}_j$ la tribu engendrée par $\cup_{i\in I_j}\mathcal{C}_i$. Alors les tribus $(\mathcal{A}_j)_{j\in J}$ sont indépendantes.

1) Vérifier que tout réel $x \in]0,1[$ s'écrit d'une manière unique sous la forme précédente, $x = 0, x_1x_2...x_k...$ avec $x_i \in \{0,1\}$ si et seulement si x n'est pas un rationnel dyadique, c'est-à-dire un rationnel de la forme $\frac{p}{2^q}$ avec $p, q \in \mathbb{N}$ et $p \leq 2^q$.
Préciser la forme des $D_n(x)$ quand x est un rationnel dyadique.

2) On va montrer que $(D_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes de même loi de Bernoulli de paramètre $1/2$. On note $\varepsilon^n = (\varepsilon_1, ...\varepsilon_n)$ un élément de $\{0,\ 1\}^n$ et $\varepsilon = (\varepsilon_1, ..., \varepsilon_n, ...)$ un élément de $\{0,\ 1\}^{\mathbb{N}}$.

2a) Montrer que $\cap_{j=1}^{n}\{D_j = \varepsilon_j\}$ est un intervalle de $[0,1]$ que l'on précisera et calculer sa probabilité. (On trouve 2^{-n}).

2b) En déduire que pour toute partie non vide J de $\{1,\ 2, ...,\ n\}$ on a

$$\mathbb{P}(\cap_{j\in J}\{D_j = \varepsilon_j\}) = \frac{1}{2^{|J|}}.$$

Conclure que les D_n sont bien indépendantes.

3) On va montrer que, pour tout $n \in \mathbb{N}^*$, la variable aléatoire R_n est de loi uniforme sur $[0,1)$ et que les variables aléatoires R_n et $(D_1, D_2, \ldots, D_n)$ sont indépendantes.

3a) Vérifier que $R_n(x) = 2^n x - \sum_{j=1}^{n} 2^{n-j} D_j(x) = \sum_{j=n+1}^{\infty} 2^{n-j} D_j(x)$ et en déduire que R_n et $(D_1, D_2, \ldots, D_n)$ sont indépendantes.

3b) Montrer que la loi de R_n est uniforme, i.e.

$$\mathbb{E}(f(R_n)) = \int_{\mathbb{R}} \mathbb{1}_{[0,1[}(y) f(y) dy.$$

3c) Déduire des deux questions précédentes que pour toute fonction borélienne positive $f : \mathbb{R} \to \mathbb{R}$ et toute partie J de $\{1,\ 2,\ ..., n\}$,

$$\mathbb{E}\left[f(R_n)\Pi_{i\in J}\mathbb{1}_{D_i=\varepsilon_i}\right] = \mathbb{P}(\cap_{i\in J}\{D_i = \varepsilon_i\}) \left[\int_{\mathbb{R}} \mathbb{1}_{[0,1[}(y) f(y) dy\right]. \tag{3.1}$$

4) En vérifiant et utilisant la relation $D_n = -R_n + 2R_{n-1}$, montrer que la suite de v.a. R_n n'est pas une suite de variables aléatoires indépendantes.

5) Soit $(\mu_j)_{j>0}$ une suite de probabilités sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On se propose de construire explicitement des variables aléatoires $(X_j)_{j>0}$ indépendantes de lois respectives $(\mu_j)_{j>0}$.

5a) Soit $(N_j)_{j>0}$ une suite de sous-ensembles infinis de $\mathbb{N}^*$ formant une partition. Pour tout $j > 0$, soit $\phi_j \in \mathbb{N}^*$ une énumération des N_j. Pour tout $j > 0$, on pose

$$Y_j = \sum_{k=1}^{\infty} \frac{1}{2^k} D_{\phi_j(k)}$$

Montrer que les variables $(Y_j)_{j>0}$ sont indépendantes.

5b) Montrer que la loi de $Y_{j,n} = \sum_{k=1}^{n} \frac{1}{2^k} D_{\varphi_j(k)}$ est la même que celle de $Z_n = \sum_{k=1}^{n} \frac{1}{2^k} D_k$ et en déduire que la loi de Y_j est uniforme sur $[0, 1]$.

5c) Soit F_j la fonction de répartition de μ_j. On définit la pseudo-inverse de F_j par

$$G_j(t) = \inf_{x \in \mathbb{R}} \{F_j(x) \geq t\}$$

On pose $X_j = G_j(Y_j)$. Montrer que la loi de X_j est μ_j (deja vu en cours) et conclure.

Esquisse de solution :
1) La forme du développement quand le nombre est dyadique est celui qui correspond à annuler les $D_n(x)$ à partir du premier entier n possible.
2.b) $\mathbb{P}(\cap_{j\in J}\{D_j = \varepsilon_j\})$ est égal à

$$\sum_{\varepsilon^{n-|J|} \in \{0,\,1\}^{J^c}} \mathbb{P}\left(\cap_{j\in J}\{D_j = \varepsilon_j\} \cap (\cap_{k\in J^c}\{D_k = \varepsilon_k^{n-|J|}\})\right) = \sum_{\varepsilon \in \{0,1\}^{J^c}} \frac{1}{2^n} 2^{n-|J|} = \frac{1}{2^{|J|}}.$$

3.a) $R_n \in \sigma(D_{n+1},\ D_{n+2}, ...)$ est indépendante de $D_1, ..., D_n$ puisque les v.a. $(D_i : i \in \mathbb{N})$ sont indépendantes (théorème des coalitions).
3.b) Remaquer que R_n a la même loi que x : ils sont tous les deux obtenus par des suites de v.a. de Bernoulli constitués en développement dyadique.
On peut aussi (Ouvrard tome II) calculer

$$\mathbb{E}\left[f(R_n)\Pi_{i=1}^n \mathbb{1}_{D_i=\varepsilon_i}\right] = \int_{\mathbb{R}} \mathbb{1}_{[0,1[}(y) f(y) dy$$

avce le changement de variable $y = 2^n x - \sum_{j=1}^n 2^{n-j}\epsilon_j$ puis sommer sur les $\epsilon_j = 0, 1$.
3.c) Utiliser 3.a) pour séparer les intégrales (par indépendance) puis 3.b). 4) D_n crée un lien entre R_n et R_{n_1} qui ne sont donc pas indépendantes. Si c'était le cas $D_n = -R_n + 2R_{n-1}$ entrainerait que $\mathbb{P}(D_n \geq 5/4) \geq \mathbb{P}(R_n \leq 1/4,\ R_{n-1} \geq 3/4) = \mathbb{P}(R_n \leq 1/4)\mathbb{P}(R_{n-1} \geq 3/4) = (1/4)(1/4) > 0$.
Ou bien utiliser que si les v.a. étaient indépendantes, leur somme admettrait aussi une densité.
5.a) On utilisera le théorème des coalitions.

Exercice 3.3. Deux exemples parlants de variables aléatoires deux à deux indépendantes et d'évènements deux à deux indépendants.

1) On considère $\Omega_1 = [0,1]^3$ muni de la mesure de Lebesgue. Montrer que les coordonnées X, Y et Z sont des variables aléatoires indépendantes.

2) On considère l'espace de probabilité $\Omega_2 = \Omega_1 \sqcup [\Omega_1 + (1,1,0)] \sqcup [\Omega_1 + (1,0,1)] \sqcup [\Omega_1 + \cup(0,1,1)]$ muni de la mesure de Lebesgue divisée par 4. Dessiner Ω_2. Montrer que l'on a $\mathbb{P}(X \in [a,b]) = (b-a)/2$ pour tout $0 \leq a \leq b \leq 2$.

3) En déduire que X, Y, Z sont encore des variables aléatoires indépendantes deux à deux. En considérant l'évènement $\{X \geq 1,\ Y \geq 1,\ Z \geq 1\}$ montrer que X, Y et Z ne sont pas indépendantes.

4) Version discrète du même exemple : $\Omega = \{0,1\}^3$ d'éléments (x,y,z) avec $x, y, z \in \{0,1\}$, muni de la probabilité vérifiant $\mathbb{P}((0,0,0)) = \mathbb{P}((1,1,0)) = \mathbb{P}((0,1,1)) = \mathbb{P}((1,0,1)) = \frac{1}{4}$ et tous les autres atomes de probabilité nulle. Montrer que les évènements $X = 1$, $Y = 1$ et $Z = 1$ sont indépendants deux à deux mais pas dans leur ensemble.

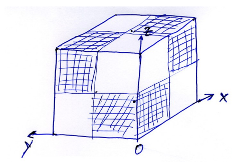

Fig. 1 – L'ensemble Ω_2. Les coordonnées ne sont pas indépendantes sur Ω_2 mais le sont deux à deux.

Exercice 3.4. On considère une urne contenant $r \geq 1$ boules rouges et $b \geq 1$ boules bleues. On tire au hasard les boules sans remise.

1) Combien existe-t-il de tirages complets possibles ?

2) On note X_n le nombre de boules rouges obtenues alors que l'on a tiré n boules. Calculer la loi de X_n.

3) Reconnaitre la loi limite de X_n quand $r \to \infty$, $r/(r+b) \to p \in]0,1[$.

4) On note $Y_k = 1$ si la $k$ième boule est rouge et $Y_k = 0$ sinon. Quelle est la loi de $(Y_1, ..., Y_{r+b})$?

5) En déduire que $Y_1, ..., Y_{r+b}$ ont même loi. Calculer la loi de Y_1.

6) Exprimer X_n en fonction de $Y_1, ..., Y_n$. Calculer $\mathbb{E}[X_n]$ et $Var(X_n)$.

7) On note S_n le nombre de boules rouges lors d'un tirage aléatoire de n boules de l'urne avec remise. Quelle est la loi de S_n ? Comparer avec la question 3).

8) Calculer $\mathbb{E}(S_n)$ et $Var(S_n)$. Comparer avec la question 6).

Exercice 3.5. Application du théorème de Fubini, Durrett R. *Probability : Theory and examples* p. 473
*R*appel : Soit F croissante sur $\mathbb{R}$, continue à droite. Alors il y a une unique mesure $\mu = dF$ sur $\mathbb{R}$ muni de la tribu borélienne telle que $\mu(]a,b]) = F(b) - F(a)$. On dit que F est une fonction de Stieltjes et μ une mesure de Stieljes. La preuve est essentiellement la même que celle qui donne l'existence de la mesure de Lebesgue.

1) Rappeler pourquoi l'ensemble des sauts d'une fonction croissante est dénombrable. Peut-on énumérer ces sauts x_n de telle sorte que $x_{n+1} \leq x_n$, $n \in \mathbb{N}$?

2) Soit $(x_i)_{i\in I}$ l'ensemble des sauts de F et α_i la valeur du saut en x_i. Montrer que

$$dF = \sum_i \alpha_i \delta_{x_i} + d\tilde{F},$$

où $\tilde{F}$ est une fonction croissante continue.

3) Soient F, μ et G, ν deux couples de Stieltjes. On note $\mu = dF$, $\nu = dG$. Montrer que $\mu \otimes \nu(\{a < x = y \leq b\}) = \sum_{x\in]a,b]} \mu(\{x\})\nu(\{x\})$.

4) Montrer que pour $-\infty < a \leq b < +\infty$,

$$(i) \quad \int_{]a,b]} (F(y) - F(a))dG(y) = (\mu \times \nu)(\{(x,y),\ a < x \leq y \leq b\});$$

5) Déduire des deux questions précédentes que

$$(ii) \quad \int_{]a,b]} F(y)dG(y) + \int_{]a,b]} G(y)dF(y) = F(b)G(b) - F(a)G(a) + \sum_{x\in]a,b]} \mu(\{x\})\nu(\{x\}).$$

(Indication : Dans $(\mu \otimes \nu)(\{(x,y),;a < x \leq y \leq b\}) + \mu \otimes \nu(\{(x,y),\ a < y \leq x \leq b\})$, la diagonale apparaît deux fois !).

6) Vérifier sur un exemple comment apparaît le second terme en posant $F(x) = G(x) = \mathbb{1}_{[0,+\infty[}(x)$ en prenant $a < 0 < b$.

$$(iii) \quad \int_{]a,b]} 2F(y)dF(y) = F^2(b) - F^2(a) \text{ si } F = G \text{ est continue.}$$

Solutions

1) Pour tout $n \in \mathbb{N}$, $\#\{i : x_i \geq 1/n\} < \infty$.

2) Classique, poser $d\tilde{F} = dF - \sum_i \alpha_i \delta_{x_i}$ et vérifier que $\tilde{F}$ n'a plus de sauts.

3) En appliquant le théorème de Fubini on a

$$\int\int \mathbb{1}_{y=x}\mu\otimes\nu(dxdy) = \int\left[\int \mathbb{1}_{y=x}(y)\nu(dy)\right]\mu(dx) = \int \nu(\{x\})\mu(dx) = \sum_x \nu(\{x\})\mu(\{x\}).$$

En effet, la deuxième intégrale est en fait une somme dénombrable puisque $\mu(\{x\})$ est nulle sauf pour x appartenant à l'ensemble dénombrable des sauts de F. On intégre donc une fonction nulle sauf en un ensemble dénombrable de points par rapport à la mesure ν.

4) Preuve de (i) :

$$\int_{]a,b]}(F(y) - F(a))dG(y) = \int_{]a,b]}\left[\int_{]a,y]} dF(x)\right]dG(y) = \int_{]a,b]}(\int_{]a,b]} \mathbb{1}_{x\leq y}(x,y)dF(x))dG(y) =$$

$$\int\int_{]a,b]^2} \mathbb{1}_{x\le y}(x,y)\mu\otimes\nu(dxdy) = \mu\otimes\nu(\{(x,y),\ a<x\le y\le b\}),$$

où l'avant dernière égalité s'obtient par le théorème de Fubini.

5) Preuve de (ii). En utilisant le (i),

$$\begin{aligned}
& \int_{]a,b]} F(y)dG(y) + \int_{]a,b]} G(y)dF(y) \\
= \ & F(a)\int_{]a,b]} dG(y) + G(a)\int_{]a,b]} dF(x) + \mu\otimes\nu(\{a<x\le y\le b\}) + (\mu\otimes\nu)(\{a<y\le x\le b\}) \\
= \ & \mu\otimes\nu(\{]a,b]^2) + F(a)(G(b)-G(a)) + G(a)(F(b)-F(a)) + \mu\otimes\nu(\{a<x=y\le b\}) \\
= \ & F(b)G(b) - F(a)G(a) + (\mu\otimes\nu)(\{a<x=y\le b\}).
\end{aligned}$$

3.6 Exercices : Lemmes de Borel-Cantelli

Quelques rappels Soit A_n une suite d'événements d'un espace de probabilités $(\Omega,\mathcal{F},\mathbb{P})$. On écrit "$A_n$, *i.o.*" (*infinitely often*), pour l'événement

$$\{A_n\ i.o.\} = \limsup_n A_n = \bigcap_n \bigcup_{k\ge n} A_k.$$

On dit que A_n se produit "infiniment souvent" si $\mathbb{P}(A_n,\ i.o.) = 1$. Avec ces notations, on rappelle les deux lemmes de Borel-Cantelli
(BC1) : Si E_n est une suite d'événements telle que $\sum_n \mathbb{P}(E_n) < \infty$, alors $\mathbb{P}(E_n,\ i.o.) = 0$.
(BC2) : Si E_n est une suite d'événements indépendants, alors

$$\sum_n \mathbb{P}(E_n) = +\infty \Rightarrow \mathbb{P}(E_n,\ i.o.) = 1.$$

Le but des deux exercices suivants est de donner deux contrexemples à la loi des grands nombres (chapitre suivant). Le premier exercice prouve que celle-ci peut être fausse si les v.a. X_n ne sont pas équidistribuées. Le deuxième prouve que la loi des grands nombres peut être fausse si $E(|X_1|) = +\infty$.

Exercice 3.6. *Un jeu qui n'est équitable qu'en apparence (Williams,* Probability with martingales *p. 228)* Soient X_1, ..., X_n des v.a.i. telles que $X_n = n^2-1$ avec probabilité n^{-2} et $X_n = -1$ avec probabilité $1-n^{-2}$. Démontrer que $\mathbb{E}(X_n) = 0$, mais que si on pose $S_n = X_1 + X_2 + ... + X_n$, alors

$$\frac{S_n}{n} \to -1.$$

On a $\mathbb{P}(X_n = n^2-1) = n^{-2}$, *et donc* $\sum_n \mathbb{P}(X_n = n^2-1) < \infty$. *Par (BC1),* $\mathbb{P}(X_n = n^2-1,\ i.o.) = 0$ *et donc* $\mathbb{P}(X_n = -1,$ *à partir d'un certain rang.*$) = 1$, *ce qui prouve le résultat.*

Exercice 3.7. *Réciproque de la loi forte des grands nombres (Williams,* Probability with martingales*, E4.6 p. 227)*

1) Soit $Z \geq 0$ une v.a. et Y sa partie entière. Montrer que

$$Y = \sum_{n \in \mathbb{N}^\star} 1\!\!1_{Z \geq n}.$$

2) Déduire que

$$\sum_{n \in \mathbb{N}^\star} \mathbb{P}(Z \geq n) \leq \mathbb{E}(Z) \leq \sum_{n \in \mathbb{N}^\star} \mathbb{P}(Z \geq n) + 1.$$

3) Soient X_n, v.a.i.i.d. telles que $\mathbb{E}(|X_n|) = +\infty$. Démontrer que

$$\sum_n \mathbb{P}(|X_n| \geq kn) = \infty \ (k \in \mathbb{N}) \text{ et que } \limsup_n \frac{|X_n|}{n} = \infty, \ p.s.$$

4) En déduire que si $S_n = X_1 + ... + X_n$, alors $\limsup \frac{|S_n|}{n} = \infty$ *p.s.*.

Solutions
1) est immédiat.
2) se déduit en utilisant les inégalités $Y \leq Z \leq Y + 1$ et en prenant l'espérance.
3) Posons $Z = X_0/k$. Alors on sait par le 2) que $\sum_n \mathbb{P}(|Z| \geq n) = +\infty$. Comme $\mathbb{P}(|X_n| \geq kn) = \mathbb{P}(|Z| \geq n)$, on en déduit que $\sum_n \mathbb{P}(|X_n| \geq kn) = +\infty$. Les événements $\{|X_n| \geq kn\}$ étant indépendants, on a par (BC2) : $|X_n| \geq kn$ i.o. et donc $\mathbb{P}(\limsup_n \frac{|X_n|}{n} \geq k) = 1$. Ceci étant vrai pour tout k, on conclut.
4) Utiliser : $\frac{S_n}{n} - \frac{n-1}{n}\frac{S_{n-1}}{n-1} = \frac{X_n}{n}$ et minorer $\frac{|S_n|}{n}$.

Exercice 3.8. Suite du modèle pour le jeu de pile ou face, Billingsley *Probability*, Lecon suite de v.a. de Bernoulli indépedantes On reprend l'exercice sur la décomposition dyiadique d'un réel $x \in [0, 1)$:

$$x = \sum_{i=1}^{\infty} \frac{D_i(x)}{2^i} = 0, D_1(x)D_2(x)D_3(x)... \quad \text{avec } D_i(x) \in \{0, 1\}.$$

On s'intéresse au nombres de zéros consécutifs $L_n(x)$ à partir du $n$ième rang dans le développement dyadique de x :

$$\forall 0 \leq i < L_n(x), D_{n+i}(x) = 0$$

1) Montrer que si la suite d'entiers r_n vérifie $\sum_{n \geq 0} 2^{-r_n} < \infty$, alors $\mathbb{P}(L_n \geq r_n \ i.o.) = 0$.

2) Montrer que si la suite croissante d'entiers r_n vérifie $\sum_{n \geq 0} 2^{-r_n}/r_n = \infty$, alors

$\mathbb{P}(L_n \geq r_n \ i.o.) = 1$
3) En déduire que $\mathbb{P}(\limsup_{n \to \infty} L_n/\log_2(n) = 1) = 1$.

Solutions
1) C'est une application de (BC1) puisque $\mathbb{P}(L_n \geq r_n) = 2^{-r_n}$.
2) On regarde la suite d'entiers u_k *définie par* $u_{k+1} = u_k + r_{u_k}$ *et on définit* $A_k = \{L_{u_k} \geq r_{u_k}\} = \{D_i = 0 : u_k \leq i < u_{k+1}\}$*. Or les* D_i *sont i.i.d et donc les* A_k *sont indépendants de probabilité* $2^{-r_{u_k}}$*. De plus*

$$\sum_{k\geq 0} 2^{-r_{u_k}} = \sum_{k\geq 0} 2^{-r_{u_k}} \frac{\sum_{u_k\leq i<u_{k+1}} 1}{r_{u_k}} \geq \sum_{k\geq 0} \sum_{u_k\leq i<u_{k+1}} \frac{2^{-r_i}}{r_i} = \infty$$

et on peut donc appliquer (BC2).
3) Comme pour tout $\epsilon > 0$, $\sum_{n\geq 1} 2^{-\log_2(n^{1+\epsilon})} = \sum_{n\geq 0} \frac{1}{n^{1+\epsilon}} < \infty$ *et donc* $\mathbb{P}(\limsup_{n\to\infty} L_n/\log_2(n) \geq 1+\epsilon) = 0$ *et en faisant tendre* ϵ *vers zéro,* $\mathbb{P}(\limsup_{n\to\infty} L_n/\log_2(n) > 1) = 0$.
De même comme $\sum_{n\geq 1} 2^{-\log_2(n)}/\log_2(n) = \sum_{n\geq 1} 1/(n\log_2(n)) = \infty$*, on obtient* $\mathbb{P}(\limsup_{n\to\infty} L_n/\log_2(n) \geq 1) = 1$.

Exercice 3.9. Limite supérieure d'une suite de variables normales indépendantes, Williams, *Probability with martingales*, E4.5 p. 227

1) Soit G une v.a. normale $\mathcal{N}(0,1)$. On pose $\varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ Démontrer que pour $x > 0$,

$$(x + x^{-1})^{-1}\varphi(x) \leq \mathbb{P}(G \geq x) \leq \frac{1}{x}\varphi(x).$$

2) Soit $(X_n)_n$ une suite de v.a.i. normales $\mathcal{N}(0,1)$. Montrer qu'avec probabilité 1, $L \leq 1$, où L est définie par

$$L := \limsup_n \frac{X_n}{\sqrt{2\log n}}.$$

3) Montrer que $\mathbb{P}(L = 1) = 1$.

4) On pose $S_n = X_1 + ... + X_n$ et on rappelle que $\frac{S_n}{\sqrt{n}}$ est encore normale $\mathcal{N}(0,1)$. Montrer que $\mathbb{P}(|S_n| < 2\sqrt{n\log n},\ ev.) = 1$.

Voir la section 14.7 de Williams pour un résultat plus précis : la loi du Logarithme Itéré. Remarquer que cela implique la loi forte des grands nombres dans ce cas particulier : $\mathbb{P}(\frac{S_n}{n} \to 0) = 1$.

Solutions 1) Pour l'inégalité de droite, utiliser $\varphi'(y) = -y\varphi(y)$ *et pour l'inégalité de gauche,* $(\frac{\varphi(y)}{y})' = -(1 + \frac{1}{y^2})\varphi(y)$.

2) Par l'inégalité de droite de la question 1), on a

$$\mathbb{P}(|X_n| \geq L\sqrt{2\log n}) \leq \frac{1}{L\sqrt{4\pi\log n}} e^{-L^2\log n} = \frac{1}{L\sqrt{4\pi\log n}} n^{-L^2}.$$

Si $L > 1$*, alors* $\sum_n \mathbb{P}(|X_n| \geq L\sqrt{2\log n}) < \infty$ *et, par (BC1),* $\mathbb{P}(|X_n| \geq L\sqrt{2\log n}$ *i.o.*$) = 0$*, et donc* $\mathbb{P}(\limsup_n \frac{|X_n|}{\sqrt{2\log n}} \leq L) = \mathbb{P}(|X_n| \leq L\sqrt{2\log n}$ *ev.*$) = 1$.

3) On par l'inégalité de gauche du 1),

$$\mathbb{P}(X_n \geq \sqrt{2\log n}) \geq (L\sqrt{2\log n} + \frac{1}{L\sqrt{2\log n}})n^{-L^2}.$$

Donc, si $L \leq 1$, *on obtient* $\sum_n \mathbb{P}(|X_n| \geq L\sqrt{2\log n}) = +\infty$, *et donc, par (BC2),* $\mathbb{P}(|X_n| \geq L\sqrt{2\log n},\ i.o.) = 1$. *Utilisant ceci, et le résultat du 2), on conclut que* $L = 1$, *p.s.*

4) Les S_n *ne sont pas indépendantes, mais on peut toujours invoquer le résultat du 2), qui n'utilise pas l'indépendance des* X_n. *On a donc*

$$\limsup_n (\frac{S_n}{\sqrt{2n\log n}}) \leq 1,\ p.s.$$

Cela implique que pour tout $C > \sqrt{2}$, $\mathbb{P}(\frac{S_n}{\sqrt{n\log n}} \leq C,\ ev.) = 1$ *et on obtient le résultat annoncé en prenant par exemple* $C = 2 > \sqrt{2}$.

Exercices 9.10, 9.11, 9.12 de Ouvrard Tome II, p. 84-85.