9.2 Indépendance

Définition 9.22 *Soit $p \in \mathbb{N}$, $p \geq 2$, $d_1, \ldots, d_p \in \mathbb{N}^\star$. soit $(\Omega, \mathcal{A})$ un espace probabilisable et, pour tout $i \in \{1, \ldots, p\}$, X_i un v.a. de dimension d_i. Les v.a. $\mathrm{X}_1, \ldots, \mathrm{X}_p$ sont dits indépendants si les tribus $\sigma(\mathrm{X}_1), \ldots, \sigma(\mathrm{X}_p)$ (engendrées par $\mathrm{X}_1, \ldots, \mathrm{X}_p$) sont indépendantes (cf définition 2.57, p. 43)*

La proposition suivante donne des opérations possibles sur l'indépendance.

Proposition 9.23 *Soit $p \in \mathbb{N}$, $p \geq 2$, $d_1, \ldots, d_p \in \mathbb{N}^\star$. soit $(\Omega, \mathcal{A})$ un espace probabilisable et, pour tout $i \in \{1, \ldots, p\}$, X_i un v.a. de dimension d_i. On suppose que les v.a. $\mathrm{X}_1, \ldots, \mathrm{X}_p$ sont indépendants. On se donne maintenant une suite strictement croissante $p_0, \ldots, p_q$ $(q \geq 2)$ t.q. $1 = p_0 < \ldots < p_q = p+1$ et, pour $i \in \{1, \ldots, q\}$, on note Y_i le v.a. $(\mathrm{X}_{p_{i-1}}, \ldots, \mathrm{X}_{p_i - 1})^t$, qui est donc un v.a. de dimension $r_i = d_{p_{i-1}} + \ldots + d_{p_i - 1}$. On a alors*

1. Les v.a. $\mathrm{Y}_1, \ldots, \mathrm{Y}_q$ sont indépendantes,

2. Si φ_i est, pour tout $i \in \{1, \ldots, q\}$, une application borélienne de $\mathbb{R}^{r_i}$ dans $\mathbb{R}^{s_i}$ (avec $s_i \in \mathbb{N}^\star$), les v.a. $\varphi_1(\mathrm{Y}_1), \ldots, \varphi_q(\mathrm{Y}_q)$ sont indépendants.

DÉMONSTRATION – le premier item est une conséquence immédiate de la proposition 2.59 (sur l'indépendance de tribus) et de la proposition 9.3. Le second item est une conséquence immédiate du premier car $\sigma(\varphi_i(\mathrm{Y}_i)) \subset \sigma(\mathrm{Y}_i)$ pour tout i. ■

Remarque 9.24 Soit $(\Omega, \mathcal{A})$ un espace probabilisable et X, Y deux v.a. (pas nécessairement de même dimension). La proposition 9.23 montre que si X et Y sont indépendants, toute composante de X est indépendante de toute composante de Y. Réciproquement, si toute composante de X est indépendante de toute composante de Y, on en déduit que X et Y sont indépendants. Ceci est encore une conséquence simple de la proposition 2.59.

On donne maintenant une généralisation de la proposition 4.59.

Proposition 9.25 *Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $n \geq 2$ et $\mathrm{X}_1, \ldots, \mathrm{X}_n$ des v.a. indépendants, de dimension $d_1, \ldots, d_n \in \mathbb{N}^\star$.*

1. Soit, pour tout $i \in \{1, \ldots, n\}$, une fonction borélienne, notée φ_i, de $\mathbb{R}^{d_i}$ dans $\overline{\mathbb{R}}_+$. On a alors :

$$\mathrm{E}(\prod_{i=1}^{n} \varphi_i(\mathrm{X}_i)) = \prod_{i=1}^{n} \mathrm{E}(\varphi_i(\mathrm{X}_i)). \tag{9.2}$$

(En convenant qu'un produit de termes est nul si l'un des termes est nul.)

2. Soit, pour tout $i \in \{1, \ldots, n\}$, φ_i une fonction borélienne de $\mathbb{R}^{d_i}$ dans $\mathbb{R}$. On suppose que $\varphi(\mathrm{X}_i)$ est intégrable pour tout $i = 1, \ldots, n$. La v.a.r. $\prod_{i=1}^{n} \varphi_i(\mathrm{X}_i)$ est alors intégrable et l'égalité (9.2) est vraie.

3. Soit, pour tout $i \in \{1, \ldots, n\}$, φ_i une fonction borélienne bornée de $\mathbb{R}^{d_i}$ dans $\mathbb{R}$. Alors, l'égalité (9.2) est vraie.

N.B. Comme dans la proposition 4.59, l'item 3 est donc une condition nécessaire suffisante pour que les v.a. $\mathrm{X}_1, \ldots, \mathrm{X}_n$ soient indépendants.

La démonstration suit pas à pas celle de la proposition 4.59, sans difficulté supplémentaire.

Remarque 9.26 Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé. La proposition 9.25 donne aussi des résultats quand les fonctions φ_i sont à la valeurs dans $\mathbb{R}^{r_i}$ avec $r_i \neq 1$. Par exemple, soit X, Y deux v.a. indépendants de dimensions d. On suppose X et Y intégrables (c'est-à-dire $\mathrm{E}(|\mathrm{X}|) < \infty$ et $\mathrm{E}(|\mathrm{Y}|) < \infty$). La v.a.r. $\mathrm{X} \cdot \mathrm{Y}$ est alors intégrable et $\mathrm{E}(\mathrm{X} \cdot \mathrm{Y}) = \mathrm{E}(\mathrm{X}) \cdot \mathrm{E}(\mathrm{Y})$.

Plus généralement, soit X est un v.a. de dimension d, Y est un v.a. de dimension $\bar{d}$, X, Y indépendants. On suppose que φ est une fonction borélienne de $\mathbb{R}^d$ dans $\mathbb{R}^r$ et ψ une fonction borélienne de $\mathbb{R}^{\bar{d}}$ dans $\mathbb{R}^r$ t.q. $\varphi \circ \mathrm{X}$ et $\psi \circ \mathrm{Y}$ soient intégrables. On a alors $\varphi \circ \mathrm{X} \cdot \psi \circ \mathrm{Y}$ intégrable et $\mathrm{E}(\varphi \circ \mathrm{X} \cdot \psi \circ \mathrm{Y}) = \mathrm{E}(\varphi \circ \mathrm{X}) \cdot \mathrm{E}(\psi \circ \mathrm{Y})$.

Enfin, si $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont des v.a. indépendants de dimension d ($d \geq 1$), la formule donnée dans la proposition 6.97 se généralise et donne :

$$\mathrm{Cov}(\mathrm{X}_1 + \ldots + \mathrm{X}_n) = \sum_{i=1}^{n} \mathrm{Cov}(\mathrm{X}_i).$$

Pour le montrer, il suffit de traiter le cas $n = 2$ (qui est traité dans l'exercice 9.7) et de faire une récurrence sur n.

On donne maintenant la généralisation de la proposition 4.61.

Proposition 9.27 *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé,* $n \geq 2$ *et* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *des v.a. de dimension* $d_1, \ldots, d_n \in \mathbb{N}^*$. *Ces v.a. sont indépendants si et seulement si on a, pour tout famille* $\{\varphi_i,\ i = 1, \ldots n\}$ *t.q.* $\varphi_i \in \mathrm{C}_c(\mathbb{R}^{d_i}, \mathbb{R})$ *pour tout* i,

$$\mathrm{E}\Big(\prod_{i=1}^{d} \varphi_i(\mathrm{X}_i)\Big) = \prod_{i=1}^{d} \mathrm{E}(\varphi_i(\mathrm{X}_i)), \tag{9.3}$$

(En convenant qu'un produit de termes est nul si l'un des termes est nul.)

Ici encore, la démonstration suit pas à pas celle de la proposition 4.61, sans difficulté supplémentaire.

Théorème 9.28 (Loi d'un couple de v.a. indépendantes) *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé.*

1. *Soit* X, Y *deux v.a.r.. Les v.a.r* X *et* Y *sont indépendantes si et seulement si la loi du v.a.* $(\mathrm{X}, \mathrm{Y})^t$ *(ou du couple* (X, Y)*) est* $\mathrm{P}_{(\mathrm{X},\mathrm{Y})^t} = \mathrm{P}_\mathrm{X} \otimes \mathrm{P}_\mathrm{Y}$.
2. *Soit* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *(*$n \geq 2$*)* n *v.a. de dimension* $d_1, \ldots, d_n \in \mathbb{N}^*$.
   On pose $\mathrm{Z} = (\mathrm{X}_1, \ldots, \mathrm{X}_n)^t$ *(Z est donc un v.a. de dimension* $d_1 + \ldots + d_n$*). Les v.a.* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *sont indépendantes si et seulement si la loi du v.a.* Z *est* $\mathrm{P}_\mathrm{Z} = \mathrm{P}_{\mathrm{X}_1} \otimes \ldots \otimes \mathrm{P}_{\mathrm{X}_n}$.

DÉMONSTRATION – La démonstration du premier item fait partie de l'exercice 9.8. le second item est une généralisation assez simple (en faisant, par exemple, une récurrence sur n pour se ramener au cas $n = 2$). ■

On utilise souvent en théorie des probabilités une suite (finie ou infinie) de v.a.r. indépendantes ayant des lois prescrites. Le théorème suivant montre qu'il existe effectivement un espace probabilisé et une suite de variables aléatoires réelles indépendantes sur cet espace ayant des lois prescrites.

Théorème 9.29 (Existence de v.a.r. indépendantes de lois prescrites)
Soit $(p_n)_{n \in \mathbb{N}^*}$ *une suite de probabilités sur* $\mathcal{B}(\mathbb{R})$. *On a alors :*

1. *Pour tout $n \in \mathbb{N}^*$, il existe un espace probabilisé, noté $(\Omega, \mathcal{A}, \mathrm{P})$, et une suite finie de v.a.r. indépendantes, notées $\mathrm{X}_1, \ldots, \mathrm{X}_n$, t.q. $\mathrm{P}_{\mathrm{X}_k} = p_k$ pour tout $k \in \{1, \ldots, n\}$.*
2. *Il existe un espace probabilisé, noté $(\Omega, \mathcal{A}, \mathrm{P})$, et une suite de v.a.r. indépendantes, notée $(\mathrm{X}_n)_{n \in \mathbb{N}^*}$ t.q. $\mathrm{P}_{\mathrm{X}_n} = p_n$ pour tout $n \in \mathbb{N}^*$.*

DÉMONSTRATION – Nous démontrons ici le premier item. Le second (plus difficile) sera admis. Soit $n \in \mathbb{N}^*$. Un raisonnement par récurrence utilisant le théorème d'existence (et unicité) de la mesure produit (théorème 7.3) permet de construire une mesure p sur $\mathcal{B}(\mathbb{R}^n)$ vérifiant, pour toute famille $\mathrm{A}_1, \ldots, \mathrm{A}_n$ de boréliens de $\mathbb{R}$:

$$p(\prod_{i=1}^{n} \mathrm{A}_i) = \prod_{i=1}^{n} p_i(\mathrm{A}_i).$$

On a donc $p = p_1 \otimes \ldots \otimes p_n$.

On prend alors $(\Omega, \mathcal{A}, \mathrm{P}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), p)$ et, pour tout $i = 1, \ldots, n$, X_i est l'application qui a $\omega \in \mathbb{R}^n$ associe sa i-ième composante. Enfin, on note $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_n)^t$, de sorte que X est un v.a. de dimension n.

Pour tout $\omega \in \mathbb{R}^n$, on a $\mathrm{X}(\omega) = \omega$, ceci prouve que $\mathrm{P}_\mathrm{X} = \mathrm{P}$. Puis, pour tout $i \in \{1, \ldots, n\}$, on a $\mathrm{X}(\omega) = \omega_i$ (où ω_i désigne la i-ième composante de ω). On en déduit que $\mathrm{P}_{\mathrm{X}_i} = p_i$. Enfin, comme $p = p_1 \otimes \ldots \otimes p_n$, on a aussi $\mathrm{P}_\mathrm{X} = p_{\mathrm{X}_1} \otimes \ldots \otimes p_{\mathrm{X}_n}$. Le théorème 9.28 donne donc que les v.a.r. $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes. ∎

On s'intéresse maintenant à la somme de variables aléatoires indépendantes.

Proposition 9.30 (Loi de la somme de v.a. indépendantes) *Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $d \geq 1$ et X, Y deux v.a. indépendants de dimension d. Alors, $\mathrm{P}_{\mathrm{X}+\mathrm{Y}} = \mathrm{P}_\mathrm{X} * \mathrm{P}_\mathrm{Y}$.*

DÉMONSTRATION – Soit φ une fonction borélienne bornée de $\mathbb{R}^d$ dans $\mathbb{R}$. Comme X et Y sont indépendants, on a $\mathrm{P}_{(\mathrm{X},\mathrm{Y})} = \mathrm{P}_\mathrm{X} \otimes \mathrm{P}_\mathrm{Y}$ (par le théorème 9.28) et donc :

$$\int_\Omega \varphi(\mathrm{X}+\mathrm{Y}) d\mathrm{P} = \int_{\mathbb{R}^{2d}} \varphi(x+y) d\mathrm{P}_{(\mathrm{X},\mathrm{Y})}(x,y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x+y) d\mathrm{P}_\mathrm{X}(x) d\mathrm{P}_\mathrm{Y}(y).$$

La définition de la convolution de mesure donne (voir la définition 7.26 et la proposition 7.27 :

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x+y) d\mathrm{P}_\mathrm{X}(x) d\mathrm{P}_\mathrm{Y}(y) = \int_{\mathbb{R}^d} \varphi(z) d(\mathrm{P}_\mathrm{X} * \mathrm{P}_\mathrm{Y})(z).$$

On en déduit que $\int_\Omega \varphi(\mathrm{X}+\mathrm{Y}) d\mathrm{P} = \int_{\mathbb{R}^d} \varphi(z) d(\mathrm{P}_\mathrm{X} * \mathrm{P}_\mathrm{Y})(z)$, c'est-à-dire $\mathrm{P}_{\mathrm{X}+\mathrm{Y}} = \mathrm{P}_\mathrm{X} * \mathrm{P}_\mathrm{Y}$. ∎

9.3 Vecteurs gaussiens

Soit $m \in \mathbb{R}$ et $\sigma > 0$. On rappelle que la loi normale $\mathcal{N}(m, \sigma^2)$ est la probabilité (sur $\mathcal{B}(\mathbb{R})$) de densité f avec, pour $x \in \mathbb{R}$,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Pour introduire les v.a. gaussiens, il est utile de définir aussi la loi normale $\mathcal{N}(m, 0)$. Cette loi normale est la probabilité (sur $\mathcal{B}(\mathbb{R})$) δ_m (appelée mesure de Dirac au point m). Comme elle vérifie, pour tout fonction φ borélienne de $\mathbb{R}$ dans $\mathbb{R}$, $\int \varphi d\delta_m = \varphi(m)$, il est facile de vérifier que $\mathcal{N}(m, 0)$ est la limite étroite, quand $\sigma \to 0$, $\sigma > 0$, de $\mathcal{N}(0, \sigma^2)$.

Définition 9.31 *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé,* $d \geq 1$ *et* X *un v.a. de dimension* d*. le v.a.* X *est un v.a. gaussien si* $u \cdot \mathrm{X}$ *est, pour tout* $u \in \mathbb{R}^d$*, une v.a.r. gaussienne.*

Remarque 9.32 Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $d > 1$ et X un v.a. de dimension d. On suppose que chaque composante de X est une v.a.r. gaussienne. Alors, le vecteur X n'est pas nécessairement gaussien. Mais, si les composantes de X sont indépendantes, le vecteur X est alors un v.a. gaussien. On peut aussi montrer que si X est un v.a. gaussien et que les composantes de X sont indépendantes deux à deux (ou si on a seulement $\mathrm{Cov}(\mathrm{X}_i, \mathrm{X}_j) = 0$ pour tout i, j t.q. $i \neq j$), alors les composantes de X sont indépendantes (voir l'exercice 10.14 pour tous ces résultats).

Les v.a. gaussiens sont les vecteurs qui suivent un loi normale multidimensionnelle, dont la définition et donnée dans la définition 9.5, à condition d'étendre convenablement la définition 9.5 au cas D non inversible. C'est l'objectif de la proposition suivante.

Proposition 9.33 *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé,* $d \geq 1$ *et* X *un v.a. de dimension* d*.*

1. *Soit* $m \in \mathbb{R}^d$ *et* D *une matrice s.d.p. (de taille* $d \times d$*). Si* $\mathrm{X} \sim \mathcal{N}(m, \mathrm{D})$*, où* $\mathcal{N}(m, \mathrm{D})$ *est définie par la définition 9.5, alors* X *est un vecteur gaussien.*
2. *Si* X *est un vecteur gaussien. On pose* $m = \mathrm{E}(\mathrm{X})$ *et* $\mathrm{D} = Cov(\mathrm{X})$*. Alors, la loi de* X *ne dépend que de* m *et* D*. Si* D *est inversible on a* $\mathrm{X} \sim \mathcal{N}(m, \mathrm{D})$ *(où* $\mathcal{N}(m, \mathrm{D})$ *est définie par la définition 9.5).*

Ceci permet de définir $\mathcal{N}(m, \mathrm{D})$ *dans le cas où* D *est seulement symétrique semi–définie positive (et* $m \in \mathbb{R}^d$*). On définit* $\mathcal{N}(m, \mathrm{D})$ *comme étant la loi d'un vecteur gaussien de dimension* d *t.q.* $m = \mathrm{E}(\mathrm{X})$ *et* $\mathrm{D} = Cov(\mathrm{X})$ *(on peut montrer qu'un tel vecteur existe).*

Le premier item est démontré dans l'exercice 10.12 et le deuxième item est démontré dans l'exercice 10.13.

Remarque 9.34 La notion de vecteur aléatoire gaussien (et de loi normale multidimensionnelle) généralise la notion de variable aléatoire gaussienne (et de loi normale). Le théorème central limite que nous avons énoncé dans la remarque 6.100 se généralise au cas vectoriel. Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $d \geq 1$ et $(\mathrm{X}_n)_{n\in\mathbb{N}^*}$ une suite de v.a. de dimension d i.i.d.. On suppose que $\mathrm{E}(|\mathrm{X}_1|^2) < \infty$. On pose $\mathrm{E}(\mathrm{X}_1) = m$, $\mathrm{D} = \mathrm{Cov}(\mathrm{X}_1)$ et, pour $n \in \mathbb{N}^*$,

$$\mathrm{Y}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\mathrm{X}_i - m).$$

La suite $(\mathrm{Y}_n)_{n\in\mathbb{N}^*}$ converge alors en loi vers tout v.a. Y dont la loi est la normale multidimensionnelle $\mathcal{N}(0, \mathrm{D})$. (Voir la définition 9.5 et la proposition 9.33 pour la définition de cette loi normale multidimensionnelle.) Ceci sera démontré au chapitre 10 (théorème 10.24) en utilisant la transformation de Fourier.