

4.4 Mesures et probabilités de densité

4.4.1 Définitions

A partir d'une mesure et d'une fonction mesurable positive, on peut définir une autre mesure de la manière suivante :

Définition 4.21 (Mesure de densité) Soient (E, \mathcal{T}, m) un espace mesuré et $f \in \mathcal{M}_+$. Pour $A \in \mathcal{T}$, on rappelle que $f 1_A$ est la fonction (de E dans $\overline{\mathbb{R}}_+$) définie par $f 1_A(x) = f(x)$ si $x \in A$ et $f 1_A(x) = 0$ si $x \in A^c$ (cette fonction appartient à \mathcal{M}_+) et on définit $\int_A f dm$ par $\int f 1_A dm$.

On définit alors $\mu : \mathcal{T} \rightarrow \overline{\mathbb{R}}_+$ par :

$$\mu(A) = \int f 1_A dm = \int_A f dm, \quad \forall A \in \mathcal{T}.$$

L'application μ ainsi définie est une mesure sur \mathcal{T} (ceci est démontré dans l'exercice 4.25), appelée mesure de densité f par rapport à m , et notée $\mu = f m$.

Proposition 4.22 Soient (E, \mathcal{T}, m) un espace mesuré, $f \in \mathcal{M}_+$ et μ la mesure de densité f par rapport à m . Alors, la mesure μ est absolument continue par rapport à la mesure m , c'est-à-dire que si $A \in \mathcal{T}$ est tel que $m(A) = 0$, alors $\mu(A) = 0$.

DÉMONSTRATION – Soit $A \in \mathcal{T}$ tel que $m(A) = 0$. On a alors $f 1_A = 0$ m -p.p. et donc $\mu(A) = \int f 1_A dm = 0$ d'après le lemme 4.12. ■

On déduit de cette proposition que la mesure de Dirac en 0, définie en (2.2), n'est pas une mesure de densité par rapport à la mesure de Lebesgue (on peut montrer que ces deux mesures sont étrangères (voir définition 2.29 et proposition 2.30).

Notons que l'on peut aussi définir des mesures signées de densité, voir la définition 6.74.

4.4.2 Exemples de probabilités de densité

Définition 4.23 (Probabilité de densité) Soit p une probabilité sur $\mathcal{B}(\mathbb{R})$, on dit que p est une probabilité de densité (par rapport à Lebesgue) s'il existe $f \in \mathcal{M}_+$ t.q. $\int f d\lambda = 1$ et $p(A) = \int f 1_A d\lambda = \int_A f d\lambda$ pour tout $A \in \mathcal{B}(\mathbb{R})$.

Les lois de probabilité sur $\mathcal{B}(\mathbb{R})$, de densité par rapport à la mesure de Lebesgue, données dans la proposition suivante seront souvent utilisées dans le calcul des probabilités. (On rappelle qu'une loi de probabilité est, par définition, une probabilité sur $\mathcal{B}(\mathbb{R})$).

Définition 4.24 (Quelques lois de densité sur $\mathcal{B}(\mathbb{R})$) On donne ici trois exemples de lois de densité.

1. Loi uniforme, $\mathcal{U}(a, b)$ Soit $a, b \in \mathbb{R}$, $a < b$, la loi uniforme sur $[a, b]$ est la loi de densité $\frac{1}{b-a} 1_{[a, b]} : p(A) = \frac{1}{b-a} \int 1_{[a, b]} 1_A d\lambda, \forall A \in \mathcal{B}(\mathbb{R})$.

2. Loi exponentielle, $\mathcal{E}(\tau)$ Soit $\tau > 0$; la loi exponentielle est définie par la densité f définie par :

$$f(x) = \begin{cases} 0 & \text{si } x < 0, \\ \tau e^{-\tau x} & \text{si } x \geq 0. \end{cases}$$

3. Loi de Gauss, $\mathcal{N}(\mu, \sigma^2)$ Soit $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*$; la loi de Gauss de paramètre (μ, σ) est définie par la densité f définie par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \text{ pour } x \in \mathbb{R}.$$

Dans le théorème 4.44, l'hypothèse de convergence p.p. de f_n vers f peut être remplacée par une hypothèse de convergence en mesure (plus précisément, avec l'hypothèse de domination donnée dans le théorème 4.44, on a même équivalence entre la convergence en mesure et la convergence dans L^1). On obtient ainsi le théorème suivant (ou seule la partie utile de cette équivalence est donnée). On rappelle aussi que, lorsque l'espace mesuré est un espace probabilisé, la convergence en mesure est appelée convergence en probabilité.

Théorème 4.45 (Convergence en mesure dominée) Soit (E, T, m) un espace mesuré. L'espace $L^1_{\mathbb{R}}(E, T, m)$ est noté L^1 . Soit $(f_n)_{n \in \mathbb{N}} \subset L^1$ et f une fonction de E dans \mathbb{R} telles que :

1. $f_n \rightarrow f$ en mesure.

2. $\exists F \in L^1$ t.q., pour tout $n \in \mathbb{N}$, $|f_n| \leq F$ p.p..

Alors $f \in L^1$ (au sens où il existe $g \in \mathcal{L}^1_{\mathbb{R}}(E, T, m)$ t.q. $f = g$ p.p.) et $f_n \rightarrow f$ dans L^1 , c'est-à-dire

$$\int |f_n - f| dm \rightarrow 0 \text{ lorsque } n \rightarrow +\infty.$$

Ceci donne aussi $\int f_n dm \rightarrow \int f dm$, lorsque $n \rightarrow +\infty$.

DÉMONSTRATION – En choisissant des représentants de f_n et f , la démonstration de ce théorème se ramène à celle de l'exercice 4.35. ■

4.9 Espérance et moments des variables aléatoires

Définition 4.54 (Espérance, moment, variance) Soient (Ω, \mathcal{A}, p) un espace probabilisé et X une variable aléatoire réelle.

1. Si $X \geq 0$ (c'est-à-dire $X(\omega) \geq 0$ pour tout $\omega \in \Omega$), on définit l'espérance $E(X)$ de la variable aléatoire X par $E(X) = \int X(\omega) dp(\omega)$.

2. Si $X \in L^1_{\mathbb{R}}(\Omega, \mathcal{A}, p)$ (c'est-à-dire $E(|X|) < +\infty$), on définit l'espérance $E(X)$ de la variable aléatoire X par :

$$E(X) = \int X(\omega) dp(\omega).$$

On définit la variance de X par $\text{Var}(X) = \sigma^2(X) = E((X - E(X))^2)$ (avec $\sigma(X) \geq 0$).

3. Pour $r \in [1, +\infty[$, le moment d'ordre r de la variable aléatoire X est l'espérance de la variable aléatoire $|X|^r$.

Définition 4.55 (Covariance) Soient (Ω, \mathcal{A}, p) un espace probabilisé et X, Y deux v.a.r. t.q. $E(X^2) < +\infty$ et $E(Y^2) < +\infty$. On définit la covariance de X et Y par : $\text{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$. (Remarque que $(X - E(X))(Y - E(Y))$ est une v.a.r. intégrable car sa valeur absolue est majorée, par exemple, par $X^2 + Y^2 + E(X)^2 + E(Y)^2$ qui est intégrable.)

On calcule rarement l'espérance d'une v.a. comme intégrale par rapport à la probabilité p ; en effet, l'espace (Ω, \mathcal{A}, p) est souvent mal connu. Le théorème 4.58 montre qu'il suffit en fait de connaître la loi de la v.a. X pour calculer son espérance (ou, plus généralement, l'espérance d'une fonction de X). On se ramène ainsi au calcul d'une intégrale sur \mathbb{R} .

Les deux inégalités suivantes découlent immédiatement du lemme 4.15 :

Lemme 4.56 (Inégalité de Markov) Soient (Ω, \mathcal{A}, p) un espace probabilisé, X une variable aléatoire réelle positive sur Ω et $\lambda \in \mathbb{R}_+^*$. On suppose que $0 < E(X) < +\infty$. Alors :

$$p(\{X \geq \lambda E(X)\}) \leq \frac{1}{\lambda}.$$

DÉMONSTRATION – Il suffit, par exemple, d'appliquer le lemme 4.15 avec $f = X$ et $t = \lambda E(X)$. ■

Lemme 4.57 (Inégalité de Bienaymé Tchebychev) Soit (Ω, \mathcal{A}, P) un espace probabilisé, X une variable aléatoire réelle sur Ω , intégrable et t.q. sa variance vérifie $0 < \sigma^2(X) < +\infty$, et $\lambda \in \mathbb{R}_+^*$. Alors :

$$P(\{|X - E(X)| \geq \lambda \sigma(X)\}) \leq \frac{1}{\lambda^2}.$$

DÉMONSTRATION – Appliquer le lemme 4.15 avec $f = |X - E(X)|^2$ et $t = \lambda \sigma(X)$. ■

Soit (Ω, \mathcal{A}, P) un espace probabilisé, X une variable aléatoire réelle sur Ω . La loi de X , notée P_X est définie par $P_X(A) = P(X^{-1}(A))$, pour tout $A \in \mathcal{B}(\mathbb{R})$. Ceci est équivalent à dire que pour tout $A \in \mathcal{B}(\mathbb{R})$, on a, avec $\varphi = 1_A$:

$$\int_{\Omega} \varphi \circ X(\omega) dP(\omega) = \int_{\mathbb{R}} \varphi(x) dP_X(x). \quad (4.11)$$

On rappelle que $\varphi \circ X$ est souvent improprement noté $\varphi(X)$, ce qui s'explique par le fait $\varphi \circ X(\omega) = \varphi(X(\omega))$ pour tout $\omega \in \Omega$. Le théorème 4.58 montre que cette égalité est vraie pour une large classe de fonctions boréliennes φ de \mathbb{R} dans \mathbb{R} ou $\overline{\mathbb{R}}_+$ (on rappelle que borélienne signifie mesurable quand les espaces sont munis de la tribu de Borel).

Théorème 4.58 (Loi image) *Soit (Ω, \mathcal{A}, P) un espace probabilisé, X une variable aléatoire réelle sur Ω et P_X la loi de la variable aléatoire X . On a alors :*

1. *L'égalité (4.11) est vraie pour toute fonction φ borélienne de \mathbb{R} dans $\overline{\mathbb{R}}_+$ et toute fonction borélienne bornée de \mathbb{R} dans \mathbb{R} .*
2. *Soit φ une fonction borélienne de \mathbb{R} dans \mathbb{R} , la fonction $\varphi \circ X$ appartient à $L^1_{\mathbb{R}}(\Omega, \mathcal{A}, P)$ si et seulement si $\varphi \in L^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$. De plus, si $\varphi \in L^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$, L'égalité (4.11) est vraie.*

DÉMONSTRATION – On remarque que (4.11) est vraie pour tout $\varphi = 1_A$, avec $A \in \mathcal{B}(\mathbb{R})$ (par définition de P_X). Par linéarité positive, (4.11) est encore vraie pour tout φ borélienne étagée positive de \mathbb{R} dans \mathbb{R} . Par convergence monotone, (4.11) est alors vraie pour tout φ borélienne de \mathbb{R} dans $\overline{\mathbb{R}}_+$. Ceci donne la première partie du premier item. En utilisant la décomposition $\varphi = \varphi^+ - \varphi^-$, on montre alors le deuxième item. Enfin, la deuxième partie du premier item vient du fait que φ est intégrable pour la probabilité P_X si φ est borélienne bornée. ■

Un produit de v.a.r. intégrables et indépendantes est une v.a.r. intégrable (ce qui est, bien sûr, faux sans l'hypothèse d'indépendance) et l'espérance de ce produit est égal au produit des espérances. Ce résultat plus général est donnée dans la proposition suivante.

Proposition 4.59 *Soit (Ω, \mathcal{A}, P) un espace probabilisé, $d > 1$ et X_1, \dots, X_d des v.a.r. indépendantes.*

1. *Soit $\varphi_1, \dots, \varphi_d$ des fonctions boréliennes de \mathbb{R} dans $\overline{\mathbb{R}}_+$. On a alors :*

$$E\left(\prod_{i=1}^d \varphi_i(X_i)\right) = \prod_{i=1}^d E(\varphi_i(X_i)). \quad (4.12)$$

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

2. *Soit $\varphi_1, \dots, \varphi_d$ des fonctions boréliennes de \mathbb{R} dans \mathbb{R} . On suppose que $\varphi_i(X_i)$ est intégrable pour tout $i = 1, \dots, d$. La v.a.r. $\prod_{i=1}^d \varphi_i(X_i)$ est intégrable et l'égalité (4.12) est vraie.*
3. *Soit $\varphi_1, \dots, \varphi_d$ des fonctions boréliennes bornées de \mathbb{R} dans \mathbb{R} . L'égalité (4.12) est vraie.*

N.B. Si X_1, \dots, X_d sont des v.a.r., le fait que (4.12) soit vraie pour toute famille $\varphi_1, \dots, \varphi_d$ de fonctions boréliennes bornées de \mathbb{R} dans \mathbb{R} est donc une condition nécessaire et suffisante pour les v.a.r. X_1, \dots, X_d soient indépendantes.

DÉMONSTRATION – Si $\varphi_1, \dots, \varphi_d$ sont des fonctions caractéristiques de boréliens de \mathbb{R} , l'égalité (4.12) est une conséquence immédiate de la définition de l'indépendance des X_i (Si $\varphi_i = 1_{A_i}$ avec $A_i \in \mathcal{B}(\mathbb{R})$, on a $E(\varphi_i(X_i)) = P(\{X_i \in A_i\}) = P(X_i^{-1}(A_i))$). Par linéarité positive, on en déduit que (4.12) est vraie si les fonctions φ_i sont (boréliennes) étagées positives (c'est-à-dire $\varphi \in \mathcal{E}_+$). Puis, par convergence monotone, on en déduit le premier item de la proposition (car toute fonction borélienne de \mathbb{R} dans $\overline{\mathbb{R}}_+$ est limite croissante d'éléments de \mathcal{E}_+).

Pour le deuxième item, on utilise (4.12) avec la fonction $x \mapsto |\varphi_i(x)|$ au lieu de la fonction φ_i (pour tout i). On montre ainsi que la v.a.r. $\prod_{i=1}^d \varphi_i(X_i)$ est intégrable. Puis, on montre (4.12) par linéarité (utilisant $\varphi_i = \varphi_i^+ - \varphi_i^-$).

Le troisième item est conséquence immédiate du deuxième (car si X est une v.a.r. et φ est une fonction borélienne bornée, la v.a.r. $\varphi(X)$ est intégrable). ■

Une conséquence de la proposition 4.59 est que XY est intégrable et $\text{cov}(X, Y) = 0$ si X, Y sont deux v.a.r. indépendantes et intégrables sur un espace probabilisé (Ω, \mathcal{A}, p) .

Pour montrer que des v.a.r. sont indépendantes, il est parfois utile de savoir qu'il suffit de montrer (4.12) lorsque les fonctions φ_i sont continues à support compact de \mathbb{R} dans \mathbb{R} . C'est l'objet de la proposition 4.61 qui se démontre à partir d'un résultat d'unicité (proposition 4.60) sur lequel nous reviendrons au chapitre 5. On note $C_c(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions continues à support compact de \mathbb{R} dans \mathbb{R} (on rappelle qu'une fonction φ de \mathbb{R} dans \mathbb{R} est à support compact s'il existe un compact K de \mathbb{R} t.q. $\varphi = 0$ sur K^c).

Proposition 4.60 Soit m et μ deux mesures sur $\mathcal{B}(\mathbb{R})$, finies sur les compacts de \mathbb{R} . On suppose que :

$$\int \varphi dm = \int \varphi d\mu \text{ pour tout } \varphi \in C_c(\mathbb{R}, \mathbb{R}).$$

Alors, $m = \mu$.

DÉMONSTRATION – Puisque m et μ sont des mesures sur $\mathcal{B}(\mathbb{R})$, finies sur les compacts, on a bien $C_c(\mathbb{R}, \mathbb{R}) \subset L^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ et $C_c(\mathbb{R}, \mathbb{R}) \subset L^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. On pose maintenant $\mathcal{C} = \{]a, b[, a, b \in \mathbb{R}, a < b\}$ et on commence par montrer que $m = \mu$ sur \mathcal{C} .

Soit $a, b \in \mathbb{R}, a < b$. Il existe une suite $(\varphi_n)_{n \in \mathbb{N}} \subset C_c(\mathbb{R}, \mathbb{R})$ t.q. $\varphi_n \uparrow 1_{]a, b[}$. En effet, il suffit de construire φ_n , pour $n \geq 2/(b-a)$, de la manière suivante :

$$\begin{aligned} \varphi_n(x) &= 0 \text{ si } x \leq a, \\ \varphi_n(x) &= n(x-a) \text{ si } a < x < a + \frac{1}{n}, \\ \varphi_n(x) &= 1 \text{ si } a + \frac{1}{n} < x < b - \frac{1}{n}, \\ \varphi_n(x) &= -n(x-b) \text{ si } b - \frac{1}{n} \leq x \leq b \\ \varphi_n(x) &= 0 \text{ si } b \leq x. \end{aligned}$$

Puis, en passant à la limite quand $n \rightarrow +\infty$ dans l'égalité $\int \varphi_n dm = \int \varphi_n d\mu$, on obtient (par convergence monotone ou par convergence dominée) $m(]a, b[) = \mu(]a, b[)$.

On conclut enfin que $m = \mu$ en utilisant, par exemple, la proposition 2.31. ■

Proposition 4.61 Soit (Ω, \mathcal{A}, P) un espace probabilisé, $d > 1$ et X_1, \dots, X_d des v.a.r. Ces v.a.r. sont indépendantes si et seulement si on a, pour toute famille $\{\varphi_1, \dots, \varphi_d\} \subset C_c(\mathbb{R}, \mathbb{R})$,

$$E\left(\prod_{i=1}^d \varphi_i(X_i)\right) = \prod_{i=1}^d E(\varphi_i(X_i)), \quad (4.13)$$

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

DÉMONSTRATION – Le fait que la condition est nécessaire est une conséquence immédiate de la proposition 4.59 car une fonction continue à support compact est borélienne et bornée.

On montre maintenant que la condition est suffisante. On suppose donc que (4.13) est vraie pour toute famille $\{\varphi_1, \dots, \varphi_d\} \subset C_c(\mathbb{R}, \mathbb{R})$ et on veut montrer que les v.a.r. X_1, \dots, X_d sont indépendantes, c'est-à-dire que pour tout $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R})$, on a :

$$\mathbb{E} \left(\prod_{i=1}^d 1_{A_i}(X_i) \right) = \prod_{i=1}^d \mathbb{E}(1_{A_i}(X_i)). \quad (4.14)$$

On rappelle en effet que

$$\mathbb{E}(1_{A_i}(X_i)) = P(X_i^{-1}(A_i)) \text{ et } \mathbb{E} \left(\prod_{i=1}^d 1_{A_i}(X_i) \right) = P \left(\bigcap_{i=1}^d X_i^{-1}(A_i) \right).$$

Pour montrer (4.14), on introduit, pour tout $1 \leq n \leq d+1$, la propriété suivante :

P_n : (4.13) est vraie si $\varphi_i = 1_{A_i}$, avec $A_i \in \mathcal{B}(\mathbb{R})$, pour $i < n$, et $\varphi_i \in C_c(\mathbb{R}, \mathbb{R})$ pour $i \geq n$.

L'hypothèse de la proposition donne que P_1 est vraie. On suppose maintenant que P_n est vraie pour un $n \in \{1, \dots, d\}$. Soit $A_i \in \mathcal{B}(\mathbb{R})$ pour $i < n$ (et $\varphi_i = 1_{A_i}$) et $\varphi_i \in C_c(\mathbb{R}, \mathbb{R})$ pour $i > n$. Pour $A_n \in \mathcal{B}(\mathbb{R})$, on pose, avec $\varphi_n = 1_{A_n}$:

$$\begin{aligned} m(A_n) &= \mathbb{E} \left(\prod_{i=1}^d \varphi_i(X_i) \right), \\ \mu(A_n) &= \prod_{i=1}^d \mathbb{E}(\varphi_i(X_i)). \end{aligned}$$

Les applications m et μ sont des mesures sur $\mathcal{B}(\mathbb{R})$. La propriété P_n montre que $\int \varphi d m = \int \varphi d \mu$ pour tout $\varphi \in C_c(\mathbb{R}, \mathbb{R})$. La proposition 4.60 montre alors que $m = \mu$ ce qui donne la propriété P_{n+1} . Par récurrence sur n , on montre ainsi que P_{d+1} est vraie, ce qui donne (4.14) et l'indépendance de X_1, \dots, X_d . ■