Leçon 263 : Variables aléatoires à densité. Exemples et applications.

On se donne un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. On considèrera des variables aléatoires à valeurs dans $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$. On notera $\langle \cdot \mid \cdot \rangle$ le produit scalaire canonique sur $\mathbf{R}^d$.

1 Définitions et premières propriétés

Définition 1. Soit μ une mesure de probabilité sur $\mathbf{R}^d$. On dit que μ admet un densité (par rapport à la mesure de Lebesgue) s'il existe f mesurable positive telle que pour tout borélien B, $\mu(B) = \int_{\mathbf{R}^d} f(x)\,\mathrm{d}x$.
On dit qu'une v.a. X est à densité si sa loi admet une densité f, c'est-à-dire si pour tout borélien B, $\mathbf{P}(X \in B) = \int_{\mathbf{R}^d} f(x)\,\mathrm{d}x$.

Proposition 2. *La densité caractérise la loi : si μ admet deux densités f et g, alors $f = g$ p.p.*

Théorème 3 (transfert)**.** *Soit X une v.a. dans $\mathbf{R}^d$ admettant une densité f par rapport à la mesure de Lebesgue. Soit $g : X(\Omega) \to \mathbf{R}$ mesurable, alors $g(X)$ est intégrable ssi $\int_{\mathbf{R}^d} |g(x)| f(x)\,\mathrm{d}x$ est fini et alors :*

$$\mathbf{E}[g(X)] = \int_{\mathbf{R}^d} g(x) f(x)\,\mathrm{d}x$$

Proposition 4. *Soient X et Y deux v.a. indépendantes, à densité, et à valeurs dans $\mathbf{R}^d$. Alors $X + Y$ est à densité et $f_{X+Y} = f_X * f_Y$.*

Exemple 5. Si X suit la loi $\Gamma(a, \lambda)$, si Y suit la loi $\Gamma(b, \lambda)$ et si X et Y sont indépendantes, alors $X + Y$ suit la loi $\Gamma(a + b, \lambda)$.

Proposition 6. *Soit (X, Y) un couple de v.a. dans $\mathbf{R}^d \times \mathbf{R}^p$ admettant une densité h. Alors X admet une densité $f(x) = \int_{\mathbf{R}^p} h(x, y)\,\mathrm{d}y$ et Y admet une densité $g(y) = \int_{\mathbf{R}^d} h(x, y)\,\mathrm{d}x$, appelées densités marginales.*

Théorème 7. *Soient X une v.a. à valeurs dans $\mathbf{R}^d$ et Y une v.a. à valeurs dans $\mathbf{R}^p$.*

1. *On suppose que X et Y admettent respectivement des densités f et g. Si X et Y sont indépendantes, alors le couple (X, Y) admet une densité $h(x, y) = f(x)g(y)$.*
2. *Si (X, Y) admet une densité h qui s'écrit sous la forme $h(x, y) = f(x)g(y)$ avec f et g des fonctions positives, alors X et Y sont indépendantes. De plus, la densité de X est proportionnelle à f et celle de Y est proportionnelle à g.*

Exemple 8. Si (X, Y) admet pour densité $h(x, y) = \frac{1}{2\pi} \exp(-\frac{x^2+y^2}{2})$, alors X et Y sont indépendantes de même loi $\mathscr{N}(0, 1)$.

Définition 9. Soit μ une mesure de probabilité sur $\mathbf{R}$, on appelle fonction de répartition de μ la fonction $F_\mu : \mathbf{R} \to [0, 1]$ définie par $F_\mu(t) = \mu(]-\infty, t])$.
Soit X une v.a. réelle, on appelle fonction de répartition de X la fonction de répartition de sa loi, c'est-à-dire $F_X(t) = \mathbf{P}(X \leqslant t)$.

Proposition 10. *qq propriétés des f.r.*

Définition 11. Soit X une v.a. On appelle fonction caractéristique de X la fonction $\varphi_X : \mathbf{R}^d \to \mathbf{C}$ définie par $\varphi(t) = \mathbf{E}[\mathrm{e}^{\mathrm{i}\langle t \mid X \rangle}]$.
Dans le cas où X admet une densité f, la fonction caractéristique est la transformée de Fourier de f.

Proposition 12. *La fonction caractéristique caractérise la loi : si deux v.a. ont la même fonction caractéristique, alors elles ont la même loi.*

Théorème 13. *Si X est une v.a. dont la fonction caractéristique est intégrable, alors X admet une densité.*

2 Lois usuelles et simulation

Définition 14.

1. On appelle loi uniforme sur $[a,b]$ la loi qui a pour densité $f(x) = \frac{1}{b-a}\mathbf{1}_{[a,b]}(x)$.
2. De façon générale, si $B \subseteq \mathbf{R}^d$ est mesurable de mesure non nulle, on appelle loi uniforme sur B la loi qui a pour densité $\frac{1}{\mathrm{mes}(B)}\mathbf{1}_B$.

Application 15. Soit μ une mesure de probabilité sur $\mathbf{R}$, de fonction de répartition F. On définit le pseudo-inverse de F par $F^{(-1)}(u) = \inf\{x \in \mathbf{R} \mid u \leqslant F(x)\}$ pour $u \in]0,1[$. Si U suit la loi uniforme sur $[0,1]$, alors $X = F^{(-1)}(U)$ suit la loi μ.

Définition 16.

1. On appelle loi normale de paramètres $m \in \mathbf{R}$ et $\sigma^2 > 0$, notée $\mathcal{N}(m,\sigma^2)$, la loi sur $\mathbf{R}$ qui a pour densité $f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(x-m)^2}{2\sigma^2})$.
2. Soient $Z_1,\dots,Z_n$ des v.a. indépendantes de même loi $\mathcal{N}(0,1)$. On appelle loi du chi-deux à n degré de liberté, notée $\chi^2(n)$, la loi de $Z_1^2+\cdots+Z_n^2$.

Définition 17.

1. On appelle loi exponentielle de paramètre $\lambda > 0$, notée $\mathcal{E}(\lambda)$, la loi sur $\mathbf{R}$ qui a pour densité $f(x) = \lambda \mathrm{e}^{-\lambda x}\mathbf{1}_{x>0}$.
2. On appelle loi gamma de paramètres $a > 0$ et $\lambda > 0$, notée $\Gamma(a,\lambda)$, la loi sur $\mathbf{R}$ qui a pour densité $f(x) = \frac{\lambda^a}{\Gamma(a)}x^{a-1}\mathrm{e}^{-\lambda x}\mathbf{1}_{x>0}$.

Remarque 18. $\mathcal{E}(\lambda) = \Gamma(1,\lambda)$ et $\chi^2(n) = \Gamma(\frac{n}{2},\frac{1}{2})$.

Proposition 19. *La loi exponentielle vérifie la propriété d'absence de mémoire : pour tous $t,s > 0$, $\mathbf{P}(X > s+t \mid X > s) = \mathbf{P}(X > t)$.*

Application 20 (Box–Muller)**.** Si Θ suit la loi uniforme sur $[0,2\pi]$ et si S suit la loi exponentielle de paramètre $\frac{1}{2}$, alors en notant $R = \sqrt{S}$, les variables $X = R\cos\Theta$ et $Y = R\sin\Theta$ sont indépendantes et de même loi $\mathcal{N}(0,1)$.

3 Vecteurs gaussiens

Définition 21. On dit qu'une v.a. X est un vecteur gaussien si pour tout $a \in \mathbf{R}^d$, $\langle X \mid a\rangle$ suit une loi gaussienne.

Exemple 22. Si $X_1,\dots,X_d$ sont des v.a. i.i.d. de loi $\mathcal{N}(0,1)$ alors $(X_1,\dots,X_d)$ est un vecteur gaussien, de moyenne 0 et de matrice de covariance I_d.

Proposition 23. *L'image d'un vecteur gaussien d'espérance m et de matrice de covariance Γ par une application affine $x \mapsto Ax+b$ et encore un vecteur gaussien, d'espérance $Am+b$ et de matrice de covariance $A\Gamma{}^{\mathrm{t}}A$.*

Proposition 24. *Soit $m \in \mathbf{R}^d$ et Γ une matrice symétrique définie positive. Alors la loi $\mathcal{N}(m,\Gamma)$ admet pour densité la fonction :*

$$f(x) = \frac{1}{(2\pi)^{d/2}}\frac{1}{\sqrt{\det\Gamma}}\exp\left(-\frac{1}{2}\langle\Gamma^{-1}(x-m) \mid x-m\rangle\right)$$

Théorème 25. *Soit X une v.a. L^2 d'espérance m et de matrice de covariance Γ. Alors X est un vecteur gaussien ssi $\varphi_X(t) = \mathrm{e}^{\mathrm{i}\langle m|t\rangle}\mathrm{e}^{-\frac{1}{2}\langle\Gamma t|t\rangle}$.*

Corollaire 26. *Si deux vecteurs gaussiens ont même espérance et même matrice de covariance, alors ils ont la même loi. On note $\mathcal{N}(m,\Gamma)$ la loi d'un vecteur gaussien d'espérance m et de matrice de covariance Γ.*

Théorème 27. *Soient $d_1,\dots,d_k$ des entiers strictement positifs de somme d. Soit $X = (X_1,\dots,X_d)$ un vecteur gaussien dont la matrice de covariance Γ est diagonale par blocs :*

$$\Gamma = \begin{pmatrix} \Gamma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \Gamma_k \end{pmatrix}$$

avec Γ_i bloc de taille d_i. Posons $Y_1 = (X_1,\dots,X_{d_1})$, $Y_2 = (X_{d_1+1},\dots,X_{d_1+d_2})$, etc. Alors les vecteurs $Y_1,\dots,Y_k$ sont gaussiens et indépendants.

Corollaire 28. *Des variables gaussiennes sont indépendantes ssi elles sont deux à deux non corrélées.*

Théorème 29 (Cochran)**.** *Soit X un vecteur gaussien d'espérance nulle et de matrice de covariance I_d. On considère $E_1 \oplus \cdots \oplus E_r$ une décomposition de $\mathbf{R}^d$ en sous-espaces orthogonaux de dimensions respectives $d_1, \ldots, d_r$. Alors les projections orthogonales $\Pi_{E_1}(X), \ldots, \Pi_{E_r}(X)$ sont des vecteurs gaussiens indépendants. De plus, pour tout $j \in \{1, \ldots, r\}$,*

$$\|\Pi_{E_j}(X)\|_2^2 \sim \chi^2(d_j)$$

4 Convergence en loi

Définition 30. On dit qu'une suite de v.a. (X_n) converge vers X en loi si pour toute fonction f continue bornée, $\mathbf{E}[f(X_n)] \to \mathbf{E}[f(X)]$.

Lemme 31 (Scheffé)**.** *Soit $(E, \mathscr{E}, \mu)$ un espace mesuré et $f, (f_n)$ des fonctions mesurables positives intégrables sur E. On suppose que :*

1. *$f_n \to f$ μ–p.p.*
2. *$\int_E f_n \, \mathrm{d}\mu \to \int_E f \, \mathrm{d}\mu$*

Alors $f_n \to f$ dans L^1.

Corollaire 32. *Soit (X_n) une suite de v.a. admettant des densités (f_n) par rapport à la mesure de Lebesgue sur $\mathbf{R}^d$. Soit X admettant une densité f telle que $f_n \to f$ p.p. Alors (X_n) converge en loi vers X.*

Remarque 33. La réciproque est fausse : sur $\mathbf{R}$, on considère les v.a. X_n de densités $f_n(x) = \big(1 + \cos(2\pi n x)\big)\mathbf{1}_{[0,1]}(x)$. Alors (X_n) converge en loi vers une v.a. de loi $\mathscr{U}([0,1])$ mais les densité (f_n) ne convergent pas.

Théorème 34 (théorème central limite)**.** *Soit $(X_n)_{n \geqslant 1}$ une suite de v.a.r. i.i.d. de carré intégrable. Notons m leur espérance et σ^2 leur variance. Notons $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Alors :*

$$\sqrt{n}\big(\overline{X}_n - m\big) \xrightarrow[n\to\infty]{\mathscr{L}} \mathcal{N}(0, \sigma^2)$$

Lemme 35. *Soit (X_n) une suite de vecteurs aléatoires et X un vecteur aléatoire. Si pour tout $a \in \mathbf{R}^d$, $\langle X_n \mid a\rangle$ converge en loi vers $\langle X \mid a\rangle$, alors (X_n) converge en loi vers X.*

Théorème 36 (TCL multidimensionnel)**.** *Soit (X_n) une suite de vecteurs aléatoires i.i.d. L^2. On note m leur espérance et Γ leur matrice de covariance. Alors :*

$$\sqrt{n}\big(\overline{X}_n - m\big) \xrightarrow{\mathscr{L}} \mathcal{N}(0, \Gamma)$$

Application 37 (théorème de Fisher)**.** Soit ν une loi sur $\{1, \ldots, q\}$ et soit (X_n) une suite i.i.d. de loi ν. Notons :

$$N_{n,j} = \sum_{i=1}^n \mathbf{1}_{\{X_i = j\}} \qquad T_n = \sum_{j=1}^q \frac{\big(N_{n,j} - n\nu(j)\big)^2}{n\nu(j)}$$

Alors :

$$T_n \xrightarrow[n\to+\infty]{\mathscr{L}} \chi^2(q-1)$$

Application 38 (intervalle de confiance asymptotique)**.** Soit $(X_n)_{n\geqslant 1}$ une suite de v.a.r. i.i.d. de carré intégrable. Notons m leur espérance et σ^2 leur variance. Soient $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ et $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2$. On se donne $\alpha \in \,]0,1[$. Soit $\hat{I}_n = \big[\overline{X}_n - q_\alpha \frac{\hat{\sigma}_n}{\sqrt{n}}; \overline{X}_n + q_\alpha \frac{\hat{\sigma}_n}{\sqrt{n}}\big]$ où q_α vérifie $\mathbf{P}(Z \leqslant q_\alpha) = 1 - \frac{\alpha}{2}$ pour Z de loi $\mathcal{N}(0,1)$. Alors :

$$\lim_{n\to\infty} \mathbf{P}(m \in \hat{I}_n) = 1 - \alpha$$

Application 39 (formule de Stirling)**.** On a l'équivalent $n! \sim \sqrt{2\pi}\, n^{n+\frac{1}{2}} \mathrm{e}^{-n}$.

Développements

1. Théorèmes de Cochran et de Fisher. **[29][37]**
2. Formule de Stirling par le TCL. **[39]**

Références

— BARBE et LEDOUX, *Probabilités.*
— BERCU et CHAFAÏ, *Modélisation stochastique et simulation.*
— CARRIEU, *Probabilité – Exercices corrigés.*
— GARET et KURTZMAN, *De l'intégration aux probabilités.*
— OUVRARD, *Probabilités 2.*
— RIVOIRARD et STOLTZ, *Statistique mathématique en action.*