

263. Variables aléatoires à densité. Exemples et applications.

Introduction : Les lois à densité permettent de modéliser certains phénomènes continus.

On se place sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, et on munit \mathbb{R} de $(\mathcal{B}(\mathbb{R}), \lambda)$.

1 Généralités

1.1 Notion de densité

Définition 1. Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est appelée densité si elle est positive et si $\int_{\mathbb{R}} f(t)dt = 1$.

Exemple 2. $f : x \mapsto \frac{1}{x^2} \mathbb{1}_{]1, +\infty[}(x)$ est une densité.

Proposition 3. La fonction $F : x \mapsto \int_{-\infty}^x f(t)dt$ définit alors une fonction de répartition.

Définition 4. On dit qu'une variable aléatoire réelle X admet la densité f si pour tout $x \in \mathbb{R}$, $\mathbb{P}(X \leq x) = \int_{-\infty}^x f(t)dt$.

Remarque 5. La densité d'une variable aléatoire n'est définie que presque partout.

Proposition 6. Si X admet la densité f , alors sa fonction de répartition F est dérivable partout où f est continue et $F'(x) = f(x)$. Réciproquement, si F est continue et \mathcal{C}^1 par morceaux, alors X admet une densité donnée par $f = F'$.

1.2 Espérance, variance et moments

Théorème 7 (de transfert). Soient X variable aléatoire réelle de densité f , et $g : \mathbb{R} \rightarrow \mathbb{R}$ mesurable. Alors la variable aléatoire $g(X)$ est intégrable ssi $\int_{\mathbb{R}} |g(t)|f(t)dt < +\infty$. Si cette intégrale converge, on a alors $\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(t)f(t)dt$.

Corollaire 8. En particulier, X a un moment d'ordre p ssi $\int_{\mathbb{R}} |t|^p f(t)dt < +\infty$, et alors $\mathbb{E}[X^p] = \int_{\mathbb{R}} x^p f(x)dx$.

Contre-exemple 9. Si X suit une loi de Cauchy, dont la densité est donnée par $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$, alors X n'admet pas de moment d'ordre 1.

Contre-exemple 10. Même lorsqu'elle est définie, la suite des moments ne caractérise pas la densité d'une variable aléatoire. Par exemple, on pose pour $x \in \mathbb{R}$, $f(x) = \frac{1}{2\pi x} e^{-\ln(x)^2/2} \mathbb{1}_{x>0}$, et pour $a \in [-1, 1]$, $f_a(x) = f(x)(1 + a \sin(2\pi \ln(x)))$. Alors f, f_a sont des densités, et si X, X_a ont pour densité respectivement f, f_a , elles ont les mêmes moments.

Proposition 11. Si X variable aléatoire de densité f , sa fonction caractéristique est $\phi_X(t) = \mathbb{E}[e^{itX}] = \int_{\mathbb{R}} f(x)e^{itx}dx$. C'est la transformée de Fourier de f .

Proposition 12. La fonction caractéristique caractérise la loi de X . De plus, si X admet un moment d'ordre p , alors ϕ_X est p fois dérivable et $\phi_X^{(p)}(t) = \int_{\mathbb{R}} x^p f(x)e^{itx}dx$.

1.3 Exemples de lois usuelles

Loi uniforme

Définition 13. Pour $a < b \in \mathbb{R}$, on dit que X suit une loi uniforme sur $[a, b]$, et on note $X \sim \mathcal{U}([a, b])$, si elle admet pour densité $\frac{1}{b-a} \mathbb{1}_{[a, b]}$.

Proposition 14. On a alors $\mathbb{E}[X] = \frac{a+b}{2}$, $\text{Var}(X) = \frac{(b-a)^2}{12}$ et $\phi_X(t) = \frac{e^{itb} - e^{ita}}{it(b-a)}$.

Loi normale

Définition 15. Pour $m \in \mathbb{R}, \sigma > 0$, on dit que X suit une loi normale $\mathcal{N}(m, \sigma^2)$ si elle admet pour densité $f : x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-m)^2/2\sigma^2}$.

Proposition 16. On a alors $\mathbb{E}[X] = m$, $\text{Var}(X) = \sigma^2$ et $\phi_X(t) = e^{itm - t^2\sigma^2/2}$.

Loi Gamma

Définition 17. On dit que X suit une loi Gamma de paramètres $a > 0, \lambda > 0$ si elle admet pour densité $f : x \mapsto \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} \mathbb{1}_{]0, +\infty[}(x)$. Si $a = 1$, on dit qu'elle suit une loi exponentielle de paramètre λ , notée $\mathcal{E}(\lambda)$.

Proposition 18. On a alors $\mathbb{E}[X] = \frac{a}{\lambda}$, $\text{Var}(X) = \frac{a}{\lambda^2}$.

Proposition 19. La transformée de Laplace est alors définie sur $] -\infty, \lambda[$ et donnée par $L_X(t) = \mathbb{E}[e^{tX}] = \left(\frac{\lambda}{\lambda - t}\right)^a$. La fonction caractéristique est donnée par $\forall t \in \mathbb{R}, \phi_X(t) = \left(\frac{\lambda}{\lambda - it}\right)^a$.

2 Utilisation de la notion de densité

2.1 Loïs marginales

Définition 20. Soit $X = (X_1, \dots, X_d)$ un vecteur aléatoire à valeurs dans \mathbb{R}^d . On dit qu'il admet la densité $f : \mathbb{R}^d \rightarrow \mathbb{R}_+$ si $\int_{\mathbb{R}^d} f d\lambda = 1$ et si $\forall x_1, \dots, x_d \in \mathbb{R}$,

$$\mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_d} f(y_1, \dots, y_d) dy_1 \dots dy_d.$$

Proposition 21. X_1 a alors $f_1 : x \mapsto \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x, y_2, \dots, y_d) dy_2 \dots dy_d$ pour densité.

Exemple 22. Si (X, Y) de loi uniforme sur le disque unité de \mathbb{R}^2 , il admet pour densité $f : (x, y) \mapsto \frac{1}{\pi} \mathbf{1}_{x^2 + y^2 \leq 1}$, et alors X admet pour densité $f_X : x \mapsto \frac{2}{\pi} \sqrt{1 - x^2} \mathbf{1}_{[-1, 1]}(x)$.

2.2 Indépendance et conditionnement

Théorème 23. Soient X, Y deux variables aléatoires réelles de densités respectives f_X, f_Y . Elles sont indépendantes si et seulement si (X, Y) admet pour densité $f : (x, y) \mapsto f_X(x) f_Y(y)$.

Proposition 24. Si X, Y variables aléatoires réelles indépendantes, de densités respectives f_X, f_Y , alors $X + Y$ admet pour densité $f_X * f_Y$.

Exemple 25. Si X, Y indépendantes et suivent une loi χ_1^2 , de densité $f : x \mapsto \frac{e^{-x/2}}{\sqrt{2\pi x}} \mathbf{1}_{]0, +\infty[}(x)$, alors $X + Y$ suit une loi $\mathcal{E}(1/2)$.

Proposition 26. Si X, Y indépendantes, alors $\phi_{X+Y} = \phi_X \phi_Y$.

Exemple 27. Si $X \sim \mathcal{N}(m, \sigma^2), Y \sim \mathcal{N}(m', \sigma'^2)$, alors $X + Y \sim \mathcal{N}(m + m', \sigma^2 + \sigma'^2)$.

Proposition 28. Soit (X, Y) vecteur aléatoire de densité f . Alors pour tout $x \in \mathbb{R}$ tel que $f_X(x) \neq 0$, $f_{Y|X=x} : y \mapsto \frac{f(x, y)}{f_X(x)}$ définit une densité.

Définition 29. On l'appelle densité conditionnelle de Y sachant $X = x$. On définit alors l'espérance conditionnelle de Y sachant $X = x$ par

$$\mathbb{E}[Y|X = x] = \int_{\mathbb{R}} y f_{Y|X=x}(y) dy.$$

Proposition 30. On a alors $\mathbb{E}[Y] = \int_{\mathbb{R}} \mathbb{E}[Y|X = x] f_X(x) dx$.

Exemple 31. Si $T = \{(x, y) \in \mathbb{R}^2, 0 < y < x < 1\}$ et (X, Y) vecteur aléatoire de densité $f : x \mapsto \frac{1}{x} \mathbf{1}_T(x, y)$, on a $f_X(x) = \mathbf{1}_{]0, 1[}(x)$ et pour $x \in]0, 1[$, $\mathbb{E}[Y|X = x] = \frac{x}{2}$.

2.3 Simulation de lois à densité

Définition 32. Soit F une fonction de répartition. On définit son inverse généralisée par $\forall u \in]0, 1[$, $F^{\leftarrow}(u) = \inf\{x, F(x) > u\}$.

Proposition 33. Si $U \sim \mathcal{U}([0, 1])$, alors $F^{\leftarrow}(U)$ admet F pour fonction de répartition.

Exemple 34. Si $F(x) = 1 - e^{-\lambda x}$ avec $\lambda > 0$ (fonction de répartition d'une loi $\mathcal{E}(\lambda)$), on a $F^{\leftarrow}(u) = -\frac{\ln(1-u)}{\lambda}$. Or si $U \sim \mathcal{U}([0, 1])$, on a aussi $1 - U \sim \mathcal{U}([0, 1])$. Ainsi, $Y := -\frac{1}{\lambda} \ln(U) \sim \mathcal{E}(\lambda)$.

Proposition 35. Soient S, Θ indépendantes de lois respectives $\mathcal{E}(1/2), \mathcal{U}([0, 2\pi])$, on pose $X = \sqrt{S} \cos(\Theta), Y = \sqrt{S} \sin(\Theta)$. Alors X, Y sont indépendantes de loi $\mathcal{N}(0, 1)$.

Remarque 36. D'après ce qui précède, on sait simuler simplement une loi $\mathcal{N}(0, 1)$ à partir d'une loi $\mathcal{U}([0, 1])$.

Proposition 37 (Méthode de rejet). Soient deux densités f, g telles qu'il existe $a \geq 1$ tel que $\forall x \in \mathbb{R}, f(x) \leq ag(x)$. Soient $(U_n)_{n \in \mathbb{N}}, (Z_n)_{n \in \mathbb{N}}$ suites de variables iid, respectivement de loi $\mathcal{U}([0, 1])$ et de densité g . Soient $N = \inf\{n \in \mathbb{N}^*, f(Z_n) > aU_n g(Z_n)\}$ et $Y = Z_N$. Alors N, Y sont indépendantes, N suit une loi géométrique de paramètre $\frac{1}{a}$ et Y admet f pour densité.

Exemple 38. Si $D = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$, on peut simuler une loi uniforme sur D en tirant X, Y indépendantes de loi $\mathcal{U}([0, 1])$ jusqu'à ce que $(X, Y) \in D$.

3 Applications

3.1 Théorème central limite

Théorème 39 (Lévy). Si $(X_n)_{n \in \mathbb{N}}, X$ sont des variables aléatoires de fonctions caractéristiques $(\phi_{X_n})_{n \in \mathbb{N}}$, et si $(\phi_{X_n})_{n \in \mathbb{N}}$ converge simplement vers ϕ_X , alors $(X_n)_{n \in \mathbb{N}}$ converge en loi vers X .

Théorème 40 (central limite). Soit $(X_n)_{n \in \mathbb{N}^*}$ suite de variables aléatoires iid dans L^2 . Alors $\sqrt{\frac{n}{\text{Var}(X_1)}} \left(\frac{S_n}{n} - \mathbb{E}[X_1] \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$

Application 41. Soit $(X_n)_{n \geq 1}$ suite de variables aléatoires iid de loi $\mathcal{B}(\theta)$ où $\theta \in \mathbb{R}$ est inconnu. On définit, pour $n \geq 1$, la moyenne empirique $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$. Alors

$\lim_{n \rightarrow \infty} \mathbb{P} \left(\theta \in \left[\bar{X}_n - \frac{q_\alpha}{\sqrt{2n}}, \bar{X}_n + \frac{q_\alpha}{\sqrt{2n}} \right] \right) \geq 1 - \alpha$, où q_α est le quantile d'ordre α de la loi $\mathcal{N}(0, 1)$.

3.2 Méthode de Monte-Carlo

Proposition 42 (Monte-Carlo). Soit $(X_n)_{n \in \mathbb{N}}$ suite de variables aléatoires iid de loi uniforme sur $[-1, 1]^d$, soit f mesurable bornée sur $[-1, 1]^d$.

Alors $\hat{I}_n := \frac{1}{n} \sum_{i=1}^n f(X_i) \xrightarrow[n \rightarrow \infty]{} I := \int_{[-1, 1]^d} f(x) dx$ en probabilité.

Remarque 43. Pour $d = 1$, on a $\mathbb{P} \left(I \in \left[\hat{I}_n - \sqrt{\frac{6}{n}} q_\alpha, \hat{I}_n + \sqrt{\frac{6}{n}} q_\alpha \right] \right) \xrightarrow[n \rightarrow +\infty]{} 1 - \alpha$.

Développements

- Étude de la loi Gamma.
- Théorème central limite et application aux intervalles de confiance.

Références

- [1] P. Barbe, M. Ledoux, PROBABILITÉ, EDP Sciences.
- [2] S. Méléard, ALÉATOIRE, École Polytechnique.
- [3] J.-Y. Ouvrard, PROBABILITÉS TOME 2, Cassini.