Leçon 260 : Espérance, variance et moments d'une variable aléatoire.

On se donne un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. On appelle variable aléatoire tout application mesurable définie sur $(\Omega, \mathscr{A}, \mathbf{P})$. Dans cette leçon on s'intéressera aux variables aléatoires réelles, c'est-à-dire à valeurs dans $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$, et aux vecteurs aléatoires, à valeurs dans $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$.

1 Espérance d'une v.a.

Définition 1.

1. Soit X une v.a.r. intégrable. On appelle espérance de X le nombre réel $\mathbf{E}(X) = \int_\Omega X \,\mathrm{d}\mathbf{P}$. On dit que X est centrée si $\mathbf{E}(X)$ est nulle.
2. Soit X une v.a.r. dans L^p, on appelle moment d'ordre p le nombre $\mathbf{E}(X^p)$. On appelle moment centré d'ordre p le nombre $\mathbf{E}[(X - \mathbf{E}X)^p]$.
3. Soit X une v.a. dans $\mathbf{R}^d$. On dit que X est de puissance p–ième intégrable si chacune de ses composantes est dans L^p. C'est équivalent à dire que $\mathbf{E}(\|X\|^p) < \infty$ où $\|\cdot\|$ est une norme quelconque sur $\mathbf{R}^d$. Si X est intégrable, son espérance est le vecteur $\mathbf{E}(X) = (\mathbf{E}(X_1), \dots, \mathbf{E}(X_d))$.

Exemple 2. Soit $A \in \mathscr{A}$, alors $\mathbf{1}_A$ est intégrable et $\mathbf{E}(\mathbf{1}_A) = \mathbf{P}(A)$.

Proposition 3 (Hölder)**.** *Soient $p, q \in [1, \infty]$ des exposants conjugués, et soient $X \in \mathrm{L}^p$ et $Y \in \mathrm{L}^q$. Alors XY est intégrable et $\mathbf{E}(|XY|) \leqslant \mathbf{E}(|X|^p)^{\frac{1}{p}} \mathbf{E}(|Y|^q)^{\frac{1}{q}}$.*

Corollaire 4. *Si $1 \leqslant p \leqslant q \leqslant +\infty$, alors $\mathrm{L}^\infty \subseteq \mathrm{L}^q \subseteq \mathrm{L}^p \subseteq \mathrm{L}^1$.*

Théorème 5 (transfert)**.** *Soit X une v.a. dans $\mathbf{R}^d$ de loi μ_X et soit $g : \mathbf{R}^d \to \mathbf{R}$ mesurable.*

1. *Si g est positive, alors $\mathbf{E}[g(X)] = \int_{\mathbf{R}^d} g \,\mathrm{d}\mu_X$.*
2. *La v.a. $g(X)$ est dans $\mathrm{L}^1(\Omega, \mathscr{A}, \mathbf{P})$ ssi g est dans $\mathrm{L}^1(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d), \mu_X)$ et dans ce cas, l'égalité précédente est vérifiée.*

Remarque 6. En particulier, l'espérance ne dépend que de la loi de X.

Corollaire 7.

1. *Soit X une v.a. discrète et $D \subseteq \mathbf{R}$ fini ou dénombrable tel que $X(\Omega) = D$. Soit $g : D \to \mathbf{R}$ mesurable, alors $g(X)$ est intégrable ssi $\sum_{x \in D} |g(x)| \mathbf{P}(X = x)$ est fini et alors :*
$$\mathbf{E}[g(X)] = \sum_{x \in D} g(x) \mathbf{P}(X = x)$$
2. *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 :*
$$\mathbf{E}[g(X)] = \int_{\mathbf{R}^d} g(x) f(x) \,\mathrm{d}x$$

Exemple 8.

1. Si $X \sim \mathscr{B}(n, p)$, alors $\mathbf{E}(X) = np$.
2. Si $X \sim \mathscr{P}(\lambda)$, alors $\mathbf{E}(X) = \lambda$.
3. Si $X \sim \mathscr{G}_{\mathbf{N}^*}(p)$, alors $\mathbf{E}(X) = \frac{1}{p}$.
4. Si $X \sim \mathscr{E}(\lambda)$, alors $\mathbf{E}(X) = \frac{1}{\lambda}$.
5. Si $X \sim \mathscr{N}(m, \sigma^2)$, alors $\mathbf{E}(X) = m$.
6. Si X suit une loi de Cauchy, X n'admet pas d'espérance.

Proposition 9 (Markov)**.** *Soit X une v.a.r. positive. Alors pour tout $t > 0$, $\mathbf{P}(X \geqslant t) \leqslant \frac{\mathbf{E}(X)}{t}$.*

Corollaire 10 (Bienaymé–Tchebytchev)**.** *Soit X une v.a.r. de carré intégrable, alors pour tout $t > 0$, $\mathbf{P}(|X - \mathbf{E}(X)| \geqslant t) \leqslant \frac{\mathrm{Var}(X)}{t^2}$.*

Application 11 (polynômes de Bernstein)**.** Soit $f \in \mathrm{C}([0,1])$, on lui associe le n–ième polynôme de Bernstein :

$$B_n[f](x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}$$

Alors $B_n[f]$ converge uniformément vers f. L'ensemble des fonctions polynomiales est donc dense dans $\mathrm{C}([0,1])$.

Proposition 12 (Jensen)**.** *Soit X une v.a.r. intégrable et soit $\varphi : \mathbf{R} \to \mathbf{R}$ convexe avec $\varphi(X)$ intégrable. Alors $\varphi\big(\mathbf{E}(X)\big) \leqslant \mathbf{E}\big(\varphi(X)\big)$.*

Proposition 13. *Deux variables aléatoires X et Y sont indépendantes ssi pour toutes fonctions mesurables bornées f et g, $\mathbf{E}[f(X)g(Y)] = \mathbf{E}[f(X)]\mathbf{E}[g(Y)]$.*

2 Variance et covariance

Définition 14.

1. Soit X une v.a.r. dans L^2. On appelle variance de X le moment centré d'ordre 2 : $\mathrm{Var}(X) = \mathbf{E}[(X - \mathbf{E}X)^2]$. On appelle écart-type de X la racine carré de la variance de X, on le note σ_X.
2. Soient X et Y des v.a.r. dans L^2, on appelle covariance de X et Y la quantité $\mathrm{Cov}(X,Y) = \mathbf{E}\big[(X-\mathbf{E}X)(Y-\mathbf{E}(Y)\big]$. Si $\mathrm{Cov}(X,Y) = 0$, on dit que X et Y sont non corrélées.
3. Soit X une v.a. dans $\mathbf{R}^d$ de carré intégrable. On appelle matrice de covariance de X la matrice $\Gamma_X = \big(\mathrm{Cov}(X_i, X_j)\big)$.

Proposition 15.

1. *L'application* Cov *est une forme bilinéaire symétrique positive sur* L^2, *sa forme quadratique associée est* Var.
2. $\mathrm{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$.
3. $\forall a, b \in \mathbf{R}$, $\mathrm{Cov}(X-a, Y-b) = \mathrm{Cov}(X,Y)$.
4. $\mathrm{Var}(X) = 0$ *ssi X est constante p.s.*
5. $|\mathrm{Cov}(X,Y)| \leqslant \sigma_X \sigma_Y$.
6. *Si X et Y sont indépendantes,* $\mathrm{Cov}(X,Y) = 0$ *et* $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Exemple 16.

1. Si $X \sim \mathscr{B}(n,p)$, alors $\mathrm{Var}(X) = np(1-p)$.
2. Si $X \sim \mathscr{P}(\lambda)$, alors $\mathrm{Var}(X) = \lambda$.
3. Si $X \sim \mathscr{G}_{\mathbf{N}^*}(p)$, alors $\mathrm{Var}(X) = \frac{1-p}{p^2}$.
4. Si $X \sim \mathscr{E}(\lambda)$, alors $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.
5. Si $X \sim \mathscr{N}(m, \sigma^2)$, alors $\mathrm{Var}(X) = \sigma^2$.
6. Si $X \sim \mathscr{N}(0,1)$ et $\varepsilon \sim \frac{\delta_1 + \delta_{-1}}{2}$ sont indépendantes, alors X et εX sont non corrélées mais ne sont pas indépendantes.

Exemple 17 (loi multinomiale)**.** On considère une urne contenant k sortes de boules. On va précéder au tirage de n boules avec remise (les tirages sont indépendants). On note p_i la probabilité de tirer une boule de type i et on note Z_i le nombre de boules de type i que l'on a tirées. La loi de $Z = (Z_1, \ldots, Z_k)$ est appelée la loi multinomiale de paramètres n et $p = (p_1, \ldots, p_k)$. Alors Z a pour espérance le vecteur np et pour matrice de covariance Γ.

$$\Gamma = \begin{bmatrix} np_1(1-p_1) & -np_1p_2 & \cdots & -p_1p_n \\ -np_1p_2 & np_2(1-p_2) & & -p_2p_k \\ \cdots & & \ddots & \vdots \\ -p_1p_k & -p_2p_k & \cdots & np_k(1-p_k) \end{bmatrix}$$

3 Fonction caractéristique et moments

Définition 18. Soit X une v.a.r. On définit la fonction caractéristique de X comme $\varphi_X(t) = \mathbf{E}(\mathrm{e}^{\mathrm{i}tX})$.

Exemple 19. Si $X \sim \mathcal{N}(m,\sigma^2)$ alors $\varphi_X(t) = \mathrm{e}^{\mathrm{i}mt - \frac{t^2\sigma^2}{2}}$.

Théorème 20. *Soit X une v.a.r. et soit φ sa fonction caractéristique.*

1. *Si X admet un moment d'ordre $n \in \mathbf{N}^*$, alors φ est de classe C^n et pour tout $1 \leqslant k \leqslant n$, $\varphi^{(k)}(t) = \mathrm{i}^k\mathbf{E}(X^k\mathrm{e}^{\mathrm{i}tX})$. En particulier, $\varphi^{(k)}(0) = \mathrm{i}^k\mathbf{E}(X^k)$.*
2. *Si φ est k fois dérivable en 0 (avec $k \geqslant 2$), alors X admet des moments jusqu'à l'ordre $2\lfloor\frac{k}{2}\rfloor$.*

Application 21. Soit X de loi $\mathcal{N}(0,1)$. Alors X admet des moments de tout ordre et ceux-ci sont donnés par :

$$\mathbf{E}(X^{2k+1}) = 0, \quad \mathbf{E}(X^{2k}) = 1 \times 3 \times 5 \times \cdots \times (2k-1)$$

Remarque 22. Les moments ne suffisent pas à caractériser une loi en général : par exemple, si Z suit la loi log-normale (c'est-à-dire $\log(Z)$ suit une loi $\mathcal{N}(0,1)$), notons Z_a la v.a. admettant pour densité $f_a(x) = f(x)(1 + a\sin(2\pi\log x))$ sur $\mathbf{R}_+^*$ où $a \in [-1,1]$ et f est la densité de Z. Alors Z et Z_a ont les mêmes moments.

Théorème 23 (des moments)**.** *Si X et Y sont deux v.a.r. à valeurs dans un intervalle borné $[a,b]$. Si $\mathbf{E}(X^k) = \mathbf{E}(Y^k)$ pour tout $k \in \mathbf{N}^*$, alors X et Y ont même loi.*

4 Théorèmes limites

Théorème 24 (loi forte des grands nombres L^2)**.** *Soit $(X_n)_{n\geqslant 1}$ une suite de v.a.r. de carré intégrable, deux à deux non corrélées, telles que* $\sup_n \mathrm{Var}(X_n) < \infty$. *Posons $S_n = X_1 + \cdots + X_n$, alors :*

$$\frac{S_n - \mathbf{E}[S_n]}{n} \xrightarrow[n\to\infty]{\text{p.s.}} 0$$

Corollaire 25. *Soit $(X_n)_{n\geqslant 1}$ une suite de v.a.r. i.i.d. de carré intégrable. Notons $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ et $m = \mathbf{E}(X_1)$. Alors $\overline{X}_n$ converge vers m p.s.*

Remarque 26. Ce dernier résultat est vrai sous l'hypothèse que les v.a. sont simplement L^1. *(admis)*

Théorème 27 (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 :*

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

Application 28 (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 $\alpha \in]0,1[$. Alors :

1. $\overline{X}_n$ et $\hat{\sigma}_n^2$ convergent p.s. vers m et σ^2. On dit que ce sont des estimateurs consistants de m et σ^2.
2. 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 :

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

Développements

1. Fonction caractéristique et moments. **[20]**
2. Loi des grands nombres L^2. **[24]**

Références

- BARBE et LEDOUX, *Probabilités.*
- CARRIEU, *Probabilité – Exercices corrigés.*
- GARET et KURTZMAN, *De l'intégration aux probabilités.*
- OUVRARD, *Probabilités 2.*