

Indépendance

Cours de Probabilités — Licence 3

Plan du chapitre

Indépendance d'événements

Indépendance de variables aléatoires

Indépendance de v.a. continues

Indépendance de v.a. discrètes

Somme de variables aléatoires indépendantes

Minimum ou maximum de v.a. indépendantes

Synthèse

Indépendance d'événements

Deux événements indépendants

Définition

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité et $A, B \in \mathcal{A}$. On dit que A et B sont **indépendants** si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Interprétation. Si $\mathbb{P}(B) > 0$, cela équivaut à $\mathbb{P}(A | B) = \mathbb{P}(A)$: savoir que B est réalisé ne modifie pas la probabilité de A .

Remarque

Si A et B sont indépendants, alors A et B^c le sont aussi :

$$\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A) \cdot \mathbb{P}(B^c).$$

Définition

Soit $A_1, \dots, A_n \in \mathcal{A}$. On dit que $A_1, \dots, A_n$ sont **indépendants** si, pour tout $p \in \llbracket 1, n \rrbracket$ et tous $1 \leq k_1 < \dots < k_p \leq n$,

$$\mathbb{P}(A_{k_1} \cap \dots \cap A_{k_p}) = \mathbb{P}(A_{k_1}) \cdots \mathbb{P}(A_{k_p}).$$

Attention

L'indépendance **deux à deux** ne suffit **pas** à obtenir l'indépendance de tous les événements ensemble !

Exemple : deux dés (mise en place)

Exemple

On lance deux dés équilibrés : $\Omega = \llbracket 1, 6 \rrbracket^2$, $\mathbb{P}$ uniforme. On considère

$A =$ “les deux dés donnent le même résultat”

$B =$ “le 1er dé donne un résultat pair”

$C =$ “le 2ème dé donne un 1”

$D =$ “le 2ème dé donne un 1 ou un 2”

$$\mathbb{P}(A) = \frac{1}{6}, \quad \mathbb{P}(B) = \frac{1}{2}, \quad \mathbb{P}(C) = \frac{1}{6}, \quad \mathbb{P}(D) = \frac{1}{3}.$$

Exemple : deux dés (indépendances deux à deux)

On vérifie

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{(2, 2), (4, 4), (6, 6)\}) = \frac{3}{36} = \frac{1}{12} = \mathbb{P}(A) \cdot \mathbb{P}(B),$$

donc A, B indépendants ; de même A, C ; A, D ; B, C ; B, D sont indépendants.

Mais...

$$\mathbb{P}(C \cap D) = \mathbb{P}(C) = \frac{1}{6} \neq \frac{1}{18} = \mathbb{P}(C) \cdot \mathbb{P}(D) :$$

C et D ne sont pas indépendants !

Exemple : deux dés (indépendance globale)

Même si A, B, C sont deux à deux indépendants, ils ne sont **pas** indépendants ensemble :

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\emptyset) = 0 \neq \frac{1}{72} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

En revanche

A, B, D **sont** indépendants : on vérifie

$$\mathbb{P}(A \cap B \cap D) = \mathbb{P}(\{(2, 2)\}) = \frac{1}{36} = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(D),$$

et les conditions pour $p = 1, 2$ ont déjà été vérifiées.

Indépendance de v.a.

Définition

Pour $1 \leq i \leq n$, soit X_i une v.a. à valeurs dans $(E_i, \mathcal{E}_i)$. On dit que $X_1, \dots, X_n$ sont **indépendantes** si, pour tous $B_i \in \mathcal{E}_i$,

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \cdots \mathbb{P}(X_n \in B_n).$$

Interprétation. Savoir quelque chose sur X_1 n'apporte aucune information sur X_2 (et vice-versa).

Là encore, l'indépendance deux à deux ne suffit pas.

Remarque

Pour $A_1, \dots, A_n \in \mathcal{A}$: les événements $A_1, \dots, A_n$ sont indépendants $\iff$ les v.a. $1_{A_1}, \dots, 1_{A_n}$ sont indépendantes.

Conséquences de l'indépendance

Si $X_1, \dots, X_n$ sont indépendantes :

- pour tout $\mathcal{I} \subset \llbracket 1, n \rrbracket$, les $(X_i)_{i \in \mathcal{I}}$ sont indépendantes (prendre $B_j = \Omega$ pour $j \notin \mathcal{I}$) ;
- pour tous $B_i \in \mathcal{E}_i$, les événements $\{X_1 \in B_1\}, \dots, \{X_n \in B_n\}$ sont indépendants.

Proposition

Pour $1 \leq i \leq n$, soit X_i une v.a. à valeurs dans $(E_i, \mathcal{E}_i)$ et $f_i: E_i \rightarrow F_i$ mesurable. Si $X_1, \dots, X_n$ sont indépendantes, alors $f_1(X_1), \dots, f_n(X_n)$ le sont aussi.

Preuve. Pour tous $B_i \in \mathcal{F}_i$:

$$\mathbb{P}(f_1(X_1) \in B_1, \dots) = \mathbb{P}(X_1 \in f_1^{-1}(B_1), \dots) = \prod_i \mathbb{P}(X_i \in f_i^{-1}(B_i)) = \prod_i \mathbb{P}(f_i(X_i) \in B_i). \quad \square$$

Exemple

Si X, Y sont réelles indépendantes, alors X^2 et $2Y + 1$ sont indépendantes.

Proposition

Pour $1 \leq i \leq n$, soit X_i à valeurs dans $(E_i, \mathcal{E}_i)$. Sont équivalents :

1. $X_1, \dots, X_n$ sont indépendantes ;
2. $P_{(X_1, \dots, X_n)} = P_{X_1} \otimes \dots \otimes P_{X_n}$;
3. pour toutes $f_i: E_i \rightarrow \mathbb{R}_+$ mesurables, $\mathbb{E} [\prod_i f_i(X_i)] = \prod_i \mathbb{E} [f_i(X_i)]$;
4. idem pour $f_i: E_i \rightarrow \mathbb{R}$ avec $\mathbb{E} [|f_i(X_i)|] < \infty$: alors $\prod_i f_i(X_i) \in L^1$ et l'égalité reste vraie.

(ii) est essentielle : indépendance $\iff$ la loi jointe est le produit des lois marginales.

Preuve : (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii)

(i) $\Rightarrow$ (ii). Pour $B_i \in \mathcal{E}_i$:

$$P_{(X_1, \dots, X_n)}(B_1 \times \dots \times B_n) = \prod_i \mathbb{P}(X_i \in B_i) = \prod_i P_{X_i}(B_i) = (P_{X_1} \otimes \dots \otimes P_{X_n})(B_1 \times \dots).$$

Conclu par le **théorème d'unicité** (les pavés engendrent la tribu produit).

(ii) $\Rightarrow$ (iii). Par transfert puis **Fubini–Tonelli** :

$$\mathbb{E} \left[\prod_i f_i(X_i) \right] = \int \prod_i f_i(x_i) \, d(P_{X_1} \otimes \dots \otimes P_{X_n}) = \prod_i \int f_i \, dP_{X_i} = \prod_i \mathbb{E} [f_i(X_i)].$$

Preuve : (iii) $\Rightarrow$ (i) et le cas (iv)

(iii) $\Rightarrow$ (i). Cas particulier avec $f_i = 1_{B_i}$: alors $\prod_i f_i(X_i) = 1_{\{X_1 \in B_1, \dots\}}$, ce qui redonne la définition.

Pour (iv). On applique (ii) $\Rightarrow$ (iii) à $|f_i|$ pour obtenir l'intégrabilité, puis on répète le calcul avec **Fubini–Lebesgue** à la place de Fubini–Tonelli. La preuve de (iv) $\Rightarrow$ (i) est identique à (iii) $\Rightarrow$ (i). □

Exemple

Soit $X \sim \mathcal{U}([0, 1])$ et $Y \sim \mathcal{C}(0, 1)$ indépendantes. Alors

$$dP_{(X,Y)}(x,y) = 1_{[0,1]}(x) \frac{1}{\pi(1+y^2)} dx dy.$$

Conséquence

Si $X, Y \in L^2$ sont indépendantes, alors $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ (par (iv), car $L^2 \subset L^1$), donc $\text{Cov}(X, Y) = 0$.

La réciproque est fausse : covariance nulle n'implique pas indépendance.

Corollaire

Si $X_1, \dots, X_n$ sont indépendantes et $0 = i_0 < i_1 < \dots < i_k = n$, alors les paquets

$$(X_{i_0+1}, \dots, X_{i_1}), (X_{i_1+1}, \dots, X_{i_2}), \dots, (X_{i_{k-1}+1}, \dots, X_{i_k})$$

sont indépendants.

Preuve. Conséquence de la caractérisation (ii) et de l'associativité du produit de mesures :

$P_{X_1} \otimes \dots \otimes P_{X_n}$ se regroupe en produit des lois jointes de chaque paquet. □

Exemple

Si $X_1, \dots, X_4$ sont indépendantes, alors $X_1^2 X_2$ et $X_3 - X_4$ sont indépendantes : $(X_1, X_2) \perp (X_3, X_4)$, puis on applique la proposition sur les fonctions de v.a. indépendantes.

V.a. continues indépendantes

Proposition

Soit $X_1, \dots, X_n$ réelles telles que $(X_1, \dots, X_n)$ soit continue, de densité

$$p_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = \prod_{i=1}^n q_i(x_i), \quad q_i: \mathbb{R} \rightarrow \mathbb{R}_+ \text{ mesurables.}$$

Alors $X_1, \dots, X_n$ sont **indépendantes** et X_i a densité $p_{X_i} = \frac{1}{c_i} q_i$, où

$$c_i = \int_{\mathbb{R}} q_i(x) dx > 0.$$

Un critère très pratique : factoriser la densité jointe suffit à conclure l'indépendance.

Par **Fubini–Tonelli** : $\prod_i \int_{\mathbb{R}} q_i(x_i) dx_i = \int_{\mathbb{R}^n} p_{(X_1, \dots, X_n)} dx_1 \cdots dx_n = 1$, donc chaque $c_i > 0$.

La proposition sur les marginales continues donne, pour chaque i ,

$$p_{X_i}(x_i) = q_i(x_i) \prod_{j \neq i} \left(\int_{\mathbb{R}} q_j(x_j) dx_j \right) = \frac{1}{c_i} q_i(x_i).$$

Enfin, $p_{(X_1, \dots, X_n)} = \prod_i q_i(x_i) = \prod_i c_i p_{X_i}(x_i) = \prod_i p_{X_i}(x_i)$ (car $\prod_i c_i = 1$), donc $P_{(X_1, \dots, X_n)} = \bigotimes_i P_{X_i}$: indépendance. □

Exemple : densité qui se factorise

Exemple

Soit (X, Y) de loi $e^{-2x}1_{\mathbb{R}_+ \times [0,2]}(x, y) dx dy$.

La densité se factorise :

$$p_{(X,Y)}(x, y) = (e^{-2x}1_{\mathbb{R}_+}(x)) \cdot (1_{[0,2]}(y)).$$

Donc X, Y sont **indépendantes**. On calcule $\int_{\mathbb{R}} 1_{[0,2]}(y) dy = 2$, donc

$$p_X(x) = 2e^{-2x}1_{\mathbb{R}_+}(x) \quad \text{et} \quad p_Y(y) = \frac{1}{2}1_{[0,2]}(y).$$

$X \sim \mathcal{E}(2)$ et $Y \sim \mathcal{U}([0, 2])$, **indépendantes**.

Exemple : le triangle, contre-exemple

Exemple

Reprenons (X, Y) uniforme sur $T = \{x \geq 0, y \geq 0, x + y \leq 1\}$, de densité $2 \mathbf{1}_T(x, y)$.

L'indicatrice $\mathbf{1}_{x+y \leq 1}$ ne semble **pas factorisable** en un produit d'une fonction de x et d'une fonction de y — cela suggère la non-indépendance, mais n'est *pas* une preuve.

Contre-exemple explicite

$$\mathbb{P}(X > 1/2, Y > 1/2) = 0 \quad (\text{car } X + Y \leq 1 \text{ p.s.})$$

mais $\mathbb{P}(X > 1/2) > 0$ et $\mathbb{P}(Y > 1/2) > 0$ (aires de triangles non nulles) : donc

$$\mathbb{P}(X > 1/2, Y > 1/2) \neq \mathbb{P}(X > 1/2) \mathbb{P}(Y > 1/2).$$

V.a. discrètes indépendantes

Proposition

Soit $X_1, \dots, X_n$ à valeurs dans des ensembles finis ou dénombrables $E_1, \dots, E_n$. Si

$$\forall x_i \in E_i, \quad \mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \cdots \mathbb{P}(X_n = x_n),$$

alors $X_1, \dots, X_n$ sont indépendantes.

La réciproque est immédiate en prenant $B_i = \{x_i\}$ dans la définition.

Il suffit de vérifier l'égalité sur les singletons !

$X = (X_1, \dots, X_n)$ prend ses valeurs dans $E_1 \times \dots \times E_n$, dénombrable : sa loi est caractérisée par ses singletons. Pour $x = (x_1, \dots, x_n)$:

$$P_X(\{x\}) = \mathbb{P}(X = x) = \prod_i \mathbb{P}(X_i = x_i) = \prod_i P_{X_i}(\{x_i\}) = (P_{X_1} \otimes \dots \otimes P_{X_n})(\{x\}).$$

Donc $P_{(X_1, \dots, X_n)} = P_{X_1} \otimes \dots \otimes P_{X_n}$: indépendance par la caractérisation (ii). □

Exemple : deux dés, différence et parité

Exemple

On relance deux dés. Soit $X(i, j) = |i - j|$ (différence absolue) et $Y = 1_B$ où $B =$ “le 1er dé donne un résultat pair”.

On vérifie, par exemple,

$$\mathbb{P}(X = 0, Y = 1) = \frac{3}{36} = \frac{6}{36} \cdot \frac{1}{2} = \mathbb{P}(X = 0)\mathbb{P}(Y = 1),$$

et de même pour $X = 1, \dots, 5$.

Astuce

Pas besoin de vérifier $\mathbb{P}(X = i, Y = 0)$ séparément : $\{X = i\} \perp \{Y = 1\} \Rightarrow \{X = i\} \perp \{Y = 1\}^c = \{Y = 0\}$. Donc X, Y sont indépendantes.

Sommes de v.a. indép.

Proposition

Soit X, Y réelles continues indépendantes. Alors $X + Y$ est continue, de densité

$$p_{X+Y}(s) = (p_X * p_Y)(s) = \int_{\mathbb{R}} p_X(x) p_Y(s - x) \, dx.$$

Méthode : fonction caractéristique d'une somme

Méthode

Si $X_1, \dots, X_n$ sont indépendantes, alors

$$\phi_{X_1+\dots+X_n}(\theta) = \mathbb{E} \left[\prod_{k=1}^n e^{i\theta X_k} \right] = \prod_{k=1}^n \mathbb{E} \left[e^{i\theta X_k} \right] = \prod_{k=1}^n \phi_{X_k}(\theta),$$

en utilisant l'indépendance (critère (iv), généralisé aux fonctions complexes).

Si $\phi_{X_1+\dots+X_n}$ s'identifie à la fonction caractéristique d'une loi connue, on a trouvé la loi de la somme !

Exemple : somme de v.a. de Poisson

Exemple

Soit $X_1, \dots, X_n$ indépendantes, $X_k \sim \mathcal{P}(\lambda_k)$. Rappel : $\phi_{X_k}(\theta) = \exp(\lambda_k(e^{i\theta} - 1))$.

$$\phi_{X_1 + \dots + X_n}(\theta)$$

Exemple : somme de v.a. de Poisson

Exemple

Soit $X_1, \dots, X_n$ indépendantes, $X_k \sim \mathcal{P}(\lambda_k)$. Rappel : $\phi_{X_k}(\theta) = \exp(\lambda_k(e^{i\theta} - 1))$.

$$\phi_{X_1 + \dots + X_n}(\theta) = \prod_{k=1}^n \exp(\lambda_k(e^{i\theta} - 1))$$

Exemple : somme de v.a. de Poisson

Exemple

Soit $X_1, \dots, X_n$ indépendantes, $X_k \sim \mathcal{P}(\lambda_k)$. Rappel : $\phi_{X_k}(\theta) = \exp(\lambda_k(e^{i\theta} - 1))$.

$$\phi_{X_1 + \dots + X_n}(\theta) = \prod_{k=1}^n \exp(\lambda_k(e^{i\theta} - 1)) = \exp\left(\left(\sum_{k=1}^n \lambda_k\right)(e^{i\theta} - 1)\right).$$

$$X_1 + \dots + X_n \sim \mathcal{P}(\sum_k \lambda_k).$$

Exemple

Soit $X_1, \dots, X_n$ indépendantes, $X_k \sim \mathcal{N}(m_k, \sigma_k^2)$. Rappel :

$$\phi_{X_k}(\theta) = \exp(im_k\theta - \frac{1}{2}\sigma_k^2\theta^2).$$

$$\phi_{X_1+\dots+X_n}(\theta) = \exp\left(i\left(\sum_k m_k\right)\theta - \frac{1}{2}\left(\sum_k \sigma_k^2\right)\theta^2\right).$$

$$X_1 + \dots + X_n \sim \mathcal{N}\left(\sum_k m_k, \sum_k \sigma_k^2\right).$$

Min/max de v.a. indép.

Méthode

Pour $X_1, \dots, X_n$ réelles indépendantes, la loi de max ou min s'obtient via la fonction de répartition. Pour $a \in \mathbb{R}$:

$$\mathbb{P}(\max(X_1, \dots, X_n) \leq a) = \prod_{k=1}^n \mathbb{P}(X_k \leq a),$$

$$\mathbb{P}(\min(X_1, \dots, X_n) \leq a) = 1 - \prod_{k=1}^n \mathbb{P}(X_k > a) = 1 - \prod_{k=1}^n (1 - F_{X_k}(a)).$$

Le max/min utilise directement l'indépendance des événements $\{X_k \leq a\}$ ou $\{X_k > a\}$.

Exemple : minimum de v.a. exponentielles

Exemple

Soit $X_1, \dots, X_n$ indépendantes, $X_k \sim \mathcal{E}(\lambda_k)$, et $M_n = \min(X_1, \dots, X_n)$.

Pour $a \geq 0$:

$$F_{M_n}(a)$$

Exemple : minimum de v.a. exponentielles

Exemple

Soit $X_1, \dots, X_n$ indépendantes, $X_k \sim \mathcal{E}(\lambda_k)$, et $M_n = \min(X_1, \dots, X_n)$.

Pour $a \geq 0$:

$$F_{M_n}(a) = 1 - \prod_{k=1}^n (1 - F_{X_k}(a))$$

Exemple : minimum de v.a. exponentielles

Exemple

Soit $X_1, \dots, X_n$ indépendantes, $X_k \sim \mathcal{E}(\lambda_k)$, et $M_n = \min(X_1, \dots, X_n)$.

Pour $a \geq 0$:

$$F_{M_n}(a) = 1 - \prod_{k=1}^n (1 - F_{X_k}(a)) = 1 - e^{-\lambda_1 a} \dots e^{-\lambda_n a}$$

Exemple : minimum de v.a. exponentielles

Exemple

Soit $X_1, \dots, X_n$ indépendantes, $X_k \sim \mathcal{E}(\lambda_k)$, et $M_n = \min(X_1, \dots, X_n)$.

Pour $a \geq 0$:

$$F_{M_n}(a) = 1 - \prod_{k=1}^n (1 - F_{X_k}(a)) = 1 - e^{-\lambda_1 a} \dots e^{-\lambda_n a} = 1 - e^{-(\lambda_1 + \dots + \lambda_n)a},$$

et $F_{M_n}(a) = 0$ pour $a < 0$ (car $M_n > 0$ p.s.).

$$M_n \sim \mathcal{E}(\lambda_1 + \dots + \lambda_n).$$

Interprétation

Si le temps de désintégration d'un atome suit $\mathcal{E}(\lambda)$ (perte de mémoire), alors le temps avant la première désintégration parmi N atomes indépendants suit

$$\mathcal{E}(N\lambda).$$

Plus il y a d'atomes, plus la première désintégration est rapide — exemple classique d'utilisation du minimum de v.a. exponentielles indépendantes.

Synthèse

- L'indépendance **deux à deux** n'implique jamais l'indépendance globale.
- Caractérisation clé : $X_1, \dots, X_n$ indépendantes $\iff P_{(X_1, \dots, X_n)} = P_{X_1} \otimes \dots \otimes P_{X_n}$.
- Pour des v.a. **continues** : factoriser la densité jointe donne l'indépendance et les marginales directement.
- Pour des v.a. **discrètes** : il suffit de vérifier l'égalité sur les singletons.
- Somme de v.a. indépendantes : **convolution** des densités, ou produit des fonctions caractéristiques.
- Min/max de v.a. indépendantes : passer par la **fonction de répartition**.

Inégalité de concentration