Plan

1. Espace de probabilité
2. Variable aléatoire
   - Variable aléatoire discrète
   - Lois discrètes usuelles
3. Couple de variables aléatoires
4. Vecteur de variables aléatoires

Cas général — Espace produit mesurable

Définition 2.1.39

Soient $(E_i, \mathcal{F}_i)$, $1 \leq i \leq d$, des espaces mesurables. On pose $E = E_1 \times \cdots \times E_d$ et on appelle **tribu produit** sur E :

$$\bigotimes_{i=1}^{d} \mathcal{F}_i = \sigma\Big(\big\{A_1 \times \cdots \times A_d : \ A_i \in \mathcal{F}_i\big\}\Big).$$

Le couple $(E, \bigotimes_{i=1}^{d} \mathcal{F}_i)$ est un **espace produit mesurable**.

- Les $A_1 \times \cdots \times A_d$ sont des **rectangles mesurables**.
- On prend la tribu *engendrée* par ces rectangles, et non leur simple famille.

Cas général — Vecteur aléatoire

Définition 2.1.40

Soient $(\Omega, \mathcal{A}, \mathbb{P}_\Omega)$ un espace de probabilité et $(\prod_{i=1}^d E_i, \bigotimes_{i=1}^d \mathcal{F}_i)$ un espace produit mesurable. Une application mesurable

$$\boxed{X : (\Omega, \mathcal{A}) \longrightarrow \left(\prod_{i=1}^{d} E_i, \bigotimes_{i=1}^{d} \mathcal{F}_i \right)}$$

est appelée un **vecteur aléatoire** de dimension d.

Puisque X prend ses valeurs dans $E_1 \times E_2 \times ... \times E_d$, donc sous forme de d-uplets, on écrit $\boxed{X = (X_1, \ldots, X_d)}$, où chaque $X_i : \Omega \to E_i$ est une variable aléatoire.

Réciproquement, si chaque X_i est mesurable, alors $(X_1, \ldots, X_d)$ est un vecteur aléatoire.

Cas discret — Vecteur aléatoire

Si $E = \prod_{i=1}^{d} E_i$ est dénombrable, on dit que $X = (X_1, \ldots, X_d)$ est un **vecteur aléatoire discret**. Dans ce cas, la tribu d'arrivée par défaut est alors $\mathcal{P}(E)$.
**Ici, on a l'égalité $\mathcal{P}(E) = \bigotimes_{i=1}^{d} \mathcal{P}(E_i)$.*

Exemple 2.1.41

On lance deux dés équilibrés : $\Omega = \{1, \ldots, 6\}^2$. Pour $(i, j) \in \Omega$, on définit

$$I(i,j) = i, \qquad J(i,j) = j, \qquad Z(i,j) = \mathbf{1}_{\{i=j\}}.$$

Le vecteur $X = (I, J, Z)$ prend ses valeurs dans $E = \{1, \ldots, 6\}^2 \times \{0, 1\}$.

Par exemple, $X(2,2) = (2,2,1)$ et $X(2,5) = (2,5,0)$.

Remarque : toutes les valeurs de E ne sont pas nécessairement atteintes.

1. Lorsque l'on munit Ω de la tribu $\mathcal{P}(\Omega)$ et E de la tribu $\mathcal{P}(E)$, on peut parler de la mesurabilité de l'application X.
2. Si l'on munit ensuite l'espace mesurable $(\Omega, \mathcal{P}(\Omega))$ d'une probabilité, on peut alors parler du caractère aléatoire de l'application mesurable X.

Cas général — Loi jointe

Définition 2.1.42

Soit $X = (X_1, \ldots, X_d)$ un vecteur aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P}_\Omega)$, à valeurs dans $(E, \bigotimes_{i=1}^d \mathcal{F}_i)$. Sa **loi jointe** est la probabilité $\mathbb{P}_X$ sur cet espace définie par

$$\mathbb{P}_X(C) = \mathbb{P}_\Omega(\{\omega \in \Omega : X(\omega) \in C\}), \qquad C \in \bigotimes_{i=1}^d \mathcal{F}_i.$$

La probabilité de base est ainsi transportée vers l'espace d'arrivée du vecteur.

Cas discret — Fonction de masse jointe

Définition 2.1.43

Pour un vecteur aléatoire discret $X = (X_1, \ldots, X_d)$, on appelle **fonction de masse jointe** :

$$\boxed{p_X(x_1, \ldots, x_d) = \mathbb{P}_X(X_1 = x_1, \ldots, X_d = x_d).}$$

Elle détermine entièrement la loi jointe : $\mathbb{P}_X(C) = \sum_{x \in C} p_X(x)$.

On retrouve la fonction de masse du couple lorsque $d = 2$.

Cas discret — Normalisation de la masse jointe

Proposition 2.1.44

La fonction de masse jointe d'un vecteur discret est positive et sa somme vaut 1 :

$$\boxed{\sum_{x_1 \in E_1} \cdots \sum_{x_d \in E_d} p_X(x_1, \ldots, x_d) = 1.}$$

Démonstration. Les événements $\{X = x\}$, pour $x \in E$, sont deux à deux disjoints et leur réunion est Ω.

Par σ-additivité,

$$\sum_{x \in E} p_X(x) = \mathbb{P}_\Omega\left(\bigsqcup_{x \in E} \{X = x\}\right) = \mathbb{P}_\Omega(\Omega) = 1.$$

Cas discret — Exemple de masse jointe

Exemple 2.1.45

Deux dés équilibrés. On reprend $X = (I, J, Z)$, avec $I(i,j) = i$, $J(i,j) = j$ et $Z(i,j) = \mathbf{1}_{\{i=j\}}$.

Pour $(a, b, z) \in \{1, \ldots, 6\}^2 \times \{0, 1\}$:

$$p_{I,J,Z}(a,b,z) = \begin{cases} \dfrac{1}{36}, & \text{si } z = \mathbf{1}_{\{a=b\}}, \\ 0, & \text{sinon.} \end{cases}$$

- Si $a = b$, seul le triplet $(a, a, 1)$ est possible.
- Si $a \neq b$, seul le triplet $(a, b, 0)$ est possible.
- Il y a exactement 36 triplets possibles de masse $1/36$.

Cas général — Lois marginales

Définition 2.1.46

Soit $X = (X_1, \ldots, X_d)$ de loi jointe $\mathbb{P}_X$. Pour un ensemble non vide d'indices $K \subseteq \{1, \ldots, d\}$, la loi du sous-vecteur $X_K = (X_i)_{i \in K}$ est appelée une **loi marginale** de X.

En notant $\pi_K : E \to \prod_{i \in K} E_i$ la projection sur les coordonnées de K, on a :

$$\boxed{\mathbb{P}_{X_K}(B) = \mathbb{P}_X\Big(\pi_K^{-1}(B)\Big), \qquad B \in \bigotimes_{i \in K} \mathcal{F}_i.}$$

En particulier : la loi de chaque X_i est une marginale de la loi jointe de X.

Cas discret — Calcul des marginales

Proposition 2.1.47

Soit $X = (X_1, \ldots, X_d)$ un vecteur discret. Pour $K \subseteq \{1, \ldots, d\}$ non vide, la fonction de masse du sous-vecteur X_K s'obtient en **sommant sur les autres coordonnées** :

$$\boxed{p_{X_K}(x_K) = \sum_{x_{K^c} \in E_{K^c}} p_X(x_K, x_{K^c}).}$$

Cas discret — Exemple de lois marginales

Exemple 2.1.48

Avec $X = (I, J, Z)$, on somme d'abord sur les valeurs de Z :

$$p_{I,J}(a, b) = \sum_{z=0}^{1} p_{I,J,Z}(a, b, z) = \frac{1}{36}.$$

Puis, en sommant sur les valeurs de J :

$$p_I(a) = \sum_{b=1}^{6} p_{I,J}(a, b) = \sum_{b=1}^{6} \frac{1}{36} = \frac{1}{6}.$$

De même, $p_J(b) = 1/6$; et $p_Z(1) = 1/6$, $p_Z(0) = 5/6$.

Cas général — Indépendance des composantes

Définition 2.1.49

Les variables aléatoires $X_1, \ldots, X_d$, définies sur le même espace de probabilité de base, sont dites **mutuellement indépendantes** si, pour tous $A_i \in \mathcal{F}_i$,

$$\mathbb{P}_\Omega\left(\bigcap_{i=1}^{d}\{\omega \in \Omega : X_i(\omega) \in A_i\}\right) = \prod_{i=1}^{d}\mathbb{P}_\Omega(\{\omega \in \Omega : X_i(\omega) \in A_i\}).$$

- Le cas $d = 2$ retrouve la définition de l'indépendance d'un couple.
- Si $d \geq 3$, l'indépendance **deux à deux** ne suffit pas, en général, à garantir l'indépendance mutuelle.
- Contrairement au cas des **événements**, il suffit ici de vérifier la factorisation pour les d indices : celle de toute sous-famille s'en déduit en prenant $A_i = E_i$ pour les indices omis.

Cas discret — Critère de factorisation

Proposition 2.1.50

Soit $X = (X_1, \ldots, X_d)$ un vecteur discret. Les variables $X_1, \ldots, X_d$ sont mutuellement indépendantes si et seulement si, pour toutes les valeurs $x_i \in E_i$,

$$\boxed{p_X(x_1, \ldots, x_d) = \prod_{i=1}^{d} p_{X_i}(x_i).}$$

Pourquoi cela suffit-il ? Pour $A_i \subseteq E_i$, on somme sur le rectangle :

$$\begin{aligned}\mathbb{P}_\Omega(X_1 \in A_1, \ldots, X_d \in A_d) &= \sum_{x_1 \in A_1} \cdots \sum_{x_d \in A_d} \prod_{i=1}^{d} p_{X_i}(x_i) \\ &= \prod_{i=1}^{d} \sum_{x_i \in A_i} p_{X_i}(x_i) = \prod_{i=1}^{d} \mathbb{P}_{X_i}(A_i).\end{aligned}$$

Cas discret — Indépendance deux à deux ou mutuelle ?

Exemple 2.1.51

Pour les deux dés, $I(i,j) = i$, $J(i,j) = j$ et $Z(i,j) = \mathbf{1}_{\{i=j\}}$.

- I et J sont indépendantes : $p_{I,J}(a,b) = 1/36$.
- I et Z sont indépendantes : $p_{I,Z}(a,1) = 1/36 = (1/6)(1/6)$, et de même pour $Z = 0$.
- Par symétrie, J et Z sont indépendantes.

Cependant, pour le triplet $(1,1,0)$:

$$\mathbb{P}_\Omega(I = 1, J = 1, Z = 0) = 0,$$
$$\mathbb{P}_\Omega(I = 1)\mathbb{P}_\Omega(J = 1)\mathbb{P}_\Omega(Z = 0) = \frac{5}{216} > 0.$$

Conclusion : les trois composantes sont indépendantes **deux à deux**, mais **pas mutuellement**.

Cas général — Fonctions de variables indépendantes

Proposition 2.1.52

Soient $X_1, \ldots, X_d$ des variables aléatoires mutuellement indépendantes, et $\Gamma_i : E_i \to G_i$ des applications mesurables.
Alors $\Gamma_1(X_1), \ldots, \Gamma_d(X_d)$ sont **mutuellement indépendantes**.

Démonstration. Pour des ensembles mesurables $B_i \subseteq G_i$,

$$\mathbb{P}_\Omega\left(\bigcap_{i=1}^{d}\{\Gamma_i(X_i) \in B_i\}\right)$$
$$= \mathbb{P}_\Omega\left(\bigcap_{i=1}^{d}\{X_i \in \Gamma_i^{-1}(B_i)\}\right) = \prod_{i=1}^{d}\mathbb{P}_\Omega(\Gamma_i(X_i) \in B_i).$$

Cas général — Conditionnement par un sous-vecteur

Définition 2.1.53

Soit $X = (X_1, \ldots, X_d)$ un vecteur aléatoire à valeurs dans

$$(E, \mathcal{F}) = \left(\prod_{i=1}^{d} E_i, \bigotimes_{i=1}^{d} \mathcal{F}_i \right),$$

de loi jointe $\mathbb{P}_X$. Soit $\varnothing \neq K \subsetneq \{1, \ldots, d\}$. On note $K^c = \{1, \ldots, d\} \setminus K$ et on pose

$$U = X_K = (X_i)_{i \in K}, \qquad V = X_{K^c} = (X_i)_{i \in K^c}.$$

On note également

$$\mathcal{F}_K = \bigotimes_{i \in K} \mathcal{F}_i, \qquad \mathcal{F}_{K^c} = \bigotimes_{i \in K^c} \mathcal{F}_i.$$

(Voir la page suivante)

Définition 2.1.54

(*La suite.*) Pour $A \in \mathcal{F}_K$ tel que $\mathbb{P}_U(A) > 0$, on définit la **loi conditionnelle de V sachant $U \in A$** par

$$\mathbb{P}_{V|U\in A}(B) = \frac{\mathbb{P}_X(V \in B,\ U \in A)}{\mathbb{P}_U(U \in A)}, \qquad B \in \mathcal{F}_{K^c}.$$

Pour A fixé, l'application $B \mapsto \mathbb{P}_{V|U\in A}(B)$ est une **probabilité** sur $(E_{K^c}, \mathcal{F}_{K^c})$, où $E_{K^c} = \prod_{i\in K^c} E_i$.

Cas discret — Masse conditionnelle

Si U et V sont discrets, on déduit de la définition générale, pour tout v :

$$p_{V|U\in A}(v) = \mathbb{P}_{V|U\in A}(\{v\})$$
$$= \frac{\mathbb{P}_X(V = v,\ U \in A)}{\mathbb{P}_U(U \in A)} = \boxed{\frac{\sum_{u\in A} p_{U,V}(u, v)}{\sum_{u\in A} p_U(u)}}.$$

En particulier, si $p_U(u) > 0$,

$$\boxed{p_{V|U=u}(v) = \frac{p_{U,V}(u, v)}{p_U(u)}.}$$

Toutes les sommes portent sur des valeurs du vecteur U.

Cas discret — Exemple de conditionnement

Exemple 2.1.55

On reprend $X = (I, J, Z)$ pour deux dés équilibrés, avec $Z = \mathbf{1}_{\{I=J\}}$. On conditionne Z par le sous-vecteur $U = (I, J)$.

Pour tout $(a, b) \in \{1, \dots, 6\}^2$:

$$p_{Z|I=a,\, J=b}(1) = \frac{p_{I,J,Z}(a, b, 1)}{p_{I,J}(a, b)} = \begin{cases} 1, & a = b, \\ 0, & a \neq b. \end{cases}$$

- Si $I = 2$ et $J = 2$, alors $Z = 1$ avec probabilité 1.
- Si $I = 2$ et $J = 5$, alors $Z = 1$ avec probabilité 0.

Contrairement au conditionnement par I seul, connaître (I, J) détermine entièrement Z.

Cas général — Indépendance et loi conditionnelle

Proposition 2.1.56

Soient U et V deux sous-vecteurs aléatoires définis sur le même espace de probabilité de base. Ils sont **indépendants si et seulement si**, pour tout événement A mesurable de l'espace de U tel que $\mathbb{P}_U(A) > 0$,

$$\boxed{\mathbb{P}_{V|U\in A} = \mathbb{P}_V.}$$

Démonstration. Pour tout ensemble mesurable B de l'espace de V :

- Si U et V sont indépendants, $\mathbb{P}_{V|U\in A}(B) = \dfrac{\mathbb{P}_U(A)\mathbb{P}_V(B)}{\mathbb{P}_U(A)} = \mathbb{P}_V(B)$.
- Réciproquement, on multiplie par $\mathbb{P}_U(A)$; si $\mathbb{P}_U(A) = 0$, les deux membres sont nuls.

Cas discret — Critère par les lois conditionnelles

Proposition 2.1.57

Soient U et V deux sous-vecteurs discrets. Ils sont indépendants si et seulement si, pour chaque u tel que $p_U(u) > 0$,

$$\boxed{p_{V|U=u}(v) = p_V(v), \qquad \text{pour tout } v.}$$

Exercices

1. Soient $X_1, \ldots, X_n$ des variables i.i.d. de loi $\mathcal{B}(p)$ et $S = X_1 + \cdots + X_n$. Déduire la fonction de masse de la loi binomiale de S.

2. Soient X et Y deux variables aléatoires indépendantes telles que

$$X \sim \mathcal{B}(n, p), \qquad Y \sim \mathcal{B}(m, p),$$

où $n, m \in \mathbb{N}^*$ et $p \in [0, 1]$. On pose $S = X + Y$, déterminer la loi de S.

3. Soient $(X_i)_{i \geq 1}$ des variables i.i.d. de loi $\mathcal{B}(p)$, avec $p \in (0, 1]$, et $T = \inf\{i \geq 1 : X_i = 1\}$. Déduire la fonction de masse de la loi géométrique de T.

4. Montrer que, lorsque n est grand, p petit et $np \approx \lambda$, la loi $\mathcal{B}(n, p)$ peut être approximée par $\mathcal{P}(\lambda)$.

   Indication : fixer $\lambda > 0$, prendre $p_n = \lambda/n$ et étudier la limite quand $n \to \infty$.

Exercice4

Lors des Jeux olympiques de Sydney, 300 médailles d'or étaient en jeu.

On adopte le modèle simplifié selon lequel le nombre de médailles d'or remportées par un pays suit une loi binomiale $\mathcal{B}(300, p)$, avec

$$p = \frac{\text{population du pays}}{\text{population mondiale}}.$$

Pour la France, on prend 50 millions d'habitants et, pour le monde, 5 milliards.

1. Déterminer p et la loi de X, nombre de médailles françaises.
2. Calculer $\mathbb{P}(X \geq 10)$ et donner une approximation de Poisson.
3. La France a remporté 13 médailles d'or. Commenter ce résultat.