

Chapitre 23

Probabilités 2 – Indépendance, conditionnement

Plan du chapitre

1	Probabilités conditionnelles	2
1.1	Définition	2
1.2	Formule des probabilités composées	3
1.3	Formule des probabilités totales	3
1.4	Formule de Bayes	4
1.5	Loi conditionnelle	4
2	Indépendance d'événements	5
2.1	Indépendance de deux événements	5
2.2	Événements mutuellement indépendants	7
3	Couples de v.a.	8
3.1	Définition, loi conjointe	8
3.2	Lois marginales	9
4	Variables aléatoires indépendantes	10
4.1	Définition	10
4.2	Indépendance de n variables aléatoires	10
4.3	Somme de n v.a. indépendantes de même loi $\mathcal{B}(p)$	12

Hypothèse

Dans tout ce chapitre, (Ω, \mathbb{P}) désigne un espace probabilisé fini (i.e. Ω désigne un univers fini et \mathbb{P} est une probabilité définie sur Ω).

E, F sont des ensembles quelconques.

1 Probabilités conditionnelles

1.1 Définition

Définition 23.1

Soit A, B deux événements de Ω . Si $\mathbb{P}(B) > 0$, on appelle probabilité conditionnelle de A sachant B le réel noté :

$$\mathbb{P}(A | B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

On note également $\mathbb{P}_B(A) = \mathbb{P}(A | B)$

$\mathbb{P}(A | B)$ représente la probabilité que A se produise sachant déjà que B s'est produit. On prendra garde au fait que la notation est trompeuse : " $A | B$ " n'est pas un événement.

Exemple 1. On tire une pièce deux fois, on note A l'événement "deux faces ont été obtenues" et B l'événement "au moins un pile a été obtenu". Alors

$$\mathbb{P}(B) = \dots \neq 0$$

Or, $A \cap B = \dots$ d'où

$$\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \dots$$

On note A' l'événement "deux piles ont été obtenus". Alors $A' \subset B$ si bien que

$$\mathbb{P}(A' | B) = \frac{\mathbb{P}(A' \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A')}{\mathbb{P}(B)} = \dots$$

Proposition 23.2

Soit B un événement de Ω tel que $\mathbb{P}(B) > 0$. Alors l'application

$$\begin{aligned} \mathbb{P}_B : \Omega &\rightarrow [0, 1] \\ A &\mapsto \mathbb{P}_B(A) = \mathbb{P}(A | B) \end{aligned}$$

est une probabilité sur Ω .

Elle hérite ainsi des propriétés vérifiées par toute probabilité :

1. $\mathbb{P}_B(\Omega) = \mathbb{P}(\Omega | B) = 1$ et $\mathbb{P}(\emptyset | B) = 0$
2. Si A, A' sont des événements disjoints, on a

$$\mathbb{P}(A \cup A' | B) = \mathbb{P}(A | B) + \mathbb{P}(A' | B)$$

- 3.

$$\mathbb{P}(\bar{A} | B) = 1 - \mathbb{P}(A | B)$$

4. Si $A \subset A' \subset \Omega$, alors

$$\mathbb{P}(A | B) \leq \mathbb{P}(A' | B)$$

5. Pour tous événements A, A' ,

$$\mathbb{P}(A \cup A' | B) = \mathbb{P}(A | B) + \mathbb{P}(A' | B) - \mathbb{P}(A \cap A' | B)$$

1.2 Formule des probabilités composées

Proposition 23.3

Pour tous événements A, B ,

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

avec la convention $\mathbb{P}(A | B)\mathbb{P}(B) = 0$ si $\mathbb{P}(B) = 0$.

Si $\mathbb{P}(B) = 0$, le réel $\mathbb{P}(A | B)$ n'a pas de sens en soi. Cependant, avec la convention, la formule reste valide : en effet si $\mathbb{P}(B) = 0$, alors $\mathbb{P}(A \cap B) = 0$ car $A \cap B \subset B$.

Proposition 23.4 (Formule des probabilités composées)

Soit A_1, \dots, A_n des événements. Alors

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 | A_1)\mathbb{P}(A_3 | A_1 \cap A_2) \dots \mathbb{P}(A_n | A_1 \cap \dots \cap A_{n-1})$$

Cette formule est particulièrement utilisée lorsque A_1, \dots, A_n forment une suite d'événements "chronologiques" qui ne sont pas indépendants.

Exemple 2. On dispose d'une urne avec 3 boules rouges et 3 boules noires. On tire successivement 3 boules, sans remise. Pour tout $i \in \llbracket 1, 3 \rrbracket$, on note A_i l'événement "le tirage numéro i donne une boule rouge". Quelle est la probabilité de tirer 3 boules rouges ?

1.3 Formule des probabilités totales

Proposition 23.5

Soit (B_1, \dots, B_n) un système complet d'événements. Alors pour tout événement A , on a

$$\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A | B_i)\mathbb{P}(B_i)$$

Démonstration. On a vu au chapitre précédent que, comme (B_1, \dots, B_n) est un système complet d'événements,

$$\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A \cap B_i) = \sum_{i=1}^n \mathbb{P}(A | B_i)\mathbb{P}(B_i)$$

d'où le résultat. □

Rappelons que si $\mathbb{P}(B_i) = 0$, alors le terme correspondant dans la somme, i.e. $\mathbb{P}(A | B_i)\mathbb{P}(B_i)$, est nul par convention.

Exemple 3. Voir exemple suivant.

1.4 Formule de Bayes

Proposition 23.6 (Formule de Bayes)

Soit A, B deux événements tels que $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$. Alors

$$\mathbb{P}(B | A) = \frac{\mathbb{P}(A | B)\mathbb{P}(B)}{\mathbb{P}(A)}$$

Démonstration. Cela résulte du fait que

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

□

La formule de Bayes est un grand classique : elle permet de “retourner” un conditionnement. Bien souvent, pour trouver $\mathbb{P}(A)$, on utilise la formule des probabilités totales.

Exemple 4. Une certaine maladie affecte une personne sur dix mille. On dispose d’un test pour détecter cette maladie. On note M l’événement “la personne testée est malade” et T l’événement “le test est positif”. On a donc

$$\mathbb{P}(M) = \frac{1}{10000} = 10^{-4}$$

- Si la personne est malade, le test détectera la maladie dans 99% des cas, c’est à dire

$$\mathbb{P}(T | M) = 99\% = 0,99$$

(c’est la proportion de vrai positif)

- Si la personne n’est pas malade, le test sera positif dans 0,1% des cas, c’est à dire

$$\mathbb{P}(T | \bar{M}) = 0,1\% = 0,001$$

Calculer la probabilité d’être malade sachant que le test est positif.

1.5 Loi conditionnelle

Rappel : si $X : \Omega \rightarrow E$ est une v.a. sur l’espace probabilisé (Ω, \mathbb{P}) , la loi de X est l’application

$$\begin{aligned} \mathbb{P}_X : \mathcal{P}(E) &\rightarrow [0, 1] \\ A &\mapsto \mathbb{P}(X \in A) = \mathbb{P}(X^{-1}(A)) \end{aligned}$$

Définition 23.7

Soit $X : \Omega \rightarrow E$ une v.a. sur l'espace probabilisé (Ω, \mathbb{P}) . Pour tout événement $B \subset \Omega$ tel que $\mathbb{P}(B) \neq 0$ on définit la loi conditionnelle de X sachant B comme étant l'application

$$\begin{aligned} \mathcal{P}(E) &\rightarrow [0, 1] \\ A &\mapsto \mathbb{P}(X \in A \mid B) = \mathbb{P}_B(X \in A) = \mathbb{P}_B(X^{-1}(A)) \end{aligned}$$

Autrement dit, cela correspond à la loi de X dans l'espace probabilisé (Ω, \mathbb{P}_B) .

On prendra garde au fait que dans la définition ci-dessus, on a $B \subset \Omega$ mais $A \subset E$. Contrairement à \mathbb{P}_X ci-dessus, il n'y a pas de notation spécifique pour la loi conditionnelle. Comme toutes les lois, elle est entièrement déterminée par sa valeur sur les singletons $\{x\}$ avec $x \in E$, c'est-à-dire par la distribution de probabilités

$$(\mathbb{P}(X = x \mid B))_{x \in E}$$

Exemple 5. On lance un dé à six faces. On pose X la v.a. égale au numéro obtenu, et B l'événement $\{X \text{ est pair}\}$. Déterminer la loi conditionnelle de X sachant B .

Proposition 23.8

Soit X, Y des v.a. sur (Ω, \mathbb{P}) à valeurs dans E et F respectivement. Alors pour tout $x \in E$

$$\mathbb{P}(X = x) = \sum_{y \in F} \mathbb{P}(X = x \mid Y = y) \mathbb{P}(Y = y)$$

avec la convention $\mathbb{P}(X = x \mid Y = y) \mathbb{P}(Y = y) = 0$ si $\mathbb{P}(Y = y) = 0$.

2 Indépendance d'événements

2.1 Indépendance de deux événements

Définition 23.9 (Indépendance)

Deux événements A, B sont dits indépendants si

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

Cela signifie concrètement que le fait que l'événement A soit réalisé ou non ne dépend pas du fait que B soit réalisé ou non.

Exemple 6. On tire une carte dans un jeu de 52 cartes. Montrer que les événements suivants sont indépendants :

$$A = \{\text{la carte est un as}\}$$

$$B = \{\text{la carte est un trèfle}\}$$

La loi étant uniforme, on a

$$\mathbb{P}(A) = \frac{4}{52} = \frac{1}{13} \quad \mathbb{P}(B) = \frac{13}{52} = \frac{1}{4} \quad \mathbb{P}(A \cap B) = \frac{1}{52}$$

On a bien $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$: les deux événements sont indépendants.

Remarque. Ne pas confondre événements indépendants et événements incompatibles. En général ce sont même des notions opposées. Si on suppose $A \cap B = \emptyset$ avec $A \neq \emptyset$ et $B \neq \emptyset$, alors

$$\mathbb{P}(A \cap B) = 0 \quad \text{mais} \quad \mathbb{P}(A)\mathbb{P}(B) \neq 0$$

donc les événements A et B ne sont pas indépendants. C'est assez naturel : si A, B sont incompatibles, cela signifie que si l'un se produit, l'autre ne peut pas se produire. Il y a donc bien une dépendance entre ces événements.

La Proposition suivante fournit un éclairage saisissant :

Proposition 23.10

Soit A, B deux événements avec $\mathbb{P}(B) > 0$. Alors A, B sont indépendants si et seulement si

$$\mathbb{P}(A | B) = \mathbb{P}(A)$$

Démonstration.

$$\mathbb{P}(A | B) = \mathbb{P}(A) \iff \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A) \iff \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

□

Cette Proposition (ainsi que la suivante) montre en particulier que si A, B sont indépendants, le fait que B soit réalisé ou non ne change pas la probabilité que A le soit aussi.

Proposition 23.11

Si A et B sont des événements indépendants, alors A et \bar{B} sont indépendants. En particulier

$$\mathbb{P}(A | \bar{B}) = \mathbb{P}(A)$$

Démonstration. Par la formule des probabilités totales, comme (B, \bar{B}) forme un système complet d'événements,

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \bar{B}) \\ &= \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(A \cap \bar{B}) \end{aligned}$$

d'où

$$\mathbb{P}(A \cap \bar{B}) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(\bar{B})$$

Donc A et \bar{B} sont indépendants.

□

2.2 Événements mutuellement indépendants

Définition 23.12

Soit A_1, \dots, A_n des événements de Ω .

- A_1, \dots, A_n sont dits indépendants 2 à 2 si pour tous $i, j \in \llbracket 1, n \rrbracket$ distincts, les événements A_i, A_j sont indépendants.
- A_1, \dots, A_n sont dits (mutuellement) indépendants si pour tout sous-ensemble I de $\llbracket 1, n \rrbracket$, on a

$$\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i)$$

L'indépendance mutuelle est plus forte que l'indépendance deux à deux.

Exemple 7. On lance une pièce deux fois. On pose les événements

$$A = \{\text{le premier lancer est pile}\}$$

$$B = \{\text{le deuxième lancer est pile}\}$$

$$C = \{\text{les deux lancers sont identiques}\}$$

Alors A, B, C sont indépendants deux à deux mais ne sont pas mutuellement indépendants.

3 Couples de v.a.

3.1 Définition, loi conjointe

Définition 23.13

Soit $X : \Omega \rightarrow E$ et $Y : \Omega \rightarrow F$ deux v.a. Alors on définit l'application notée

$$(X, Y) : \Omega \rightarrow E \times F$$

$$\omega \mapsto (X(\omega), Y(\omega))$$

L'application (X, Y) est appelée un couple de v.a.

Définition 23.14 (Loi conjointe)

Soit $X : \Omega \rightarrow E$ et $Y : \Omega \rightarrow F$ deux v.a. La loi conjointe du couple (X, Y) est la loi du couple (X, Y) . C'est donc l'application

$$\mathbb{P}_{(X, Y)} : \mathcal{P}(E \times F) \rightarrow [0, 1]$$

$$S \mapsto \mathbb{P}((X, Y) \in S)$$

Notation. On a donc $S \subset E \times F$ dans la définition ci-dessus. En général, on considère des parties S de la forme $S = A \times B$ avec $A \subset E$ et $B \subset F$. Dans ce cas, on note

$$\mathbb{P}(X \in A, Y \in B) := \mathbb{P}((X, Y) \in A \times B)$$

$$= \mathbb{P}(\{X \in A\} \cap \{Y \in B\})$$

et pour tous $x \in E$ et $y \in F$,

$$\mathbb{P}(X = x, Y = y) := \mathbb{P}(\{X = x\} \cap \{Y = y\})$$

La virgule prend donc ici la valeur d'un "et". Toutefois, on continuera d'écrire $\mathbb{P}(A \cap B \cap C)$ et non $\mathbb{P}(A, B, C)$.

Proposition 23.15

La loi conjointe est entièrement déterminée par la distribution de probabilités

$$(\mathbb{P}(X = x, Y = y))_{x \in E, y \in F}$$

On peut donc se contenter de s'intéresser uniquement aux probabilités de cette forme.

Exemple 8. On lance deux dés à 4 faces et on note X la valeur minimale, Y celle la valeur maximale. L'univers correspondant est $\Omega = \llbracket 1, 4 \rrbracket^2$ muni de la probabilité uniforme. Déterminer la loi conjointe de (X, Y) .

On peut synthétiser les résultats sous la forme d'un tableau :

$\mathbb{P}(X = \dots, Y = \dots)$	$Y = 1$	$Y = 2$	$Y = 3$	$Y = 4$
$X = 1$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$X = 2$	0	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$
$X = 3$	0	0	$\frac{1}{16}$	$\frac{1}{8}$
$X = 4$	0	0	0	$\frac{1}{16}$

3.2 Lois marginales

Définition 23.16

Pour tout couple (X, Y) de v.a., la loi de X et la loi de Y sont appelées des lois marginales du couple (X, Y) .

Ainsi pour un couple (X, Y) , on dispose de 3 lois :

1. La loi conjointe $\mathbb{P}_{(X,Y)}$ déterminée par les valeurs de $\mathbb{P}(X = x, Y = y)$.
2. La loi marginale \mathbb{P}_X déterminée par les valeurs de $\mathbb{P}(X = x)$.
3. La loi marginale \mathbb{P}_Y déterminée par les valeurs de $\mathbb{P}(Y = y)$.

À partir de la loi conjointe, on peut déterminer les lois marginales :

Proposition 23.17

Soit $X : \Omega \rightarrow E$ et $Y : \Omega \rightarrow F$ deux v.a. Alors

$$\forall x \in E \quad \mathbb{P}(X = x) = \sum_{y \in F} \mathbb{P}(X = x, Y = y)$$

Exemple 9. Déterminons les lois marginales de l'exemple précédent. Pour cela on reprend le tableau précédent : il suffit de faire la somme sur chaque ligne et chaque colonne pour avoir les lois marginales.

$\mathbb{P}(X = \dots, Y = \dots)$	$Y = 1$	$Y = 2$	$Y = 3$	$Y = 4$	$\mathbb{P}(X = \dots)$
$X = 1$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	
$X = 2$	0	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	
$X = 3$	0	0	$\frac{1}{16}$	$\frac{1}{8}$	
$X = 4$	0	0	0	$\frac{1}{16}$	
$\mathbb{P}(Y = \dots)$					Total : 1

Remarque. Comme on l’a dit, la connaissance de la loi conjointe suffit à déterminer les lois marginales. La réciproque est fautive : par exemple ci-dessus, si on ne connaît que les lois marginales, on a accès à la somme de chaque ligne et chaque colonne, i.e. 8 informations. Cela ne suffit pas à “reconstruire” les 16 valeurs du tableau qui correspondent à la loi conjointe !

4 Variables aléatoires indépendantes

4.1 Définition

Définition 23.18

Deux v.a. $X : \Omega \rightarrow E$ et $Y : \Omega \rightarrow F$ sont dites indépendantes si pour toutes parties $A \subset E$ et $B \subset F$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants. Autrement dit,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

On note alors $X \perp\!\!\!\perp Y$. Une formulation équivalente est que

$$\forall x \in E \quad \forall y \in F \quad \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

Autrement dit, lorsque les v.a. X, Y sont indépendantes, il suffit de connaître les lois marginales, à savoir les familles $(\mathbb{P}(X = x))_{x \in E}$ et $(\mathbb{P}(Y = y))_{y \in F}$ pour déterminer la loi conjointe.

Exemple 10. Les variables X, Y de l’exemple 9 ne sont pas indépendantes : en effet

$$\mathbb{P}(X = 4, Y = 1) = 0 \quad \text{et} \quad \mathbb{P}(X = 4)\mathbb{P}(Y = 1) = \frac{1}{16} \times \frac{1}{16} \neq 0$$

Exemple 11. On lance une pièce deux fois. Les v.a. X et Y qui correspondent au résultat de chaque lancer sont indépendantes.

Proposition 23.19

Soit $X : \Omega \rightarrow E$ et $Y : \Omega \rightarrow F$ deux v.a. indépendantes. Alors pour tout $x \in E$ et pour tout $y \in F$, si $\mathbb{P}(Y = y) > 0$, on a

$$\mathbb{P}(X = x \mid Y = y) = \mathbb{P}(X = x)$$

Proposition 23.20

Soit $X : \Omega \rightarrow E$ et $Y : \Omega \rightarrow F$ deux v.a. Soit f, g des fonctions définies sur E et F respectivement. Alors

$$X \perp\!\!\!\perp Y \implies f(X) \perp\!\!\!\perp g(Y)$$

4.2 Indépendance de n variables aléatoires

Définition 23.21

Soit X_1, \dots, X_n des v.a. à valeurs dans des ensembles E_1, \dots, E_n respectivement. Les v.a. X_1, \dots, X_n sont

dites (mutuellement) indépendantes si

$$\forall (x_1, \dots, x_n) \in E_1 \times \dots \times E_n \quad \mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{k=1}^n \mathbb{P}(X_k = x_k)$$

Une définition équivalente reviendrait à remplacer “ $X_k = x_k$ ” par “ $X_k \in A_k$ ”, où A_k est une partie quelconque de E_k .

Enfin, on peut également définir la notion de v.a. indépendantes deux à deux. Comme pour les événements, l'indépendance mutuelle entraîne l'indépendance deux à deux mais la réciproque est fautive.

Proposition 23.22

Soit X_1, \dots, X_n des v.a. à valeurs dans des ensembles E_1, \dots, E_n respectivement. Soit f_1, \dots, f_n des fonctions définies sur E_1, \dots, E_n respectivement. Si (X_1, \dots, X_n) sont (mutuellement) indépendantes, alors les v.a.

$$f_1(X_1), \dots, f_n(X_n)$$

sont également (mutuellement) indépendantes.

Exemple 12. Soit X, Y, Z, T quatre v.a.r. indépendantes. Alors

$$X^2, e^Y, \ln(1 + |Z|), \arctan(T)$$

sont des v.a.r. indépendantes

Lemme 23.23 (Lemme des coalitions)

Soit X_1, \dots, X_n des v.a. à valeurs dans des ensembles E_1, \dots, E_n respectivement. Soit $p_1, \dots, p_r \in \llbracket 1, n \rrbracket$ tels que $1 \leq p_1 < p_2 < \dots < p_r < n$.

- Soit f_1 une fonction définie sur $E_1 \times \dots \times E_{p_1}$, de sorte que $f_1(X_1, \dots, X_{p_1})$ ait un sens.
- Soit f_2 une fonction définie sur $E_{p_1+1} \times \dots \times E_{p_2}$, de sorte que $f_2(X_{p_1+1}, \dots, X_{p_2})$ ait un sens.
- ...
- Soit f_{r+1} une fonction définie sur $E_{p_r+1} \times \dots \times E_n$, de sorte que $f_{r+1}(X_{p_r+1}, \dots, X_n)$ ait un sens.

Alors les v.a.

$$f_1(X_1, \dots, X_{p_1}), \quad f_2(X_{p_1+1}, \dots, X_{p_2}), \quad \dots \quad f_{r+1}(X_{p_r+1}, \dots, X_n)$$

sont (mutuellement) indépendantes.

Exemple 13. Soit X, Y, Z, T quatre v.a.r. indépendantes. Alors

$$e^X Y^2 \quad \text{et} \quad \arctan(ZT) \quad \text{sont indépendantes}$$

$$\sqrt{|X|} \quad \text{et} \quad \cos(Y^Z) \quad \text{sont indépendantes}$$

4.3 Somme de n v.a. indépendantes de même loi $\mathcal{B}(p)$

Proposition 23.24

Soit X_1, \dots, X_n des v.a. (mutuellement) **indépendantes**. On suppose que X_1, \dots, X_n suivent toutes la même loi de Bernoulli de paramètre $p \in [0, 1]$, ce qu'on peut noter $X_1 \sim \dots \sim X_n \sim \mathcal{B}(p)$. Alors la v.a. $Y = X_1 + \dots + X_n$ suit une loi binomiale de paramètre (n, p) , i.e. $Y \sim \mathcal{B}(n, p)$.

Preuve "moche" mais compréhensible. Comme X_1, \dots, X_n sont à valeurs dans $\{0, 1\}$, la v.a. Y est à valeurs dans $\llbracket 0, n \rrbracket$. Soit $k \in \llbracket 0, n \rrbracket$, calculons $\mathbb{P}(Y = k)$. L'événement $\{Y = k\}$ est réalisé quand exactement k v.a. parmi X_1, \dots, X_n valent 1.

Supposons que les k variables qui valent 1 soient les k premières, à savoir X_1, \dots, X_k . Alors par indépendance,

$$\begin{aligned} \mathbb{P}(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0) &= \prod_{i=1}^k \mathbb{P}(X_i = 1) \prod_{i=k+1}^n \mathbb{P}(X_i = 0) \\ &= \prod_{i=1}^k p \prod_{i=k+1}^n (1-p) \\ &= p^k (1-p)^{n-k} \end{aligned}$$

On peut vérifier que peu importe le choix des k variables parmi X_1, \dots, X_n , on retombe sur la même valeur $p^k (1-p)^{n-k}$. De plus, les événements obtenus par deux choix différents sont incompatibles. Cela permet d'écrire

$$\begin{aligned} \mathbb{P}(Y = k) &= \underbrace{\mathbb{P}(\dots\dots\dots)}_{\text{un choix possible de } k \text{ v.a. mises à 1, les autres à 0}} + \underbrace{\mathbb{P}(\dots\dots\dots)}_{\text{un autre choix}} + \dots \\ &= p^k (1-p)^{n-k} + p^k (1-p)^{n-k} + \dots \end{aligned}$$

Enfin, il y a autant de choix possibles que de façons de choisir k v.a. parmi n , càd $\binom{n}{k}$. Ainsi,

$$\mathbb{P}(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Cela conclut la preuve. □

Exemple 14. Si on lance n fois une pièce, et qu'on compte le nombre de piles obtenus, alors la v.a. correspondante suit une loi binomiale $\mathcal{B}\left(n, \frac{1}{2}\right)$.