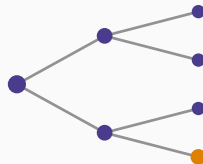


Probabilités

Chapitre 7 : Conditionnement et indépendance



Rétrécir l'univers

La règle des probabilités composées

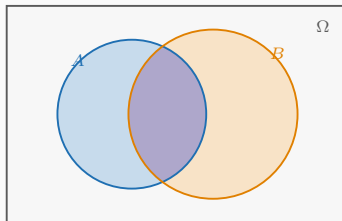
L'indépendance

Les probabilités totales

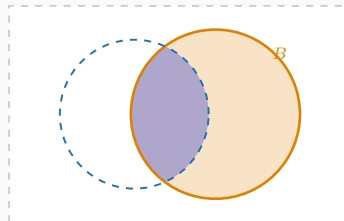
Ce qu'il faut retenir

Rétrécir l'univers

Ce que fait une information nouvelle



avant : l'univers est Ω



après : on *sait* que B est réalisé,
l'univers devient B

Conditionner, c'est **rétrécir l'univers**. La partie de A qui survit est $A \cap B$, et on la rapporte à la nouvelle taille :

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Probabilité conditionnelle

Soit B un événement de probabilité non nulle. La probabilité de A sachant B est

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

La formule dit ce que fait la figure

Apprendre que B est réalisé **remplace l'univers** : B devient le nouveau tout.

Il faut donc, d'une part, ne garder de A que ce qui survit dans B — c'est $A \cap B$ — et d'autre part rapporter cette part à la nouvelle taille $P(B)$, pour que le total revienne à 1.

C'est la distribution conditionnelle du cours de statistique

Dans un tableau de contingence, on lisait le profil-ligne en divisant chaque case par le total de sa ligne :

$$f_{j|i} = \frac{n_{ij}}{n_{i\bullet}}$$

Remplacez les effectifs par des probabilités : vous obtenez $P(A | B) = P(A \cap B)/P(B)$. La même opération, sous un autre nom.

Sur le dé

On lance un dé et l'on apprend que le résultat est pair. Quelle est la probabilité qu'il soit supérieur à 3?

$$P(> 3 \mid \text{pair}) = \frac{P(\{4, 6\})}{P(\{2, 4, 6\})} = \frac{2/6}{3/6} = \frac{2}{3}$$

Sans l'information, la réponse était $P(\{4, 5, 6\}) = 1/2$. L'information a modifié la réponse.

La règle des probabilités composées

Probabilités composées

$$P(A \cap B) = P(B) P(A \mid B) = P(A) P(B \mid A)$$

Pour trois événements :

$$P(A \cap B \cap C) = P(A) P(B \mid A) P(C \mid A \cap B)$$

C'est la formule de l'arbre

Un arbre se lit en **multipliant le long d'un chemin** : la probabilité de la première branche, puis celle de la deuxième *sachant* la première, et ainsi de suite.

La règle des probabilités composées n'est que l'écriture algébrique de ce geste.

Deux cartes défectueuses

Un lot de 20 pièces contient 4 défectueuses. On en tire deux successivement, sans remise.

$$P(\text{deux défectueuses}) = \underbrace{\frac{4}{20}}_{\text{première}} \times \underbrace{\frac{3}{19}}_{\text{seconde, sachant la première}} = \frac{12}{380} \approx 0,032$$

Le sans-remise est un conditionnement

La deuxième fraction n'est pas $4/20$: le stock a changé. C'est très exactement ce que dit $P(B \mid A)$.

Les chapitres 4 et 5 comptaient ces situations; ce chapitre les calcule directement, sans passer par le dénombrement.

L'indépendance

Événements indépendants

A et B sont **indépendants** lorsque l'information n'apporte rien :

$$P(A \mid B) = P(A) \quad \Longleftrightarrow \quad \boxed{P(A \cap B) = P(A) \times P(B)}$$

La seconde écriture est celle qu'il faut utiliser

Elle est **symétrique** — elle ne privilégie ni A ni B — et elle reste valable même si $P(B) = 0$.

C'est aussi elle qui se généralise : n événements sont indépendants si la probabilité de toute intersection est le produit des probabilités.

Le piège du chapitre 2, réglé

Incompatible et indépendant sont contraires

	$A \cap B$	Relation
Incompatibles	$\emptyset$	l'un exclut l'autre
Indépendants	$P(A)P(B)$	l'un n'apprend rien sur l'autre

Deux événements incompatibles ne sont jamais indépendants

Si A et B sont incompatibles et de probabilités non nulles :

$$P(A \mid B) = \frac{P(\emptyset)}{P(B)} = 0 \neq P(A)$$

Apprendre que B s'est produit a tout appris sur A : il ne s'est pas produit. C'est le contraire même de l'indépendance.

Elle se postule, elle ne se vérifie pas

Dans un exercice, l'indépendance est presque toujours **donnée par l'énoncé** : lancers successifs, crédits d'emprunteurs sans lien, pannes de machines distinctes.

Dans la réalité, c'est une **hypothèse de modélisation** — souvent commode, parfois fausse, et toujours à interroger.

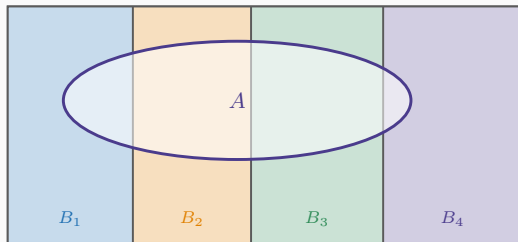
Ce que l'indépendance a coûté en 2008

Les modèles de risque de crédit supposaient les défauts largement indépendants. Cette hypothèse rendait les portefeuilles très sûrs : $0,04^{10}$ est un nombre minuscule.

Mais une récession frappe tous les emprunteurs **en même temps**. Les défauts sont corrélés, et le risque réel était sans commune mesure avec le risque calculé. **Toute l'erreur tenait dans une hypothèse d'indépendance.**

Les probabilités totales

Découper pour calculer



Les B_j forment un **système complet** : ils ne se chevauchent pas et recouvrent tout Ω .

L'événement A est alors découpé en morceaux, un par B_j :

$$P(A) = \sum_j P(B_j) P(A \mid B_j)$$

C'est la formule des probabilités totales : on calcule $P(A)$ cas par cas, puis on rassemble.

Formule des probabilités totales

Si $B_1, \dots, B_k$ forment un système complet, alors pour tout événement A :

$$P(A) = \sum_{j=1}^k P(B_j) P(A \mid B_j)$$

Cas particulier le plus fréquent : $P(A) = P(B)P(A \mid B) + P(\overline{B})P(A \mid \overline{B})$.

C'est une moyenne pondérée

$P(A)$ est la **moyenne des probabilités conditionnelles**, pondérée par le poids de chaque cas.

Exactement la structure de $\bar{x} = \sum f_i x_i$ — la moyenne générale d'une population découpée en sous-populations. Encore une fois, la même formule.

Trois fournisseurs

Un distributeur s'approvisionne à 50 % chez F_1 , 30 % chez F_2 , 20 % chez F_3 . Les taux de défaut respectifs sont 2 %, 5 % et 8 %.

$$\begin{aligned}P(\text{défaut}) &= 0,50(0,02) + 0,30(0,05) + 0,20(0,08) \\ &= 0,010 + 0,015 + 0,016 = 0,041\end{aligned}$$

Le taux global est de 4,1 %.

Ce que le chiffre global cache

F_3 ne représente que 20 % des volumes, mais fournit $0,016/0,041 = 39\%$ des pièces défectueuses.

Le taux global de 4,1 % ne dit rien de cela. **C'est la décomposition, non le total, qui indique où agir.**

Ce qu'il faut retenir

Six points

1. Conditionner, c'est **rétrécir l'univers** : $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$.
2. C'est la **distribution conditionnelle** de la statistique descriptive, transposée.
3. **Probabilités composées** : $P(A \cap B) = P(B)P(A \mid B)$ – la règle de lecture d'un arbre.
4. **Indépendance** : $P(A \cap B) = P(A)P(B)$; elle se postule et doit toujours être interrogée.
5. Incompatible **n'est pas** indépendant : c'est le contraire.
6. **Probabilités totales** : $P(A) = \sum_j P(B_j)P(A \mid B_j)$, une moyenne pondérée sur les cas.

La suite

Nous savons calculer $P(\text{alerte} \mid \text{défaut})$ – la performance d'un test. Mais un banquier reçoit l'alerte et veut $P(\text{défaut} \mid \text{alerte})$.

Ces deux nombres n'ont rien à voir. Le chapitre 8 explique comment passer de l'un à l'autre – et pourquoi les confondre est l'erreur de raisonnement la plus répandue.