

Probabilidad Condicional, Independencia y Bayes

Clase 2 — 14.310x: Data Analysis for Social Scientists

B.Sc. Cristian Lazo Quispe

BREIT (Brescia Institute of Technology) · Aporta — Grupo Breca

Programa SDS (Statistics and Data Science, MITx) · **BREIT** (Brescia Institute of Technology), iniciativa de Aporta —Grupo Breca— en alianza con MIT IDSS.

Basado en MIT 14.310x (CC BY-NC-SA 4.0).

✉ clazoq@uni.pe · [in linkedin.com/in/cristian-lazo-quispe](https://www.linkedin.com/in/cristian-lazo-quispe) · github.com/CristianLazoQuispe

¿Qué veremos hoy?

- 1 La idea: condicionar es reescalar el mundo
- 2 Independencia
- 3 Ley de probabilidad total
- 4 Regla de Bayes
- 5 Diagramas de árbol
- 6 Manos a la obra

Contenido

- 1 La idea: condicionar es reescalar el mundo
- 2 Independencia
- 3 Ley de probabilidad total
- 4 Regla de Bayes
- 5 Diagramas de árbol
- 6 Manos a la obra

¿Cómo cambia una probabilidad con información nueva?

La pregunta de hoy

¿Cómo cambia la probabilidad de un evento cuando **sabemos** que ocurrió otro? Esa es la **probabilidad condicional (conditional probability)**.

- Al saber que ocurrió B , nuestro mundo ya no es S : ahora es B .
- Condicionar es **reescalar**: medir qué fracción de ese mundo más pequeño corresponde a A .
- De aquí salen las herramientas de toda la inferencia: independencia, ley de probabilidad total y regla de Bayes.

Definición y regla del producto

Definición 1 (Probabilidad condicional)

La probabilidad de A **dado** B (conditional on B), con $\mathbb{P}(B) > 0$:

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

Reordenando se obtiene la **regla del producto (multiplication rule)**:

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

Demo: reescalar el espacio

Problema. En una encuesta, $\mathbb{P}(\text{formal y con seguro}) = 0.35$ y $\mathbb{P}(\text{formal}) = 0.50$. Entre los trabajadores formales, ¿qué fracción tiene seguro?

Solución. $\mathbb{P}(\text{seguro} | \text{formal}) = \frac{0.35}{0.50} = 0.70$.

Contenido

- 1 La idea: condicionar es reescalar el mundo
- 2 Independencia**
- 3 Ley de probabilidad total
- 4 Regla de Bayes
- 5 Diagramas de árbol
- 6 Manos a la obra

Cuando saber B no cambia nada

Definición 2 (Independencia)

A y B son *independientes* (*independent*) si

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

Equivalentemente (si $\mathbb{P}(B) > 0$): $\mathbb{P}(A | B) = \mathbb{P}(A)$. Saber que ocurrió B no cambia la probabilidad de A .

Independiente $\neq$ mutuamente excluyente

Son ideas *opuestas*, no sinónimos. Si A y B son disjuntos y ambos tienen probabilidad positiva, saber que ocurrió B *garantiza* que A no ocurrió: ¡eso es máxima dependencia! Disjuntos con prob. positiva nunca son independientes.

Demo: ¿independencia en una tabla de contingencia?

Problema

Se encuestó a 200 personas sobre si **participan** en un programa social y si **conocen sus derechos**. ¿Son independientes “participar” y “conocer”?

	Conoce	No	Total
Participa	70	30	100
No particip.	40	60	100
Total	110	90	200

Solución (regla del producto)

$$\mathbb{P}(\text{particip.}) = \frac{100}{200} = 0.5, \quad \mathbb{P}(\text{conoce}) = \frac{110}{200} = 0.55.$$

Si fueran indep.: $\mathbb{P}(\text{ambos}) = 0.5 \times 0.55 = 0.275$ (55 pers.).

Pero la celda real es $\frac{70}{200} = 0.35 \neq 0.275$.

No son independientes: participar se asocia con conocer más los derechos.

Contenido

- 1 La idea: condicionar es reescalar el mundo
- 2 Independencia
- 3 Ley de probabilidad total**
- 4 Regla de Bayes
- 5 Diagramas de árbol
- 6 Manos a la obra

Divide y vencerás para probabilidades

Ley de probabilidad total (law of total probability)

Si $B_1, \dots, B_k$ *particionan* el espacio (disjuntos y exhaustivos, cada uno con prob. positiva), entonces para cualquier evento A :

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

Intuición: promediamos A sobre escenarios disjuntos, *ponderando* por la probabilidad de cada uno.

Demo: dos oficinas

Problema. La oficina 1 procesa el 60 % de los trámites (error 5 %) y la oficina 2 el 40 % restante (error 10 %). ¿Probabilidad de que un trámite tenga error?

Solución. $\mathbb{P}(\text{error}) = 0.05 \cdot 0.60 + 0.10 \cdot 0.40 = 0.07$.

Contenido

- 1 La idea: condicionar es reescalar el mundo
- 2 Independencia
- 3 Ley de probabilidad total
- 4 Regla de Bayes**
- 5 Diagramas de árbol
- 6 Manos a la obra

Invertir el condicionamiento

Regla de Bayes (Bayes' rule)

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

Bayes convierte “probabilidad del síntoma dado el estado” en “probabilidad del estado dado el síntoma”. Partimos de una creencia previa (*prior*) $\mathbb{P}(B)$ y la actualizamos a una posterior (*posterior*) $\mathbb{P}(B | A)$ tras observar A .

$$\text{prior } \mathbb{P}(B) \xrightarrow{\text{observamos } A} \text{posterior } \mathbb{P}(B | A)$$

Bayes: con partición y con evidencia adicional

Regla de Bayes (con una partición $\{A_1, A_2, \dots\}$)

$$\mathbb{P}(A_i | B) = \frac{\mathbb{P}(A_i) \mathbb{P}(B | A_i)}{\sum_j \mathbb{P}(A_j) \mathbb{P}(B | A_j)}$$

Bayes con evidencia adicional C

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

Es el *mismo* Bayes, pero **dentro del mundo donde ya sabemos C** : todas las probabilidades se condicionan además a C .

Demo: la prueba de tamizaje y la tasa base

Problema

Una enfermedad afecta al 1 % de la población. Un test detecta correctamente al enfermo el 90 % de las veces (sensibilidad) y da falso positivo en el 8 % de los sanos. Si una persona da positivo, ¿cuál es la probabilidad de que esté **realmente enferma**?

Solución. Sea E enfermo y $+$ positivo. Por la regla de Bayes:

$$\mathbb{P}(E | +) = \frac{\mathbb{P}(+ | E) \mathbb{P}(E)}{\mathbb{P}(+)} = \frac{0.90 \cdot 0.01}{0.90 \cdot 0.01 + 0.08 \cdot 0.99} = \frac{0.009}{0.0882} \approx 0.102.$$

Cuidado con la tasa base (base rate)

Aunque el test “acierta” mucho, ¡solo el 10 % de los positivos está enfermo! La baja prevalencia (1 %) domina el resultado.

Contenido

- 1 La idea: condicionar es reescalar el mundo
- 2 Independencia
- 3 Ley de probabilidad total
- 4 Regla de Bayes
- 5 Diagramas de árbol**
- 6 Manos a la obra

El árbol: multiplicar por camino, sumar entre caminos

Un **diagrama de árbol (tree diagram)** organiza el experimento por etapas: cada rama lleva una probabilidad condicional, y un camino completo se calcula *multiplicando* sus ramas. Sumar caminos que terminan en el mismo evento es la ley de probabilidad total.

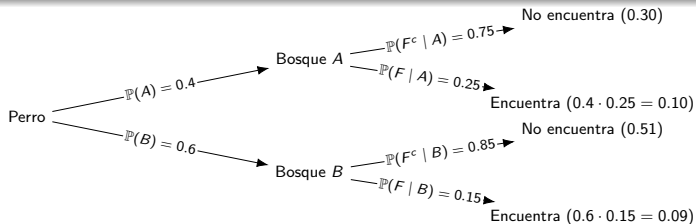
Demo: el perro perdido de Oscar

El perro está en el bosque A con prob. 0.4 o en el B con prob. 0.6. Buscando *un día* en el bosque correcto, se encuentra con prob. 0.25 si está en A y 0.15 si está en B (en el bosque equivocado, 0). ¿Dónde conviene buscar el primer día? ¿Y si busca en A y *no* lo encuentra?

Árbol de probabilidad: el perro de Oscar

Problema

El perro está en el bosque A con $\mathbb{P}(A) = 0.4$ o en el B con $\mathbb{P}(B) = 0.6$. Sea F el evento “lo encuentra” al buscar *un día* en el bosque correcto: $\mathbb{P}(F \mid A) = 0.25$ y $\mathbb{P}(F \mid B) = 0.15$. Organiza el experimento en un árbol.



Cada rama lleva una probabilidad *condicional*; multiplica a lo largo de un camino.

La conclusión del árbol

(a) ¿Dónde buscar el primer día?

Buscar en A : lo encuentra con prob. $0.4 \times 0.25 = 0.10$. Buscar en B : $0.6 \times 0.15 = 0.09$.
Como $0.10 > 0.09$, conviene buscar en el **bosque** A .

(b) Buscó en A y no lo encontró: ¿sigue estando en A ?

Por Bayes (no encontrarlo buscando en A tiene prob. 0.75 si está en A , y 1 si está en B):

$$\mathbb{P}(A \mid \text{no enc. en } A) = \frac{0.4 \times 0.75}{0.4 \times 0.75 + 0.6 \times 1} = \frac{0.30}{0.90} = \frac{1}{3}.$$

La búsqueda fallida *baja* la creencia de que esté en A : de 0.4 a $1/3$.

Contenido

- 1 La idea: condicionar es reescalar el mundo
- 2 Independencia
- 3 Ley de probabilidad total
- 4 Regla de Bayes
- 5 Diagramas de árbol
- 6 Manos a la obra**

¿Y ahora?

- Revisa el **PDF de teoría** para el detalle y más ejemplos.
- Practica con el **PDF de ejercicios** (3 niveles: Basic · Core · Challenge).
- Verifica con el **PDF de soluciones** paso a paso.
- Recuerda: **condicional** reescala el mundo, el **árbol** multiplica por camino y suma entre caminos, y **Bayes** convierte prior en posterior.

¡A practicar!