Modelos de probabilidad y axiomas

Clase 2 (parte 2 de 2): modelos discretos, continuos y aditividad contable

6.431x — Probability: The Science of Uncertainty and Data

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 6.431x / MIT OCW RES.6-012 (CC BY-NC-SA 4.0).

clazoq@uni.pe · linkedin.com/in/cristian-lazo-quispe · github.com/CristianLazoQuispe

Dónde quedamos (Clase 1)

La receta, en una línea

Un **modelo de probabilidad** = un espacio muestral Ω + una **ley de probabilidad** $\mathbb{P}(\cdot)$ que asigna un número a cada evento $A \subseteq \Omega$.

Los tres axiomas

1. **No negatividad**: $\mathbb{P}(A) \geq 0$.
2. **Normalización**: $\mathbb{P}(\Omega) = 1$.
3. **Aditividad**: si $A \cap B = \varnothing$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Lo que ya dedujimos de ellos

- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, $\mathbb{P}(\varnothing) = 0$, $\mathbb{P}(A) \leq 1$.
- Monotonía: $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.
- Inclusión–exclusión: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
- Cota de Boole: $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$.
- De Morgan para "al menos uno": $(A_1 \cup \cdots \cup A_n)^c = A_1^c \cap \cdots \cap A_n^c$.

Lo que falta

Todo eso son *reglas*. Hoy toca **construir leyes concretas** —discretas y continuas— y ver hasta dónde llega el tercer axioma.

¿Qué veremos hoy? (Unit 1 — Lecture 1, parte 2 de 2)

1. Modelos discretos
2. Modelos continuos
3. Aditividad contable y sutilezas
4. Interpretaciones de la probabilidad
5. Cierre

Contenido

1. Modelos discretos
2. Modelos continuos
3. Aditividad contable y sutilezas
4. Interpretaciones de la probabilidad
5. Cierre

Modelo discreto: sumar los pedacitos

Ley discreta

Si Ω es finito o numerable y $A = \{s_1, \ldots, s_k\}$, la aditividad da

$$\mathbb{P}(A) = \mathbb{P}(s_1) + \mathbb{P}(s_2) + \cdots + \mathbb{P}(s_k).$$

Basta conocer la probabilidad de **cada resultado individual**.

Caso uniforme (*discrete uniform law*)

Si Ω tiene n resultados **igualmente probables**:

$$\mathbb{P}(A) = \frac{\text{número de elementos de } A}{n}.$$

Aquí —y **solo** aquí— calcular probabilidades es **contar**.

Cuidado

"Igualmente probable" es una **hipótesis del modelo**, no una ley de la naturaleza. Si el dado está cargado, la fórmula de contar **no** aplica.

Demo: Ω discreto pero *infinito*

Problema

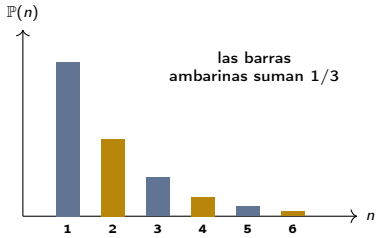
Lanzo una moneda justa hasta que salga la primera cara; n es el número de lanzamientos. Entonces $\Omega = \{1, 2, 3, \dots\}$ y $\mathbb{P}(n) = 1/2^n$. ¿Cuánto vale $\mathbb{P}(\{2, 4, 6, \dots\})$, es decir, "$n$ es par"?

Solución

Los eventos $\{2\}, \{4\}, \dots$ son disjuntos, así que se suman (aditividad contable):

$$\mathbb{P}(\text{par}) = \tfrac{1}{2^2} + \tfrac{1}{2^4} + \tfrac{1}{2^6} + \cdots$$
$$= \tfrac{1}{4}\left(1 + \tfrac{1}{4} + \tfrac{1}{4^2} + \cdots\right) = \tfrac{1}{4} \cdot \frac{1}{1 - \frac{1}{4}} = \tfrac{1}{3}.$$

Y el modelo es legítimo porque la masa total es 1: $\sum_{n \geq 1} 2^{-n} = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{2}} = 1$. ✓



Masa geométrica: cada barra es la mitad de la anterior.

Contenido

1. Modelos discretos
2. Modelos continuos
3. Aditividad contable y sutilezas
4. Interpretaciones de la probabilidad
5. Cierre

Modelo continuo: la probabilidad es *área*

Problema

Se elige un punto (x, y) al azar y "uniformemente" en el cuadrado unitario $\Omega = [0,1] \times [0,1]$: la probabilidad de un subconjunto es su **área**. Calcula $\mathbb{P}(x + y \leq 1/2)$.

Solución

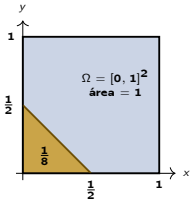
La región es el triángulo de vértices $(0,0)$, $(1/2,0)$, $(0,1/2)$:

$$\mathbb{P} = \text{área} = \tfrac{1}{2} \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{8} = 0.125.$$

Chequeo de normalización: el área total del cuadrado es 1. ✓

Lo raro (y clave)

Cada **punto** individual tiene área 0, o sea $\mathbb{P}\big((x,y)\big) = 0 \ldots$ ¡y sin embargo *algún* punto sale! **Probabilidad 0 no significa imposible.** Por eso en el mundo continuo ya no se puede "sumar resultado por resultado".



Contenido

1. Modelos discretos
2. Modelos continuos
3. Aditividad contable y sutilezas
4. Interpretaciones de la probabilidad
5. Cierre

El tercer axioma, en su versión fuerte

Aditividad contable (*countable additivity*)

Si $A_1, A_2, A_3, \ldots$ es una **sucesión** (infinita pero **numerable**) de eventos **disjuntos** dos a dos, entonces

$$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \cdots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) + \cdots$$

- Es lo que nos permitió sumar la serie geométrica de la moneda.
- "Numerable" significa que se pueden **enumerar**: 1.°, 2.°, 3.°, ...
- Para uniones **no numerables** el axioma **no dice nada** —y de hecho falla.

Demo: por qué no vale para uniones no numerables

Problema

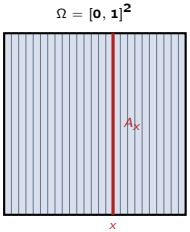
Sea Ω el cuadrado unitario y, para cada real $x \in [0,1]$, sea $A_x = \{(x,y) : y \in [0,1]\}$ la **recta vertical** que pasa por x. Esas rectas son disjuntas y su unión es todo Ω. ¿Se pueden sumar?

Solución

Cada recta tiene área cero: $\mathbb{P}(A_x) = 0$ para todo x. Pero

$$\mathbb{P}\Big(\bigcup_{x\in[0,1]} A_x\Big) = \mathbb{P}(\Omega) = 1 \;\neq\; \sum_{x} \mathbb{P}(A_x) = 0.$$

No hay contradicción: hay **tantas** rectas que no se pueden poner en una lista $1.^\circ, 2.^\circ, 3.^\circ, \ldots$ ($[0,1]$ no es numerable), y el axioma **solo se aplica a sucesiones**.



Cada recta pesa 0; "todas juntas" pesan 1. La unión es *no numerable*.

Sutilezas matemáticas (para dormir tranquilo)

- En un Ω continuo **no** se le puede asignar probabilidad a *cualquier* subconjunto: existen conjuntos patológicos (no medibles) para los que "área" no tiene sentido.
- La salida formal: la ley de probabilidad se define solo sobre una familia bien portada de eventos (σ-álgebra), que incluye todo lo que uno escribe en la práctica (intervalos, rectángulos, uniones e intersecciones numerables...).
- **Para este curso**: no vas a tropezar con esos conjuntos. Basta saber que la teoría está bien fundada y que el ingrediente delicado es la **numerabilidad**.

Qué llevarte

Si tus eventos se construyen con uniones, intersecciones y complementos **numerables** de cosas razonables, todo funciona.

Contenido

1. Modelos discretos
2. Modelos continuos
3. Aditividad contable y sutilezas
4. Interpretaciones de la probabilidad
5. Cierre

¿Qué *significa* un número de probabilidad?

Frecuentista (*frequentist*)

$\mathbb{P}(A)$ es la **fracción de veces** que ocurre A si repites el experimento muchísimas veces. Natural para dados, control de calidad, encuestas.

Grado de creencia (*belief / bayesiana*)

$\mathbb{P}(A)$ mide **cuánta confianza** tienes en A, y se actualiza con datos. Necesaria cuando el evento **no se puede repetir**.

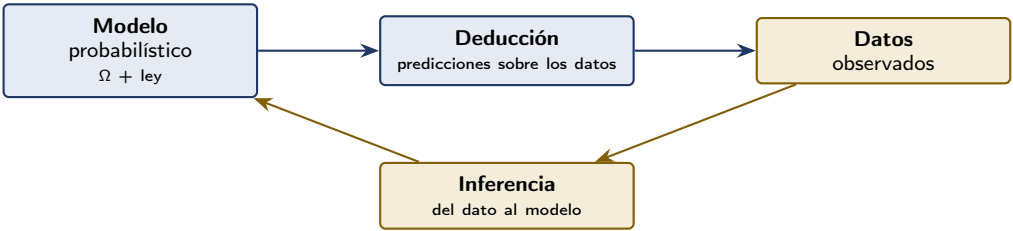
Demo: dos frases, dos lecturas

- "La moneda sale cara con probabilidad 1/2": en 10 000 lanzamientos, cerca de la mitad son caras. *Frecuentista*.
- "Hay 70 por ciento de probabilidad de que llueva mañana": mañana ocurre **una sola vez**, no se puede repetir. Es un **grado de creencia** calibrado con datos históricos. *Bayesiana*.

Lo importante

Las dos escuelas usan **los mismos tres axiomas**. Discrepan en la *interpretación*, no en las reglas de cálculo.

El rol de la probabilidad en todo el programa



- **Hacia la derecha** (Units 1–6): dado un modelo, ¿qué datos esperar?
- **Hacia la izquierda** (Units 7 en adelante): dados los datos, ¿qué modelo es creíble? Eso es **inferencia**, y es a donde va el curso.

Contenido

1. Modelos discretos
2. Modelos continuos
3. Aditividad contable y sutilezas
4. Interpretaciones de la probabilidad
5. Cierre

Kahoot de repaso (acumulativo: Clases 1 y 2)

Jugamos ahora

20 preguntas que cubren **toda la Unit 1 / Lecture 1**: conjuntos y De Morgan, espacio muestral, axiomas, propiedades derivadas, modelos discretos y continuos, aditividad contable e interpretaciones.

- Cada pregunta tiene una **trampa**: el error clásico está entre las opciones.
- No corras: primero decide **si los eventos son disjuntos**, después calcula.
- En los modelos continuos, **dibuja la región** antes de integrar o contar área.

Resumen de las Clases 1 y 2

Lo que hay que saber hacer

1. Construir Ω: mutuamente excluyente, exhaustivo, con la granularidad que pide la pregunta.
2. Escribir el evento como conjunto (y usar De Morgan para "al menos uno").
3. Verificar los tres axiomas: $\mathbb{P} \geq 0$, $\mathbb{P}(\Omega) = 1$, aditividad **si son disjuntos**.
4. Deducir: $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, monotonía, inclusión–exclusión, cota de Boole.
5. Discreto: sumar resultados (series geométricas incluidas). Continuo: calcular áreas.
6. Saber que la aditividad contable exige una **sucesión**, no un continuo de eventos.
7. Leer un número de probabilidad en sus dos sentidos: frecuencia relativa y grado de creencia.

Siguiente clase

Unit 2 — **Conditioning and independence**: probabilidad condicional, regla de Bayes e independencia. Todo se apoyará en los axiomas de estas dos clases.

¡Gracias!

Clase 2 — Modelos de probabilidad y axiomas, parte 2 (6.431x, Unit 1)

clazoq@uni.pe · linkedin.com/in/cristian-lazo-quispe · github.com/CristianLazoQuispe

BREIT (Brescia Institute of Technology) · Aporta — Grupo Breca · en alianza con MIT IDSS

Basado en MIT 6.431x / MIT OCW RES.6-012 (Bertsekas & Tsitsiklis), CC BY-NC-SA 4.0.