Modelos de probabilidad y axiomas

Clase 1 (parte 1 de 2): conjuntos, espacio muestral y los axiomas

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

La plataforma oficial: MIT Learn

Dónde vive el curso

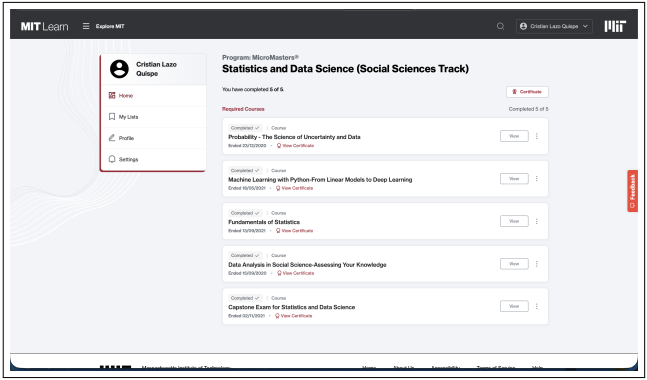
Todo el material oficial —videos, ejercicios, exámenes y certificado— está en **MIT Learn**:

`learn.mit.edu/dashboard`

- Al entrar verás el programa **MicroMasters® en Statistics and Data Science**.
- Nuestro curso es el primero de la ruta: **Probability — The Science of Uncertainty and Data** (6.431x).
- El *dashboard* lleva tu avance curso por curso y libera el certificado al aprobar.

Cómo encajan estas clases

Estas sesiones son de **repaso y práctica**: refuerzan lo de MIT Learn, no lo reemplazan. Ve el video allá y trae tus dudas acá.



Dashboard de MIT Learn: los cursos del MicroMasters SDS.

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

1. Repaso de conjuntos (sets)
2. Espacio muestral y eventos
3. Los tres axiomas
4. Propiedades que se deducen
5. Cierre

La receta de todo el curso

Un *modelo de probabilidad* tiene exactamente dos piezas

1. El **espacio muestral** Ω (*sample space*): la lista de todo lo que puede pasar.
2. Una **ley de probabilidad** (*probability law*): un número $\mathbb{P}(A)$ para cada evento A.

- Los **axiomas** dicen qué reglas debe cumplir ese número. Son solo **tres**.
- Todo lo demás del curso —condicional, variables aleatorias, TLC, inferencia— sale de ahí.
- Hoy no hay fórmulas que memorizar: hay que aprender a **modelar**.

La idea en una frase

Probabilidad no es adivinar el futuro: es **repartir una unidad de creencia** (el 1) entre todos los resultados posibles, respetando tres reglas.

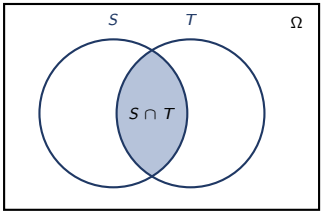
Contenido

1. Repaso de conjuntos (sets)
2. Espacio muestral y eventos
3. Los tres axiomas
4. Propiedades que se deducen
5. Cierre

El lenguaje: conjuntos

- Un conjunto es una colección de elementos distintos: $x \in S$.
- **Unión** $S \cup T$: los que están en S, en T, o en ambos.
- **Intersección** $S \cap T$: los que están en ambos.
- **Complemento** S^c: los que *no* están en S (dentro de Ω).
- **Disjuntos**: $S \cap T = \varnothing$ (no se pisan).

Con índices: $x \in \bigcup_n S_n$ si $x \in S_n$ para **algún** n;
$x \in \bigcap_n S_n$ si $x \in S_n$ para **todo** n.



Las dos leyes que más vas a usar

Distributiva

$$S \cap (T \cup U) = (S \cap T) \cup (S \cap U) \qquad S \cup (T \cap U) = (S \cup T) \cap (S \cup U)$$

De Morgan

$$\Big(\bigcup_n S_n\Big)^c = \bigcap_n S_n^c \qquad \Big(\bigcap_n S_n\Big)^c = \bigcup_n S_n^c$$

Cómo leerla en castellano (esto es lo que hay que recordar)

- "**No** ocurre **ninguno**" = "**todos** los complementos ocurren".
- "**No** ocurren **todos**" = "**alguno** falla".

Traducir el enunciado a esta forma suele ser *la* dificultad del problema; la cuenta viene después.

Demo: “al menos uno” se ataca por el complemento

Problema

Un sistema tiene 3 componentes. Sea A_i el evento “el componente i **falla**”. Escribe con conjuntos: (a) “falla al menos uno”; (b) “no falla ninguno”; (c) “falla exactamente uno”.

Solución

(a) Al menos uno $= A_1 \cup A_2 \cup A_3$.

(b) Ninguno $= A_1^c \cap A_2^c \cap A_3^c$. Por De Morgan es justo $(A_1 \cup A_2 \cup A_3)^c$: el complemento de (a). **Por eso** $\mathbb{P}(\text{al menos uno}) = 1 - \mathbb{P}(\text{ninguno})$.

(c) Exactamente uno $= (A_1 \cap A_2^c \cap A_3^c) \cup (A_1^c \cap A_2 \cap A_3^c) \cup (A_1^c \cap A_2^c \cap A_3)$, y los tres pedazos son **disjuntos** (no pueden pasar a la vez).

Regla práctica

Si el enunciado dice “al menos”, casi siempre conviene calcular el **complemento**.

Contenido

1. Repaso de conjuntos (sets)
2. Espacio muestral y eventos
3. Los tres axiomas
4. Propiedades que se deducen
5. Cierre

Ω: la lista de todo lo que puede pasar

Definición

El **espacio muestral** Ω es el conjunto de todos los resultados posibles de un experimento. Un **evento** (*event*) es un **subconjunto** de Ω.

Para que Ω sea un modelo válido, la lista debe ser:

- **Mutuamente excluyente** (*mutually exclusive*): ocurre **uno solo** de los elementos.
- **Colectivamente exhaustiva** (*collectively exhaustive*): siempre ocurre **alguno**.
- Con la **granularidad correcta** (*right granularity*): ni tan fina que sea inmanejable, ni tan gruesa que borre lo que te interesa.

Granularidad: el arte del modelado

Si lanzo una moneda y quiero estudiar la moneda, $\Omega = \{C, S\}$ basta. Si registro "cara y llueve", "cara y no llueve", ... el modelo sigue siendo válido, pero cargo detalle irrelevante. Si en cambio *sí* me interesa la lluvia, el modelo grueso ya **no sirve**: la granularidad la fija **la pregunta**.

Demo: modelar dos lanzamientos de un dado de 4 caras

Problema

Lanzo dos veces un dado de 4 caras. Propón Ω y expresa como evento "el **menor** de los dos resultados es 2".

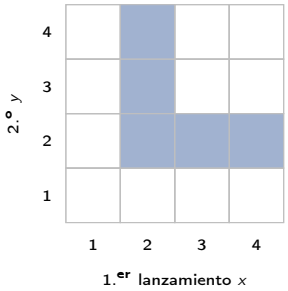
Solución

$\Omega = \{(x, y) : x, y \in \{1, 2, 3, 4\}\}$, con $|\Omega| = 4 \times 4 = 16$ pares. Es mutuamente excluyente (sale *un* par) y exhaustivo (siempre sale alguno).

El evento es

$$M_2 = \{(2,2), (2,3), (2,4), (3,2), (4,2)\}, \quad |M_2| = 5.$$

Si el dado es justo, cada par vale $1/16$, así que $\mathbb{P}(M_2) = 5/16$.



Los 16 resultados; sombreado, el evento $\min(x, y) = 2$.

Contenido

1. Repaso de conjuntos (sets)
2. Espacio muestral y eventos
3. Los tres axiomas
4. Propiedades que se deducen
5. Cierre

Los axiomas de probabilidad

Una ley de probabilidad asigna $\mathbb{P}(A)$ a cada evento $A \subseteq \Omega$ cumpliendo

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

Intuición: la probabilidad es *masa*

Imagina 1 kilo de arena repartido sobre Ω. $\mathbb{P}(A)$ es cuánta arena cae dentro de A.

- No negatividad: no existe arena negativa.
- Normalización: el total es exactamente 1 kilo.
- Aditividad: si dos montones **no se pisan**, el peso junto es la suma de los pesos.

El error clásico

La aditividad **exige** que sean disjuntos. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ es **falso** si A y B se solapan: estarías contando dos veces la parte común.

Demo: ¿es una ley de probabilidad válida?

Problema

Sobre $\Omega = \{a, b, c\}$ se proponen tres leyes. ¿Cuáles son válidas?

(i) $(0.2,\ 0.5,\ 0.3)$ (ii) $(0.5,\ 0.6,\ -0.1)$ (iii) $(0.2,\ 0.3,\ 0.4)$

Solución: se chequea axioma por axioma

- **(i)** Todas ≥ 0 ✓ y $0.2 + 0.5 + 0.3 = 1$ ✓ $\Rightarrow$ **válida**.
- **(ii)** Suma $= 1$ ✓ pero $\mathbb{P}(c) = -0.1 < 0 \Rightarrow$ viola **no negatividad**.
- **(iii)** Todas ≥ 0 ✓ pero suman $0.9 \neq 1 \Rightarrow$ viola **normalización** (falta masa: hay escenarios sin contar).

Con la ley (i): $\mathbb{P}(\{a, c\}) = \mathbb{P}(a) + \mathbb{P}(c) = 0.2 + 0.3 = 0.5$, usando aditividad porque $\{a\}$ y $\{c\}$ son disjuntos.

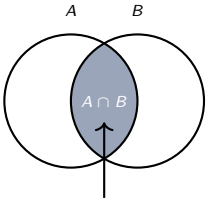
Contenido

1. Repaso de conjuntos (sets)
2. Espacio muestral y eventos
3. Los tres axiomas
4. Propiedades que se deducen
5. Cierre

Todo esto sale *solo* de los tres axiomas

Consecuencias

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ y $\mathbb{P}(A) \leq 1$.
2. $\mathbb{P}(\varnothing) = 0$.
3. **Monotonía**: si $A \subseteq B$, entonces $\mathbb{P}(A) \leq \mathbb{P}(B)$.
4. **Aditividad finita**: si $A_1, \ldots, A_k$ son disjuntos dos a dos, $\mathbb{P}(A_1 \cup \cdots \cup A_k) = \sum_{i=1}^{k} \mathbb{P}(A_i)$.
5. **Inclusión–exclusión**: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
6. **Cota de la unión** (Boole): $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$.
7. Descomposición disjunta: $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(A^c \cap B) + \mathbb{P}(A^c \cap B^c \cap C)$.



Por eso hay que restarlo una vez: eso *es* inclusión–exclusión.

Demo: inclusión–exclusión en una encuesta

Problema

En un hogar limeño, la probabilidad de que se consuma café es 0.60; té, 0.40; y ambos, 0.25. ¿Cuál es la probabilidad de que se consuma **al menos una** de las dos bebidas? ¿Y **ninguna**? ¿Y **solo café**?

Solución paso a paso

Sea C "café" y T "té": $\mathbb{P}(C) = 0.60$, $\mathbb{P}(T) = 0.40$, $\mathbb{P}(C \cap T) = 0.25$.

$$\mathbb{P}(C \cup T) = \mathbb{P}(C) + \mathbb{P}(T) - \mathbb{P}(C \cap T) = 0.60 + 0.40 - 0.25 = 0.75.$$

$$\mathbb{P}(\text{ninguna}) = \mathbb{P}\big((C \cup T)^c\big) = 1 - 0.75 = 0.25.$$

$$\mathbb{P}(\text{solo café}) = \mathbb{P}(C) - \mathbb{P}(C \cap T) = 0.60 - 0.25 = 0.35.$$

Chequeo: $0.35 + 0.15 + 0.25 + 0.25 = 1$ (solo café, solo té, ambos, ninguna). ✓

Trampa

Sumar $0.60 + 0.40 = 1.00$ y concluir "seguro toma algo" cuenta dos veces a quien toma ambas. Sumar probabilidades **solo** vale para eventos disjuntos.

Contenido

1. Repaso de conjuntos (sets)
2. Espacio muestral y eventos
3. Los tres axiomas
4. Propiedades que se deducen
5. Cierre

Resumen de la Clase 1

Lo que hay que saber hacer

1. Manejar el lenguaje de conjuntos: unión, intersección, complemento, disjuntos, y las dos leyes de De Morgan.
2. Construir Ω: mutuamente excluyente, exhaustivo, con la granularidad que pide la pregunta.
3. Escribir el evento como subconjunto de Ω (y usar De Morgan para "al menos uno").
4. Verificar los tres axiomas: $\mathbb{P} \geq 0$, $\mathbb{P}(\Omega) = 1$, aditividad **si son disjuntos**.
5. Deducir de ellos: $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, monotonía, inclusión–exclusión y la cota de Boole.

El error que más caro cuesta

Sumar $\mathbb{P}(A) + \mathbb{P}(B)$ sin preguntarse antes si A y B son disjuntos.

Siguiente clase (Clase 2)

Seguimos en la misma Unit 1 — Lecture 1

Ya tenemos las reglas del juego (los axiomas). Falta **construir leyes de probabilidad concretas** y ver hasta dónde llegan esas reglas:

- **Modelos discretos**: sumar los pedacitos, incluso cuando Ω es infinito (series geométricas).
- **Modelos continuos**: la probabilidad como **área**.
- **Aditividad contable**: el tercer axioma en su versión fuerte, y por qué no vale para uniones no numerables.
- **Interpretaciones**: qué *significa* un número de probabilidad.

Para llegar listo

Resuelve la hoja de ejercicios de esta clase: los niveles Fundamentos y Núcleo son exactamente lo que usaremos como base.

¡Gracias!

Clase 1 — Modelos de probabilidad y axiomas, parte 1 (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.