Probabilidad

Introducción

10/03/2023

Lara Sánchez Peña[1]

UTDT 2023 - MEC

[1] Basado en las notas de Andrea Rotnitzky

Presentación de la materia

- Programa - contenidos y criterios de aprobación
- Ayudante: Maximiliano Mendoza Greco, maximendozagreco@gmail.com
- Teórica: Lara Sánchez Peña, lara.spena@gmail.com

Interpretación frecuentista vs. bayesiana de la probabilidad

- **Paradigma frecuentista**: Se le atribuye probabilidad sólo a resultados de circunstancias **azarosas**. Se conceptualiza la posibilidad de repeticiones infinitas de idénticas circunstancias azarosas. La probablidad de un resultado se intepreta como la **frecuencia relativa** con la que ocurriría en las hipotéticas infinitas repeticiones.
- **Paradigma bayesiano**: La probabilidad se entiende como una medida de **incertidumbre** acerca de la veracidad de cualquier proposición. No requiere de la existencia de un proceso aleatorio. Hay dos grandes vertientes filosóficas:
  - interpersonal o lógica (JM Keynes, Rudolf Camap y H Jeffreys) considera que ante la misma evidencia, la probabilidad debe valer lo mismo para todos los individuos
  - subjetiva (F. Ramsey, Bruno de Finetti, LJ Savage) considera a la probabilidad como una medida de incertidumbre personal.

Def. de probabilidad según Laplace

Pierre-Simon Laplace (1749-1821) propuso la slguiente definición del concepto de probabilidad: consideremos un **experimento aleatorio** que tiene un **número finito** de resultados posibles y **equiprobables**

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$$

entonces la probabilidad de un evento $A \subset \Omega$ se define por

$$P(A) = \frac{\text{casos favorables}}{\text{casos posibles}} = \frac{\#(A)}{\#(\Omega)}$$

Notemos que el aparente azar en este ejemplo del dado, se debe en realidad a nuestra ignorancia. Porque la mecánica clásica (Newtoniana) nos dice que el movimiento del dado es en realidad un proceso completamente determinístico (no aleatorio). Y si conociéramos la posición y velocidad iniciales y las fuerzas que actúan, podríamos calcular (en principio) en forma exacta, cómo se va a mover el dado.

Def. de probabilidad según Kolmogorov

Un tipo diferente de azar, mucho más fundamental, aparece en la mecánica cuántica, una de las teorías fundamentales de la física moderna. Esta teoría postula que el azar es un componente esencial e irreductible de la naturaleza a nivel microscópico. Así por ejemplo, no podemos predecir con exactitud donde vamos a encontrar un electrón, sino solamente calcular la probabilidad de que el electrón esté en una cierta región del espacio.

El problema de tener espacios muestrales no necesariamente finitos hace que no se le pueda asignar probabilidad a cualquier evento que consideremos, es decir, a cualquier subconjunto de un espacio muestral Ω.

Modelo probabilístico de Kolmogorov

Un modelo probabilístico se compone de:

- **Espacio muestral Ω**: el conjunto de todos los resultados posibles del experimento.
- **Un evento es un subconjunto del espacio muestral**. Muchas veces, en vez de escribirlo como un conjunto lo enunciaremos con palabras.
- Una **ley de probabilidad**: asigna a un evento $A \subseteq \Omega$ de posibles resultados un número no negativo $P(A)$.
- La lista de eventos a los que se les puede asignar probabilidad se suele conocer como Σ, que es una σ-**álgebra**.[2]

Si tiramos un dado y queremos estudiar en qué cara cae, podemos considerar el evento A "el resultado es par". Matemáticamente, Ω = {⚀, ⚁, ⚂, ⚃, ⚄, ⚅} y A = {⚁, ⚃, ⚅}.

[2] Una σ-álgebra de Ω es una familia de subconjuntos de Ω que es cerrada bajo complementos, uniones e intersecciones finitas o numerables. Σ contiene a $\varnothing$ y Ω. No vamos a ahondar en los detalles porque requiere conocimientos sobre Teoría de la Medida.

Relación entre eventos y variables aleatorias

$$\begin{tikzcd} B \subseteq V \arrow[r, leftrightarrow] \arrow[d] & A \subseteq \Omega \arrow[d] \\ P(X \in B) \arrow[r, leftrightarrow] & P(A) \end{tikzcd}$$

donde V es el soporte de una variable aleatoria X.
Por ejemplo: Consideremos,

- $A = \{\text{cara}\}$, $\Omega = \{\text{cara}, \ \text{cruz}\}$
- $B = \{1\}$, $V = \{0, 1\}$
- X es una variable que toma el valor 1 si sale cara y 0 si no.

Antes de calcular probabilidades de eventos, queremos describir dichos eventos. Para ello, es necesario apelar a teoría de conjuntos. Luego, más adelante en la cursada, veremos cómo codificar dichos eventos a través de los valores que toma una variable aleatoria.

Conjuntos

Un **conjunto** es una colección de objetos, llamados **elementos** del conjunto.

- Si S es un conjunto y x es un elemento de S, escribimos $x \in S$.
- Si x no es un elemento de S, escribimos $x \notin S$.
- Llamamos al conjunto que no tiene elementos **conjunto vacío**, denotado por $\varnothing$.

Ejemplos

1. Si el experimento consiste en tirar una moneda y ver de qué lado cae, el espacio muestral Ω_M es un conjunto $\Omega_M = \{H, T\}$.
2. Si el experimento consiste en tirar un dado y ver de qué cara cae, el espacio muestral Ω_D es un conjunto $\Omega_D = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$.

Notación: Los conjuntos se describen con llaves: $\{\cdots\}$. *Eso implica que el orden de los elementos listados no es relevante.*

Conjuntos

- Los conjuntos pueden describirse por **extensión** (dando una lista de sus elementos) o por **comprensión** (dando una característica que describe unívocamente sus elementos).
- Decimos que A es un **subconjunto** de B si todo elemento de A es un elemento de B:

$$A \subset B \Leftrightarrow x \in A \Rightarrow x \in B$$

- En particular, si $A \subset B$ y $B \subset A$ diremos que $A = B$.
- Llamamos **conjunto universal** Ω al conjunto que contiene todos los objetos de nuestro interés. Una vez fijado Ω, solo consideraremos los conjuntos A que sean subconjuntos de Ω.
- En el contexto probabilístico Ω es el espacio muestral y los eventos cuyas probabilidades querramos calcular son subconjuntos de Ω.
- **Todas las operaciones que veremos a continuación entre eventos** (subconjuntos de Ω), **construyen nuevos eventos a los que les podremos asignar probabilidad porque pertenecen a Σ.**

Cardinalidad de conjuntos

Clasificamos conjuntos según su **cardinal** (cantidad de elementos que tienen):

1. finitos
2. infinitos numerables (contables)
3. infinitos no numerables (no contables)

Describa los siguientes espacios muestrales Ω y clasifíquelos según su cardinal:

1. El conjunto de resultados de tirar dos monedas idénticas simultáneamente.
2. El conjunto de resultados de tirar una moneda sucesivamente dos veces.
3. El conjunto de resultados de duración **en días enteros** de una bombilla eléctrica.
4. El conjunto de posibles puntos donde podría caer un rayo en una región determinada.

Álgebra (operaciones) de conjuntos

- El **complemento** de un conjunto A, respecto del conjunto universal Ω, es el conjunto:

$$A^c = \{x \in \Omega : x \notin A\}$$

- La **unión** de dos conjuntos A y B es el conjunto de todos los elementos que pertenecen a A o a B (o a ambos):

$$A \bigcup B = \{x \in \Omega : x \in A \text{ ó } x \in B\}$$

- La **intersección** de dos conjuntos A y B es el conjunto de todos los elementos que pertenecen a A y a B:

$$A \bigcap B = \{x \in \Omega : x \in A \text{ y además } x \in B\}$$

Para practicar: Sea $\Omega = [0, 1]$. Sea $A_n = \{x \in \Omega /$ dígito decimal n-ésimo es $9\}$, con $n \in \mathbb{N}$. Cómo describiría el conjunto de los números en Ω que tienen dígitos 9: 1) en al menos un dígito? 2) en todos los dígitos? 3) ningún dígito? 4) en el tercer pero no en el cuarto dígito?

Álgebra (operaciones) de conjuntos

En algunos casos, consideraremos la unión o intersección de finito (o infinitos) conjuntos.

Ejemplo: sean $A_1, A_2, \ldots, A_n, \cdots \subseteq \Omega$ conjuntos indexados por $n \in \mathbb{N}$.

- La unión de los conjuntos $A_1, A_2, \ldots$ se define como

$$\bigcup_{n\in\mathbb{N}} A_n = A_1 \bigcup A_2 \bigcup \cdots = \{x \in \Omega : x \in A_n \text{ para } \textbf{al menos un } n \in \mathbb{N}\}$$

- La intersección de los conjuntos $A_1, A_2, \ldots$ se define como

$$\bigcap_{n\in\mathbb{N}} A_n = A_1 \bigcap A_2 \bigcap \cdots = \{x \in \Omega : x \in A_n \text{ para } \textbf{todo } n \in \mathbb{N}\}$$

Álgebra (operaciones) de conjuntos

- La **diferencia** de conjuntos entre A y B es el conjunto que contiene a los elementos de A que no están en B.

$$A \setminus B = A \cap B^c$$

- La **diferencia simétrica** entre dos conjuntos A y B es un conjunto de elementos que pertenecen solamente a uno de los dos conjuntos.

$$A \triangle B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$$

Álgebra (operaciones) de conjuntos

- $\cap$ y $\cup$ son operaciones **conmutativas** (no importa el orden):

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

- $\cap$ y $\cup$ son operaciones **asociativas** (se pueden aplicar muchas veces):

$$(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C$$
$$(A \cap B) \cap C = A \cap (B \cap C) = A \cap B \cap C$$

- propiedad **distributiva** entre $\cap$ y $\cup$ (cómo combinar ambas operaciones):

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

Álgebra (operaciones) de conjuntos

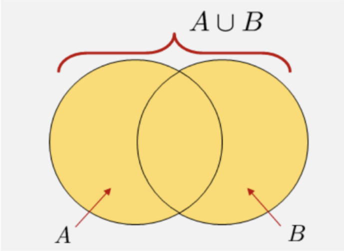
- **Leyes de De Morgan** (cómo tomar complemento a $\cap$ y $\cup$):
  1. $\left(\bigcup_{n\in\mathbb{N}} A_n\right)^c = \bigcap_{n\in\mathbb{N}} A_n^c$
  2. $\left(\bigcap_{n\in\mathbb{N}} A_n\right)^c = \bigcup_{n\in\mathbb{N}} A_n^c$

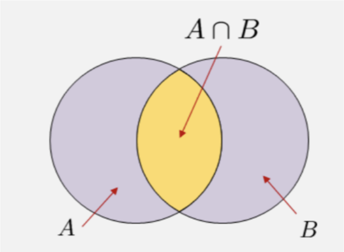
  ¿Cuál es la intuición de los resultados anteriores? Utilice como referencia el ejemplo donde $\Omega = [0,1]$ en la slide 9. Note que los resultados también valen si se toman uniones o intersecciones finitas.

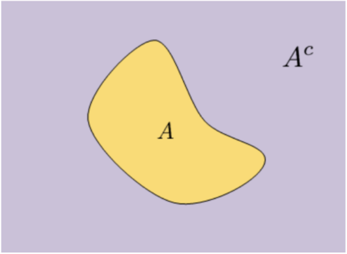
  Veamos, por ejemplo, cómo demostrar que vale 1) mostrando que vale la doble inclusión.

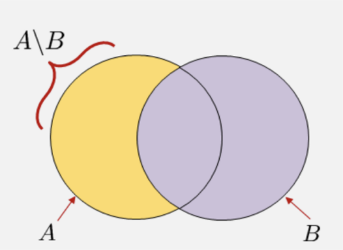
$$x \in \left(\bigcup_{n\in\mathbb{N}} A_n\right)^c \Leftrightarrow x \notin \bigcup_{n\in\mathbb{N}} A_n \Leftrightarrow x \notin A_n \;\forall\, n \in \mathbb{N} \Leftrightarrow x \in A_n^c \;\forall\, n \in \mathbb{N} \Leftrightarrow x \in \bigcap_{n\in\mathbb{N}} A_n^c$$

Álgebra (operaciones) de conjuntos





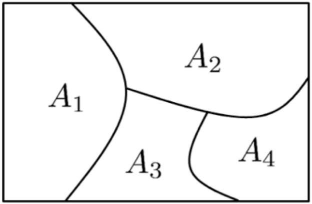




Partición de Ω

- Diremos que A y B son conjuntos **disjuntos** si $A \cap B = \varnothing$.
- Diremos que la colección de conjuntos A_i con $i \in I$ forma una **partición** de un conjunto Ω si
  - Son disjuntos: $A_i \bigcap A_j = \varnothing, \quad \forall i \neq j$, donde $i, j \in I$
  - S es la unión de toda la colección: $\Omega = \dot{\bigcup\limits_{i \in I}} A_i$.

<u>Observaciones:</u>



- El conjunto I puede ser finito. Por ejemplo, en el gráfico, $I = \{1, 2, 3, 4\}$.
- El conjunto I puede ser infinito numerable. Por ejemplo, $I = \mathbb{N}$.
- El conjunto I puede ser infinito no numerable. Por ejemplo, $I = \mathbb{R}$.

Partición de Ω

Para practicar:

Decida si, en cada uno de los ítems, los conjuntos A_i (con $i \in I$) forman particiones de $\Omega = \{\text{alumnos de Proba de UTDT}\}$ si

(a) $A_n = \{\text{personas con } n \text{ años}\}$ donde $n \in \mathbb{N}$.[3]

(b) $A_n = \{\text{personas cuyo DNI que termina en } n\}$ donde $n \in \{0, \dots, 9\}$.

(c) $A_p = \{\text{juega al fútbol con la pierna } p\}$, donde $p \in \{I, D\}$.

[3] Notemos que, cuando nos preguntan nuestra edad, respondemos nuestra edad medida en años con un número entero positivo.

Axiomas de probabilidad

Definimos la ley de probabilidad axiomáticamente, es decir, a partir de **supuestos** que suponemos ciertos:

1. **Axioma de no negatividad:** para todo evento A se cumple que $P(A) \geq 0$.
2. **Axioma de aditividad:** Si A y B son conjuntos disjuntos, entonces la probabilidad de su unión cumple

   $$P(A \uplus B) = P(A) + P(B)$$

   En general, si el espacio muestral tiene **numerables** elementos y $A_1, A_2, \dots$ es una sucesión de eventos disjuntos, la probabilidad de su unión verifica

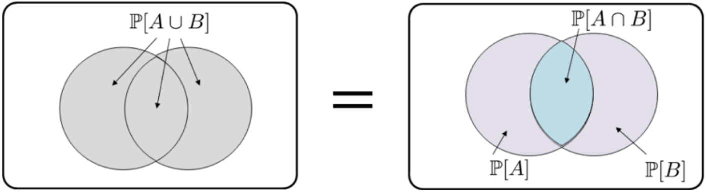
   $$P(A_1 \uplus A_2 \uplus \dots) = P(A_1) + P(A_2) + \dots$$

3. **Axioma de normalización:** $P(\Omega) = 1$.

Propiedades de una ley de probabilidad

A partir de los axiomas podemos **deducir** (demostrar) varias propiedades:

1. $P(\varnothing) = 0$
2. $P(A^c) = 1 - P(A)$
3. $A \subset B \Rightarrow P(A) \leq P(B)$
4. $P\left(A \bigcup B\right) = P(A) + P(B) - P\left(A \bigcap B\right)$



5. $P\left(A \bigcup B\right) \leq P(A) + P(B)$
6. $P\left(A \bigcup B \bigcup C\right) = P(A) + P\left(A^c \bigcap B\right) + P\left(A^c \bigcap B^c \bigcap C\right)$ (P.I.E.)

Modelos discretos

- Diremos que un modelo es **discreto** si su espacio muestral Ω es finito o infinito numerable.
  Si el espacio muestral es discreto, entonces la ley de probabilidad queda **completamente determinada** por las probabilidades de los eventos que consisten de un solo elemento.

- En particular, la probabilidad de cualquier evento $A = \{a_1, \ldots, a_n\}$ es la suma de las probabilidades de sus diferentes elementos:

$$P(\{a_1, \ldots, a_n\}) = P(\{a_1\}) + P(\{a_2\}) + \cdots + P(\{a_n\})$$

- Si el espacio muestral consiste en n resultados que son **equiprobables** (cada uno tiene la misma probabilidad de ocurrir), entonces

$$P(A) = \frac{\text{cantidad de elementos de A}}{n} = \frac{\#A}{\#\Omega}$$

Ejemplo: eventos en modelos discretos

Consideremos el experimento de tirar una vez un **dado equilibrado** y observar el resultado de la cara superior.

- $\Omega = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}$
- $\Sigma = \{\{\text{⚀}\}, \ldots, \{\text{⚅}\}, \{\text{⚀}, \text{⚁}\}, \ldots, \{\text{⚀}, \text{⚂}, \text{⚄}\}, \{\text{⚁}, \text{⚃}, \text{⚅}\}, \{\text{⚀}, \text{⚃}, \text{⚅}\}, \ldots\}$

Describimos los siguientes eventos por extensión o comprensión:

- $A_1 = \{\text{⚀}, \text{⚂}, \text{⚄}\} = \{\text{resultado impar}\}$
- $A_2 = \{\text{⚃}, \text{⚄}, \text{⚅}\} = \{\text{sale más que 3}\}$

Notemos que

- $A_1 \cap A_2 = \{\text{⚄}\}$
- $A_1 \cup A_2 = \{\text{⚁}\}^c$
- $A_1 \setminus A_2 = A_1 \cap A_2^c = \{\text{⚀}, \text{⚂}\}$

En modelos discretos, si queremos listar todos los elementos de Σ (los eventos de Σ, que son conjuntos) podemos considerar uniones, intersecciones y complementos (finitas o numerables) de eventos puntuales.

Ejemplo - espacio equiprobable

Consideremos el experimento de tirar un **dado equilibrado** una vez, queremos saber la probabilidad de obtener un 4. Ya sabemos que tiene que dar $1/6$. Pero, ¿por qué?

$$
\begin{aligned}
1 \underbrace{=}_{\text{ax. 3}} P(\Omega) = P\left(\{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}\right) &= P(\{\epsdice{1}\} \cup \{\epsdice{2}\} \cup \cdots \cup \{\epsdice{6}\}) \\
&\underbrace{=}_{\text{ax. 2}} P(\{\epsdice{1}\}) + \cdots + P(\{\epsdice{6}\})
\end{aligned}
$$

Como el dado está **equilibrado**, es **equiprobable** que salga cualquier cara del dado como resultado y por lo tanto:

$$
P(\{\epsdice{1}\}) = \cdots = P(\{\epsdice{6}\})
$$

entonces

$$
1 = P(\Omega) = 6 \cdot P(\{\epsdice{4}\}) \Rightarrow P(\{\epsdice{4}\}) = \frac{1}{6}
$$

Más ejemplos - espacios equiprobables

Para cada uno de los siguientes ejemplos, determine Ω y su cardinal:

1. Ahora tiramos un **dado equilibrado** cuatro veces y queremos saber la probabilidad de obtener **exactamente** un 6 y que salga en la primera tirada.
2. Tiramos tres veces una moneda equilibrada, ¿cuál es la probabilidad de obtener **a lo sumo** una vez cara?
3. De una urna con bolillas indistinguibles, una blanca y dos rojas, extraemos una al azar. ¿Cuál es la probabilidad de que sea roja?
4. Tiramos dos monedas equilibradas al mismo tiempo, ¿cuál es la probabilidad de que sus resultados sean distintos?

Modelos continuos

Diremos que un modelo es **continuo** si el espacio muestral Ω es algún subconjunto de la recta real.

Observación

Para modelos continuos, puede ocurrir que las probabilidades de eventos puntuales no lleguen a caracterizar las probabilidades de cualquier evento en Σ.

Ejemplo:

Una rueda de la fortuna de longitud 2π se gira. Los valores del borde de la rueda pertenecen a $\Omega = [0, 2\pi)$. Suponiendo que la rueda es justa, es razonable suponer que todos los resultados son equiprobables. Sin embargo, ¿cuál es la probabilidad de un evento que consiste de un solo elemento?