Probabilidades y Estadística

Ramiro Dibur

2025

Índice general

Prefacio

Estas son mis notas de Probabilidades y Estadística del segundo cuatrimentre de 2025. Las escribo más que nada para estudiar yo, pero las publico por si le llegan a ser útil a alguien.

1

Nociones básicas

Formalizamos algunas conceptos de probabilidad que vienen de la intuición.

1.1 Modelo probabilístico

Consideremos un experimento con distintos posibles resultados.

Definición 1.1. El ***espacio muestral*** de un experimento es el conjunto de posibles resultados del experimento.

Usualmente denotamos un espacio muestral con Ω.

Observación 1.2. Todo resultado corresponde con un único elemento $\omega \in \Omega$.

Veamos algunos ejemplos.

Ejemplo 1.3. Consideremos el siguiente experimento:

1. Se tira un dado balanceado de 6 caras.
2. Se graba el resultado.

En este caso, el espacio muestral es el conjunto

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Cabe aclarar que no importa de qué manera escribimos los resultados siempre y cuando la correspondencia con el resultado sea clara. Por ejemplo, podríamos haber definido el espacio muestral como

$$\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}.$$

Ejemplo 1.4. Consideremos el siguiente experimento:

1. Se tira una moneda 3 veces.
2. Se graba el resultado.

El espacio muestral es

$$\Omega = \{CCC, CCS, \dots, SSC, SSS\}.$$

Nótese que Ω se puede escribir como $\{C, S\}^3$.

Ejemplo 1.5. Consideremos el siguiente experimento:

1. Se elije un habitante de Buenos Aires al azar.
2. Se mide su altura en metros.

El espacio muestral podría ser

$$\Omega = \mathbb{R}.$$

Uno podría argumentar que el espacio muestral debería ser

$$\Omega = [0, 3],$$

ya que es imposible que alguien mida $-1\,\mathrm{m}$ o $100\,\mathrm{m}$. Sin embargo, lo único que nos interesa es que, al medir a alguien, caiga dentro de Ω.

1.2 Eventos

Definición 1.6. Sea Ω un espacio muestral. Un ***evento*** es un subconjunto de Ω.

Veámoslo en algunos ejemplos.

Ejemplo 1.7. Consideramos el experimento del ejemplo 1.3. El conjunto

$$A = \{\text{el resultado es un número par}\} = \{2, 4, 6\}$$

es un evento dado que $A \subseteq \Omega$.

Por ahora, usemos la noción intuitiva de probabilidades.

La probabilidad se le asigna a un evento, no a un resultado. Por ejemplo, cuando decimos

$$P(⚂) = \frac{1}{6}$$

en realidad queremos decir

$$P(\{⚂\}) = \frac{1}{6}.$$

No obstante, por practicidad acudiremos a la primera notación.

Usualmente calculamos la probabilidad de un evento de la siguiente manera:

$$P(A) = \frac{\#\text{ casos donde sucede } A}{\#\text{ casos totales}}.$$

Veamos por qué esto no es generalizable.

Ejemplo 1.8. Consideremos el siguiente experimento:

1. Se tiran 2 dados balanceados de 6 caras.

2. Se suman los números de las caras.

3. Se graba el resultado.

Un espacio muestral podría ser

$$\Omega = \{2, 3, \dots, 12\}.$$

Sin embargo, $P(2) \neq \frac{1}{10}$. Esto se puede resolver tomando el espacio muestral

$$\Omega = \{⚀⚀, ⚀⚁, \dots, ⚅⚄, ⚅⚅\}.$$

Por lo tanto, para todo resultado $\omega \in \Omega$,

$$P(\omega) = \frac{\#\text{ casos donde sucede } A}{\#\text{ casos totales}} = \frac{1}{36}.$$

A partir de este ejemplo surge una definición.

Definición 1.9. Sea Ω un espacio muestral. Si Ω es finito y todos sus elementos tienen la misma probabilidad, decimos que Ω es un ***espacio muestral equiprobable***.

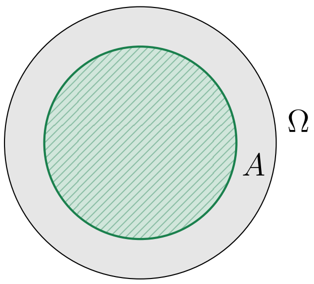
Sin embargo, hasta ahora únicamente tratamos con espacios muestrales finitos. Veamos qué pasa con los infinitos.

Ejemplo 1.10. Consideremos el siguiente experimento:

1. Se elije un punto del disco unitario $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ al azar.

2. Se graba el resultado.

Sea $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq \frac{1}{2}\}$.



Por lo que

$$P(A) = \frac{\text{área}(A)}{\text{área}(\Omega)}.$$

1.3 Eventos aleatorios

No podemos definirle una probablidad a todos los eventos de un espacio muestral.

Definición 1.11. Un evento al cual le podemos definir una probabilidad es llamado un ***evento aleatorios***.

Agregamos algunas reglas adicionales.

Definición 1.12. Llamamos $\mathcal{F}$ a una familia de eventos a los cuales podemos calcularles su probabilidad si cumple los siguientes axiomas:

(F1) $\Omega \in \mathcal{F}$.

(F2) Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$.

(F3) Si $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$, entonces $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$.

Observación 1.13. Si una familia cumple los axiomas (F1), (F2) y (F3), entonces se llama una σ-álgebra de conjuntos.

Ejemplo 1.14. En el ejemplo 1.3, el espacio muestral

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

tiene a la familia de eventos $\mathcal{F} = \mathcal{P}(\Omega)$ que se les puede asignar una probabilidad.

Veamos algunas propiedades que podemos deducir.

Proposición 1.15. *Sea Ω un espacio muestral y $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ una σ-álgebra. Las siguientes proposiciones son verdaderas:*

1. $\varnothing \in \mathcal{F}$.
2. *Si* $\{A_n\}_{1\leq n\leq N} \subseteq \mathcal{F}$, *entonces* $\bigcup_{1\leq n\leq N} A_n \in \mathcal{F}$.
3. *Si* $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$, *entonces* $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{F}$. *(También la versión finita.)*
4. *Si* $A, B \in \mathcal{F}$, *entonces* $A \setminus B \in \mathcal{F}$.

Demostración. (*1.*) Dado que $\Omega \in \mathcal{F}$ por (F1), obtenemos que $\varnothing = \Omega^c \in \mathcal{F}$ por (F2).

(*2.*) Sea $\{A_n\}_{1\leq n\leq N} \subseteq \mathcal{F}$. Consideremos la familia $\{B_n\}_{n\in\mathbb{N}}$ tal que

$$B_n = \begin{cases} A_n & \text{si } 1 \leq n \leq N, \\ \varnothing & \text{si } n \geq N. \end{cases}$$

Entonces, por (F3),

$$\bigcup_{1\leq n\leq N} A_n = \bigcup_{n\in\mathbb{N}} B_n \in \mathcal{F}.$$

(*3.*) Sea $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$. Por (F3),

$$\bigcap_{n\in\mathbb{N}} A_n = \left(\bigcup_{n\in\mathbb{N}} A_n^{\mathrm{c}}\right)^{\mathrm{c}} \in \mathcal{F}.$$

(*4.*) Sean $A, B \in \mathcal{F}$. Dado que B^{c},

$$A \setminus B = A \cap B^{\mathrm{c}} \in \mathcal{F}.$$

□

Ejemplo 1.16. Consideremos el siguiente experimento:

1. Se elije un número real del intervalo $[0, 1]$ al azar.
2. Se graba el resultado.

Sea $\Omega = [0, 1]$ el espacio muestral y sea $\mathcal{F}$ la familia de eventos a los cuales les podemos asignar una probabilidad. Para un intervalo $[a, b]$ la probabilidad se puede calcular como

$$P([a, b]) = b - a.$$

Aplicando las propiedades de la proposición 1.15, podemos deducir que en $\mathcal{F}$ están los eventos:

- Los intervalos abiertos y cerrados.
- Uniones e intersecciones numerables de cerrados y/o abiertos.
- Los puntos $\{x\}$ con $x \in [0, 1]$.
- Los números racionales $\mathbb{Q}$.

¿Cuál es la probabilidad de $\mathbb{Q}$? Basta con tomar $\{B_m\} = \{B(q_m, \frac{\varepsilon}{2^{m+1}})\}_{m\in\mathbb{N}}$ y ver que

$$\begin{aligned} P\left(\bigcup_{m\in\mathbb{N}} B_m\right) &\leq \sum_{m\in\mathbb{N}} P(B_m) \\ &\leq \sum_{m\in\mathbb{N}} \frac{\varepsilon}{2^m} \\ &\leq \varepsilon. \end{aligned}$$

Tomando $\varepsilon \to 0$, obtenemos que $P(\mathbb{Q}) = 0$.

1.4 Definición de probabilidad

La *idea de Laplace* de probabilidad consta en que la probabilidad de un evento es el límite de la frecuencia con la que sucede cuando la cantidad de ensayos tiende a infinito.

Por otro lado, está la *axiomatización de Kolmogorov*:

Definición 1.17. Sea Ω un espacio muestra y $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ una σ-álgebra. Una función probabilidad es una función $P : \mathcal{F} \to [0,1]$ que cumple los siguientes axiomas:

(P1) $P(\Omega) = 1$.

(P2) $P(A) \geq 0$ para todo $A \in \mathcal{F}$.

(P3) Si $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ una familia de eventos disjuntos, entonces $P(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=1}^{\infty} P(A_n)$.

Con esto podemos definir un espacio de probabilidad.

Definición 1.18. Sea Ω un espacio muestral, $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ una familia de eventos y $P : \mathcal{F} \to [0,1]$ una función probabilidad. Entonces, la terna $(\Omega, \mathcal{F}, P)$ es un ***espacio de probabilidad***.

Probamos algunos resultados inmediatos.

Proposición 1.19. *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Entonces, las siguientes proposiciones son verdaderas:*

1. $P(\varnothing) = 0$.
2. *Si A y B son eventos disjuntos, entonces $P(A \cup B) = P(A) + P(B)$.*
3. *Si $\{A_n\}_{1\leq n\leq N}$ es una familia de eventos disjuntos, entonces $P(\bigcup_{1\leq n\leq N} A_n) = \sum_{n=1}^{N} P(A_n)$.*
4. $P(A^c) = 1 - P(A)$.
5. *Si $A \subseteq B$, entonces $P(A) \leq P(B)$.*
6. $P(\bigcup_{n\in\mathbb{N}} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$.

Demostración. (*1.*) Consideremos la familia $\{\Omega, \varnothing, \ldots\}$. Por (P3),

$$\begin{aligned} P(\Omega) &= P(\Omega \cup \varnothing \cup \cdots) \\ &= P(\Omega) + \underbrace{P(\varnothing) + \cdots}_{=0} \end{aligned}$$

(*2.*) La propiedad sale utilizando (P3) y tomando la familia $\{A, B, \varnothing, \ldots\}$.
(*3.*) Se prueba por inducción y ussando la proposición anterior.
(*4.*) Podemos escribir como unión disjunta $\Omega = A \cup A^c$. Y con la propiedad *2.* obtenemos que

$$1 = P(\Omega) = P(A) + P(A^c),$$

entonces

$$P(A^c) = 1 - P(A).$$

(*5.*) Como $B = (B \setminus A) \cup A$,

$$P(A) \leq P(B \setminus A) + P(A) = P(B).$$

(*6.*) Sea $\{A_n\}_{n\in\mathbb{N}}$ una familia de eventos. Consideremos $\{B_n\}_{n\in\mathbb{N}}$ tal que

$$B_n = A_n \setminus \bigcup_{1\leq k\leq n-1} A_k.$$

Como los eventos B_n son disjuntos dos a dos,

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = P\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n=1}^{\infty} P(B_n) \leq \sum_{n=1}^{\infty} P(A_n).$$

□

Ejemplo 1.20. Consideremos el siguiente experimento:

1. Se tira una moneda balanceada hasta que salga cara.
2. Se anota la cantidad total de lanzamientos.

El espacio muestral es

$$\Omega = \mathbb{N},$$

con

$$P(1) = \frac{1}{2}, \quad P(i) = \frac{1}{2^i}, \quad i \geq 1.$$

Ahora pensémoslo con infinitas tiradas de moneda.

Ejemplo 1.21. Consideremos el experimento del ejemplo 1.20. Entonces, el espacio muestral es

$$\Omega = \{0,1\}^{\mathbb{N}}.$$

Sea $C_k = \{\text{la } k\text{-ésima moneda es cara}\}$ y definimos

$$A_n = \{\text{las primeras } n \text{ monedas son cara}\} = \bigcap_{k=1}^{n} C_k.$$

Entonces, $P(A_n) = \frac{1}{2^n}$. Consideremos

$$A = \{\text{sale siempre cara}\} = \bigcap_{n\in\mathbb{N}} A_n.$$

Por lo tanto $A \subseteq A_n$ para todo $n \in \mathbb{N}$. Así, $P(A) \leq P(A_n) = \frac{1}{2^n} \to 0$ y $P(A) = 0$.

Veamos otro evento:

$$B = \{\text{a partir de algún momento salen sólo caras}\}.$$

Es decir, existe $n_0 \in \mathbb{N}$ tal que para todo $k \geq n_0$ la moneda k es cara. En notación,

$$B = \bigcup_{n_0\in\mathbb{N}} \bigcap_{k=n_0}^{\infty} C_k.$$

Para cada n_0 fijo,

$$\begin{aligned} P\left(\bigcap_{k=n_0}^{\infty} C_k\right) &= \lim_{m\to\infty} P\left(\bigcap_{k=n_0}^{m} C_k\right) \\ &= \lim_{m\to\infty} \frac{1}{2^{m-n_0+1}} \\ &= 0. \end{aligned}$$

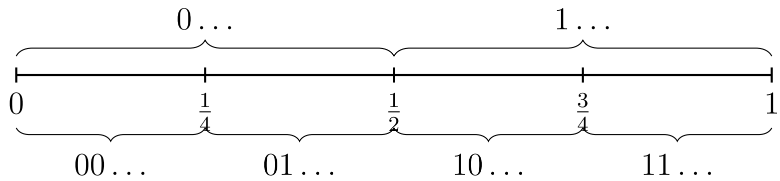
Entonces, por subaditividad numerable,

$$P(B) \leq \sum_{n_0\in\mathbb{N}} 0 = 0.$$

Podemos mirar a $\omega = (\omega_k)_{k\in\mathbb{N}} \in \{0,1\}^{\mathbb{N}}$ como la expansión binaria de un número real $x \in [0,1]$:

$$x = \sum_{k=1}^{\infty} \frac{\omega_k}{2^k}.$$

De esta forma, elegir una secuencia infinita de lanzamientos es equivalente a elegir un número real en $[0,1]$.



Proposición 1.22. *Sea $\{A_n\}_{n\in\mathbb{N}}$ una familia de eventos con $A_n \subseteq A_{n+1}$ para todo n. Entonces*

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} P(A_n).$$

Demostración. Definimos $B_n = A_n \setminus A_{n-1}$ (con $A_0 = \varnothing$). Entonces

$$\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n \qquad \text{y} \qquad B_i \cap B_j = \varnothing \text{ si } i \neq j.$$

Como $P(B_n) = P(A_n) - P(A_{n-1})$, tenemos

$$\begin{aligned} P\left(\bigcup_{n\in\mathbb{N}} A_n\right) &= \sum_{n=1}^{\infty} P(B_n) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \Big(P(A_n) - P(A_{n-1})\Big) \\ &= \lim_{N\to\infty} \Big(P(A_N) - P(A_0)\Big). \end{aligned}$$

Como $A_0 = \varnothing$, queda

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{N\to\infty} P(A_N).$$

$\square$

Corolario 1.23. *Tomando complementos se obtiene que, si $A_n \supseteq A_{n+1}$, entonces*

$$P\left(\bigcap_{n\in\mathbb{N}} A_n\right) = 1 - P\left(\bigcup_{n\in\mathbb{N}} A_n^{c}\right) = 1 - \lim_{n\to\infty} P(A_n^{c}) = \lim_{n\to\infty} P(A_n).$$

2

Probabilidad condicional

2.1 Probabilidad condicional

Ya vimos cómo calcular probabilidades de eventos. Ahora queremos formalizar qué significa calcular probabilidades *condicionadas* a que cierto evento ya ocurrió.

Ejemplo 2.1. Tiro dos dados y sumo los resultados. Consideremos los eventos

$$A = \{\text{la suma es } 5\},$$
$$B = \{\text{el segundo dado es par}\}.$$

El espacio muestral es

$$\Omega = \{(a, b) \mid a, b \in \{1, \ldots, 6\}\}.$$

En particular,

$$A = \{(1, 4), (4, 1), (2, 3), (3, 2)\}.$$

Luego,

$$P(A \mid B) = \frac{\#(A \cap B)}{\#B} = \frac{2}{18} = \frac{1}{9}.$$

El cálculo anterior nos motiva a introducir la definición general de probabilidad condicional.

Definición 2.2. Dados A y B eventos, con $P(B) \neq 0$, la ***probabilidad condicional*** de A dado B se define como

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

Veamos cómo aplica la definición en un ejemplo.

Ejemplo 2.3. Consideremos una urna con 9 bolitas de la siguiente forma:

- 5 naranjas,
- 4 violetas, de las cuales 3 tienen una cruz.

Denotemos $N = \{\text{naranja}\}$, $V = \{\text{violeta}\}$, $C = \{\text{tiene cruz}\}$. Entonces,

$$P(N) = \frac{5}{9}, \qquad P(N \mid V) = \frac{1}{3}, \qquad P(C \mid V) = \frac{1}{2}.$$

A partir de la definición, podemos verificar fácilmente que $P(\,\cdot \mid B)$ cumple los axiomas de probabilidad.

Proposición 2.4. *Sea B un evento con $P(B) \neq 0$. Entonces la aplicación*

$$A \mapsto P(A \mid B)$$

define una probabilidad sobre el espacio muestral condicionado.

Demostración. Para todo B con $P(B) > 0$:

(P1) $P(\Omega \mid B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$

(P2) Como P es una probabilidad, $P(A \mid B) \geq 0$.

(P3) Si $\{A_n\}_{n\in\mathbb{N}}$ son eventos disjuntos dos a dos,

$$\begin{aligned} P\left(\bigcup_{n\in\mathbb{N}} A_n \,\middle|\, B\right) &= \frac{P(\bigcup_{n\in\mathbb{N}}(A_n \cap B))}{P(B)} \\ &= \frac{\sum_{n\in\mathbb{N}} P(A_n \cap B)}{P(B)} \\ &= \sum_{n\in\mathbb{N}} \frac{P(A_n \cap B)}{P(B)} \\ &= \sum_{n\in\mathbb{N}} P(A_n \mid B). \end{aligned}$$

□

Observación 2.5. Es importante aclarar que $A \mid B$ no es un evento. Por lo tanto, expresiones como $P(A \mid B \mid C)$ no tienen sentido.

Proposición 2.6. *Para todo par de eventos A, B con $P(B) \neq 0$, se cumple*

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

Este resultado es útil para calcular probabilidades conjuntas a partir de probabilidades condicionales.

Ejemplo 2.7. Sacamos dos cartas de un mazo de 52 cartas sin reemplazo. Sea

$$\begin{aligned} A_1 &= \{\text{la primera carta es trébol}\}, \\ A_2 &= \{\text{la segunda carta es trébol}\}. \end{aligned}$$

Entonces,

$$\begin{aligned} P(A_1) &= \frac{13}{52}, \\ P(A_2 \mid A_1) &= \frac{12}{51}. \end{aligned}$$

Aplicando la proposición anterior,

$$P(A_1 \cap A_2) = P(A_1)\, P(A_2 \mid A_1) = \frac{13}{52} \cdot \frac{12}{51}.$$

Proposición 2.8. *Para eventos $A_1, A_2, \ldots, A_n$ se cumple*

$$\begin{aligned} &P(A_1 \cap A_2 \cap \cdots \cap A_n) \\ &\qquad = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap \cdots \cap A_{n-1}). \end{aligned}$$

Idea de demostración. La demostración se hace fácilmente por inducción en n. □

2.2 Probabilidad total

La noción de probabilidad condicional nos permite descomponer probabilidades en términos de eventos disjuntos.

Proposición 2.9 (Probabilidad total en dos partes)**.** *Sea $B \subseteq \Omega$ un evento. Entonces, para todo $A \subseteq \Omega$,*

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^{c}) \\ &= P(A \mid B)\, P(B) + P(A \mid B^{c})\, P(B^{c}). \end{aligned}$$

Más en general, si $\{A_i\}_{i \in I}$ es una partición numerable de Ω, se cumple la *ley de la probabilidad total*:

$$P(B) = \sum_{i \in I} P(B \mid A_i)\, P(A_i).$$

Ejemplo 2.10. Consideremos el siguiente experimento:

1. En una caja hay 3 monedas: una con dos caras y dos equilibradas.
2. Se elige una moneda al azar.
3. Se lanza la moneda.

Definamos los eventos

$$\begin{aligned} C &= \{\text{sale cara}\}, \\ A &= \{\text{la moneda elegida es la de dos caras}\}. \end{aligned}$$

Entonces,

$$\begin{aligned} P(C) &= P(C \mid A)\, P(A) + P(C \mid A^{\mathrm{c}})\, P(A^{\mathrm{c}}) \\ &= 1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} \\ &= \frac{2}{3}. \end{aligned}$$

Este cálculo es un primer ejemplo de la *fórmula de la probabilidad total*, que nos permite descomponer la probabilidad de un evento en función de una partición del espacio.

Ejemplo 2.11. Ahora consideremos otro experimento. Lanzamos un dado. Sea

$$\begin{aligned} A &= \{\text{saco un } 5\}, \\ B &= \{\text{saco un } 7\}, \\ C &= \{\text{ni } 5 \text{ ni } 7\} = (A \cup B)^{\mathrm{c}}. \end{aligned}$$

Supongamos que ganar depende de estos eventos: si ocurre A se gana siempre, si ocurre B nunca se gana, y si ocurre C se gana con cierta probabilidad.

Denotemos $G = \{\text{ganar}\}$. Entonces, por probabilidad total,

$$P(G) = P(G \mid A)\, P(A) + P(G \mid B)\, P(B) + P(G \mid C)\, P(C).$$

Como $P(G \mid A) = 1$, $P(G \mid B) = 0$, y $P(A) = P(B) = \frac{1}{6}$, $P(C) = \frac{4}{6}$, se obtiene

$$P(G) = 1 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + P(G \mid C)\, \frac{4}{6}.$$

Si además sabemos que $P(G \mid C) = \frac{1}{2}$, queda

$$P(G) = \frac{1}{6} + \frac{1}{2} \cdot \frac{4}{6} = \frac{6}{15}.$$

Estos ejemplos muestran cómo las probabilidades condicionales y la probabilidad total se combinan para calcular probabilidades en situaciones más complejas.

2.3 Independencia de eventos

La independencia de eventos se relaciona íntimamente con la probabilidad condicional. En esencia, si la probabilidad de un evento A dado un evento B no cambia, entonces son independientes.

Definición 2.12. Decimos que dos eventos A y B son independientes si

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

Observación 2.13. Si $P(A) = 0$ o $P(A) = 1$, entones es independiente de cualquier evento.

Esta noción la podemos generalizar a más eventos.

Definición 2.14. Decimos que los eventos $A_1, A_2, \ldots, A_n$ son independientes si, para todo $I \subseteq \{1, 2, \ldots, n\}$,

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).$$

Y para familias de eventos:

Definición 2.15. Decimos que una familia de eventos $\mathcal{A}$ es independiente si todo subconjunto finito de $\mathcal{A}$ es independiente.

3

Variables aleatorias y distribuciones

3.1 Definición

Una variable aleatoria es una característica numérica de un experimento.

Ejemplo 3.1. Algunos ejemplos de variables aleatorias:

- Tiro una moneda 5 veces y X es la cantidad de veces que sale cara.
- Tiro dos dados y X es la suma de las caras.
- Elijo un $w \in [0,1]$ y $X(w) = w^2$.

Ahora sí, veamos la definición.

Definición 3.2. Dado $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, una ***variable aleatoria*** es una función $X : \Omega \to \mathbb{R}$ tal que, para todo $a \in \mathbb{R}$,

$$X^{-1}(-\infty, a] \in \mathcal{F}.$$

Es decir, siempre podemos calcular $P(X \leq a)$.

Observación 3.3. Si $\mathcal{F} = \mathcal{P}$, entonces la condición siempre se cumple.

La σ-álgebra más chica que contiene a todas las semirrectas se llama la σ-álgebra de Borel.

3.2 Funciones de variables aleatorias

Sea $g : \mathbb{R} \to \mathbb{R}$. ¿Cuando es $g(X)$ una variable aleatoria? Requerimos que $g^{-1}(B) \in \mathcal{B}$ para todo $B \in \mathcal{B}$, es decir g es *medible Borel*.

Algunas funciones medibles son: las continuas y las monótonas. Además, la suma, el producto y la división de funciones medibles Borel son medibles Borel.

3.3 Función de distribución de una variable aleatoria

Definición 3.4. Dada X una variable aleatoria, definimos su función de distribución como $F_x : \mathbb{R} \to \mathbb{R}$ tal que

$$F_X(t) = P(X \leq t) = P_x(-\infty, t].$$

Algunos ejemplos.

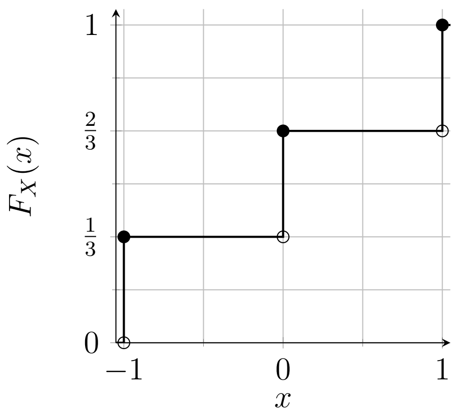
Ejemplo 3.5. Sea X una variable aleatoria tal que

$$P(X = -1) = P(X = 0) = P(X = 1) = \frac{1}{3}.$$

Podemos calcular su función de distribución F_X:

$$F_X(x) = \begin{cases} 0 & \text{si } x < -1, \\ \frac{1}{3} & \text{si } -1 \leq x < 0, \\ \frac{2}{3} & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Así es el gráfico de F_X.

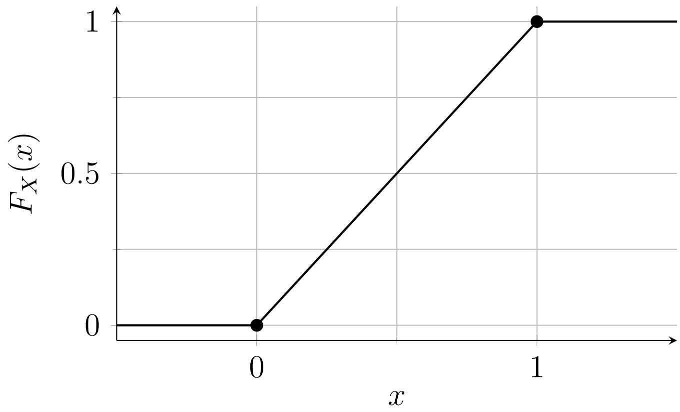


Consideremos otra variable aleatoria X dada por

$$P_X(A) = \int_{A \cap [0,1]} dx.$$

Es decir, $X \sim \text{Uniforme}(0, 1)$. Su función de distribución es

$$F_X(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$



Algunas propiedades de las funciones de distribución.

Proposición 3.6. *Si F_X es una función de distribución, entonces valen:*

(D1) F_X *es creciente.*

(D2) $\lim_{t\to+\infty} F_X(t) = 1$ *y* $\lim_{t\to-\infty} F_X(t) = 0$

(D3) F_X *es continua a derecha. Es decir,*

$$\lim_{t\to t_0^+} F_X(t) = F_X(t_0).$$

Demostración. (D1) Si $a \le b$, entonces $\{X \le a\} \subseteq \{X \le b\}$. Por lo tanto,

$$\begin{aligned} P(X \le a) &\le P(X \le b) \\ F_X(a) &\le F_X(b). \end{aligned}$$

(D2) Si $a_n \nearrow +\infty$, entonces

$$(-\infty, a_n] \subseteq (-\infty, a_{n+1}].$$

Consideramos $\mathbb{R} = \bigcup_{n\in\mathbb{N}}(-\infty, a_n]$. Entonces,

$$\lim_{n\to\infty} F_X(a_n) = P_X\left(\bigcup_{n\in\mathbb{N}}(-\infty, a_n]\right) = P_X(\Omega) = 1.$$

Para probar límite, simplemente consideramos la secuencia decreciente.

(D3) Sea $a_n \searrow t_0$. La intersección $\bigcap_{n\in\mathbb{N}}(\infty, a_n] = (-\infty, t_0]$. Entonces,

$$\begin{aligned} \lim_{n\to\infty} F_X(t_n) &= \lim_{n\to\infty} P_X(\infty, a_n] \\ &= P_X(-\infty, t_0] \\ &= F_X(t_0). \end{aligned}$$

□

Es más, no lo vemos, pero las condiciones (D1), (D2) y (D3) suficientes para determinar si F_X es una función de distribución.

3.4 Tipos de variables aleatorias

Primero una definición.

> **Definición 3.7.** Si X e Y son variables aleatorias con $F_X = F_Y$, entonces decimos que X se distribuye como Y. Esto se escribe como $X \sim Y$.

Clasificamos las variables aleatorias en cuatro grupos:

1. **Discretas.** X es discreta si existe un conjunto numerable $A \subseteq \mathbb{R}$ tal que $P_X(A) = 1$. En este caso, la función de distribución F_X es ***escalonada***.

2. **Absolutamente continuas.** X es absolutamente continua si existe una función $f_X : \mathbb{R} \to [0, \infty)$, llamada ***densidad de probabilidad***, tal que
$$F_X(t) = \int_{-\infty}^{t} f_X(x)\,dx.$$
En este caso, la probabilidad de un intervalo se calcula como
$$P(a \leq X \leq b) = \int_a^b f_X(x)\,dx.$$

3. **Mixtas.** X es mixta cuando su distribución tiene una parte discreta y una parte absolutamente continua. Es decir, existe una descomposición
$$P_X = \alpha P_d + (1-\alpha)P_c, \qquad 0 < \alpha < 1,$$
donde P_d es una medida discreta y P_c una medida absolutamente continua.

4. **Otras.** Existen distribuciones que no encajan en las categorías anteriores (por ejemplo, distribuciones ***singulares*** como la de Cantor). En este curso no nos vamos a preocupar por ellas.

3.5 Variables aleatorias discretas

Una definición previa:

> **Definición 3.8.** Llamamos el ***rango*** de una variable aleatoria X al conjunto
> $$R_X = \{t \in \mathbb{R} \mid P_X(t) > 0\}.$$

Veamos distintos tipos de variables aleatorias discretas:

1. **Bernoulli.** Sea $p \in [0, 1]$. Una variable $X \sim \mathrm{Bernoulli}(p)$ toma valores en $\{0, 1\}$ con
$$P(X = 1) = p, \qquad P(X = 0) = 1 - p.$$

2. **Binomial.** Sea $n \in \mathbb{N}$ y $p \in [0,1]$. Una variable $X \sim \text{Binomial}(n,p)$ representa la cantidad de éxitos en n ensayos independientes de Bernoulli(p). Su rango es $\{0,1,\ldots,n\}$ y

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad 0 \leq k \leq n.$$

3. **Geométrica.** Sea $p \in (0,1]$. Una variable $X \sim \text{Geom}(p)$ modela el número de ensayos hasta obtener el primer éxito. Su rango es $\{1,2,3,\ldots\}$ y

$$P(X = k) = (1-p)^{k-1} p, \qquad k \geq 1.$$