Capítulo 1. Introducción a la Probabilidad

Índice

Uno de los propósitos de la ciencia es enunciar leyes que expliquen el comportamiento de los fenómenos del mundo real. En general, dichas leyes pueden clasificarse en dos tipos:

- Determinísticas. Son aquellas que predicen con absoluta certeza. Son típicas de las ciencias físicas; por ejemplo, la relación entre la temperatura, el volumen y la presión de un gas.
- Aleatorias. Se caracterizan porque realizan predicciones con incertidumbre asociada. Son comunes en las ciencias sociales aunque también aparecen en física, biología y otras ciencias. Por ejemplo, el número de errores que una persona comete en una determinada tarea.

Este curso trata sobre los fundamentos de los modelos aleatorios con el objetivo de que el lector pueda aplicarlos en ciencias sociales.

1. Espacio muestral y sucesos

Un experimento aleatorio es aquel cuyos resultados no pueden predecirse con absoluta certeza, como el lanzamiento de una moneda o un dado. En el contexto de las ciencias sociales, una gran parte de los procedimientos de recogida de datos pueden representarse como experimentos aleatorios. Por ejemplo, el resultado de la aplicación de un examen a un grupo de individuos, la cantidad de dinero que una persona ahorra a lo largo de varios meses, el tiempo necesario para resolver una tarea, el número de veces que se realiza una tarea en un periodo fijo de tiempo, etc.

> *Suceso elemental.* Cada uno de los posibles resultados de un experimento aleatorio se denomina *suceso elemental.* Los sucesos elementales se indican mediante letras mayúsculas: A, B, etc.

Ejemplo 1. Se pide a un sujeto con dislexia que lea en alto cinco palabras. Un posible suceso es $A = \{\text{no cometer ningún error en la lectura}\}$, otro suceso es $B = \{\text{cometer un error en la primera palabra}\}$, $C = \{\text{cometer un error en la segunda palabra}\}$, etc.

Ejemplo 2. Como parte de un test de inteligencia, se pide a un sujeto que resuelva un rompecabezas y no se fija tiempo límite. Se mide el tiempo que tarda en completarlo. Un posible resultado es $A = \{\text{tardar 20 segundos}\}$, otro sería $B = \{\text{tardar 55 segundos}\}$, etc. En este ejemplo, a diferencia del anterior, el número de posibles resultados diferentes es infinito.

> *Espacio muestral.* El conjunto de todos los posibles resultados de un experimento aleatorio se denomina *espacio muestral* y se designa por Ω.

La definición de Ω en un problema concreto depende de cual sean los aspectos que desean analizarse. Una distinción básica es entre espacio muestral finito o infinito.

Ejemplo 3. Supongamos que pedimos a un sujeto con dislexia que lea en alto cinco palabras y contabilizamos el número de palabras leídas correctamente. Este experimento aleatorio tiene un espacio muestral finito, $\Omega = \{0,\ 1,\ 2,\ 3,\ 4,\ 5\}$.

Ejemplo 4. Supongamos que una persona debe realizar una prueba deportiva y se cuenta el tiempo en segundos que tarda en completarla. El espacio muestral es

$$\Omega = \{0, 1, 2, \dots\}$$

Este es un espacio muestral infinito numerable. Es decir, tiene infinitos elementos que pueden contarse.

Ejemplo 5. Supongamos que medimos el tiempo que tarda en realizarse una prueba deportiva con absoluta precisión. El tiempo sería una variable que toma como valores los números reales no negativos. El espacio muestral del experimento es:

$$\Omega = \{t;\ t \geq 0\}$$

Se trata de un espacio muestral infinito no numerable. Es decir, tiene infinitos elementos que, por las propiedades de los números reales, no pueden contarse.

A partir de los sucesos elementales, es posible crear *sucesos compuestos* utilizando las siguientes leyes de composición:

- *Unión de sucesos.* Dados dos sucesos A y B, el suceso unión, $A \cup B$, es el suceso que se realiza cuando ocurre al menos uno de los dos sucesos, A o B.
- *Intersección de sucesos.* Dados dos sucesos A y B, el suceso intersección, $A \cap B$, es el suceso que se realiza cuando ocurren ambos sucesos, A y B.
- *Negación o complementación de sucesos.* Dado un suceso A, el suceso contrario, $\overline{A}$, se realiza si no ocurre A.

Ejemplo 6. Supongamos que una persona responde a una escala de actitudes que consta de dos ítems con tres opciones de respuesta cada uno (A: acuerdo, N: neutral, D: desacuerdo). El espacio muestral es el conjunto de todos los posibles resultados:

$$\Omega = \{AA, AN, AD, NA, NN, ND, DA, DN, DD\}$$

donde AA significa estar de acuerdo con el primer y con el segundo ítem, AN significa estar de acuerdo con el primero y neutral en el segundo, etc. A partir de Ω, es posible definir el suceso compuesto *Estar de acuerdo con el primer ítem*, cuyos elementos son: $A_1 = \{AA, AN, AD\}$. Del mismo modo, el suceso compuesto *Estar de acuerdo con el segundo ítem* se define del modo: $A_2 = \{AA, NA, DA\}$. Algunos ejemplos de aplicación de las leyes de composición sobre estos subconjuntos son:

- El conjunto $C = A_1 \cup A_2$; significa estar de acuerdo con el primer ítem o estarlo con el segundo. Es decir:

$$C = \{AA, AN, NA, AD, DA\}$$

- El conjunto $D = A_1 \cap A_2$ significa estar de acuerdo con el primer ítem y también con el segundo. Contiene los sucesos elementales que están contenidos simultáneamente en A_1 y A_2:

$$D = \{AA\}$$

- El conjunto $E = \overline{C}$ significa no estar de acuerdo con alguno de los ítems. Contiene los elementos:

$$E = \{NN, ND, DN, DD\}$$

Se dice que dos sucesos son *excluyentes* si su intersección es el conjunto vacío; es decir si $A \cap B = \emptyset$.

Ejemplo 7. Supongamos que se lanza un dado. El espacio muestra es

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

Algunos sucesos compuestos son:

- Obtener menos de tres: $A = \{1, 2\}$.
- Obtener par: $B = \{2, 4, 6\}$
- Obtener un múltiplo de tres: $C = \{3, 6\}$
- Obtener un número primo: $D = \{2, 3, 5\}$.

El suceso *obtener un número par y múltiplo de tres* se define

$$B \cap C = \{6\}$$

El suceso *obtener un número par o primo* es

$$B \cup D = \{2, 3, 4, 5, 6\}$$

Los sucesos A y C son excluyentes porque no tienen ningún elemento en común, no existe ningún múltiplo de tres que sea menor de tres. Es decir, la intersección de A y C es el conjunto vacío, $A \cap C = \emptyset$.

2. Probabilidad

En un experimento aleatorio no es posible determinar con exactitud cual será el resultado, únicamente podemos de determinar cómo de verosímiles son los resultados. Intuitivamente, la probabilidad indica el grado de confianza en que ocurra cada suceso. Existen diferentes interpretaciones del concepto de probabilidad que pretenden dar sentido a esta idea. Algunas de ellas son las siguientes:

- *Probabilidad clásica.* Se define la probabilidad como el cociente entre el número de casos favorables y el número total de casos, siempre que todos ellos sean igualmente posibles. Por ejemplo, si se lanza un dado, la probabilidad de cada cara es $1/6$. Según esta definición, la probabilidad depende de la geometría, las propiedades físicas, o la construcción del objeto en cuestión.

- *Probabilidad frecuentista.* La probabilidad frecuentista se basa conceptualmente en un recuento de resultados si se realizara empíricamente el procedimiento aleatorio un número indefinido de veces. La definición frecuentista es adecuada para aquellos fenómenos para los que no puede determinarse la probabilidad al modo clásico, pero sí es posible realizar repeticiones del experimento aleatorio para contar la frecuencia del suceso. Por ejemplo, si estamos analizando las dificultad de las preguntas de un examen, no es posible obtener la probabilidad de acertar cada una a partir de la *geometría* del problema; en cambio, es posible aplicar el examen a una muestra amplia y hacer un recuento de la tasa de acierto de cada pregunta. En general, si un experimento aleatorio tiene N resultados diferentes y la frecuencia de ocurrencia de un suceso A es n, la probabilidad de A es

$$P(A) = \lim_{N \to \infty} \frac{n}{N} \ .$$

- *Probabilidad subjetiva.* La probabilidad no se concibe como una propiedad del experimento sino como un grado de creencia personal que el individuo tiene acerca del resultado del mismo. Esta definición se ha utilizado, por ejemplo, en economía cuando se trata de fenómenos que no se pueden repetir en veces sucesivas. Por ejemplo, si hablamos de la probabilidad de que la economía mejore el próximo trimestre, dicho fenómeno ocurrirá solo una vez, por lo que no es posible hacer un recuento de veces en que se produce la mejoría. En ocasiones se utilizan las apuestas para medir la probabilidad subjetiva. Por ejemplo, supongamos que una persona está dispuesta a apostar como máximo una cantidad de X euros acerca de la ocurrencia de un suceso, en caso de que dicho suceso ocurra la persona cobraría Y euros; entonces, la probabilidad subjetiva que tiene esta persona acerca de la ocurrencia del suceso es $p = X/Y$.

Estas definiciones pretender dar sentido al concepto de probabilidad, estableciendo el modo en que puede calcularse. Además, la probabilidad se define desde un punto de vista estrictamente matemático mediante los denominados axiomas de Kolmogorov.

Probabilidad Una medida de probabilidad es una función que asigna a cada suceso un valor numérico en el intervalo [0, 1] y que cumple las propiedades:

1. $P(A) \geq 0$ para todo suceso A.
2. $P(\Omega) = 1$.
3. Si $A_1, A_2, \ldots, A_n$ es un conjunto de sucesos excluyentes, entonces $P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$.

De esta definición axiomática de probabilidad se derivan algunas consecuencias importantes:

1. La probabilidad del suceso $\overline{A}$, complementario de A, es:

$$P(\overline{A}) = 1 - P(A)$$

2. La probabilidad del suceso imposible es cero:

$$P(\emptyset) = 0$$

3. Si $A \subset B$ (A está incluido o es un subconjunto de B), entonces $P(A) \leq P(B)$.
4. *Regla de la suma.* Dados dos sucesos cualesquiera, la probabilidad de que ocurra uno u otro es la probabilidad de su unión, y se calcula

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Ejemplo 8. Según la definición clásica de probabilidad, en el lanzamiento de un dado la probabilidad de cada cara es 1/6. Entonces, la probabilidad de algunos sucesos compuestos es:

- Obtener menos de tres: $P(A) = P(\{1, 2\}) = 2/6$.
- Obtener par: $P(B) = P(\{2, 4, 6\}) = 3/6$
- Obtener un múltiplo de tres: $P(C) = P(\{3, 6\}) = 2/6$
- Obtener un número par y múltiplo de tres: $P(B \cap C) = P(\{6\}) = 1/6$

Consideremos el suceso *obtener un número par o un múltiplo de tres*, los sucesos elementales que lo forman son:

$$D = B \cup C = \{2, 3, 4, 6\}$$

La probabilidad del suceso D se obtiene por la regla de la suma:

$$
\begin{aligned}
P(B \cup C) &= P(B) + P(C) - P(B \cap C) \\
&= \frac{3}{6} + \frac{2}{6} - \frac{1}{6} \\
&= \frac{4}{6}.
\end{aligned}
$$

En este ejemplo vemos por qué en la regla de la suma se resta la probabilidad de la intersección. El elemento 6 está contenido tanto en B como en C. Por tanto, al sumar $P(B) + P(C)$ la probabilidad del valor 6 aparece dos veces, $P(B)+P(C) = P(\{2,4,6\})+P(\{3,6\}) = P(\{2\})+P(\{3\})+P(\{4\})+2P(\{6\})$. Por eso es necesario restar la probabilidad de la intersección, $P(B \cap C) = P(\{6\})$, para obtener el resultado correcto.

Ejemplo 9. En un casino tienen una ruleta con 52 números, los 26 primeros son rojos y otros tantos negros. Un jugador apuesta a que sale negro y otro a que sale impar.

- La probabilidad de que gane el primer jugador es $P(G_1) = 26/52 = 1/2$.
- La probabilidad de que gane el segundo jugador es $P(G_2) = 1/2$.
- Para que ganen ambos tiene que salir un número negro e impar. La probabilidad de que esto suceda es $P(G_1 \cap G_2) = 13/52$.
- De acuerdo con la regla de la suma, la probabilidad de que gane alguno de ellos es: $P(G_1 \cup G_2) = P(G_1) + P(G_2) - P(G_1 \cap G_2) = 26/52 + 26/52 - 13/52 = 39/52$.

3. Probabilidad marginal, condicionada e independencia

La *probabilidad marginal* de un suceso, $P(A)$, representa la probabilidad de que ocurra A cuando los posibles resultados del experimento aleatorio son los que se incluyen en el espacio muestral, Ω.

La *probabilidad condicionada* indica la probabilidad de que se de un de suceso sabiendo que se ha dado otro. La probabilidad condicionada puede verse como una reducción en el espacio muestral. Si se toma un subconjunto de Ω y se define un nuevo espacio muestral B, entonces $P(A \mid B)$ representa la probabilidad del suceso A cuando los posibles resultados del experimento aleatorio son los contenidos en el subconjunto B. La probabilidad de A condicionada a B se define

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

Ejemplo 10. Supongamos que suceso $A = \{2,4,6\}$ consiste en obtener par al lanzar un dado. Como Ω contiene los números del 1 al 6 entonces la probabilidad marginal de A es

$$P(A) = \frac{3}{6} = \frac{1}{2}.$$

Si se define el suceso $B = \{1,2,3\}$ consistente en obtener tres o menos, entonces el suceso $A \mid B = \{2\}$ significa obtener un número par cuando la extracción se realiza desde el conjunto B. Aplicando la fórmula de la probabilidad condicionada:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{3/6} = \frac{1}{3} \ .$$

En este ejemplo la probabilidad marginal de A es superior a su probabilidad condicionada, es decir, es más probable obtener un número par al lanzar un dado que obtener par sabiendo que el resultado ha sido menor de cuatro.

Ejemplo 11. Continuando con el ejemplo 9, la probabilidad de que gane el segundo jugador sabiendo que la ha ganado el primero es:

$$\begin{aligned}
P(G_2 \mid G_1) &= \frac{P(G_2 \cap G_1)}{P(G_1)} \\
&= \frac{13/52}{26/52} \\
&= \frac{1}{2}\ .
\end{aligned}$$

Es decir, en este caso la probabilidad marginal es igual a la condicionada, $P(G_2) = P(G_2 \mid G_1)$.

Desarrollando la expresión de la probabilidad condicionada se obtiene la denominada *regla del producto*, que indica cual es la probabilidad de la intersección de dos sucesos:

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

Independencia de sucesos. Cuando dos sucesos, A y B, son estadísticamente independientes, el saber que se ha dado uno no aporta información acerca de la ocurrencia del otro. Matemáticamente esto significa que la probabilidad marginal de un suceso es igual a su probabilidad condicionada, $P(A \mid B) = P(A)$. La probabilidad de A es igual tanto si ocurre B como si no. En consecuencia, cómo $P(A \cap B) = P(A \mid B)P(B)$, en caso de independencia

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

Ejemplo 12. Supongamos que al lanzar un dado se definen los sucesos A = *obtener par = {2, 4, 6}* y B = *obtener primo = {2, 3, 5}*. Ambos tienen la misma probabilidad marginal, *P(A) = P(B) = 1/2*. Sin embargo no son sucesos independientes. La intersección de ambos es $A \mid B = \{2\}$, por lo que

$$P(A \mid B) = \frac{1/6}{1/2} = \frac{1}{3}$$

Cómo $P(A) \neq P(A \mid B)$ los sucesos no son independientes. Saber que se ha dado B disminuye la probabilidad de obtener A porque el conjunto de los números primos contiene tres elementos y solo uno de ellos es par.

Ejemplo 13. En el juego de la ruleta, la probabilidad de los sucesos G_1 = negro y G_2 = impar son independientes. Como se ha visto en el ejemplo 9,$P(G_2 \mid G_1) = P(G_2) = 1/2$. Además, $P(G_2 \cap G_1) = P(G_1)P(G_2) = 1/4$.

Continuando con el ejemplo, supongamos que un tercer jugador apuesta a que sale un número primo, es decir, $G_3 = \{2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47\}$. Entonces

- La probabilidad de que gane el tercer jugador es $P(G_3) = 15/52$.
- La probabilidad de que ganen simultáneamente el segundo y tercer jugador es $P(G_2 \cap G_3) = 14/52$.
- La probabilidad de que gane el tercer jugador sabiendo que ha ganado el segundo es:

$$\begin{aligned}
P(G_3 \mid G_2) &= \frac{P(G_3 \cap G_2)}{P(G_2)} \\
&= \frac{14/52}{26/52} \\
&= \frac{7}{13}\ .
\end{aligned}$$

  Por tanto, $P(G_3 \mid G_2) \neq P(G_3)$ y los sucesos G_2 y G_3 no son independientes.

- La probabilidad de que gane el segundo jugador sabiendo que ha ganado el tercero es:

$$
\begin{aligned}
P(G_2 \mid G_3) &= \frac{P(G_3 \cap G_2)}{P(G_3)} \\
&= \frac{14/52}{15/52} \\
&= \frac{14}{15} \ .
\end{aligned}
$$

4. Ley de la probabilidad total

La ley de la probabilidad total permite calcular la probabilidad marginal de un suceso cuando el espacio muestral está dividido en varios subconjuntos excluyentes entre sí. Supongamos que el espacio muestral está dividido en varios subconjuntos B_1, B_2, …, B_J, de modo que la unión de todos ellos es el espacio muestral, $B_1 \cup B_2 \cup \cdots \cup B_J = \Omega$. Además, $P(A, B_j)$ indica la probabilidad de la intersección de un suceso A con B_j, donde $j = 1, \ldots, J$. Entonces la probabilidad marginal de A es

$$
P(A) = \sum_{j=1}^{J} P(A, B_j).
$$

Ejemplo 14 . Continuando con el ejemplo del dado, supongamos que B_1 son los números primos, B_2 son los no primos (1, 4 y 6) y A son los números pares. Entonces la intersección de A y B_1 son los números pares y primos, es decir $A \cap B_1 = \{2\}$. El conjunto de números que son pares pero no son primos es $A \cap B_2 = \{4, 6\}$. La probabilidad de estos dos conjuntos es

$$
\begin{aligned}
P(A, B_1) &= \frac{1}{6} \\
P(A, B_2) &= \frac{2}{6}
\end{aligned}
$$

Por tanto la ley de la probabilidad total nos dice que la probabilidad marginal de encontrar un número par es

$$
P(A) = P(A, B_1) + P(A, B_2) = \frac{1}{6} + \frac{2}{6} = \frac{1}{2}.
$$

La ley de la probabilidad total puede parecer una forma demasiado enrevesada de calcular la probabilidad de un suceso, y de hecho lo es en nuestro sencillo ejemplo del dado. Sin embargo es necesaria para modelos estadísticos más complejos en los que resulta natural su aplicación.

5. Teorema de Bayes

El teorema de Bayes, en su presentación moderna, es la combinación de la ley de la probabilidad condicionada con la ley de la probabilidad total. Según el teorema de Bayes la probabilidad de A condicionada a B es

$$
P(B_j \mid A) = \frac{P(A, B_j)}{P(A)} = \frac{P(A \mid B_j) P(B_j)}{\sum_{j'=1}^{J} P(A, B_{j'})}.
$$

Ejemplo 15 En el lanzamiento del dado, queremos calcular la probabilidad de obtener un número sea primo sabiendo que es par. El resultado es

$$P(B_1 \mid A) = \frac{P(A, B_1)}{P(A)} = \frac{1/6}{1/6 + 2/6} = \frac{1}{3}.$$

A su vez, la probabilidad de que no sea primo condicionada en par es

$$P(B_2 \mid A) = \frac{P(A, B_2)}{P(A)} = \frac{2/6}{1/6 + 2/6} = \frac{2}{3}.$$

Ejemplo 16. El número de accidentes de tráfico registrados en una determinada ciudad ha sido 120, que se reparten durante los siete días de una semana del modo 15, 16, 15, 18, 22, 24, 10. Además, el número de accidentes que han requerido la intervención de los servicios médicos ha sido 40, repartidos del modo 5, 4, 6, 4, 10, 7, 4. Queremos saber

- La probabilidad del suceso $A =$ intervención de los servicios médicos
- La probabilidad de que un accidente haya ocurrido un lunes sabiendo que han tenido que intervenir los servicios médicos.

De acuerdo con la ley de la probabilidad total, los sucesos $B_1, \ldots, B_7$ representan la ocurrencia de un accidente en cada uno de los días de la semana. Entonces, la probabilidad de que intervengan los servicios médicos un lunes es

$$P(A \cap B_1) = P(A \mid B_1)P(B_1) = \frac{5}{15}\frac{15}{120} = \frac{5}{120}$$

Del mismo modo, la probabilidad de intervención en cada uno de los demás días es

$$\begin{aligned}
P(A \cap B_2) &= \frac{4}{120} \\
P(A \cap B_3) &= \frac{6}{120} \\
P(A \cap B_4) &= \frac{4}{120} \\
P(A \cap B_5) &= \frac{10}{120} \\
P(A \cap B_6) &= \frac{7}{120} \\
P(A \cap B_7) &= \frac{4}{120}
\end{aligned}$$

Entonces, la probabilidad marginal de A es

$$P(A) = \sum_{j=1}^{7} P(A \cap B_j) = \frac{40}{120} = \frac{1}{3}$$

La probabilidad de que un accidente ocurriera un lunes sabiendo que tuvieron que intervenir los servicios médicos es

$$P(B_1 \mid A) = \frac{P(A \cap B_1)}{P(A)} = \frac{5/120}{1/3} = 0,125$$

6. Ejercicios

Ejercicio 1. Un sujeto responde a dos preguntas de verdadero o falso. Asumiendo que la probabilidad de acertar cada pregunta es π:

1. Indicar el espacio muestral
2. Obtenga la probabilidad de acertar 0, 1 o 2 preguntas.

Ejercicio 2. Sean dos sucesos A y B. Sabiendo que $P(A) = 0{,}5$ y $P(A \cup B) = 0{,}6$:

- Indicar el valor de $P(B)$ si A y B son excluyentes.
- Indicar el valor de $P(B)$ si A y B son independientes.

Ejercicio 3. Supongamos que un sujeto resuelve tres problemas. La probabilidad de resolver un problema correctamente es π.

1. Escriba el conjunto Ω.
2. Obtenga la probabilidad de cada suceso en Ω.
3. Obtenga la probabilidad del suceso $D \equiv$ cometer un error en la última palabra.
4. Obtenga la probabilidad del suceso $F \equiv$ "obtener un solo acierto".
5. Obtenga la probabilidad $P(D \mid F)$.

Ejercicio 4. Se dispone de dos urnas con la siguiente composición:

1. Urna I: 2 bolas blancas y 3 bolas negras.
2. Urna II: 6 bolas blancas y 4 bolas negras.

Se realiza el experimento aleatorio consistente en lanzar una moneda. Si sale C se obtiene una bola de la urna I, si sale X se obtiene una bola de la urna II.

- Describir el espacio muestral Ω.
- Sea $A \equiv$ obtener blanco. Obtenga $P(A)$.
- Sea $B \equiv$ obtener Xen el lanzamiento de la moneda. ¿Cuanto vale $P(B \mid A)$ y $P(A \cup B)$?

Ejercicio 5. Supongamos que lanzamos un dado y se definen los sucesos $A \equiv$ obtener par, $B_1 \equiv$ obtener 3 o menos y $B_2 \equiv$ obtener 4 o más.

1. Obtenga la probabilidad de de A según la ley de la probabilidad total.
2. Calcule $P(B_1 \mid A)$ utilizando el teorema de Bayes.
3. Indique si A y B_1 son sucesos independientes.