1 Probabilidad condicional

Definición 1.1 Asumiendo $P(B) > 0$, definimos la probabilidad condicional de A, dado que B a ocurrido como sigue

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Para cualquier evento B fijo tal que $P(B) > 0, \mathbb{P}(\cdot|B)$ es una función de probabilidad (es decir, que cumple los tres axiomas de probabilidad). En particular $\mathbb{P}(A|B) \geq 0, \mathbb{P}(\Omega|B) = 1$ y si $A_1, A_2, \ldots$ son disjuntos, entonces $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i|B) = \sum_{i=1}^{\infty} \mathbb{P}(A_i|B)$. Pero esto no es cierto en general que $\mathbb{P}(A|B \cup C) = \mathbb{P}(A|B) + P(A|C)$.

Las reglas de la probabilidad se aplican a los eventos de la izquierda de la barra $|$. En general, no se cumple que $P(A|B) = P(B|A)$.

Hay mucha confusión con esto todo el tiempo. Por ejemplo, la probabilidad de que tengas puntos (manchas) dado que tienes sarampión es 1, pero la probabilidad de que tengas sarampión, dado que tienes puntos (manchas) no es 1. En este caso, la diferencia entre $\mathbb{P}(A|B)$ y $\mathbb{P}(B|A)$ es obvia, pero hay casos en los que es menos obvio. Este error sucede con bastante frecuencia en los casos legales que a veces se llaman *falacias del fiscal*.

Si $\mathbb{P}(A) > 0$ y $\mathbb{P}(B) > 0$, tenemos

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A).$$

Intuitivamente, $\mathbb{P}(A|B)$ es la probabilidad de que A ocurra suponiendo que ocurrió el evento B. De acuerdo con esa intuición, la probabilidad condicional tiene las siguientes propiedades

- Si $B \subset A$ entonces $\mathbb{P}(A|B) = \mathbb{P}(B)/\mathbb{P}(B) = 1$.
- Si $A \cap B = \emptyset$ entonces $\mathbb{P}(A|B) = 0/\mathbb{P}(B) = 0$.
- $A \subset B$ entonces $\mathbb{P}(A|B) = \mathbb{P}(A)/\mathbb{P}(B)$.
- La probabilidad condicional, puede ser vista como una función de probabilidad

  $$\mathbb{P}_A(E) = \mathbb{P}(E|A)$$

  y todas las propiedades y intuiciones que se aplican a la función probabilidad se aplican a $\mathbb{P}_A$ también.
- Asumiendo que el evento A ocurrió, $\mathbb{P}_A$ generalmente tiene mejores habilidades de pronóstico que $\mathbb{P}$.

Como se mencionó anteriormente, las probabilidades condicionales son usualmente intuitivas. El siguiente ejemplo de (Feller, 1968), sin embargo, muestra una situación contra-intuitiva que implica probabilidades condicionales. Esto demuestra que la intuición no debe ser un sustituto de la computación rigurosa.

Ejemplo 1.1 Consideramos dos familias con dos hijos donde la probabilidad de género de cada niño es simétrica $(1/2)$. Seleccionamos una familia al azar y consideramos el espacio muestral que describe el género de los niños $\Omega = \{MM, MF, FM, FF\}$. Asumimos un modelo clásico, implicando que las probabilidades de todos los eventos elementales son $1/4$.

Definimos el evento en que ambos hijos en la familia son niños como $A = \{MM\}$, el evento que una familia tiene un niño es $B = \{MF, FM, MM\}$ y el evento que el primer hijo es un niño como $C = \{MF, MM\}$.

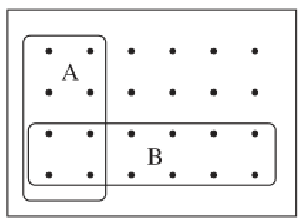
Dado que el primer hijo es un niño, la probabilidad de que ambos hijos sean niños es

$$\mathbb{P}(A|C) = \mathbb{P}(A \cap C)/\mathbb{P}(C) = \mathbb{P}(A)/\mathbb{P}(C) = (1/4)/(1/2) = 1/2.$$

Esto coincide con nuestra intuición. Dado que la familia tiene un niño, la probabilidad de que ambos hijos sean niños es contraintuitiva

$$\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B) = \mathbb{P}(A)/\mathbb{P}(B) = (1/4)/(3/4) = 1/3.$$

Ejemplo 1.2 Observamos en la siguiente figura, que $\mathbb{P}(A \cap B) = 1/6$, que $\mathbb{P}(B) = 1/2$ y que $\mathbb{P}(A|B) = (1/6)/(1/2) = 1/3$, como se esperaba. Sabemos que B ha ocurrido y que el evento A ocurrirá si uno de los cuatro resultados en $A \cap B$ se elige entre los 12 igualmente probable.



Podemos generalizar la ecuación de la probabilidad condicional a múltiples eventos. Sea $A_i, i = 1, 2, \ldots, k$ para k eventos para el cual $\mathbb{P}(A_1 \cap A_2 \cap \cdots A_k) > 0$. Entonces

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots A_k) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1}).$$

Ejemplo 1.3 En una clase de nivel de posgrado de primer año de 60 estudiantes, diez estudiantes son estudiantes de pregrado.

Calculemos la probabilidad de que tres estudiantes elegidos al azar sean estudiantes de pregrado. Para ello sea A_1 el evento en el cual el primer estudiante elegido es un estudiante de pregrado, A_2 el evento en que el segundo estudiante sea de pregrado y así sucesivamente. Luego de la ecuación anterior tenemos

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2)$$

reemplazando obtenemos

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \frac{10}{60} \times \frac{9}{59} \times \frac{8}{58} = 0.003507.$$

Ejemplo 1.4 Un test médico para una enfermedad D tiene salidas + y -

Salidas	D	D^c
+	.009	0.099
-	.001	0.891

Desde la definición de probabilidad condicional,

$$\mathbb{P}(+|D) = \frac{\mathbb{P}(+\cap D)}{\mathbb{P}(D)} = \frac{.009}{.009+.001} = .9$$

y

$$\mathbb{P}(-|D^c) = \frac{\mathbb{P}(-\cap D^c)}{\mathbb{P}(D^c)} = \frac{.891}{.891+.099} \approx .9$$

Al parecer, la prueba es bastante exacta. Personas enfermas producen un positivo del 90 por ciento de la veces y las personas sanas producen una negativa sobre 90 por ciento de las veces . Supongamos que una persona se hace una prueba y obtiene un resultado positivo. ¿ Cuál es la probabilidad de que esa persona tiene la enfermedad? Muchas personas responden que $0,90$, sin embargo la respuesta correcta es

$$\mathbb{P}(D|+) = \frac{\mathbb{P}(+\cap D)}{\mathbb{P}(+)} = \frac{.009}{.009+.099} \approx .08$$

Ejemplo 1.5 Una pequeña empresa de software emplea 4 programadores. El porcentaje de código escrito por cada programador y el porcentaje de errores en su código se muestran en la siguiente tabla. El código es seleccionado al azar.

Programador	% de código escrito	% de errores en código
1	15	5
2	20	3
3	25	2
4	40	1

1. Antes de que el código sea examinado, ¿ cuál es la probabilidad de que fue escrito por el

```
Programador 1? .15
Programador 2? .20
Programador 3? .25
Programador 4? .40
```

Aplicando la definición de probabilidad condicional tenemos

$$\mathbb{P}(\text{Prog1} \mid \text{no error}) = \frac{\mathbb{P}(\text{Prog1} \cap \text{no error})}{\mathbb{P}(\text{no error})}$$
$$\mathbb{P}(\text{Progr1} \cap \text{no error}) = 0.15 * 0.95 = 0.1425$$

$\mathbb{P}$(no error)

```
> 0.15*0.95 + 0.20 *0.97 + 0.25 * 0.98 + 0.40 * 0.99

[1] 0.9775
```

$\mathbb{P}(\text{Progr1}|\text{no error}) = 0.1425/0.9775 = 0.1457428.$

$\mathbb{P}(\text{Progr2}|\text{no error})$

```
> 0.20*0.97/0.97775

[1] 0.1984147
```

$\mathbb{P}(\text{Progr3}|\text{no error})$

```
> 0.25*0.98/0.97775

[1] 0.2505753
```

$\mathbb{P}(\text{Progr4}|\text{no error})$

```
> 0.40*0.99/0.97775

[1] 0.4050115
```

2 Eventos independientes

Definición 2.1 Dos eventos A y B son independientes si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Para tres eventos, A, B y C definimos los eventos independientes si

$$\begin{aligned} \mathbb{P}(A \cap B) &= \mathbb{P}(A)\mathbb{P}(B) \\ \mathbb{P}(A \cap C) &= \mathbb{P}(A)\mathbb{P}(C) \\ \mathbb{P}(B \cap C) &= \mathbb{P}(A)\mathbb{P}(C) \;\; \text{y} \\ \mathbb{P}(A \cap B \cap C) &= \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) \end{aligned}$$

Las tres primeras de estas condiciones establecen que los eventos son independientes dos a dos , por lo que llamamos a estos eventos que satisfacen estas tres condiciones como `independientes por pares`. El ejemplo siguiente mostrará que los eventos que satisfacen estas tres condiciones pueden no satisfacer la cuarta condición, de modo que la independencia por pares no determina la independencia.

Para un número finito de eventos $A_1, A_2, \ldots, A_n$ son independientes si

$$\mathbb{P}(A_1 \cap \cdots A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n)$$

y son independientes por pares si cada par $A_i, A_j, i \neq j$ son independientes.

Si generalizamos un poco, dado un conjunto de eventos $\{A_i : i \in I\}$ estos son independientes si

$$\mathbb{P}\Big(\bigcap_{i \in J} A_i\Big) = \prod_{i \in J} \mathbb{P}(A_i)$$

para cada subconjunto J de I.

Existe cierta confusión entre eventos independientes y eventos mutuamente exclusivos. A menudo se habla de estos como `no tener ningún efecto sobre el otro`, pero eso no es una caracterización precisa en ambos casos. Debes tener en cuenta que aunque los eventos mutuamente exclusivos no pueden ocurrir juntos, los eventos independientes deben ser capaces de ocurrir juntos.

Supongamos que ni $\mathbb{P}(A)$ ni $\mathbb{P}(B)$ es 0 y que A y B son mutualmente exclusivos. Entonces $\mathbb{P}(A \cap B) = 0 \neq \mathbb{P}(A)\mathbb{P}(B)$. Así A y B no son independientes. Esto es equivalente a la declaración que si A y B no pueden ser mutualmente exclusivos.

Ejemplo 2.1 Una moneda es lanzada 4 veces. Consideremos los eventos A donde la primera moneda muestra cara; B la tercera moneda muestra sello; y C hay un número igual de caras y sellos. ¿ Esos son eventos independientes ?.

Supongamos que el espacio muestral consiste de 16 puntos que muestran los lanzamientos. El espacio muestral, indicando los eventos que ocurren en cada punto, es como sigue

Puntos	Eventos
C C C C	A
C C C S	A
C C S C	A,B
S C C C	
C S C C	A
C C S S	A,B ,C
C S C S	A,C
S C C S	C
S C S C	B,C
C S S C	A,B,C
S S C C	C
S S S C	B
S S C S	
S C S S	B
C S S S	A,B
S S S S	B

Entonces $\mathbb{P}(A) = 1/2$ y $\mathbb{P}(B) = 1/2$, mientras que C consiste de 6 puntos con exactamente dos caras y dos sellos. Así $P(C) = 6/16 = 3/8$. Ahora $\mathbb{P}(A \cap B) = \dfrac{4}{16} = \dfrac{1}{4} = \mathbb{P}(A)\mathbb{P}(B); \mathbb{P}(A \cap C) = \dfrac{3}{16} = \mathbb{P}(A)\mathbb{P}(C)$ y $\mathbb{P}(B \cap C) = \dfrac{3}{16} = \mathbb{P}(B)\mathbb{P}(C)$. Por tanto los eventos A, B y C son independientes por pares.

El evento $A \cap B \cap C$ consiste de dos puntos $CSSC$ y $CCSS$ con probabilidad $2/16 = 1/8$. Así $\mathbb{P}(A \cap B \cap C) \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(B)$ y así A, B y C no son independientes.

La siguiente definición generaliza la independencia de una colección arbitraria de eventos, indexados por un conjunto (potencialmente infinito) Θ.

Definición 2.2 Múltiples eventos $A_\theta, \theta \in \Theta$ son independientes por pares si cada par de eventos es independiente. Múltiples eventos $A_\theta, \theta \in \Theta$ son independientes si para cada $k > 0$ y para cada k-subconjunto de distintos eventos $A_{\theta_1}, \ldots, A_{\theta_k}$, tenemos

$$\mathbb{P}(A_{\theta_1} \cap \ldots \cap A_{\theta_k}) = \mathbb{P}(A_{\theta_1}) \cdots \mathbb{P}(A_{\theta_k})$$

La independencia puede surgir de dos maneras distintas. Por ejemplo al lanzar una moneda dos veces, asumimos que los lanzamientos son independientes, lo que se `traduce`, en que las monedas `no tienen memoria del primer lanzamiento`. En otro caso, derivamos la independencia verificando $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Por ejemplo, en el lanzamiento de un dado, sea $A = \{2, 4, 6\}$ y sea $B = \{1, 2, 3, 4\}$, entonces

$A \cap B = \{2,4\}, \mathbb{P}(A \cap B) = 2/6 = \mathbb{P}(A)\mathbb{P}(B) = (1/2)(2/3)$ y así A y B son independientes. En este casono asumimos que A y B son independientes.

Supongamos que A y B son eventos disjuntos, cada uno con probabilidad positiva. ¿ Pueden ser independientes? No. Esto se debe a que $\mathbb{P}(A)\mathbb{P}(B) > 0$ y $\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0$. Excepto en este caso especial, no hay manera de revisar la independencia mirando los conjuntos en un diagrama de Venn .

Ejemplo 2.2 Lanzamos una moneda 10 veces. Sea el evento $A =$ `al menos una cara`. Sea T_j el evento que salga sello en el j-ésimo lanzamiento. Entonces

$$\begin{aligned}
\mathbb{P}(A) &= 1 - \mathbb{P}(A^C) \\
&= 1 - \mathbb{P}(\text{todos los sellos}) \\
&= 1 - \mathbb{P}(T_1 T_2 \cdots T_{10}) \\
&= 1 - \mathbb{P}(T_1)\mathbb{P}(T_2) \cdots \mathbb{P}(T_{10}) \text{ usando independencia} \\
&= 1 - \left(\frac{1}{2}\right)^{10} \approx .999.
\end{aligned}$$

Ejemplo 2.3 Sean dos personas que se turnan para encestar una pelota de baloncesto. La persona 1 tiene éxito con probabilidad 1/3, la persona 2 tiene éxito con probabilidad 1/4. ¿ Cuál es la probabilidad de que la persona 1 tenga éxito antes de la persona 2?. Sea E el evento de interés. Sea A_j el evento en que el primer éxito es de la persona 1 y que se produce en el número de lanzamiento j. Los conjuntos A_i son disjuntos y forman una partición para E, así

$$\mathbb{P}(E) = \sum_{j=1}^{\infty} \mathbb{P}(A_j)$$

Ahora $\mathbb{P}(A_1) = 1/3$. A_2 ocurre si tenemos la secuencia persona 1 pierde, persona 2 pierde y la persona 1 tiene éxito. Luego $\mathbb{P}(A_2) = (2/3)(3/4)(1/3) = (1/2)(1/3)$. Siguiendo esa lógica tenemos que $\mathbb{P}(A_j) = (1/2)^{j-1}(1/3)$. Así

$$\mathbb{P}(E) = \sum_{j=1}^{\infty} \frac{1}{3}\left(\frac{1}{2}\right)^{j-1} = \frac{1}{3}\sum_{j=1}^{\infty}\left(\frac{1}{2}\right)^{j-1} = \frac{2}{3}.$$

Proposición 2.1 Si A y B son eventos independientes, entonces $\mathbb{P}(A|B) = \mathbb{P}(A)$. También para algún par de eventos A y B

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A)$$

Ejemplo 2.4 Se lanza dos cartas de una baraja, sin reemplazo. Sea A el evento de que el primera carta es el as de tréboles y sea B el evento de que la segunda corta es la reina de diamantes. Entonces $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A) = (1/52) \times (1/51)$.

```
> (1/51) * (1/52)

[1] 0.0003770739
```

Proposición 2.2 Si A, B son independientes, entonces son independientes los eventos A^c, B, los eventos A, B^c y los eventos A^c, B^c.

3 Referencias

1. Introduction to Probability and Statistics, Lecture 4 Bayes Law KC Border 2016.
2. Probability (chapter 1) All of Statistics A concise course in Statistical Inference Larry Wassermann Springer 2004.