

Tema 64

Probabilidad compuesta. Probabilidad condicionada. Probabilidad total. Teorema de Bayes

64.1 Probabilidad condicionada

Mediante un espacio probabilístico damos una formulación matemática al fenómeno aleatorio que estemos observando. Parece por tanto razonable que si observamos algo que aporte información a nuestro fenómeno aleatorio, ésta deba alterar el espacio probabilístico de partida.

Por ejemplo, la extracción de una bola de una urna con tres bolas blancas y dos negras, puede formalizarse con un espacio probabilístico en el que los sucesos elementales sean las cinco bolas y donde la probabilidad sea uniforme sobre estos cinco sucesos elementales, es decir, igual a $\frac{1}{5}$.

Si extraemos una bola de la urna, es decir, si observamos el suceso A bola negra, y no la devolvemos a la urna, es razonable que el espacio probabilístico cambie en el sentido no sólo de que ahora ya habrá únicamente cuatro sucesos elementales, sino que además la función de probabilidad deberá cambiar en orden a recoger la información que la observación del suceso A nos proporcionó.

Es decir, en el nuevo espacio probabilístico deberá hablarse de probabilidad condicionada al suceso A , de forma que se recojan hechos tan evidentes como que ahora la probabilidad (condicionada) de obtener negra se habrá reducido y habrá aumentado la de blanca.

Definición 1 Sea (E, \mathcal{A}, P) un espacio de probabilidad y $A \subset E$ un suceso con $P(A) > 0$. Para cualquier otro suceso B se define la probabilidad de B

condicionada a A como:

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

Construyamos la σ -álgebra cuyo conjunto fundamental es A .

Proposición 2 $\mathcal{A}_A = \{A \cap B : B \in \mathcal{A}\} \subset \mathcal{A}$ es una σ -álgebra

Demostración:

(i) $\emptyset \in \mathcal{A}_A$, $E \in \mathcal{A}_A$

(ii) Veamos que si $C \in \mathcal{A}_A$ entonces $\bar{C} \in \mathcal{A}_A$

Como $C \in \mathcal{A}_A$ entonces $C = A \cap B$ y por tanto $\bar{C} = A \cap \bar{B}$, luego $\bar{C} \in \mathcal{A}_A$

(iii) Sea $\{C_k\}_{k=1}^{\infty} \subset \mathcal{A}_A$. Tenemos que demostrar que $\bigcup_{k=1}^{\infty} C_k \in \mathcal{A}_A$

Como cada $C_k \in \mathcal{A}_A$ se tiene que $C_k = A \cap B_k$ y por tanto

$$\bigcup_{k=1}^{\infty} C_k = \bigcup_{k=1}^{\infty} (A \cap B_k) = A \cap \left(\bigcup_{k=1}^{\infty} B_k \right)$$

luego como $B_k \in \mathcal{A} \Rightarrow \bigcup_{k=1}^{\infty} B_k \in \mathcal{A} \Rightarrow \left(A \cap \left(\bigcup_{k=1}^{\infty} B_k \right) \right) = \bigcup_{k=1}^{\infty} C_k \in \mathcal{A}_A$.

C.Q.D. \square

Veamos ahora que la probabilidad condicionada es una auténtica probabilidad.

Proposición 3 La función de conjunto

$$P_A : \mathcal{A}_A \rightarrow [0, 1]$$

$$C = A \cap B \longrightarrow P_A(C) = \frac{P(C)}{P(A)} = \frac{P(A \cap B)}{P(A)}$$

es una medida de probabilidad.

Demostración:

(1) $\forall C \in \mathcal{A}_A \quad i \quad P_A(C) \geq 0$?

Como $A \cap B \subset A$ tenemos que $0 \leq P(A \cap B) \leq P(A)$, luego dividiendo esta desigualdad por $P(A)$ resulta:

$$0 \leq \frac{P(A \cap B)}{P(A)} \leq \frac{P(A)}{P(A)} = 1$$

es decir, $0 \leq P_A(C) \leq 1$

(2) $i \quad P_A(A) = 1$?

$$P_A(A) = \frac{P(A \cap A)}{P(A)} = 1$$

(3) $i \quad P_A\left(\bigcup_{k=1}^{\infty} C_k\right) = \sum_{k=1}^{\infty} P_A(C_k) \quad \forall C_k \cap C_j = \emptyset \text{ con } k \neq j$?

Tomamos $\bigcup_{k=1}^{\infty} C_k = A \cap \left(\bigcup_{k=1}^{\infty} B_k \right)$. Entonces:

$$\begin{aligned} P_A \left(\bigcup_{k=1}^{\infty} C_k \right) &= P \left(\bigcup_{k=1}^{\infty} C_k / A \right) = \frac{P \left(A \cap \left(\bigcup_{k=1}^{\infty} B_k \right) \right)}{P(A)} = \frac{P \left(\bigcup_{k=1}^{\infty} (A \cap B_k) \right)}{P(A)} = \\ &= \frac{\sum_{k=1}^{\infty} P(A \cap B_k)}{P(A)} = \sum_{k=1}^{\infty} \frac{P(A \cap B_k)}{P(A)} = \sum_{k=1}^{\infty} P_A(A \cap B_k) = \sum_{k=1}^{\infty} P_A(C_k) \quad C.Q.D. \square \end{aligned}$$

Corolario 4 (E, \mathcal{A}_A, P_A) es un espacio de probabilidad.

64.2 Fórmula de la probabilidad compuesta

Teorema 5 (Fórmula de la probabilidad compuesta): Sean A_1, \dots, A_n n -sucesos ordenados. Entonces:

$$P \left(\bigcap_{k=1}^n A_k \right) = P(A_1) P(A_2 / A_1) P(A_3 / A_1 \cap A_2) \dots P(A_n / A_1 \cap \dots \cap A_{n-1})$$

Demostración:

Lo demostraremos por inducción sobre n :

$n = 2$:

$$P_{A_1}(A_2) = \frac{P(A_1 \cap A_2)}{P(A_1)} = P(A_2 / A_1) \Rightarrow$$

$$\Rightarrow P(A_1 \cap A_2) = P(A_1) P(A_2 / A_1)$$

$n = 3$:

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1) P(A_2 \cap A_3 / A_1) = \\ &= P(A_1) \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1)} = P(A_1) \frac{P(A_1 \cap A_2) P(A_3 / A_1 \cap A_2)}{P(A_1)} = \\ &= P(A_1) P(A_2 / A_1) P(A_3 / A_1 \cap A_2) \end{aligned}$$

Hipótesis de inducción: lo suponemos cierto para $n - 1$, es decir:

$$\begin{aligned} &P \left(\bigcap_{k=1}^{n-1} A_k \right) = \\ &= P(A_1) P(A_2 / A_1) P(A_3 / A_1 \cap A_2) \dots P(A_{n-1} / A_1 \cap \dots \cap A_{n-2}) \end{aligned}$$

Vamos a demostrarlo para n :

$$\begin{aligned}
 P(A_1 \cap \dots \cap A_{n-1} \cap A_n) &= P(A_1) P(A_1 \cap \dots \cap A_n / A_1) = \\
 &= P(A_1) \frac{P(A_1 \cap \dots \cap A_n)}{P(A_1)} = P(A_1) \frac{P(A_1 \cap \dots \cap A_{n-1}) P\left(A_n / \bigcap_{j=1}^{n-1} A_j\right)}{P(A_1)} = \\
 &= P(A_1) P(A_2 / A_1) P(A_3 / A_1 \cap A_2) \dots P(A_n / A_1 \cap \dots \cap A_{n-1}) \quad C.Q.D. \square
 \end{aligned}$$

64.3 Independencia aleatoria

Definición 6 Sean $A, B \in \mathcal{A}$ dos sucesos de probabilidades no nulas. Se dice que el suceso B es independiente del suceso A (independiente estocásticamente o aleatoriamente) sii

$$P(B/A) = P(B)$$

Se da frecuentemente el caso de que se confunden sucesos incompatibles con sucesos independientes. Observemos para evitar dicha confusión que los sucesos incompatibles son los más dependientes que existen, puesto que la ocurrencia de uno de ellos proporciona la máxima información: el otro suceso no va a ocurrir.

Vamos a caracterizar la independencia:

Teorema 7 Una condición necesaria y suficiente para que B sea independiente de A es que $P(A \cap B) = P(A)P(B)$.

Demostración:

\Rightarrow) Como B es independiente de A se tiene que $P(B/A) = P(B)$ y por tanto:

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = P(B) \Rightarrow P(A \cap B) = P(A)P(B)$$

\Leftarrow) Como $P(A \cap B) = P(A)P(B)$ resulta:

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B) \quad C.Q.D. \square$$

Enunciamos a continuación algunas propiedades relativas a la independencia estocástica de sucesos, pero antes necesitamos un par de resultados previos:

Lema 8 La independencia estocástica es una propiedad recíproca.

Demostración:

Sean A y B dos sucesos de probabilidades no nulas y supongamos que B es independiente de A . Entonces

$$P(B/A) = P(B)$$

Vamos a demostrar que A es independiente de B .

Se tiene que:

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

es decir, A es independiente de B . C.Q.D. \square

Como consecuencia de esta propiedad, si A es independiente de B se dirá que A y B son (estocásticamente) independientes.

La idea de independencia es que ninguno de los dos sucesos interviene en la verificación del otro.

Lema 9 Sean A y B dos sucesos con probabilidades no nulas. Entonces:

$$P(\bar{B}/A) = 1 - P(B/A)$$

Demostración:

$$\begin{aligned} P(\bar{B}/A) &= \frac{P(A \cap \bar{B})}{P(A)} \stackrel{(*)}{=} \frac{P(A - (A \cap B))}{P(A)} = \frac{P(A) - P(A \cap B)}{P(A)} = \\ &= 1 - \frac{P(A \cap B)}{P(A)} = 1 - P(B/A) \end{aligned}$$

donde en $(*)$ hemos tenido en cuenta que $A \cap \bar{B} \stackrel{(1)}{=} A - (A \cap B)$.

Vamos a demostrar la igualdad (1) : ¿ $A \cap \bar{B} = A - (A \cap B)$?

$$\begin{aligned} A - (A \cap B) &= A \cap (\overline{A \cap B}) = A \cap (\bar{A} \cup \bar{B}) = \\ &= (A \cap \bar{A}) \cup (A \cap \bar{B}) = \emptyset \cup (A \cap \bar{B}) = A \cap \bar{B} \end{aligned}$$

como queríamos demostrar. \square

Remark 10

$$P(B/A) \neq P(B/\bar{A})$$

Para demostrar la afirmación hecha en la nota anterior, vamos a demostrar el siguiente resultado.

Proposición 11 Sean A y B dos sucesos con probabilidades no nulas. Entonces:

$$P(B/A) = P(B/\bar{A}) \Leftrightarrow A \text{ y } B \text{ son independientes}$$

Demostración:

$$\begin{aligned}
P(B/A) = P(B/\bar{A}) &\Leftrightarrow \frac{P(A \cap B)}{P(A)} = \frac{P(\bar{A} \cap B)}{P(\bar{A})} \Leftrightarrow \\
&\Leftrightarrow P(\bar{A}) P(A \cap B) = P(\bar{A} \cap B) P(A) \Leftrightarrow \\
&\Leftrightarrow (1 - P(A)) P(A \cap B) = P(\bar{A} \cap B) P(A) \Leftrightarrow \\
P(A \cap B) - P(A) P(A \cap B) - P(\bar{A} \cap B) P(A) &= 0 \Leftrightarrow \\
&\Leftrightarrow P(A \cap B) - P(A) P(A \cap B) - P(B) P(A) - P(A \cap B) P(A) = 0 \Leftrightarrow \\
&\Leftrightarrow P(A \cap B) = P(A) P(B) \Leftrightarrow A \text{ y } B \text{ son independientes}
\end{aligned}$$

donde en (*) hemos tenido en cuenta que $\bar{A} \cap B = B - (A \cap B)$ y que

$$P(\bar{A} \cap B) = P(B - (A \cap B)) = P(B) - P(A \cap B)$$

ya que $A \cap B \subset B$. C.Q.D. \square

Proposición 12 Si $A, B \in \mathcal{A}$ son independientes, entonces también lo son:

- (1) A y \bar{B}
- (2) \bar{A} y B
- (3) \bar{A} y \bar{B}

Demostración:

(1) Sabemos que $P(A \cap B) = P(A) P(B)$ ya que A y B son independientes, y tenemos que demostrar que $P(A \cap \bar{B}) = P(A) P(\bar{B})$. Se tiene que:

$$\begin{aligned}
P(A \cap \bar{B}) &= P(A) P(\bar{B}/A) = P(A) [1 - P(B/A)] = \\
&= P(A) [1 - P(B)] = P(A) P(\bar{B})
\end{aligned}$$

es decir, A y \bar{B} son independientes.

(2) Se tiene que:

$$\begin{aligned}
P(\bar{A} \cap B) &= P(B) P(\bar{A}/B) = P(B) [1 - P(A/B)] = \\
&= P(B) [1 - P(A)] = P(B) P(\bar{A}) = P(\bar{A}) P(B)
\end{aligned}$$

es decir, \bar{A} y B son independientes.

(3) Se tiene que:

$$\begin{aligned}
P(\bar{A} \cap \bar{B}) &= P(\bar{A}) P(\bar{B}/\bar{A}) = P(\bar{A}) [1 - P(B/\bar{A})] = \\
&= P(\bar{A}) [1 - P(B)] = P(\bar{A}) P(\bar{B})
\end{aligned}$$

es decir, \bar{A} y \bar{B} son independientes. C.Q.D.□

Definimos ahora los conceptos necesarios para poder enunciar el teorema de la probabilidad total y como consecuencia el teorema de Bayes.

Definición 13 Se dice que los sucesos $A_1, \dots, A_n \in \mathcal{A}$ son totalmente independientes o mutuamente independientes si verifican:

- (•) $P(A_i \cap A_j) = P(A_i)P(A_j) \quad \forall i \neq j$
- (•) $P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k) \quad \forall i \neq j \neq k$
- ...
- (•) $P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i)$

Definición 14 Sea (E, \mathcal{A}, P) un espacio de probabilidad. Se llama sistema completo de sucesos a cualquier familia $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$ que verifique:

- (a) $\forall n \neq m \quad P(A_n \cap A_m) = 0$
- (b) $\sum_{n=1}^{\infty} P(A_n) = 1$

Proposición 15 Sea (E, \mathcal{A}, P) un espacio de probabilidad y $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$ una partición de E , es decir,

- (i) $A_n \cap A_m = \emptyset \quad \forall n \neq m$
- (ii) $\bigcup_{n=1}^{\infty} A_n = E$

Entonces, $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$ es un sistema completo de sucesos.

Demostración:

- (a) Como $A_n \cap A_m = \emptyset \Rightarrow P(A_n \cap A_m) = P(\emptyset) = 0 \quad \forall n \neq m$
- (b) Se tiene que $\bigcup_{n=1}^{\infty} A_n = E$ y por tanto:

$$\sum_{n=1}^{\infty} P(A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = P(E) = 1 \quad \text{C.Q.D.}\square$$

Teorema 16 (Teorema de la probabilidad total): Sea $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A}$ una partición de E , es decir, una sucesión tal que

- (i) $A_n \cap A_m = \emptyset \quad \forall n \neq m$
- (ii) $\bigcup_{n=1}^{\infty} A_n = E$

y supongamos que $P(A_n) > 0 \quad \forall n \in \mathbb{N}$. Entonces, $\forall B \in \mathcal{A}$ se tiene:

$$P(B) = \sum_{n=1}^{\infty} P(A_n)P(B/A_n)$$

Demostración:

Sea $B = B \cap E = B \cap \left(\bigcup_{n=1}^{\infty} A_n \right) = \bigcup_{n=1}^{\infty} (A_n \cap B)$. Entonces:

$$P(B) = P\left(\bigcup_{n=1}^{\infty} (A_n \cap B)\right) = \sum_{n=1}^{\infty} P(A_n \cap B) \stackrel{(*)}{=} \sum_{n=1}^{\infty} P(A_n) P(B/A_n)$$

donde en (*) hemos usado la fórmula de la probabilidad compuesta. C.Q.D.□

Corolario 17 (Teorema de Bayes o de la probabilidad inversa, 1764):

Sea $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A}$ una partición de E , es decir, una sucesión tal que

$$(i) A_n \cap A_m = \emptyset \quad \forall n \neq m$$

$$(ii) \bigcup_{n=1}^{\infty} A_n = E$$

y supongamos que $P(A_n) > 0 \quad \forall n \in \mathbb{N}$. Entonces, $\forall B \in \mathcal{A}$ se tiene:

$$P(A_n/B) = \frac{P(A_n) P(B/A_n)}{\sum_{n=1}^{\infty} P(A_n) P(B/A_n)}$$

Demostración:

$$P(A_n/B) = \frac{P(A_n \cap B)}{P(B)} \stackrel{(*)}{=} \frac{P(A_n) P(B/A_n)}{\sum_{n=1}^{\infty} P(A_n) P(B/A_n)}$$

donde en (*) hemos usado el teorema de la probabilidad total. C.Q.D.□

Las probabilidades $P(A_n)$ se llaman probabilidades a priori, las probabilidades $P(B/A_n)$ se llaman verosimilitudes y las probabilidades $P(A_n/B)$ se llaman probabilidades a posteriori.

Así, el teorema de Bayes consiste en calcular las probabilidades a posteriori, conocidas las probabilidades a priori.

Acerca de la interpretación del teorema de Bayes, si los A_n son considerados como hipótesis sobre alguna cuestión en estudio y tenemos una cierta información, subjetiva o no, acerca de dichas hipótesis A_n , reflejada ésta por las $P(A_n) > 0$, supongamos se produce un suceso B asociado a un experimento aleatorio (usualmente provocado por el experimentador para obtener más información), el teorema de Bayes nos proporciona el medio para obtener la alteración, provocada por B , en nuestras hipótesis A_n , dándonos las probabilidades a posteriori $P(A_n/B)$. Usualmente se continuará el proceso provocando otro suceso C (es decir, obteniendo otro nuevo resultado de nuestro experimento aleatorio considerado) y utilizando ahora las $P(A_n/B)$ antes calculadas, como probabilidades a priori.