Teoría Bayesiana

Julián Jiménez-Cárdenas[1]

[1]Universidad Nacional de Colombia, Bogotá.

juojimenezca@unal.edu.co

σ-álgebra I

Definición (Experimento Aleatorio)

Un experimento se dice aleatorio si su resultado no se puede determinar de antemano.

Definición (Espacio de Muestra)

El conjunto Ω de todos los posibles resultados de un experimento aleatorio se llama espacio de muestra. Un elemento $\omega \in \Omega$ se llama resultado o muestra. Ω se dice discreto si es finito o contable.

σ-álgebra II

Definición (σ-álgebra)

Tome $\Omega \neq \emptyset$. Una colección $\Im$ de subconjuntos de Ω se llama $\sigma-$álgebra sobre Ω si:

1. $\Omega \in \Im$,
2. *Si $A \in \Im$, entonces $A^c \in \Im$ y,*
3. *Si $A_1, A_2, \cdots \in \Im$, entonces $\bigcup_{i=1}^{\infty} A_i \in \Im$.*

Los elementos de $\Im$ se llaman eventos.

Teorema

Si $\Omega \neq \emptyset$ y $\Im_1, \Im_2, \ldots$ son $\sigma-$álgebras sobre Ω, entonces $\bigcap_{i=1}^{\infty} \Im_i$ es una σ-álgebra sobre Ω.

σ-álgebra III

Demostración.

Como $\Omega \in \Im_j$, para $j = 1, 2, \dots$, $\Omega \in \bigcap_{j=1}^{\infty} \Im_j$. Si $A \in \bigcap_{j=1}^{\infty} \Im_j$, $A \in \Im_j$, para $j = 1, 2, \dots$, de modo que $A^c \in \Im_j$, y $A^c \in \bigcap_{j=1}^{\infty} \Im_j$. Por último, si

$$A_1, A_2, \dots \in \bigcap_{j=1}^{\infty} \Im_j,$$

para todo $j = 1, 2, \dots$, $A_1, A_2, \dots \in \Im_j$, de modo que

$$\bigcup_{i=1}^{\infty} A_i \in \Im_j \text{ y } \bigcup_{i=1}^{\infty} A_i \in \bigcap_{j=1}^{\infty} \Im_j.$$

$\square$

σ-álgebra IV

Definición (σ-álgebra generada)

Tome $\Omega \neq \emptyset$ y $\mathcal{A}$ como una colección de subconjuntos de Ω. Si $\mathcal{M} := \{\Im : \Im$ es una $\sigma -$ álgebra sobre Ω que contiene a $\mathcal{A}\}$,

$$\sigma(\mathcal{A}) := \bigcap_{\Im \in \mathcal{M}} \Im$$

es la $\sigma-$álgebra más pequeña sobre Ω que contiene a $\mathcal{A}$. Esta $\sigma-$álgebra se conoce como σ-álgebra generada por $\mathcal{A}$.

Definición (Espacio de medida)

Tome $\Omega \neq \emptyset$ y sea $\Im$ una σ-álgebra sobre Ω. La pareja $(\Omega, \Im)$ se llama espacio de medida.

σ-álgebra V

$\emptyset$ es el evento imposible. Ω es el evento seguro y $\{\omega\}$, con $\omega \in \Omega$ es un evento simple. Decimos que el evento A ocurre después de llevar a cabo el experimento aleatorio si se obtiene un resultado en A, esto es, A ocurre si el resultado es algún $\omega \in A$.

1. El evento $A \cup B$ ocurre si y sólo si A ocurre, B pasa, o ambos ocurren.
2. El evento $A \cap B$ ocurre si y sólo si A y B ocurren a la vez.
3. El evento A^c ocurre si y sólo si A no ocurre.
4. El evento $A - B$ ocurre si y sólo si A ocurre pero B no ocurre.

Definición (Eventos mutuamente excluyentes)

Dos eventos A y B se dicen mutuamente excluyentes si $A \cap B = \emptyset$.

Espacio de probabilidad I

Definición (Frecuencia relativa)

Para cada evento A, el número $f_r(A) := \frac{n(A)}{n}$ se llama la frecuencia relativa de A, donde $n(A)$ indica el número de veces que ocurre A en n repeticiones del experimento aleatorio.

Cuando $n \to \infty$, se puede hablar de la probabilidad de que ocurra el evento A, normalizada de 0 a 1. La formalización de este concepto se encuentra en la idea del espacio de probabilidad.

Espacio de probabilidad II

Definición (Espacio de probabilidad)

Tome $(\Omega, \Im)$ *como un espacio de medida. Una función real* P *sobre* $\Im$ *que satisface las siguientes condiciones:*

1. $P(A) \geq 0$ *para todo* $A \in \Im$ *(no negativa),*
2. $P(\Omega) = 1$ *(normalizada) y,*
3. *si* $A_1, A_2, \ldots$ *son eventos mutuamente excluyentes en* $\Im$*, esto es, si*

$$A_i \cap A_j = \emptyset \text{ para todo } i \neq j, \text{ entonces}$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i),$$

se llama medida de probabilidad sobre $(\Omega, \Im)$*. La tripleta* $(\Omega, \Im, P)$ *se llama espacio de probabilidad.*

Espacio de probabilidad III

Teorema

Si $(\Omega, \Im, P)$ *es un espacio de probabilidad, entonces*

1. $P(\emptyset) = \emptyset$.
2. *Si* $A, B \in \Im$ *y* $A \cap B = \emptyset$, *entonces* $P(A \cup B) = P(A) + P(B)$.
3. *Para todo* $A \in \Im$, $P(A^c) = 1 - P(A)$.
4. *Si* $A \subseteq B$, *entonces* $P(A) \leq P(B)$ *y* $P(B - A) = P(B) - P(A)$. *En particular,* $P(A) \leq 1$ *para todo* $A \in \Im$.
5. *Para todo* $A, B \in \Im$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Espacio de probabilidad IV

Teorema

6. *Tome* $\{A_n\}_n \subseteq \Im$ *como una sucesión creciente, esto es,* $A_n \subseteq A_{n+1}, \forall n \in \mathbb{N}$; *entonces*

$$P\left(\lim_{m\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n), \; \textit{donde} \; \lim_{n\to\infty} A_n := \bigcup_{i=1}^{\infty} A_i.$$

7. *Tome* $\{A_n\}_n \subseteq \Im$ *como una sucesión decreciente, esto es,* $A_n \supseteq A_{n+1}, \forall n \in \mathbb{N}$; *entonces*

$$P\left(\lim_{m\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n), \; \textit{donde} \; \lim_{n\to\infty} A_n := \bigcap_{i=1}^{\infty} A_i.$$

Espacio de probabilidad V

Demostración.

1. $1 = P(\Omega \cup \emptyset \cup \emptyset \cup \cdots) = P(\Omega) + P(\emptyset) + P(\emptyset) + \cdots = 1 + P(\emptyset) + \cdots \implies P(\emptyset) = 0$.
2. $P(A \cup B) = P(A \cup B \cup \emptyset \cup \emptyset \cup \cdots) = P(A) + P(B)$.
3. $P(A) + P(A^c) = P(A \cup A^c) = P(\Omega) = 1 \implies P(A^c) = 1 - P(A)$.
4. Si $A \subseteq B$, $B = A \cup (B - A)$, de modo que $P(B) = P(A) + P(B - A)$. Como $P \geq 0$, $P(B) \geq P(A)$ y $P(B - A) = P(B) - P(A)$. Si $B = \Omega$, $P(A) \leq 1$.
5. Use el hecho de que $A \cup B = [A - (A \cap B)] \cup [B - (A \cap B)] \cup [A \cap B]$.

□

Espacio de probabilidad VI

Demostración.

6. Tome la sucesión $C_1 = A_1, C_2 = A_2 - A_1, \ldots, C_r = A_r - A_{r-1}, \ldots$. Es claro que

$$\bigcup_{i=1}^{\infty} C_i = \bigcup_{i=1}^{\infty} A_i.$$

Más aún, como $C_i \cap C_j = \emptyset \ \forall i \neq j$, se sigue que

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{n=1}^{\infty} C_n\right) = \sum_{n=1}^{\infty} P(C_n) = \lim_{n\to\infty} \sum_{k=1}^{n} P(C_k)$$

$$= \lim_{n\to\infty} P\left(\bigcup_{k=1}^{n} C_k\right) = \lim_{n\to\infty} P\left(A_n\right).$$

7. Tome la sucesión $\{B_n = A_n^c\}_n$ y aplique el resultado anterior.

Notas I

Aplicando el teorema anterior de forma inductiva, para algunos eventos $A_1, A_2, \ldots, A_n \in \Im$:

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots \cup A_n) = &\sum_{i=1}^{n} P(A_i) - \sum_{i_1<i_2} P(A_{i_1} \cap A_{i_2}) + \cdots \\ &+ (-1)^{r+1} \sum_{i_1<i_2<\cdots<i_r} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r}) \\ &+ \cdots + (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots \cap A_n). \end{aligned}$$

Tome $(\Omega, \Im, P)$ como un espacio de probabilidad con Ω finito o contable y $\Im = \mathbb{P}(\Omega)$. Tome $\emptyset \neq A \in \Im$. Es claro que

$$A = \bigcup_{\omega \in A} \{\omega\}, \text{ de modo que}$$

Notas II

$$P(A) = \sum_{\omega \in A} P(\omega), \text{ donde } P(\omega) := P(\{\omega\}).$$

Así, P queda completamente definido por $p_j := P(\omega_j)$, donde $\omega_j \in \Omega$. El vector $|\Omega|-$dimensional $p := (p_1, p_2, \dots)$ satisface las siguientes condiciones:

- $p_j \geq 0$ y
- $\sum_{j=1}^{\infty} p_j = 1$.

Un vector que satisface las anteriores condiciones se llama **vector de probabilidad**.

Introducción

Tome B como un evento cuya opción de ocurrir debe ser medida bajo la suposición de que otro evento A fue observado. Si el experimento se repite n veces bajo las mismas circunstancias, entonces la frecuencia relativa de B bajo la condición A se define como

$$f_r(B|A) := \frac{n(A \cap B)}{n(A)} = \frac{\frac{n(A \cap B)}{n}}{\frac{n(A)}{n}} = \frac{f_r(A \cap B)}{f_r(A)}, \text{ si } n(A) > 0.$$

Esto motiva la siguiente definición

Probabilidad Condicional I

Definición (Probabilidad condicional)

Tome $(\Omega, \Im, P)$ como un espacio de probabilidad. Si $A, B \in \Im$, con $P(A) > 0$, entonces la probabilidad del evento B bajo la condición A se define como sigue

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

El siguiente teorema provee algunas propiedades de la probabilidad condicional.

Probabilidad Condicional II

Teorema (Medida de probabilidad condicional)

Tome $(\Omega, \Im, P)$ *como un espacio de probabilidad y* $A \in \Im$, *con* $P(A) > 0$. *Entonces:*

1. $P(\cdot|A)$ *es una medida de probabilidad sobre* Ω *centrada en* A, *esto es,* $P(A|A) = 1$.
2. *Si* $A \cap B = \emptyset$, *entonces* $P(B|A) = 0$.
3. $P(B \cap C|A) = P(B|A \cap C)P(C|A)$ *si* $P(A \cap C) > 0$.
4. *Si* $A_1, A_2, \ldots, A_n \in \Im$, *con* $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$, *entonces*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Probabilidad Condicional III

Demostración.

1. Las tres propiedades de una medida de probabilidad deben ser verificadas.
    1. Claramente, $P(B|A) \geq 0$ para todo $B \in \Im$.
    2. $P(\Omega|A) = \frac{P(\Omega \cap A}{P(A)} = \frac{P(A)}{P(A)} = 1$. También se tiene que $P(A|A) = 1$.
    3. Tome $A_1, A_2, \cdots \in \Im$ una sucesión de conjuntos disyuntos. Entonces

$$P\left(\bigcup_{i=1}^{\infty} A_i|A\right) = \frac{P\left(A \cap \bigcup_{i=1}^{\infty} A_i\right)}{P(A)} = \frac{P\left(\bigcup_{i=1}^{\infty} A \cap A_i\right)}{P(A)}$$

$$= \sum_{i=1}^{\infty} \frac{P(A \cap A_i)}{P(A)} = \sum_{i=1}^{\infty} P(A_i|A).$$

□

Probabilidad Condicional IV

Demostración.

2. Si $A \cap B = \emptyset$, $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(\emptyset)}{P(A)} = 0$.
3. $P(B \cap C|A) = \frac{P(A \cap B \cap C)}{P(A)} = \frac{P(B \cap C \cap A)}{P(A \cap C)} \frac{P(C \cap A)}{P(A)} = P(B|A \cap C)P(C|A)$.
4. $P(A_1 \cap \cdots \cap A_n) = \frac{P(A_1 \cap \cdots \cap A_n)}{P(A_1 \cap \cdots \cap A_{n-1})} \frac{P(A_1 \cap \cdots \cap A_{n-1})}{P(A_1 \cap \cdots \cap A_{n-2})} \cdots \frac{P(A_1 \cap A_2)}{P(A_1)} P(A_1) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$.

$\square$

Los siguientes resultados son vitales para aplicaciones posteriores.

Probabilidad Condicional V

Teorema (Teorema de probabilidad total)

Tome $A_1, A_2, \ldots$ *como una partición finita o contable de* Ω, *esto es,* $A_i \cap A_j = \emptyset, \forall i \neq j$ *y* $\bigcup_{i=1}^{\infty} A_i = \Omega$, *tal que* $P(A_i) > 0$, *para todo* $A_i \in \Im$. *Entonces, para todo* $B \in \Im$*:*

$$P(B) = \sum_i P(B|A_i)P(A_i).$$

Probabilidad Condicional VI

Demostración.

Observe que

$$B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{\infty} A_i \right) = \bigcup_{i=1}^{\infty} B \cap A_i,$$

de modo que

$$P(B) = P\left(\bigcup_{i=1}^{\infty} B \cap A_i \right) = \sum_{i=1}^{\infty} P(B \cap A_i) = \sum_{i=1}^{\infty} P(B|A_i)P(A_i).$$

$\square$

Probabilidad Condicional VII

Como corolario del teorema anterior, se obtiene un resultado conocido como **regla de Bayes**, que constituye la base para la **teoría Bayesiana**.

Corolario (Regla de Bayes)

Tome $A_1, A_2, \ldots$ *como una partición finita o contable de* Ω *con* $P(A_i) > 0$*, para todo* i*; entonces, para todo* $B \in \Im$ *con* $P(B) > 0$:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_j P(B|A_j)P(A_j)}, \forall i.$$

Probabilidad Condicional VIII

Demostración.

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_j P(B|A_j)P(A_j)}.$$

Con la partición $A_1 = A, A_2 = A^c$ se obtiene la forma usual de la regla de Bayes. □

Definición (Distribuciones a priori y a posteriori)

Tome $A_1, A_2, \ldots$ como una partición finita o contable de Ω, con $P(A_i) > 0$, para todo i. Si $P(B) > 0$, con $B \in \Im$, entonces $\{P(A_n)\}_n$ se llama distribución a priori (antes de que B ocurra), y $\{P(A_n|B)\}_n$ se llama distribución a posteriori (después de que B ocurra).

Probabilidad Condicional IX

Algunas veces, la ocurrencia de un evento B no afecta la probabilidad de un evento A, es decir,

$$P(A|B) = P(A).$$

En este caso, se dice que el evento A es independiente del evento B. Esto motiva la siguiente definición.

Definición (Eventos independientes)

Dos eventos A y B se dicen independientes si y sólo si

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

Si esta condición no se tiene, se dice que los eventos son dependientes.

Referencias

Arunachalam V. Dharmaraja S. Blanco, L.
Introduction to Probability and Stochastic Processes with Applications.
Wiley, 2012.