TEMA 64: PROBABILIDAD COMPUESTA. PROBABILIDAD CONDICIONADA. PROBABILIDAD TOTAL. TEOREMA DE BAYES

Índice

1. INTRODUCCIÓN

El origen del estudio de la teoría de probabilidad se construyó sobre los juegos de azar, como se puede observar de la famosa correspondencia entre B. Pascal y P. Fermat.

En el s. XVIII T. Bayes desarrolló un marco en el que se podía actualizar las probabilidades de ocurrencia de un suceso de acuerdo a nueva información, sentando las bases del razonamiento Bayesiano. Este, hoy en día se aplica en estadística, inteligencia artificial ...

Para el desarrollo de este tema usaremos los conceptos del tema 63. Además, la importancia del ~~&~~ mismo en el contexto educativo es vital para nuestros alumnos, puesto que permite entender que "la información es poder". En otras palabras, la nueva información tiene un impacto en las probabilidades de un suceso y, por tanto, afecta nuestra toma de decisiones.

2. Probabilidad condicionada

Imaginemos que participamos en una rifa compuesta por 100 boletos. Entonces, si A

$$A := \{\text{ganamos el premio}\},$$

es claro que $P(A) = \frac{1}{100}$

Sin embargo, nos dicen por la mañana del día que se muestra el boleto ganador, que dicho boleto acaba en 5. Entonces,

¿ tenemos las mismas probabilidades de ganar ?

Así, razonamos

- Si nuestro boleto NO acaba en 5 entonces

$$P(A \mid \underset{\text{nuestro boleto}}{\text{no acaba en 5}}) = 0$$

- Sí, nuestro boleto SÍ acaba en 5 entonces

$$P(A \mid \text{sí acaba en } 5) = \frac{1}{10}$$

~~Obser~~ Supongamos que nuestro boleto es el 35. Entonces

$$A := \{\text{gana el premio el boleto } 35\}$$

$$B := \{\text{el boleto ganador acaba en } 5\}$$

$$P(A \mid B) = \frac{1}{10} \neq \frac{\overset{\cancel{\text{Además}}}{P(A \cap B)}}{P(B)} = \frac{1/100}{1/10} = \frac{1}{10}$$

Este ejemplo motiva la siguiente definición

Def^n.- Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad con A y B dos sucesos con $P(B) > 0$ entonces se define la probabilidad de A condicionada a B mediante la expresión

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1)$$

Nota.- Es fácil ver que se le llama probabilidad, puesto que cumple los 3 axiomas de Kolmogorov.

En la siguiente sección veremos los resultados más importantes de este tema.

3. TEOREMAS

3.1. TEOREMA DE FACTORIZACIÓN

A partir de la definición de probabilidad condicionada, podemos deducir por (1) que

$$P(A \cap B) = P(A|B) \cdot P(B) \quad (2)$$

El Teorema de Factorización extiende este resultado para cualquier intersección FINITA ! de sucesos.

$\boxed{\text{TEOREMA}}$ Consideremos $A_1, A_2, \ldots, A_n$ tales que $P(\bigcap_{i=1}^{n} A_i) > 0$, entonces

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P\left(A_n \Big| \bigcap_{i=1}^{n-1} A_i\right) \cdot P\left(A_{n-1} \Big| \bigcap_{i=1}^{n-2} A_i\right) \ldots P(A_1)$$

$\boxed{D}$ La demostración es sencilla por inducción.
El caso inicial si $\boxed{n=2}$ lo tenemos en (2).
Supongamos cierto para $n-1 \Rightarrow n$.

Veamos que

$$P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right) = \frac{P\left(\bigcap_{i=1}^{n} A_i\right)}{P\left(\bigcap_{i=1}^{n-1} A_i\right)}, \text{ entonces}$$

por $\boxed{HI}$ sabemos que

$$P\left(\bigcap_{i=1}^{n-1} A_i\right) = P\left(A_{n-1} \mid \bigcap_{i=1}^{n-2} A_i\right) \cdots P(A_2 \mid A_1)P(A_1)$$

Luego despejando

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right) \cdots P(A_2 \mid A_1)P(A_1)$$

$\square$

Ejemplo.- Tenemos una urna con
5 bolas BLANCAS
4 " NEGRAS

5 blancas
4 negras

Si realizamos 3 extracciones SIN Reemplazamiento ¿Cuál es la probabilidad de obtener B_1, B_2, N_3?

Es decir en la 1^a extracción → blanca
2^a " → blanca
3^a " → negra
$\Rightarrow$

$\Rightarrow P(B_1 \cap B_2 \cap N_3)$? Aplicando el T^a anterior

$$P(N_3 \cap B_2 \cap B_1) = P(N_3 \mid B_2 \cap B_1) \cdot P(B_2 \mid B_1) \cdot P(B_1)$$

$$= \frac{\cancel{4}^{2}}{7} \cdot \frac{\cancel{4}}{\cancel{8}} \cdot \frac{5}{9} = \frac{10}{63}$$

3.2. TEOREMA de la PROBABILIDAD TOTAL

El teorema de probabilidad total permite calcular la probabilidad de un suceso considerando distintos escenarios posibles.

$\boxed{\text{Nota pedagógica}}$ Los diagramas en árbol son una herramienta visual clave para representar este tipo de problemas con escenarios. El uso de ellos en la enseñanza es fundamental para estructurar el razonamiento y descomponer los eventos en distintos caminos.

Construir, analizar y estudiar dichos diagramas ayuda a comprender lo que es una partición y, por tanto, la idea de probabilidad condicionada.

<u>Def^n</u>.- Diremos que $A_1, —, A_n$ constituyen una partición del espacio muestral Ω si

$$\bigcup_{i=1}^{n} A_i = \Omega \quad y \quad A_i \cap A_j = \emptyset \quad \forall i \neq j$$

Además, supondremos que

$$P(A_i) > 0 \quad \forall i$$

$\boxed{\text{TEOREMA}}$ (prob. total) Sea $A_1, —, A_n$ una partición de Ω y B un suceso. Entonces

$$P(B) = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)$$

$\boxed{D}$ Sea B un suceso. Como $A_1, —, A_n$ una partición de Ω entonces

$$B = \bigcup_{i=1}^{n} (B \cap A_i) \quad \text{pero} \quad (B\cap A_i) \cap (B \cap A_j) = \emptyset$$

por ser partición entonces como son disjuntos

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) \underset{(1)}{=} \sum_{i=1}^{n} P(B|A_i) P(A_i)$$

Ejemplo. – En una clase de Bachillerato preguntamos por el consumo de drogas, pero para garantizar la anonimidad razonamos de la siguiente manera.
Elegimos 100 estudiantes al azar y para garantizar la anonimidad: ~~razonam~~

- Cada estudiante saca un número del 1 al 100 de una urna
- Si su número está del 1-70 responde a: ¿Has consumido drogas?

$P(p\,del) = \frac{7}{10}$

- Si su número 71-100 responde a ¿es par tu última cifra del DNI?

$P(DN) = \frac{3}{10}$

Entonces nuestro objetivo es conocer

$$P(\text{sí} \mid \text{pregunta delicada})$$

Conocemos el número total de síes, es decir sabemos $P(\text{sí})$ ⊛, entonces por el Tª Probabilidad total ⊛ Supongamos 25 síes

$$P(\text{sí}) = P(\text{sí} \mid p.\text{delicada}) \cdot P(p.\text{delicada}) + P(\text{sí} \mid p.\text{DNI}) \cdot P(p.\text{DNI})$$

Podemos estimar que $P(\text{sí} \mid p.\text{DNI}) \simeq 0.5$ entonces

$$P(\text{sí} \mid p.\text{delicada}) = \frac{P(\text{sí}) - P(\text{sí} \mid p.\text{DNI}) \cdot P(p\,\text{DNI})}{P(p.\text{delicada})}$$

$$\simeq \frac{\cancel{P(\text{sí})}\ 0{,}25 - 0{,}5 \cdot 0{,}3}{0.7} \simeq \frac{0{,}1}{0{,}7} \simeq \frac{1}{7}$$

3.3. TEOREMA DE BAYES

En ocasiones, lo que nos interesa es conocer la probabilidad ~~de~~ asociada a cada elemento de la partición dado que ha ocurrido un suceso B. Veamos un ejemplo en el que se entiende mejor lo que queremos decir

Ejemplo. - Tenemos 3 urnas que contienen bolas blancas (B) y negras (N)

U_1	U_2	U_3
3B, 1N	2B, 2N	1B, 3N

Luego queremos conocer la probabilidad de haberla escogido al azar sabiendo que $A = \{\text{extraemos bola B}\}$.

En otras palabras ¿qué U_i tiene más posibilidades de haber sido escogida?

Este tipo de ejemplos son los que motivan el T$^{\underline{a}}$ de Bayes

TEOREMA Supongamos $A_1, \ldots, A_n$ partición de Ω y consideramos B suceso con $P(B) > 0$

entonces

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{i=1}^{n} P(B|A_i) P(A_i)}$$

D Por (1) sabemos que $P(A_i|B) = \dfrac{P(A_i \cap B)}{P(B)}$

Ahora como por el T$^{\underline{a}}$ Prob total y T$^{\underline{a}}$ Factorización sabemos

$$P(B) = \sum^{n} P(B|A_i) P(A_i) \quad ; \quad P(A_i \cap B) = P(B|A_i) \cdot P(A_i)$$

tenemos el resultado

Ejemplo (continuación)

Resolvamos el anterior ejemplo aplicando Tª Bayes

$$P(U_1 | A) = \frac{P(A|U_1) \cdot P(U_1)}{\sum_{i=1}^{3} P(A|U_i) \cdot P(U_i)}$$

Como U_i se escoge al azar $P(U_i) = \frac{1}{3}$ $\forall i = 1, 2, 3$.

Entonces

$$P(U_1 | A) = \frac{\frac{3}{4} \cdot \cancel{\frac{1}{3}}}{\frac{3}{4} \cdot \cancel{\frac{1}{3}} + \frac{2}{4} \cdot \cancel{\frac{1}{3}} + \frac{1}{4} \cdot \cancel{\frac{1}{3}}} = \frac{\frac{3}{4}}{\frac{6}{4}} = \frac{1}{2}$$

$$P(U_2 | A) = \frac{\frac{2}{4}}{\frac{6}{4}} = \frac{1}{3}$$

$$P(U_3 | A) = \frac{\frac{1}{4}}{\frac{6}{4}} = \frac{1}{6}$$

Además,

$$\sum_{i=1}^{3} P(U_i | A) = \frac{1}{2} + \frac{1}{3} + \frac{1}{6} = \frac{3+2+1}{6} = 1$$

Así, la U_1 es la más probable de haber sido escogida.

4. Independencia de sucesos

En esta sección contestaremos a la pregunta que se pueden formular en las aulas

¿toda información modifica nuestras probabilidades?

Pues bien, la respuesta es <u>NO</u>. Dicha información debe tener una relación que afecte a nuestro suceso sujeto a estudio. En otro caso, no tendría valor dicha nueva información. Formalmente se dice que A y B son suceso independientes si ni A ni B aportan información sobre la ocurrencia del otro.

<u>Defn.</u>- Sean A y B dos sucesos. Decimos que A y B son independientes si

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

<u>Nota</u>.- Veamos que si sabemos que A y B son independientes entonces

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

$\Rightarrow$ B no aporta información sobre la ocurrencia de A.

<u>Ejemplos</u>.- Si lanzamos dos veces un dado y/o una moneda conocer el resultado del 1r lanzamiento no aporta información sobre el siguiente.

5. CONCLUSIONES

Este tema tiene un vínculo muy estrecho con el perfil de salida que pretendemos en nuestros alumnos, puesto que estimula el razonamiento analítico, así como visualizar posibles resultados bajo ciertos escenarios más verosímiles ~~o no~~ que otros.

Además, este aspecto último permite un nexo de unión con la educación emocional, ya que el ser humano establece como verosímiles escenarios que no lo son..

BIBLIOGRAFÍA

1h 30'