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

TIEMPO: 63 — 62

Esquema

1) Introducción
   1.1) Antigüedad
   1.2) Comienzos modernos
   1.3) s.XX y actualidad
2 Probabilidad condicionada
   2.1) Introducción
   2.2) Definición: probabilidad condicionada
   2.3) Proposición
   2.4) Teorema (probabilidad compuesta)
   2.5) Teorema (probabilidad producto)
3) Dependencia de sucesos
   3.1) Definición: dependencia + independencia
   3.2) Teorema
   3.3) Propiedades
   3.4) Definición: independientes + mutuamente independientes
4) Probabilidad producto
   4.1) Introducción
   4.2) Definición: espacio + probabilidad producto
   4.3) Definición: independencia
   4.4) Construcción de P
   4.5) Ejemplos
      4.5.1) Esquema de contagio
      4.5.2) Fiabilidad funcional
5) Teorema de Bayes
   5.1) Definición: sistema completo de sucesos
   5.2) Teorema de la Probabilidad Total
      5.2.1) Notas
   5.3) Teorema de Bayes
      5.3.1) Ejemplo

1) Introducción:

▷ Ya las civilizaciones egipcia, griega y romana utilizaban el astrágalo, hueso vulgarmente llamado "taba", situado ne la parte superior del tarso, para realizar juegos de azar. Diferentes juegos de azar eran muy conocidos en la civilización china y los dados más antiguos que se han hallado proceden de Mesopotamia y tienen unos 5.000 años (aunque se sabe que ya en la Prehistoria se practicaban juegos de azar).

▷ Comienzos modernos: las cartas aparecen en Europa en el s. XV y teorizaciones sobre el azar ya se encuentran en Demócrito y Aristóteles. Aún y así, el cálculo de probabilidades es relativamente reciente; Cardano publica en el s. XVI "*Liber de ludo alea*", es decir, "*El libro de los juegos de azar*" que se considera la primera publicación sobre el cálculo de probabilidades. Galileo, en su obra "*Consideraciones sobre el juego de los dados*" utiliza el concepto de equiprobabilidad. En el s. XVII Antoine Gomband, más conocido como el Caballero de la Mére, propone a Pascal el siguiente problema:"al lanzar 24 veces un par de dados, ante la aparición de por lo menos un seis doble en las 24 tiradas, ¿es lo mismo apostar la misma cantidad a favor que en contra?"

▷ En relación a este problema y otros de la misma índole se inicia una relación epistolar entre Pascal y Fermat que, aparte de considerarse el inicio del Cálculo de Probabilidades, permiten a Huygens escribir el primer texto conocido dedicado en exclusividad a la probabilidad: "*Sobre los razonamientos relativos a los juegos de dados*", en el s. XVII.

▷ Recién nacido el Cálculo de Probabilidades, a finales del s. XVII surgen las compañías de seguros y de forma sistemática se inicia la recopilación de gran diversidad de datos tales como: nacimientos, defunciones, accidentes, incendios, enfermedades,... de tal suerte que la Probabilidad empieza a ser una herramienta indispensable para diversos campos de la actividad humana y ya en el s. XX ésta se convierte en algo esencial incluso para la interpretación física del universo.

▷ s.XX: es más, el las últimas décadas se han desarrollado de forma notable el número de aplicaciones de la Probabilidad y la Estadística. Estas aplicaciones se encuentran en los más diversos campos: industria (control de calidad), comercio (teoría de la decisión), sociedad (encuestas de opinión), economía (índices de vida), etc.

▷ Con respecto a las ciencias ya no es suficiente con desarrollar el pensamiento determinista y es necesario prestar atención al pensamiento probabilístico para incidir sobre campos como el de la herencia genética, procesos radioactivos, la robótica,...

2) Probabilidad condicionada:

$\triangleright$ Se trata de estudiar cómo cambia la probabilidad de un suceso cuando se tiene conocimiento de que se ha producido otro. Supongamos un suceso $A \in \mathcal{A}$ tal que $P(A) > 0$; de forma intuitiva se entenderá como la probabilidad del suceso B condicionado a A a la probabilidad de que suceda B una vez haya ocurrido A. Dicha probabilidad vendrá dada por una medida de cómo B está implicado con A y dependerá, por tanto, de la probabilidad de A.

$\triangleright$ **Definición**: sea un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ y un suceso $A \in \mathcal{A}$ tal que $P(A) > 0$. Se llama probabilidad de un suceso B condicionada por el suceso A y lo denotamos por $P(B/A)$ al número dado por: $P(B/A) = \dfrac{P(B \cap A)}{P(A)}$

$\triangleright$ **Proposición**: dado $A \in \mathcal{A}$ con $P(A) > 0$, $(\Omega, \mathcal{A}, P(\bullet/A))$ es un espacio de probabilidad.

Dem. I) $P(B/A) \geq 0$. Es claro pues $P(A) > 0$ y $P(A \cap B) \geq 0$

II) $P(\Omega/A) = 1$ pues $P(\Omega/A) = \dfrac{P(\Omega \cap A)}{P(A)} = \dfrac{P(A)}{P(A)} = 1$

III) Dada $\{B_j\}$ incompatibles se tiene que:

$$P(\sum_{n=1}^{\infty} B_n/A) = \frac{P(\sum_{n=1}^{\infty} B_n \cap A)}{P(A)} = \sum_{n=1}^{\infty} \frac{P(B_n \cap A)}{P(A)} = \sum_{n=1}^{\infty} P(B_n/A)$$

$\square$

$\triangleright$ **Teorema (probabilidad compuesta)**: dados A y B con $P(A), P(B) > 0$ se tiene que:

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

Dem. No es más que aplicar la fórmula de la probabilidad condicionada a $P(A/B)$ y $P(B/A)$

$\square$

$\triangleright$ **Teorema**: sean $A_1, \ldots, A_n \in \mathcal{A}$. Entonces se verifica que, si $P(A_1 \cap \ldots \cap A_{n-1}) > 0$:

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \cdot \ldots \cdot P(A_n/A_1 \cap \ldots \cap A_{n-1})$$

Dem. $P(A_1 \cap \ldots \cap A_n) = P(A_1 \cap \ldots \cap A_{n-1}) \cdot P(A_n/A_1 \cap \ldots \cap A_{n-1})$ $\square$

3) Dependencia de sucesos:

▷ **Definición**: se dice que el suceso B depende de A si $P(B/A) \neq P(B)$

▷ El ser dependiente es una relación con las propiedades reflexivas y simétrica si $P(B) \neq 0, 1$.

▷ **Definición**: dados dos sucesos A, B con $P(A) > 0$ diremos que B es independiente de A cuando $P(B/A) = P(B)$, es decir, el que se verifique B no depende de si se verifica A o no.

▷ **Teorema**: la condición necesaria y suficiente para que el suceso B sea independiente de A es que $P(A \cap B) = P(A) \cdot P(B)$

Dem. $\Longrightarrow$: si B independiente de A tenemos:

A, B independientes $\longrightarrow P(B/A) = \dfrac{P(A \cap B)}{P(A)} = P(B) \longrightarrow P(A \cap B) = P(A) \cdot P(B)$

$\Longleftarrow$:

Si $P(A \cap B) = P(B) \cdot P(A)(> 0) \longrightarrow P(B/A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{P(A) \cdot P(B)}{P(A)} = P(B)$

□

▷ Propiedades:

1) Si B es independiente de A y $P(B) > 0 \longrightarrow A$ es independiente de B
2) Si $A, B \in \mathcal{A}$ con $P(A), P(B) > 0$ lo anterior es una equivalencia.
3) Si A y B son independientes, también lo son A y $\overline{B}$. **Dem.** $A \cap \Omega = A \cap (B \cup \overline{B}) =$
$= (A \cap B) \cup (A \cap \overline{B})$ y como son independientes: $P(A) = P(A \cap B) + P(A \cap \overline{B}) =$
$= P(A) \cdot P(B) + P(A \cap \overline{B}) \longrightarrow P(A \cap \overline{B}) = P(A)(1 - P(B)) = P(A) \cdot P(\overline{B})$
4) Si A y B son independientes, también lo son $\overline{A}$ y $\overline{B}$. **Dem.** $P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) =$
$= 1 - P(A \cup B) = 1 - [P(A) + P(B) - P(A \cap B)] = 1 - P(A) - P(B) \cdot [P(A) \cdot P(B)] =$
$= 1 - P(A) - P(B)[1 - P(A)] = P(\overline{A}) \cdot [1 - P(B)] = P(\overline{A}) \cdot P(\overline{B})$
5) Si A y B son incompatibles con $P(A), P(B) > 0$, entonces no pueden ser independientes.

▷ **Definición**: dados tres sucesos A, B, C decimos que son mutuamente/conjuntamente independientes si se cumple que:

1) $P(A \cap B) = P(A) \cdot P(B)$
2) $P(A \cap C) = P(A) \cdot P(C)$
3) $P(B \cap C) = P(B) \cdot P(C)$

▷ **Definición**: dados tres sucesos A, B, C decimos que son independientes si se cumplen **1)**, **2)**, **3)** y, además:

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

▷ *Nota*: independencia implica mutuamente independientes pero no al contrario. Por ejemplo: sean dos dados, uno rojo y otro blanco. Sean A = sacar impar en el blanco. B = sacar impar en el rojo y C = que la suma de ambos sea par.
$P(A) = P(B) = P(C) = \dfrac{1}{2}$
$P(A \cap B) = P(A \cap C) = P(B \cap C) = P(A) \cdot P(B) = P(A) \cdot P(C) = P(B) \cdot P(C) = \dfrac{1}{4}$
$P(A \cap B \cap C) = P(A \cap B) \cdot P(C/(A \cap B)) = \dfrac{1}{4} \neq \dfrac{1}{8} = P(A) \cdot P(B) \cdot P(C)$

▷ Se pueden extender estos conceptos a familias de "n" sucesos, donde la condición de independientes es que:

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \cdot \ldots \cdot P(A_n)$$

Y las otras condiciones:

$$P(A_1 \cap \ldots \cap A_k) = \prod_{j=1}^{k} P(A_j),\ 1 \leq k < n \text{ para mutuamente independientes}$$

4) Probabilidad producto:

▷ Para un cierto experimento aleatorio se tienen que analizar las condiciones de realización del mismo así como sus posibles resultados. Algunas veces se consideran experimentos sucesivos y las condiciones de uno cualquiera pueden ser o no influidas por los resultados de los precedentes. Esto sugiere la dependencia o no entre sucesos de más de un experimento aleatorio. Se conoce como experimento compuesto al formado por más de un experimento. Así, tendremos que trabajar en más de un espacio de probabilidad $(\Omega,\ \mathcal{A},\ P)$ de manera que se pueda calcular la probabilidad de que se presenten ciertos sucesos correspondientes a cada uno de los espacios de probabilidad considerados.

▷ **Definición**: dados dos espacios de probabilidad $(\Omega_1,\ \mathcal{A}_1,\ P_1)$ y $(\Omega_2,\ \mathcal{A}_2,\ P_2)$, $A_i \in \mathcal{A}_1$ y $B_j \in \mathcal{A}_2$, llamamos espacio producto cartesiano $\Omega_1 \times \Omega_2$ al formado a partir de los sucesos elementales de la forma $(A_i,\ B_j)$. Tomando en $\Omega_1 \times \Omega_2$ una σ-álgebra $\mathcal{A}$ y una ley de probabilidad P, obtendremos un espacio de probabilidad $(\Omega_1 \times \Omega_2,\ \mathcal{A},\ P)$ que se conoce como espacio de probabilidad producto.

▷ **Definición**: se dice que los experimentos aleatorios o los espacios de probabilidad $)\Omega_1 \times \Omega_2)$ y $(\Omega_2,\ \mathcal{A}_2,\ P_2)$ son independientes si se verifica que:

$$P(A_i, B_j) = P_1(A_i) \cdot P_2(B_j), \forall A_i \in \mathcal{A}_1, \forall B_j \in \mathcal{A}_2$$

Es decir, los resultados del segundo experimento no dependen de lo que suceda en el primero.

▷ Dados dos espacios de probabilidad y dado $(\Omega_1 \times \Omega_2,\ \mathcal{A})$, nos planteamos cómo definir la probabilidad P para que $(\Omega_1 \times \Omega_2,\ \mathcal{A},\ P)$ sea un espacio de probabilidad.

a) Los experimentos son independientes: $P(A_i, B_j) = P_1(A_i) \cdot P_2(B_j)$

b) Los experimentos son dependientes: $P(A_i, B_j) = P_1(A_i) \cdot P_2(B_j/A_i)$

Si tenemos $\Omega_1, \ldots, \Omega_n$ espacios de probabilidad, tendremos n-uplas $(A_1, \ldots, A_n)$ donde $A_j \in \mathcal{A}_j$ y la ultima definición nos sirve para definir P en $(\Omega_1 \times \ldots \times \Omega_n,\ \mathcal{A})$.

▷ Por ejemplo:

1) Esquema de contagio: sea una urna con bolas rojas (R) y blancas (B). Se trata de representar el llamado esquema de contagio con aplicaciones en epidemología, accidentes laborales, etc. La composición de la urna no variará si la probabilidad de contagio se mantiene constante en el tiempo. Pero, a veces, el hecho de que haya un accidente hace que aumente o disminuya la probabilidad de nuevos accidentes. Sean las bolas rojas los accidentes. Extraemos una bola de un color y se devuelve con "α" bolas de ese color y "β" del otro. Entonces, la probabilidad de que la primera sea roja es $\dfrac{R}{R+B}$, de que lo sea la segunda es $\dfrac{R+\alpha}{R+\alpha+B+\beta}$ y así sucesivamente.

   a) El muestreo con reemplazamiento es: $\alpha = \beta = 0$

   b) El muestreo sin reemplazamiento es: $\alpha = -1,\ \beta = 0$

   c) El esquema de contagio de Polya es: $\alpha > 0,\ \beta = 0$

   d) El modelo de Ehrenfest para estudiar el cambio de color entre dos cuerpos aislados es: $\alpha = -1,\ \beta = 1$

   e) El modelo de una campaña de seguridad es: $\alpha = 0,\ \beta > 0$

2) <u>Fiabilidad funcional</u>: se conoce como fiabilidad de un sistema A a la probabilidad $P(A)$ de que el sistema funcione. El sistema se compondrá de subsistemas A_j con probabilidad funcional $P(A_j)$.

   a) Si los subsistemas están acoplados en serie implica que A funcionará $\longleftrightarrow A_j$ funciona $\forall j$, luego $A = \bigcap_{j=1}^{n} A_j$. Si los suponemos independientes: $P(A) = \prod_{j=1}^{n} P(A_j)$

   b) Si los subsistemas están acoplados en paralelo implica que A funcionará si lo hace al menos uno de los subsistemas, luego $A = \bigcup_{j=1}^{n} A_j$. Si los suponemos independientes:

   $$P(A) = 1 - P(\bigcap_{j=1}^{n} \overline{A_j}) = 1 - \prod_{j=1}^{n} \overline{P(A_j)} = 1 - \prod_{j=1}^{n}(1 - P(A_j))$$

3) <u>El problema del Caballero de la Mére</u>: $A_1 =$ sacar un seis doble en dos dados. $P(A_1) = \frac{1}{36} \Longrightarrow$ No sacarlo: $P(\overline{A_1}) = \frac{35}{36}$.

   No sacar, al menos, un seis doble en 24 tiradas: $\left(\frac{35}{36}\right)^{24} \simeq 0,51 \Longrightarrow$ Es mejor apostar en contra.

5) Teorema de Bayes:

▷ **Definición**: decimos que $A_1, \ldots, A_n \in \mathcal{A}$ (no vacíos) es un sistema completo de sucesos si se cumple que:

1) $A_i \cap A_j = \emptyset, \forall i \neq j$

2) $\bigcup_{j=1}^{n} A_j = \Omega$

▷ **Teorema (probabilidad total)**: dados $B \in \Omega$ y $\{A_1, \ldots, A_n\}$ un sistema completo de sucesos, se tiene que: $P(B) = \sum_{j=1}^{n} P(A_j) \cdot P(B/A_j)$

Dem. $P(B) = P(B \cap \Omega) = P(B \cap (A_1 \cap \ldots \cap A_n)) = P((B \cap A_1) \cup \ldots \cup (B \cap A_n)) = P(\sum_{j=1}^{n} B \cap A_j) =$ $\{independientes\} = \sum_{j=1}^{n} P(A_j) \cdot P(B/A_j)$

□

▷ *Nota*: si tenemos $\{A_n\}_{n \in N}$ un conjunto numerable de sucesos tales que ninguno es vacío, $A_i \cap A_j = \emptyset$ y $P(\bigcup_{n=1}^{\infty} A_n) = 1$ también obtenemos la misma tesis. Este teorema nos dice que, sabiendo todos los posibles sucesos que conforman el experimento, podemos saber la probabilidad de cada posible resultados antes de realizar el experimento (es decir, "a priori" sabemos qué probabilidad hay de cada resultado). La pregunta al revés sería: conocido el resultado, ¿de qué suceso elemental proviene (suponiendo que tengamos una partición/sistema completo de sucesos)? La respuesta a esta pregunta es:

▷ **Teorema de Bayes**: dados $B \in \Omega$ y $\{A_1, \ldots, A_n\}$ un sistema completo de sucesos tal que $P(A_j) > 0, \forall j$. Entonces, se cumple que:

$$P(A_j/B) = \frac{P(A_j) \cdot P(B/A_j)}{\sum_{k=1}^{n} P(A_k) \cdot P(B/A_k)}$$

Dem. $P(A_j/B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(A_j) \cdot P(B/A_j)}{P(B)} = \{Teorema\ de\ la\ Probabilidad\ Total\} =$

$$= \frac{P(A_j) \cdot P(B/A_j)}{\sum_{k=1}^{n} P(A_k) \cdot P(B/A_k)}$$

□

▷ **Nota**: el resultado también es cierto en el caso infinito numerable.

▷ Por ejemplo: sabemos que sumamos **2** al lanzar cierto número de dados. ¿Cuál es la probabilidad de haber jugado con dos dados?

Sea A_j = haber jugado con j dados. Como $1 \leq j \leq 2$, $P(A_j) = 1/2$

Sea B = sumar 2, $\longrightarrow P(B/A_1) = 1/6$ y $P(B/A_2) = 1/6^2$

Luego: $P(A_2/B) = \frac{P(A_2) \cdot P(B/A_2)}{P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2)} = \frac{\frac{1}{2} \cdot \frac{1}{6^2}}{\frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{6^2}} = \frac{1}{7}$