Introducción a la Probabilidad

Clase 2

Alvarez - Arceo - Aurucis -Casparian - Castro - Chan - Howlin - Maulhardt- Spano -Stein - Calvo

Outline

LA NECESIDAD

Cuando sabemos que B ocurrió y queremos saber..

La probabilidad de que ocurra un evento A, se dice que queremos calcular la probabilidad de A condicional a B. Simbólicamente:

$$P(A/B)$$

Supongamos que hay diez bebes recién nacidos por parto normal de los cuales cuatro son varones y quince bebes nacidos por cesárea de los cuales siete son varones en la nursery del hospital.

La Representación

Esta información podría representarse en el siguiente diagrama:

Tabla: Nursery

	varón	mujer	
Parto	4	6	10
cesárea	7	8	15
total	11	14	25

Estamos interesados en saber:¿Qué probabilidad hay

de que haya sido parto normal sabiendo que es varón?.

La Definición

Como de los (11) varones, (4) nacieron por parto normal . Esta probabilidad es:

$$\boldsymbol{P(P/V)} = 4/11$$

Esta probabilidad depende de las siguientes tres probabilidades:

- $P(P) = 10/25 \rightarrow$ probabilidad de que haya nacido por parto normal.
- $P(V) = 11/25 \rightarrow$ probabilidad de que sea varón.
- $P(PV) = P(P \cap V) = 4/25 \rightarrow$ probabilidad de que sea varón y haya nacido por parto normal.

La Definición

P(A/B) se lee “probabilidad de A condicional a B” y se define como el cociente entre la probabilidad conjunta de A con B, y la probabilidad de B (**siempre que B no sea el evento imposible**). Simbólicamente:

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

con $P(B) > 0$. Para nuestro ejemplo:

$$P(P/V) = P(P \cap V)/P(V) = \frac{\frac{4}{25}}{\frac{11}{25}} = 4/11$$

Otro Ejemplo

En un estudio se intenta asociar el consumo de café con el insomnio. Para ello se consideran 200 estudiantes de la FRBA y se analiza si consumen café o no y si presentan problemas de insomnio o no. Del total de estudiantes el 40% refirieron problemas de insomnio, el 70% consume café y 35 estudiantes tienen insomnio pero no consumen café.

La representación del ejemplo

Veamos cómo construimos el diagrama de Carrol para nuestro ejemplo:
En primera instancia anotamos los datos del enunciado, recordando que el 40% de 200 es 80, el 70% de 200 es 140.

	consume café	**no consume café**	totales
duerme bien	$(CC \cap DB)$	$(\overline{CC} \cap DB)$	(DB)
insomnio	$(CC \cap I)$	35	80
totales	140	$(\overline{CC})$	200

Completamos la tabla

Utilizando las diferencias o complementos:

	consume café	**no consume café**	totales
duerme bien	95	25	120
insomnio	45	35	80
totales	140	60	200

Las Preguntas

que nos hacemos son, cuál es la probabilidad de que un alumno:

(a) tenga insomnio y consuma café?

(b) consuma café.

(c) tenga insomnio o consuma café.

(d) tenga insomnio sabiendo que consume café.

(e) tenga insomnio sabiendo que no consume café.

Respondamos ahora a las preguntas:

(a) presente insomnio y consuma café?
$P(CC \cap I) = 45/200$

(b) consuma café.
$P(CC) = 140/200$

Las Respuestas

(c) tenga insomnio o consuma café.
$P(CC \cup I) = P(CC) + P(I) - P(CC \cap I) =$
$140/200 + 80/200 - 45/200 = 175/200$

(d) tenga insomnio sabiendo que consume café.
$P(I/CC) = P(I \cap CC)/P(CC) = 45/140$

(e) tenga insomnio sabiendo que no consume café.
$P(I/\overline{CC}) = P(I \cap \overline{CC})/P(\overline{CC}) = 35/60$

Outline

Regla del Producto

De la definición de probabilidad condicional, se desprende la llamada regla o teorema del producto.

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

Notemos que también puede expresarse:

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

Cómo sería la expresión de esta regla para tres conjuntos?

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

Outline

Ejemplo Ilustrativo

Supongamos que cierta prueba para detectar la presencia de una enfermedad en un individuo, da resultado positivo (detecta la presencia de la enfermedad) en un individuo enfermo con probabilidad 0.99 y en un individuo sano con probabilidad 0.02 (falso positivo).
Es decir que, dicha prueba no señala la presencia de la enfermedad en un individuo sano con probabilidad 0.98 y no la señala en un individuo enfermo con probabilidad 0.01 (falso negativo).

Se sabe además, el 10% de la población padece esta enfermedad.

Las Preguntas

Nos preguntamos ahora:

(a) Si se selecciona al azar un individuo de esta población, se le aplica la prueba y el resultado es positivo

   i) ¿Cuál es la probabilidad de que realmente padezca la enfermedad?

   ii) ¿Cuál es la probabilidad de que no la padezca?

(b) Si se selecciona un individuo de esta población, se le aplica la prueba y el resultado es negativo,

   i) ¿Cuál es la probabilidad de que realmente no padezca la enfermedad?

   ii) ¿Cuál es la probabilidad de que el test no haya detectado la presencia de la enfermedad?

Los Eventos de Interés

Nuestros eventos en este ejemplo son:

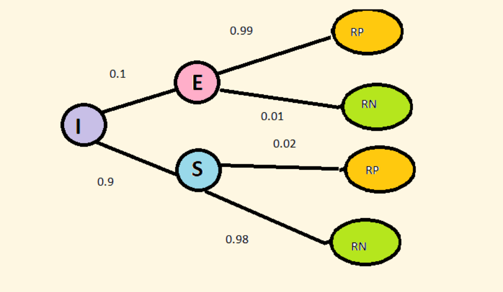
E: “la persona padece esa enfermedad”

S : “la persona no padece la enfermedad”

RP: “el resultado de la prueba es positivo”.

RN: “el resultado de la prueba es negativo”

Construyamos un diagrama de árbol para representar la situación, colocando todas las posibilidades y visualizando la estructura del problema:



Calculamos las probabilidades conjuntas

Utilizando la regla del producto podemos calcular las probabilidades conjuntas:

$$P(E \cap RP) = P(RP/E)P(E) = 0.99 \times 0.1 = 0,099$$

producto de las probabilidades de la rama rosa-naranja

$$P(E \cap RN) = P(RN/E)P(E) = 0.01 \times 0.1 = 0,001$$

producto de las probabilidades de la rama rosa-verde

$$P(S \cap RP) = P(RP/S)P(S) = 0.02 \times 0.9 = 0,018$$

producto de las probabilidades de la rama celeste - naranja

$$P(S \cap RN) = P(RN/S)P(S) = 0.98 \times 0.9 = 0,882$$

producto de las probabilidades de la rama celeste-verde

Calculemos ahora las probabilidades condicionales

$$P(E/RP) = P(E \cap RP)/P(RP) = \frac{0,099}{(0,099 + 0,018)} = 0,8461$$

$$P(S/RP) = P(S \cap RP)/P(RP) = \frac{0,018}{(0,099 + 0,018)} = 0,1538$$

Puede observarse que dado un resultado positivo, es más probable que el paciente esté enfermo, veremos luego que esto depende de la incidencia de la enfermedad!.

$$P(E/RN) = P(E \cap RN)/P(RN) = \frac{0.001}{0.001 + 0.882} = 0.00113$$

Outline

Independencia

Un concepto muy importante dentro del campo de la estadística es la independencia.

Eventos Independientes

Diremos que dos eventos son independientes cuando la ocurrencia de uno de ellos no afecta la probabilidad de ocurrencia del otro. Formalmente:

$$\textbf{A y B son Independientes} \Leftrightarrow P(A/B) = P(A)$$

Se desprende de esta definición

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

siendo $P(B) > 0$ Algunos autores dan a esta última igualdad entidad de definición de independencia.

Veamos un Ejemplo

En una fábrica se lápices los hay blandos y duros de color rojo y de color negro. La producción de hoy se presenta en la siguiente tabla:

	Rojos	**Negros**	
Duros(HH)	**10**	**20**	**30**
Blandos(BB)	12	24	36
	22	44	66

Nos interesa saber si el conocer el color nos da información respecto de la dureza...es decir si color informa algo respecto de dureza.

Análisis de la Pregunta

Para responder esta pregunta veamos qué relación hay entre las probabilidades

$$P(HH/R) \text{ y } P(HH/N)$$

$$P(HH/R) = \frac{10}{22} = \frac{5}{11}$$

$$P(HH/N) = \frac{20}{44} = \frac{5}{11}$$

...más aún.. $P(HH) = \frac{30}{66} = \frac{5}{11}$

En este caso conocer el color no nos aporta información respecto de la dureza...porque **son independientes!** Es decir:

$$P(HH) = P(HH/R) = P(HH/N)$$

Cambiemos un poco la estructura de la producción

En la anterior la proporción de rojos es la misma entre los duros y los blandos y la proporción de negros también...mientras que en esta ya no...

	Rojos	Negros	
Duros(HH)	**10**	**20**	**30**
Blandos(BB)	24	12	36
	34	32	66

Ahora nos informa algo el color respecto de la dureza?

Analicemos este nuevo esquema

Ahora

$$P(HH/R) = \frac{10}{34} = \frac{5}{17} \text{ pero } P(HH/N) = \frac{20}{32} = \frac{5}{8} \text{ y}$$

$$P(HH) = \frac{30}{66} = \frac{5}{11}$$

Es decir que si queremos un lápiz duro nos conviene sacar un negro...eso es que el color nos da información respecto de la dureza..entonces **NO SON INDEPENDIENTES**!!!
Saber el color nos da información respecto de la textura en este caso!! Los colores no tienen la misma proporción en ambas texturas o bien las texturas no tienen la misma distribución en ambos colores.

Extensión del Concepto de Independencia a Tres Eventos

Dados 3 eventos A, B y C asociados a un mismo espacio muestral

Si se verifica que: $P(A \cap B) = P(A)P(B)$, $P(C \cap B) = P(C)P(B)$ y $P(A \cap C) = P(A)P(C)$ se dice que son independientes de a pares

Si además se verifica que:

$P(A \cap B \cap C) = P(A)P(B)P(C)$, entonces se dice que son mutuamente independientes

Veamos un ejemplo

El espacio muestral asociado a un experimento aleatorio tiene cuatro puntos:$S = \{a, b, c, d\}$ Definamos los siguientes tres eventos:

$$A = \{a, b\} \qquad B = \{a, c\} \qquad C = \{a, d\}$$

Vale que:

$$P(A \cap B) = P(A)P(B) \qquad P(C \cap B) = P(B)P(C)$$

$$P(A \cap C) = P(A)P(C) \qquad = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$$

pero no vale: $P(A \cap B \cap C) = P(A)P(B)P(C)$ de modo que son independientes de a pares pero no son mutuamente independientes.

Veamos un segundo Ejemplo

Ahora el espacio muestral asociado a un experimento aleatorio tiene ocho puntos:

$$S = \{a, b, c, d, e, f, g, h\}$$

$$A = \{a, b, c, d\} \qquad B = \{a, c, f, h\} \qquad C = \{a, d, g, h\}$$

Para los siguientes tres eventos, vale la independencia de a pares igual que en el ejemplo anterior, pero:

$$P(A \cap B) = P(A)P(B) \qquad P(C \cap B) = P(B)P(C)$$

$$P(A \cap C) = P(A)P(C) \qquad = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$$

también vale que $P(A \cap B \cap C) = P(A)P(B)P(C) = \frac{1}{8}$

Outline

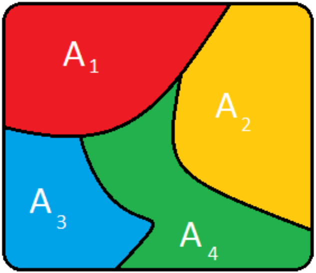
Partición de un Espacio Muestral

Decimos que una familia de subconjuntos $\{A_i\}_{1 \leq i \leq n}$ del espacio muestral S es una partición cuando satisface las siguientes tres condiciones:

Ninguna parte es vacía: $A_i \neq \emptyset$ para $1 \leq i \leq n$

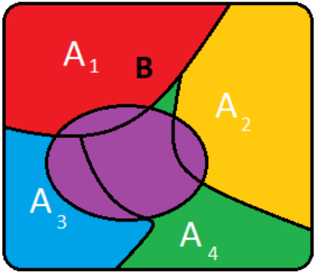
Las partes son disjuntas: $A_i \cap A_j = \emptyset$ para $i \neq j$

Las partes cubren al espacio muestral: $\cup_{1 \leq i \leq n} A_i = S$



Teorema de Probabilidad Total

Dado un espacio muestral S, un evento B asociado a este espacio muestral y una partición $\{A_i\}_{1\leq i\leq n}$

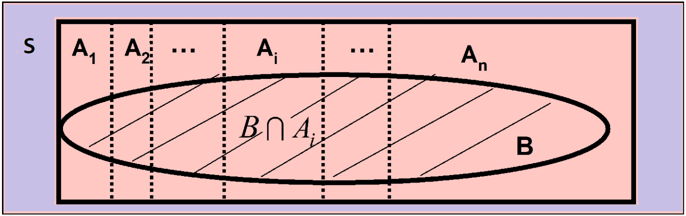


entonces vale que $P(B) = \sum_{i=1}^{n} P(B/A_i)P(A_i)$

Demostración

$P(B) = P(B \cap S)$ por propiedad de conjuntos
$P(B) = P(B \cap \cup_{i=1}^{n} A_i)$ por definición de partición
$P(B) = P(\cup_{i=1}^{n} (A_i \cap B))$ por propiedad de conjuntos
$P(B) = \sum_{i=1}^{n} P(A_i \cap B)$ por ax-3 de probabilidad
$P(B) = \sum_{i=1}^{n} P(B/A_i)P(A_i)$ por regla del producto

queda entonces demostrado!



Ejemplo

En un negocio se reciben artículos de tres proveedores que denominamos A, B y C. De los artículos que provee A el 20% tiene fallas, en B ese porcentaje es el 10% y en C el 8%. Sabemos también que A provee el 50% de la producción y que B y C proveen igual proporción. Queremos saber qué proporción de artículos están fallados.

Los datos son:

$P(A) = 0.5, P(B) = P(C) = 0.25$
$P(F/A) = 0.2, P(F/B) = 0.1$ y $P(F/C) = 0.08.$

Ejemplo (continuación)

Entonces por la regla del producto $P(F \cap A) = 0.2 * 0.5 = 0.1$, $P(F \cap B) = 0.1 * 0.25 = 0.025$ y $P(F \cap C) = 0.08 * 0.25 = 0.02$

	A	B	C	
F	0.1	0.025	0.02	0.145
NoF	0.4	0.225	0.23	0.855
	0.5	0.25	0.25	1

Luego los fallados son:

$P(F) = P(F \cap A) + P(F \cap B) + P(F \cap C) = 0.1 + 0.025 + 0.02 = 0.145$

Outline

Teorema de Bayes o Probabilidad de las Causas

Enunciado

Bajo las mismas hipótesis del Teorema de Probabilidad Total, la probabilidad de una causa es:

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

Este teorema revolucionó el campo de la probabilidad!!

Demostración

$$P(A_j/B) = \frac{P(B \cap A_j)}{P(B)}$$ por la definición de probabilidad condicional

$$P(A_j/B) = \frac{P(B \cap A_j)}{\sum_{i=1}^{n} P(B/A_i)P(A_i)}$$ por el Teorema de Probabilidad Total

$$P(A_j/B) = \frac{P(B/A_j)P(A_j)}{\sum_{i=1}^{n} P(B/A_i)P(A_i)}$$ por la regla del producto.

queda entonces demostrado.

Ejemplo

Retomando el ejemplo del negocio podría interesarnos saber si el producto pertenece al proveedor A sabiendo que está fallado.

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

$$P(A/F) = \frac{0.1}{0.145} = \frac{0.5 * 0.2}{0.5 * 0.2 + 0.25 * 0.1 + 0.25 * 0.08} \approx 0.6896$$

$$P(A/F) = \frac{P(A) * P(F/A)}{P(A) * P(F/A) + P(B) * P(F/B) + P(C) * P(F/C)}$$

Revisión Teórica

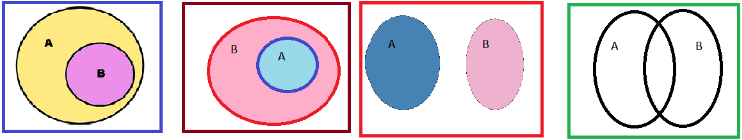
Dados dos sucesos no imposibles A y B, cuáles de las siguientes alternativas son factibles:

(i) p (A) sea mayor que p (A/B)

(ii) p (A) sea igual que p (A/B)

(iii) p (A) sea menor que p (A/B)

Juegue con distintas opciones!!



Justifique en forma gráfica, analítica o con ejemplos cada una de las respuestas.

Revisión Teórica

Indique en cuáles de los siguientes casos se ha utilizado el concepto de probabilidad condicional:

(a) Que salga un seis al tirar el dado es más probable si se sabe que ha salido un número par que si no se tiene información previa.

(b) Los caninos tienen mayor probabilidad de sobrevivir a un accidente automovilístico que los felinos.

(c) En los días de lluvia el consumo de helado es menos probable.

(d) La chance de que un equipo mecánico requiera reparación es superior cuando el equipo tiene mucho uso.

Revisión Teórica

Si A y B son dos sucesos no vacíos, incluidos en un espacio muestral S, cuyas probabilidades son 0,40 y 0,30 respectivamente, entonces la/s alternativa/s que no es posible es:

(a) que A y B sean mutuamente excluyentes

(b) que la probabilidad conjunta de A y de B sea positiva

(c) que A y B sean complementarios

Si A y B son dos sucesos no vacíos, incluidos en un espacio muestral S, tales $A \subseteq B$, entonces cuáles de las siguientes afirmaciones son posibles?

(a) A y B son independientes

(b) A y B son incompatibles

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

(d) $P(A/B) > P(A)$

(e) $P(B/A) = 1$

Completar la Frase de modo tal que resulte correcta

(a) Dados dos sucesos A y B no vacíos incluidos en un espacio muestral S, son .... cuando la ocurrencia de uno no impide la... del otro.

(b) La teoría clásica de probabilidad es válida solamente en espacios de ....

(c) La teoría axiomática de probabilidad se fundamenta en ....axiomas, que postulan que la probabilidad de cualquier evento es......................., la probabilidad del ....es igual a 1 y la probabilidad de la unión de eventos mutuamente excluyentes es igual a ....

(d) La noción frecuencial de probabilidad requiere de la realización de..................................del experimento, por lo cual se la suele llamar probabilidad a ....

muchas gracias y