Tema 3: PROBABILIDAD - II

Definición axiomática y algunas propiedades útiles. Probabilidad condicionada e independencia de sucesos. Regla de la multiplicación. Tablas de contingencia. Combinatoria.

Biología sanitaria 2017/18. Universidad de Alcalá

M. Marvá. Actualizado: 2017-09-25

Definición axiomática de probabilidad (sección 3.3.2 libro)

Se llama **función de probabilidad** a cualquier función $P : E \to [0, 1]$ que asigna a cada suceso A un valor numérico $P(A)$ y que verifica las siguientes reglas:

1. Si E es el espacio muestral,
$$P(E) = 1$$
2. Para cualquier evento A,
$$0 \leq P(A) \leq 1$$
3. Si A y B son disjuntos (incompatibles) entonces
$$P(A \cup B) = P(A) + P(B)$$

Puedes imaginar la probabilidad de un subconjunto como su “tamaño” relativo con respecto al total (suceso seguro)

Algunas propiedades de la función de probabilidad:

$P(A') = 1 - P(A)$

Ejemplo: $X =$ resultado de lanzar un dado,

$$P(X \leq 2) = 1 - P(X > 2) = 1 - 4/6 = 2/6$$

$P(\emptyset) = 0$

Ejemplo: $Z =$ día de la semana,

$$P((Z = lunes) \cap (Z = viernes)) = 0$$

Si $A \subset B$ entonces $P(A) < P(B)$ y $P(B) = P(A) + P(B \cap A')$

Ejemplo:

Dados A y B, entonces $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Ejemplo: se aplica un test diagnóstico rápido a 20 enfermos (da positivo en 18) y 30 sanos (da positivo en 5). Se selecciona un individuo al azar, ¿cuál es la probabilidad de esté enfermo o haya dado positivo?

Notación: D^+ = enfermo (D de *diesase*),

$+$ = positivo (en el test),

Cálculo:

$$P(D^+ \text{ u } +) = P(D^+) + P(+) - P(D^+ \text{ n } +) =$$

= 20/50 + 23/50 - 18/50 = 25/50

Una forma de organizar los datos anteriores es con **una tabla de contingencia**

Cada celda es una intersección

	D^+	D^-
$+$	18	5
$-$	2	25

⇒ Añadir los "márgenes" ⇒

	D^+	D^-	Total
$+$	**18**	5	**23**
$-$	2	25	28
Total	**20**	30	50

Por cierto. . .

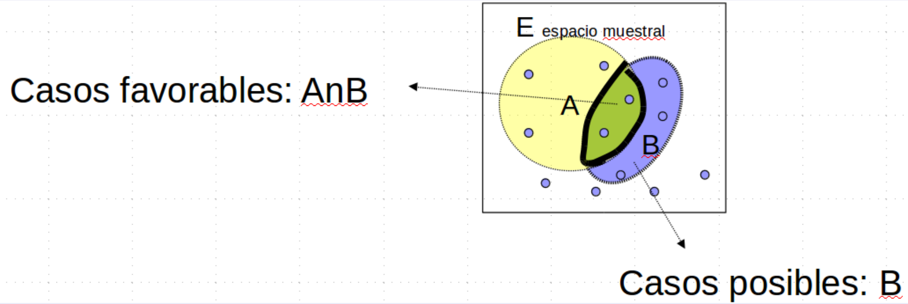
Se sabe que alguien ha dado positivo en el test, ¿cuál es la probabiliad de que esté realmente enfermo?

Probabilidad condicionada (sección 3.4 libro)

Se llama

- probabilidad de A condicionada a B,
- probabilidad de A suponiendo/sabiendo que ha sucedido B

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



Error frecuentíííííííísimo: confundir probabilidad condicionada e intersección

En ambos casos medimos la intersección, pero. . .

- En $P(A \cap B)$ con respecto a $P(E) = 1$
- En $P(A|B)$ con respecto a $P(B)$

Ejemplo

	D^+	D^-	Total
$+$	**18**	5	**23**
$-$	2	25	27
Total	20	30	**50**

- Por un lado $P(D^+ \cap +) = 18/50$
- Por otro lado $P(D^+|+) = \dfrac{18/50}{23/50} = 18/23$

Dos eventos A y B son independientes si, y sólo si,

$$P(A|B) = P(A) \Leftrightarrow P(A \cap B) = P(A)P(B)$$

Ejemplo

	Enfermo	Sano	
D+	18	5	23
D-	**2**	25	**27**
Total	**20**	30	**50**

Los eventos "Enfermo" y "Dar negativo" no son independientes porque

$$P(Enfermo \cap D-) = \frac{2}{50}$$

$$P(Enfermo)P(D-) = \frac{20}{50}\frac{27}{50}$$

Dos sucesos disjuntos (incompatibles) nunca son independientes.

La regla de la multiplicación:

De la definición de probabilidad condicionada:

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

de ahí se deduce

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

Ejemplo: toma de muestras

Hay 10 individuos para un ensayo clínico; los nombramos por 1, 2, . . . , 10 para anonimizarlos. ¿Cuál es la probabilidad de que salgan, en este orden: el 1, el 3 y el 7? Llamamos A_1, A_3, A_7 a los eventos correspondientes. Entonces:

$$P(A_1 \cap A_3 \cap A_7) = P(A_1)P(A_3|A_1)P(A_7|A_1 \cap A_3)$$

Sirve la regla de Laplace

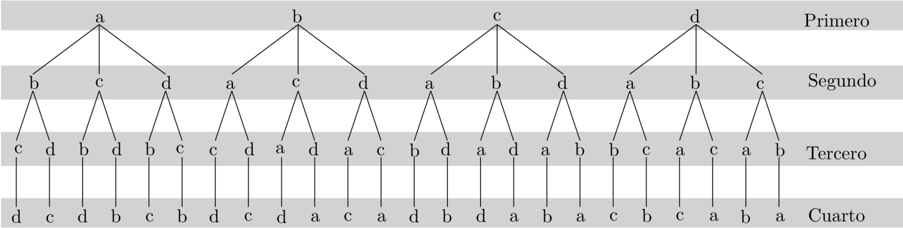
- Sin reemplazamiento: $P(A_1)P(A_3|A_1)P(A_7|A_1 \cap A_3) = \frac{1}{10}\frac{1}{9}\frac{1}{8}$ NO independientes.
- Con reemplazamiento: $P(A_1)P(A_3|A_1)P(A_7|A_1 \cap A_3) = \frac{1}{10}\frac{1}{10}\frac{1}{10}$ independientes.

COMBINATORIA (sección 3.6 del libro)

Permutaciones de n elementos:

es cada una de las formas en que se pueden ordenar los n elementos

Ejemplo: ¿de cuántas formas puedes ordenar pueden hacer cola 4 personas? ¡el orden importa! Las llamaremos a, b, c, d



Hay un total de $24 = 4 \cdot 3 \cdot 2 \cdot 1 = 4!$

El número de permutaciones de n elementos es

$$Per(n) = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 3 \cdot 2 \cdot 1 = n!$$

Variaciones de n elementos de orden k:

es cada una de las formas en que se puede ordenar k de los n elementos dados.

Ejemplo: Si 7 personas hacen cola, ¿de cuántas formas pueden ocupar los 3 primeros puestos? ¡el orden importa, pero no usamos todos los elementos!

Hay un total de

$$7 \cdot 6 \cdot 5 = \frac{7!}{(7-3)!}$$

El número número de variaciones de orden k de n elementos

$$V(n,k) = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

Combinación de n elementos tomados de k en k (o combinación de orden k de n elementos)

: es cada uno de los subconjuntos formado por k de los n elementos dados. es cada una de las formas en que se puede ordenar k de los n elementos dados.

Ejemplo: Dispones de 7 elementos {A, B, C, D, E, F, G}. ¿Cuántos subconjuntos de 3 de estos elementos puedes hacer (no imortan el orden)?

1. Si importara el orden, sería un V(10, 3)
2. En el recuento anterior, cada trío de elementos aparece per(3)=3!=6 veces

Para eliminar esas repeticiones (porque el orden no importa)

$$\frac{V(10,3)}{Per(3)}$$

El número de combinaciones n elementos tomados de k en k:

$$C(n,k) = C_n^k = \frac{n!}{k!(n-k)!}$$

Otras fórmulas de la combinatoria

Variaciones con repetición de n elementos, tomados de k en k

Si se permite que cada elemento aparezca tantas veces como se quiera, entonces:

$$\mathrm{VRep}(n, k) = n^k$$

Ejemplo: con los 3 elementos $[A, G, T]$, tomados de dos en dos, y permitiendo repeticiones obtenemos las

$$\mathrm{VRep}(3, 2) = 3^2 = 9$$

permutaciones con repetición que pueden verse en la siguiente tabla

1	2	3	4	5	6	7	8	9
A	A	A	G	G	G	T	T	T
A	G	T	A	G	T	A	G	T

Otras fórmulas de la combinatoria

Permutaciones con repetición de n elementos

El número de permutaciones que se pueden formar con m objetos entre los cuales hay n_1 iguales entre sí, otros n_2 iguales entre sí,. . . , y finalmente n_k iguales entre sí, es:

$$\text{PerRep}(n_1, n_2, \cdots, n_k) = \frac{m!}{n_1!n_2! \cdots n_k!}$$

Observa que ha de ser: $m = n_1 + n_2 + \cdots + n_k$

Ejemplo: Por ejemplo, si tenemos la lista $[A, A, G, G, T]$, es decir $m = 5$, $n_1 = 2$ (hay dos repeticiones de A), $n_2 = 2$ y $n_3 = 1$, entonces hay:

$$\text{PerRep}(2, 2, 1) = \frac{5!}{2! \cdot 2! \cdot 1!} = 30.$$