ESTADÍSTICA PARA TOD@S: ADAPTACIÓN A LA DIVERSIDAD FUNCIONAL

TEORÍA DE LA PROBABILIDAD: TEOREMA DE BAYES Y TEOREMA DE LA PROBABILIDAD TOTAL

UNIVERSIDAD PABLO DE OLAVIDE

ÁREA DE ESTADÍSTICA E IO

Autoras:

Nieves Aquino Llinares

Mª del Pilar Moreno Navarro

UNIVERSIDAD PABLO DE OLAVIDE

OCTUBRE DE 2022

ISBN: 978-84-09-44465-6

Introducción

Esta publicación forma parte de un proyecto más ambicioso que se denomina ***Estadística para tod@s: adaptación a la diversidad funcional*** y a través del cual las autoras, conscientes de la falta de material específico en estadística adaptado a diversas discapacidades, quieren poner su granito de arena para alcanzar una universidad más inclusiva, con la realización de píldoras formativas y publicaciones de apoyo al estudio accesibles a diversas discapacidades.

En concreto este anexo es un complemento al vídeo formativo **"Teorema de Bayes y teorema de la probabilidad total",** adaptado a personas con discapacidad auditiva. Estos teoremas son herramientas imprescindibles para el cálculo de probabilidades pues sirven para calcular probabilidades de sucesos con cierta información previa.
Además, para personas con discapacidad visual esta píldora docente cuenta con material de apoyo en soporte informatizado Edico disponible en el Servicio bibliográfico de ONCE.

La cápsula docente puede visualizarse en el siguiente enlace:

Cápsula formativa: Teorema de Bayes y Teorema de la Probabilidad Total

1ª DIAPOSITIVA **Minuto: 00:10**

Bienvenidos a la serie "Estadística para tod@s: adaptación a la diversidad funcional", cuyas autoras son Nieves Aquino Llinares y Mª del Pilar Moreno Navarro, profesoras del área de Estadística e Investigación Operativa de la Universidad Pablo de Olavide. Su objetivo es hacer una Universidad más inclusiva con la creación de píldoras formativas de material estadístico adaptadas a diversas discapacidades.

2ª DIAPOSITIVA **Minuto: 00:31**

Concretamente, en este vídeo hablaremos sobre conceptos básicos de la teoría de la probabilidad explicando los fundamentos teóricos del teorema de Bayes y el teorema de la probabilidad total, así como sus aplicaciones con ejemplos sencillos.

3ª DIAPOSITIVA **Minuto: 00:42**

Comencemos definiendo un sistema completo de sucesos como aquel conjunto de sucesos que no tienen ningún elemento en común, es decir son mutuamente excluyentes y que además su unión cubre todo el espacio muestral o lo que es lo mismo, forman un sistema exhaustivo de sucesos.

Para comprenderlo bien imaginemos una tarta la cual es cortada en 4 trozos diferentes, pero no tienen porqué tener el mismo tamaño. Los trozos de tarta no se mezclan entre sí y si unimos otra vez los trozos volvemos a tener la tarta completa.

Esa sería la idea de un sistema completo de sucesos.

4ª DIAPOSITIVA **Minuto: 00:42**

De forma general, si un experimento aleatorio tenemos n sucesos A_1, A_2, A_3, ..., A_n, estos constituyen un sistema completo de sucesos si y solo si son incompatibles 2 a 2, es decir, la intersección entre dos de los sucesos cualesquiera siempre es el conjunto vacío y además la unión de todos los sucesos es el espacio muestral:

$$\mathrm{Ai} \cap \mathrm{Aj} = \emptyset \text{ para cualquier i, j}$$

$$\mathrm{A1} \cup \mathrm{A2} \cup \cdots \cup \mathrm{An} = \mathrm{E}$$

5ª DIAPOSITIVA **Minuto: 01:02**

Contar con un sistema completo de sucesos nos ayuda a calcular probabilidades de otros sucesos más complejos pues cualquier suceso B puede descomponerse en componentes de dicho sistema.

Imaginemos que la tarta anterior tiene en su perímetro exterior, como adorno, nata montada y a nosotros no interesa conocer la probabilidad de tomar tarta sin nata montada. En este caso cortaríamos de los cuatro trozos iniciales que formaban el sistema completo de sucesos la parte con nata y para saber la probabilidad de tomar un trozo de tarta sin nata, suceso B, realizaríamos los siguientes análisis:

(Trozo de tarta 1 sin nata) U (Trozo de tarta 2 sin nata) U (Trozo de tarta 3 sin nata) U (Trozo de tarta 4 sin nata)

De forma general sería como calcular B como la unión de todas las intersecciones de B con los sucesos del sistema, es decir,

$B = (B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3) \cup (B \cap A_4)$

Dividimos el problema en varios problemas más sencillos y fáciles de calcular haciendo honor al dicho popular "Divide y vencerás".

6ª DIAPOSITIVA **Minuto: 01:20**

Este es el fundamento del teorema de la probabilidad total que nos dice que si conocemos la probabilidad de B en cada uno de los sucesos que componen el sistema completo $A_1, A_2, A_3, ..., A_n$, entonces podemos calcular B como la suma de las probabilidades de cada una de las intersecciones de B con los elementos del sistema.

Si hemos descompuesto **B** = ($B \cap A_1$) U ($B \cap A_2$) U … U ($B \cap A_n$) entonces la probabilidad de B es P(B) = $P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_n)$

Si continuamos con el ejemplo de la tarta esto anterior sería equivalente a calcular la probabilidad de tomar un trozo de tarta sin nata como la suma de las probabilidades de tomar tarta sin nata en cada uno de los 4 trozos en los cuales dividimos la tarta.

¿Pero cómo calculamos la probabilidad de cada una de esas intersecciones? Por ejemplo $P(B \cap A_1)$, pues de una manera muy sencilla utilizando la fórmula de la probabilidad condicionada que nos dice que:

$$P(B|A_1) = \frac{P(\mathrm{B} \cap A_1)}{P(A_1)}$$

Y despejando la probabilidad de la intersección : $P(\mathrm{B} \cap A_1) = P(B|A_1)P(A_1)$

Si hacemos esto con cada una de las intersecciones de B con los sucesos del sistema tenemos que la probabilidad del suceso complejo B sería igual a:

$P(B) = P(B|A_1)\ P(A_1) + P(B|A_2)\ P(A_2) + \ldots + P(B|A_n)\ P(A_n)$

Por tanto el teorema de la probabilidad total es muy útil para calcular probabilidades complejas, como el suceso B, a través de las intersecciones de dicho suceso con los sucesos elementales que forman el sistema completo de sucesos.

7ª DIAPOSITIVA **Minuto: 01:51**

¿Y si nos interesase saber la probabilidad de tomar el trozo de tarta número 1 sabiendo que no tiene nata? Es decir, y si queremos saber la probabilidad de uno de los sucesos elementales sabiendo que ha sucedido el suceso complejo B? Es decir, $P(Ai|B)$

Pues el Teorema de Bayes nos dice que, si conocemos la probabilidad de B en cada una de los sucesos elementales del sistema completo, entonces si ocurre B, podemos calcular la probabilidad de que ocurra A_i como:

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

Donde la probabilidad de B puede calcularse previamente utilizando el teorema de la probabilidad total visto anteriormente.

8ª DIAPOSITIVA **Minuto: 02:27**

Estudiemos los conceptos vistos con un ejemplo sencillo y completo.

Pensemos que un fabricante que cuenta con 3 plantas distintas, diseminadas por todo el territorio nacional, donde fabrica su producto estrella, microchips. En la planta 1, que denominamos P_1, fabrica 500 unidades al mes. En la planta 2, P_2, produce 2000 unidades y por último, en la planta 3, P_3, fabrica 1500 unidades. Esta información ya nos orienta sobre la probabilidad de fabricar el microchip en cada una de las 3 plantas al ser un sistema completo de sucesos.

$P(P_1)=\frac{500}{4000}=\frac{1}{8}$; $P(P_2)=\frac{2000}{4000}=\frac{1}{2}$; $P(P_3)=\frac{1500}{4000}=\frac{3}{8}$

Nótese además que el sistema de plantas del fabricante forma un sistema completo de sucesos, no se fabrica un mismo microchip en varias plantas y la unión de las producciones de todas las plantas son la producción total del fabricante.

9ª DIAPOSITIVA **Minuto: 02:56**

Introducimos una nueva característica de los microchips en el problema y es que estos puedan ser defectuosos. Este suceso sería el equivalente a B en la teoría.

La probabilidad de producir un microchip defectuoso en cada una de las 3 plantas es 0,02 en la planta 1, 0,015 en la planta 2 y 0,03 en la planta 3.

Para ayudar a aplicar de forma más sencilla el teorema de la probabilidad total y el teorema de Bayes estos datos pueden organizarse en forma de árbol.

10ª DIAPOSITIVA **Minuto: 03:10**

Del fabricante de microchips saldrían 3 ramas, una para cada una de las plantas.

Produce microchips en la planta 1 con una probabilidad de 1/8, en la planta 2 con probabilidad de ½ y en la planta 3 con una probabilidad de 3/8.

La probabilidad de que el microchip sea de la planta 1 es 1/8 pues fabrica 500 de los 4000 totales. Este sería el suceso P1. De este nodo intermedio salen 2 ramas, una para indicar que el microchip puede ser defectuoso y otra para indicar que puede ser correcto. En este caso, partiendo de la planta 1 tendríamos que la probabilidad de que un microchip sea defectuoso es igual a 0.02 por lo que la probabilidad de su contrario, que sea correcto asciende a 0,98.

Igualmente, del nodo de la planta 2 salen 2 ramas, defectuosos con probabilidad de 0,015 y correctos, con probabilidad 0,985, que es igual a (1-0,015).

Por último, de la tercera rama que sale del fabricante de microchips y llega al nodo de la planta 3 salen otras 2 ramas. Los microchips de la planta 3, que se producen con una probabilidad de 3/8 son defectuosos con una probabilidad de 0,03 y correctos con probabilidad de 0,97.

A esta forma de representar el problema de probabilidades, tras partir de un sistema completo de sucesos se le denomina problema en forma de árbol.

11ª DIAPOSITIVA **Minuto: 03:50**

Este mismo problema de cálculo de probabilidades se puede expresar en forma de tabla donde las filas son las plantas de fabricación de microchips o sucesos que componen el sistema completo de sucesos y en las columnas se tiene la segunda característica estudiada en todos ellos como es que sean defectuosos o correctos.

Así pues, en la primera fila tenemos los microchips que se fabrican en la planta 1, 500 unidades, pero diferenciando los que son defectuosos y los que son correctos. Como sabíamos que el 2% eran defectuosas se conoce que se han fabricado 500*0,02=10 unidades defectuosas en la planta 1. Como el resto son correctos se tiene que 500-10=490 son los correctos.

Esto mismo haríamos con los microchips de la planta 2, donde conocemos que 30 son defectuosos y 1970 son correctos y con los de la planta 3, de los cuales, de los 1500 fabricados, 45 son defectuosos y 1455 son correctos.

	Defectuosos	Correctos	
Planta 1	0,02*500=10	500-10=490	500
Planta 2	0,015*2000=30	2000-30=1970	2000
Planta 3	0,03*1500=45	1500-45=1455	1500
	85	3915	4000

Planteado el problema, ¿cuál es la probabilidad de que el microchip instalado en un móvil por el fabricante sea defectuoso?

Ya hemos visto las formas, árbol y tabla, en las que podemos plantear el problema, ahora somos nosotros los que tenemos que decidir qué planteamiento nos parece más cómodo para calcular la probabilidad solicitada.

12ª DIAPOSITIVA **Minuto: 04:36**

Para resolver la duda del fabricante sobre cuál es la probabilidad de instalar un microchip que sea defectuoso, independientemente de la planta en la que se haya fabricado utilizaremos el teorema de la probabilidad total que hemos estudiado.

Los sucesos que componen el sistema son las plantas de producción, P_1, P_2 y P_3 y el suceso **D** es “microchip defectuoso”.

13ª DIAPOSITIVA **Minuto: 04:36**

Si utilizamos la tabla de contingencia bastaría dividir el número de microchips defectuosos entre el total de microchips por lo que la probabilidad solicitada sería 85 unidades fabricadas defectuosas en total dividido entre las 4000 unidades que se hacen en las 3 plantas.

$$P(D) = \frac{85}{4000} = 0{,}02125$$

Esto hace que la probabilidad de seleccionar un microchip en general, entre todos los fabricados, tenga una probabilidad de 0,02125 de ser defectuoso.

14ª DIAPOSITIVA **Minuto: 05:10**

Si calculamos la probabilidad solicitada a través del problema de árbol debemos sumar todos los caminos que nos lleven de la fábrica al nodo final "defectuoso".

En este caso se observa de forma más clara el teorema de la probabilidad total.

P(microchip defectuoso)= P(planta1 ∩ defectuoso) + P(planta2 ∩ defectuoso) + P(planta3 ∩ defectuoso)

Cada una de estas probabilidades son los caminos que parten de la fábrica al nodo de defectuoso.

P(planta1 ∩ defectuoso)=(1/8)*0,02=$P(P_1)*P(P_1/D)$=0,0025

P(planta2 ∩ defectuoso)=(1/2)*0,015=0,0075

P(planta3 ∩ defectuoso)=(3/8)*0,03=0,01125

Para calcular la probabilidad de ser defectuoso basta con ir sumando estos caminos 0,0025+0,0075+0,01125=**0,02125**

Por tanto, como hemos podido comprobar, independientemente del planteamiento del problema en forma de árbol o en forma de tabla, el teorema de la probabilidad total nos dice que la probabilidad de seleccionar un microchip al azar, independientemente de la planta en la que se haya fabricado, y que este sea defectuoso asciende a 0.02125.

15ª DIAPOSITIVA **Minuto: 05:24**

Ahora el fabricante ha seleccionado un microchip de su almacén central y observa que es defectuoso, ¿qué probabilidad hay de que haya sido fabricado en la planta 1?

Nosotros conocemos la probabilidad de que sea defectuoso sabiendo que se ha fabricado en la planta 1 pero nos piden una probabilidad diferente. Nos piden la probabilidad de ser de la planta 1 pero conociendo que se ha dado el suceso D “ser defectuoso”.

16ª DIAPOSITIVA **Minuto: 05:24**

Aplicando el teorema de Bayes podemos calcularla pues:

$$P(P1|D) = \frac{P(D \cap P1)}{P(D)} = \frac{P(D|P1)P(P1)}{P(D)} = \frac{0{,}02 \cdot \frac{1}{8}}{0{,}02125} = \frac{2}{17}$$

Hay que destacar que la probabilidad de ser defectuoso, en general, es la probabilidad calculada en el apartado anterior a partir del teorema de la probabilidad total y no debe nunca confundirse con la probabilidad de ser defectuoso sabiendo que el microchip se ha fabricado en la planta 1.

Utilizando la tabla de contingencia, el cálculo es más sencillo. Si sabemos que hay 85 microchip defectuosos en total y que 10 son los que se han producido en la planta 1, la probabilidad de ser de la planta 1 sabiendo de antemano que el microchip es defectuoso será 10/85=2/17, mismo resultado anterior.

17ª DIAPOSITIVA **Minuto: 05:48**

Por último, el fabricante desea conocer la probabilidad de que un microchip que ya ha sido instalado en un móvil tenga 2 características, sea de la planta 2 y además sea defectuoso. Ahora el fabricante no conoce nada con antelación del microchip. La probabilidad solicitada es la que se obtiene de una intersección entre los sucesos “planta 2” y “defectuoso”.

18ª DIAPOSITIVA **Minuto: 05:48**

Siguiendo el esquema de árbol, y como lo que nos interesa es calcular la probabilidad de una intersección P(planta 2 ∩ defectuoso) debemos calcular la multiplicación de todas las probabilidades de los nodos por los que pasa dicho camino, es decir, “planta 2” y “defectuoso”. De esta forma, para calcular la probabilidad solicitada multiplicamos ½ por 0,015 obteniéndose que el resultado es 0,0075.

$$P(P2 \cap D) = \frac{1}{2} \cdot 0{,}015 = 0{,}0075$$

Según el planteamiento de la tabla de contingencia, si hay 30 microchips defectuosos que salen de la planta 2 y hay un total de 4000 microchips fabricados por la empresa, la probabilidad de seleccionar un microchip defectuoso que se ha fabricado en la planta 2 de los 4000 que hay en total será 30/4000=0,0075.

17ª DIAPOSITIVA **Minuto: 04:36**

Muchas gracias por la atención.

Esperamos que la serie "Estadística para tod@s" sea un granito más para alcanzar una Universidad más inclusiva.