Auxiliar 3

Probabilidades totales y teorea de Bayes.

Profesor: Vicente Acuña
Auxiliares: Sebastián López, Bruno Hernández

L1. Una empresa de celulares tiene dos máquinas A y B. El 54 % de los celulares producidos son hechos por la máquina A y el resto por la máquina B. No todos los celulares producidos están en buen estado.

La proporción de celulares defectuosos hechos por A es 0.2 y por B es 0.5.

a) ¿Cuál es la probabilidad de que un celular sea defectuoso?

b) ¿Cuál es la probabilidad de que, sabiendo que un celular es defectuoso, proceda de la máquina A?

L2. Se tiene una caja con **n** bolas blancas y **1** bola negra y se dispone de un dado perfecto de 6 caras, es decir cada una de las caras tiene igual probabilidad de aparecer.

Se tira el dado, si sale par se sacan dos bolas de la caja (sin reposición) y si sale impar se saca sólo una. Calcule la probabilidad que en el dado haya salido un **6** sabiendo que la bola negra fue extraída.

L3. Para predecir el tiempo, un día es clasificado como seco o lluvioso. Por experiencia se sabe que la probabilidad de que un día sea igual al anterior se asume constante e igual a p.

a) si el 1 de abril es seco con probabilidad β, muestre que la probabilidad de que el n-ésimo día del año (contando a partir del 1 de abril) sea seco (P_n) queda dado por:

$$P_n = \left[\left(\beta - \frac{1}{2}\right)(2p-1)^{n-1}\right] + \frac{1}{2}$$

b) Si el 16 de abril está seco, calcule la probabilidad de que el 14 de abril también lo haya estado. Para esto considere $\beta = 1$, $p = \frac{9}{10}$.

Pauta

L1. Antes de empezar a calcular la primera pregunta hagamos un resumen de la información proporcionada por el enunciado y escribámosla en términos de probabilidades. En primer lugar debemos definir los eventos que usaremos:

- A_x: es el evento de que el celular X proceda de la máquina A.
- B_x: es el evento de que el celular X proceda de la máquina B.
- X_d: es el evento de que el celular X salga defectuoso.

Como nos dicen que el 54 % de los celulares producidos son hechos por la máquina A, entonces podemos afirmar que para cualquier celular X se tiene que:

$$\mathbb{P}(A_x) = 0{,}54$$

Y, por lo tanto, como los celulares sólo pueden proceder de A o de B:

$$\mathbb{P}(B_x) = 1 - \mathbb{P}(A_x) = 0{,}46$$

Además nos dicen que que la proporción de celulares defectuosos hechos por A es de 0.2. En este caso podemos notar que, adicional a que el celular salió defectuoso, nos dan la información de que el celular fue producido por la máquina A, esto corresponde a una probabilidad condicionada ("Si el celular procede de A, entonces tiene una proporcion[1] de 0.2 de ser defectuoso"). En términos matemáticos:

$$\mathbb{P}(X_d|A_x) = 0{,}2$$

Análogamente con B:

$$\mathbb{P}(X_d|B_x) = 0{,}5$$

a) Nos piden calcular la probabilidad de que un celular cualquiera salga defectuoso $\mathbb{P}(X_d)$, como para cada celular tenemos casos diferentes ocuparemos la fórmula de probabilidades totales, separando los casos en que es proveniente de la máquina A o de la máquina B:

$$\mathbb{P}(X_d) = \mathbb{P}(X_d|A_x)\mathbb{P}(A_x) + \mathbb{P}(X_d|B_x)\mathbb{P}(B_x) = (0{,}2 \times 0{,}54) + (0{,}5 \times 0{,}46)$$

$$\mathbb{P}(X_d) = 0{,}338$$

b) Ahora nos preguntan cuál es la probabilidad de que un celular proceda de la máquina A sabiendo que es defectuoso. Como nos dan información adicional a la probabilidad, debemos condicionar. Es decir, debemos calcular $\mathbb{P}(A_x|X_d)$. Para ello ocuparemos el teorema de Bayes para probabilidades condicionales[2]. Bayes nos permite invertir la condicionalidad, haciendo uso de la parte anterior calculamos:

$$\mathbb{P}(A_x|X_d) = \frac{\mathbb{P}(X_d|A_x)\mathbb{P}(A_x)}{\mathbb{P}(X_d)} = \frac{0{,}2 \times 0{,}54}{0{,}338} = 0{,}32$$

[1] Note la diferencia entre el dato entregado en forma de porcentaje versus aquel dato que es "proporción", ambos significan una probabilidad pero la forma de expresar la información es diferente.

[2] Al momento de definir la condicionalidad no confundirse con el orden, uno condiciona siempre por información extra que posea.

L2. Definimos las variables a ocupar:

- D: es el resultado de un lanzamiento del dado, es decir alguno de $\{1, 2, 3, 4, 5, 6\}$.
- B: es el color de una bola sacada de la caja, puede ser b o n, es decir *Blanco* o *Negro*.

Notemos que nos dan información que condiciona la probabilidad ("Sabiendo que la bola negra fue extraída"), entonces lo que nos piden calcular es $\mathbb{P}(D = 6|B = n)$. Ocupando el Teorema de Bayes se tiene que:

$$\mathbb{P}(D = 6|B = n) = \frac{\mathbb{P}(B = n|D = 6)\mathbb{P}(D = 6)}{\mathbb{P}(B = n)}$$

Calculemos estas probabilidades por separados:
Para calcular la probabilidad de sacar la bola negra $\mathbb{P}(B = n)$ ocuparemos la fórmula de probabilidades totales, separando los casos en que el dado dio un resultado Par o uno $Impar$:

$$\mathbb{P}(B = n) = \mathbb{P}(B = n|D = \{par\})\mathbb{P}(D = \{par\}) + \mathbb{P}(B = n|D = \{impar\})\mathbb{P}(D = \{impar\})$$

Sabemos que el dado no está cargado y hay la misma cantidad de números pares $\{2, 4, 6\}$ que números impares $\{1, 3, 5\}$, por lo tanto:

$$\mathbb{P}(D = \{par\}) = \mathbb{P}(D = \{impar\}) = \frac{3}{6} = \frac{1}{2}$$

Mientras que $\mathbb{P}(B = n|D = \{impar\})$ corresponde a la probabilidad de sacar la bola negra si en el dado salió un valor impar. Recordemos que si el dado arrojaba un impar, entonces sacábamos sólo una bola de la caja, de entre las $n + 1$ total de bolas. Por lo tanto:

$$\mathbb{P}(B = n|D = \{impar\}) = \frac{1}{n+1}$$

Del mismo modo; si $\mathbb{P}(B = n|D = \{par\})$ es la probabilidad de sacar la bola negra dado que el dado resultó ser un número par y, por ende, tengo la opción de sacar 2 bolas de la caja.
Para calcular esta probabilidad sacaremos casos favorables y casos totales. Los casos totales representan todos los pares de bolas que puedo sacar de la caja con $n+1$ bolas, es decir; todo subconjunto de 2 bolas de un total de $n + 1$, lo que se cuenta con una combinatoria: $\binom{n+1}{2}$.
Mientras que los casos favorables son todos los pares de bolas en la que una corresponde a la bola negra; si tengo la bola negra, quiere decir que la segunda es blanca y tego n opciones para elegirla, por lo tanto tengo n casos favorables.

En conclusión:

$$\mathbb{P}(B = n|D = \{par\}) = \frac{n}{\binom{n+1}{2}} = \frac{n}{\frac{(n+1)!}{(n-1)!\,2!}} = \frac{n(n-1)!\,2!}{(n+1)!} = \frac{2}{n+1}$$

Sólo hace falta reemplazar para conocer $\mathbb{P}(B=n)$.

$$\mathbb{P}(B=n) = (\frac{2}{n+1} \times \frac{1}{2}) + (\frac{1}{n+2} \times \frac{1}{2}) = \frac{3}{2}\left(\frac{1}{n+1}\right)$$

$\mathbb{P}(D=6)$ es simplemente la probabilidad de que en el dado salga un 6. Como sabemos, el dado no está cargado, por lo tanto todos los valores tienen la misma probabilidad de uno entre seis:

$$\mathbb{P}(D=6) = \frac{1}{6}$$

Por último; $\mathbb{P}(B=n|D=6)$, la probabilidad de sacar la la bola negra dado que el dado arrojó un 6, es un dato conocido, pues 6 es número par. Entonces:

$$\mathbb{P}(B=n|D=6) = \mathbb{P}(B=n|D=\{par\}) = \frac{2}{n+1}$$

Finalmente, reemplazamos todos los datos en la expresión inicial:

$$\mathbb{P}(D=6|B=n) = \frac{\mathbb{P}(B=n|D=6)\mathbb{P}(D=6)}{\mathbb{P}(B=n)} = \frac{\left(\frac{2}{n+1}\right)\frac{1}{6}}{\frac{3}{2}\left(\frac{1}{n+1}\right)} = \frac{\frac{1}{3}}{\frac{3}{2}} = \frac{2}{9}$$

$$\mathbb{P}(D=6|B=n) = \frac{2}{9}$$

L3. Definimos las variables que usaremos:

- X_n: es la variable que indica si el día n está seco (s) o lluvioso (l).

a) Como tenemos una expresión que depende de un número natural ocuparemos el método de inducción para probarla. Partiremos comprobando que el caso base $n=1$ es válido: si $n=1$ quiere decir que partimos desde el primer día de la cuenta, osea el 1 de abril, lo que (por enunciado) tiene probabilidad β de estar seco:

$$P_1 = \beta = \beta - \frac{1}{2} + \frac{1}{2} = \left(\beta - \frac{1}{2}\right) + \frac{1}{2}$$

$$= \left(\beta - \frac{1}{2}\right)1 + \frac{1}{2} = \left(\beta - \frac{1}{2}\right)(2p-1)^{1-1} + \frac{1}{2}$$

Que es exactamente la expresión evaluada en 1.[3]

Entonces asumimos que la expresión es cierta para todos los días, desde el primero hasta el día $n-1$. Luego, por probabilidades totales:

$$P_n = \mathbb{P}(X_n = s) = \mathbb{P}(X_n = s|X_{n-1} = s)\mathbb{P}(X_{n-1} = s) + \mathbb{P}(X_n = s|X_{n-1} = l)\mathbb{P}(X_{n-1} = l)$$

Las probabilidadas condicionadas que aparecen en la expresión indican la probabilidad del pronóstico del tiempo dado los posibles casos del día anterior. Por enunciado, la

[3] En rigor, es necesario que $p \neq \frac{1}{2}$ pues no son válidas las expresiones de la forma 0^0. Pero ese caso es descartado pues P_n sería constante.

probabilidad de que dos días consecutivos tengan el mismo pronóstico es p, por lo tanto:

$$\mathbb{P}(X_n = s|X_{n-1} = s) = p$$

$$\mathbb{P}(X_n = s|X_{n-1} = l) = 1 - p$$

Además $\mathbb{P}(X_{n-1} = s)$ es la probabilidad de que el día $n-1$ esté seco, por lo tanto:

$$\mathbb{P}(X_{n-1} = s) = P_{n-1}$$

$$\mathbb{P}(X_{n-1} = l) = 1 - P_{n-1}$$

Reemplazando en la expresión anterior:

$$P_n = \mathbb{P}(X_n = s) = \mathbb{P}(X_n = s|X_{n-1} = s)\mathbb{P}(X_{n-1} = s) + \mathbb{P}(X_n = s|X_{n-1} = l)\mathbb{P}(X_{n-1} = l)$$

$$= pP_{n-1} + (1-p)(1-P_{n-1}) = (2p-1)P_{n-1} + 1 - p$$

Por nuestra hipótesis inductiva, P_{n-1} puede escribirse como la fórmula evaluada en el término $n-1$, por lo tanto lo reemplazo:

$$P_n == (2p-1)P_{n-1} + 1 - p = (2p-1)\left(\left[\left(\beta - \frac{1}{2}\right)(2p-1)^{n-2}\right] + \frac{1}{2}\right) + 1 - p$$

$$= \left[\left(\beta - \frac{1}{2}\right)(2p-1)^{n-1}\right] + \frac{1}{2}(2p-1) + 1 - p = \left[\left(\beta - \frac{1}{2}\right)(2p-1)^{n-1}\right] + p - \frac{1}{2} + 1 - p$$

$$= \left[\left(\beta - \frac{1}{2}\right)(2p-1)^{n-1}\right] + \frac{1}{2}$$

Así mostramos que la expresión es válida para n si asumimos que lo es para $n-1$. Queda demostrado el paso inductivo.

b) Para esta parte ocuparemos el Teorema de Bayes, esto es:

$$\mathbb{P}(X_{14} = s|X_{16} = s) = \frac{\mathbb{P}(X_{16} = s|X_{14} = s)\mathbb{P}(X_{14} = s)}{\mathbb{P}(X_{16} = s)}$$

Por la parte anterior sabemos que: $\mathbb{P}(X_{14} = s) = P_{14}$ y $\mathbb{P}(X_{16} = s) = P_{16}$. Entonces sólo nos basta con calcular $\mathbb{P}(X_{16} = s|X_{14} = s)$ que es la probabilidad de que el día 16 haya estado seco sabiendo que el día 14 estuvo seco. Podemos notar que esta probabilidad ya no depende de β pues ahora no nos sirve de nada la información sobre el día 1, sino que sólo depende si los días entre el 16 y el 14 fueron iguales o diferentes:

Si partimos del hecho de que $X_{14} = s$ entonces podemos afirmar que $\mathbb{P}(X_{14} = s) = 1$, por lo tanto, condicionando sobre la condición:

$$\mathbb{P}(X_{16} = s|X_{14} = s) =$$

$$\mathbb{P}(X_{16} = s|X_{15} = s \wedge X_{14} = s)\mathbb{P}(X_{15} = s \wedge X_{14} = s) + \mathbb{P}(X_{16} = s|X_{15} = l \wedge X_{14} = s)\mathbb{P}(X_{15} = l \wedge X_{14} = s)$$

Considerando que $\mathbb{P}(X_{14} = s) = 1$, y ocupando la definición de probabilidad condicionada, podemos ocupar el siguiente trucazo:

$$\mathbb{P}(X_{15} = s \wedge X_{14} = s) = \frac{\mathbb{P}(X_{15} = s \wedge X_{14} = s)}{\mathbb{P}(X_{14} = s)}\mathbb{P}(X_{14} = s)$$

$$= \mathbb{P}(X_{15} = s|X_{14} = s)\mathbb{P}(X_{14} = s) = p \times 1 = p$$

De la mima forma, obtenemos que:

$$\mathbb{P}(X_{15} = l \wedge X_{14} = s) = 1 - p$$

Por otro lado, $\mathbb{P}(X_{16} = s|X_{15} = s \wedge X_{14} = s) = p$, pues el día 16 tiene el mismo pronóstico del día 15. Así mismo $\mathbb{P}(X_{16} = s|X_{15} = l \wedge X_{14} = s) = 1-p$ pues el día 16 tiene pronóstico diferente[4] al día 15. Reemplazando todo en nuestra fórmula de probabilidades totales:

$$\mathbb{P}(X_{16} = s|X_{14} = s) =$$

$$\mathbb{P}(X_{16} = s|X_{15} = s \wedge X_{14} = s)\mathbb{P}(X_{15} = s \wedge X_{14} = s) + \mathbb{P}(X_{16} = s|X_{15} = l \wedge X_{14} = s)\mathbb{P}(X_{15} = l \wedge X_{14} = s)$$

$$= p \times p + (1-p) \times (1-p) = p^2 + (1-p)^2$$

Otra forma (menos formal) de calcular $\mathbb{P}(X_{16} = s|X_{14} = s)$ es con el siguiente esquema:

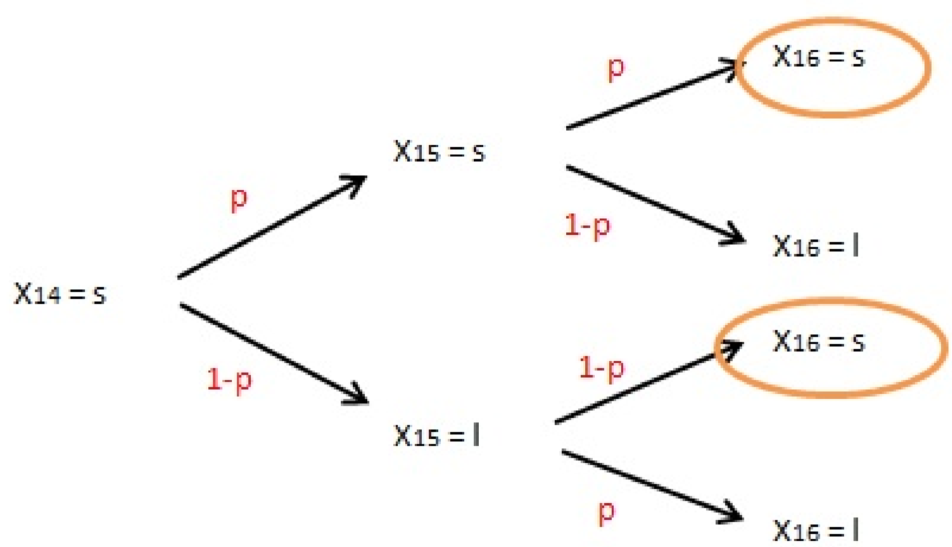

Figura 1: Esquema

[4] Note que estas probabilidades sólo dependen del día anterior y no de cómo haya estado hace 2 o 3 días antes.

Donde los valores en las flechas son las probabilidades dependiendo si los días tienen el mismo pronóstico o no.
En un óvalo se indican los estados en que $X_{16} = s$, para calcular la probabilidad total de $\mathbb{P}(X_{16} = s | X_{14} = s)$ calculamos la probabilidad de cada camino que termina en $X_{16} = s$ y los sumamos. La probabilidad de los camino es sólamente multiplicar las probabilidades de cada paso del camino. Entoces el camino $(X_{14} = s) \to (X_{15} = s) \to (X_{16} = s)$ tiene probabilidad p^2. La probabilidad del camino $(X_{14} = s) \to (X_{15} = l) \to (X_{16} = s)$ es $(1-p)^2$. Entonces:

$$\mathbb{P}(X_{16} = s | X_{14} = s) = p^2 + (1-p)^2$$

Finalmente[5], reemplazamos este valor en la fórmula del teorema de bayes del comienzo:

$$\mathbb{P}(X_{14} = s | X_{16} = s) = \frac{\mathbb{P}(X_{16} = s | X_{14} = s)\mathbb{P}(X_{14} = s)}{\mathbb{P}(X_{16} = s)}$$

$$= \frac{(p^2 + (1-p)^2)P_{14}}{P_{16}}$$

[5] Ustedes pueden reemplazar $\beta = 1$ y $p = \frac{9}{10}$ si eso les aliviana los cálculos.