

# SOL3070 Análisis de Datos Categóricos

## Tarea corta 2, respuestas

Ponderacion: 6% de la nota final del curso.

### Problema 1:

Tras escuchar en las noticias sobre un brote de una nueva enfermedad llamada “COVID”, decides realizarte un test para determinar si estás contagiado con el virus o no. Luego de recibir tus resultados, te informan que has dado positivo. Recientemente, leíste un artículo en [Nature](#) donde se detalla que:

- La “sensibilidad” del test es del (85%) (probabilidad de que una persona contagiada obtenga un test positivo).
- La tasa de falsos positivos del test es del (10%) (probabilidad de que una persona no contagiada obtenga un test positivo).
- La prevalencia de la enfermedad es del (10%) (probabilidad de que una persona esté contagiada con el virus).

### Pregunta 1:

¿Cuál es la probabilidad de que realmente estés contagiado con el “COVID” si recibes un test positivo?

### Solución 1:

Usamos el teorema de Bayes para resolver el problema:

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

Donde:

- $\mathbb{P}(A)$  es la prevalencia de la enfermedad: (0.10)
- $\mathbb{P}(B)$  es la probabilidad de obtener un test positivo
- $\mathbb{P}(B|A)$  es la sensibilidad del test: (0.85)

Para calcular  $\mathbb{P}(B)$  usamos la ley de probabilidad total:

$$\mathbb{P}(B) = \mathbb{P}(A) \times \mathbb{P}(B|A) + \mathbb{P}(A') \times \mathbb{P}(B|A')$$

Donde  $\mathbb{P}(A')$  es la probabilidad de no estar contagiado y  $\mathbb{P}(B|A')$  es la probabilidad de obtener un falso positivo. Calculamos:

$$\mathbb{P}(B) = 0.10 \times 0.85 + 0.90 \times 0.10$$

$$\mathbb{P}(B) = 0.085 + 0.090$$

$$\mathbb{P}(B) = 0.175$$

Usando estos valores en la fórmula de Bayes, obtenemos:

$$\mathbb{P}(A|B) = \frac{0.85 \times 0.10}{0.175}$$

$$\mathbb{P}(A|B) \approx 0.4857$$

Por lo tanto, hay aproximadamente un (48.57%) de probabilidad de que realmente estés contagiado con el “COVID”.

### Problema 2:

Un aplicado sociólogo toma una muestra aleatoria de aproximadamente un 10 millones de personas, todas de la misma población (o sea, con idénticos parámetros) y reporta los hallazgos en la siguiente tabla:

	Contagiado	No Contagiado
Test Positivo	800060	899986
Test Negativo	200015	8099877

### Pregunta 2:

Usando los datos provistos por el sociólogo, **estima**  $\mathbb{P}(\text{Contagiado}|\text{Test Positivo})$  sin usar el teorema de Bayes.

### Solución 2:

Para calcular la probabilidad de estar contagiado dado que el resultado de la prueba es positivo dividimos el número de verdaderos positivos por el total de resultados positivos:

$$\hat{\mathbb{P}}(\text{Contagiado}|\text{Test Positivo}) = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Falsos Positivos}}$$

En R:

```
# Probabilidad de estar contagiado dado que el test es positivo

P_Contagiado_dado_Test_Positivo =
  tabla["Test Positivo", "Contagiado"] /
  (tabla["Test Positivo", "Contagiado"] + tabla["Test Positivo", "No Contagiado"])

print(P_Contagiado_dado_Test_Positivo)
```

```
[1] 0.4706108
```