UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO

Departamento de Ciencias Básicas y Modelado

Maestría en Modelado y Simulación –
Maestría en Analítica de Datos

Modelos probabilísticos y análisis estadístico

Parte 3: Estimación de parámetros, Inferencia estadística bayesiana y modelos jerárquicos

Tema 2: Probabilidad condicional, Regla de Bayes

Profesor: Javier Riascos Ochoa, PhD
(javier.riascos@utadeo.edu.co)

CONTENIDO

Contenido

1. Introducción: Modelos e inferencia
2. Inferencia estadística clásica y estimación de parámetros.
   - Método de momentos.
   - Método de máxima verosimilitud.
3. Probabilidad condicional.
   - Teorema de probabilidad total.
   - Regla de Bayes.
   - Actualización bayesiana.
4. Inferencia estadística bayesiana.
5. Modelos jerárquicos bayesianos.

UTADEO
UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO

Probabilidad condicional

Probabilidad condicional: Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. si $A, B \in \mathcal{F}$ con $P(A) > 0$, entonces se define la probabilidad del evento B bajo la condición de A como:

$$P(B \mid A) := \frac{P(A \cap B)}{P(A)} = \frac{P(A, B)}{P(A)} = \frac{P(A\ y\ B)}{P(A)}$$

Se lee:

"Probabilidad condicional de B, condicionada al evento A" o *"Probabilidad condicional de B, dado A"*

"Probabilidad de B dado A"

En otras palabras:

Es la probabilidad de que ocurra *B* sabiendo que ha ocurrido *A*.

Probabilidad condicional

Probabilidad condicional:

También se puede escribir de la siguiente forma:

$$\boldsymbol{P(B|A) := \frac{P(A,B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}}$$

Demostración:

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

$$P(A,B) = P(A|B)P(B)$$

Probabilidad condicional

Ejemplo en R: Fumar... ¿principal causa de cáncer de pulmón?

Ver código *ProbCond.R*

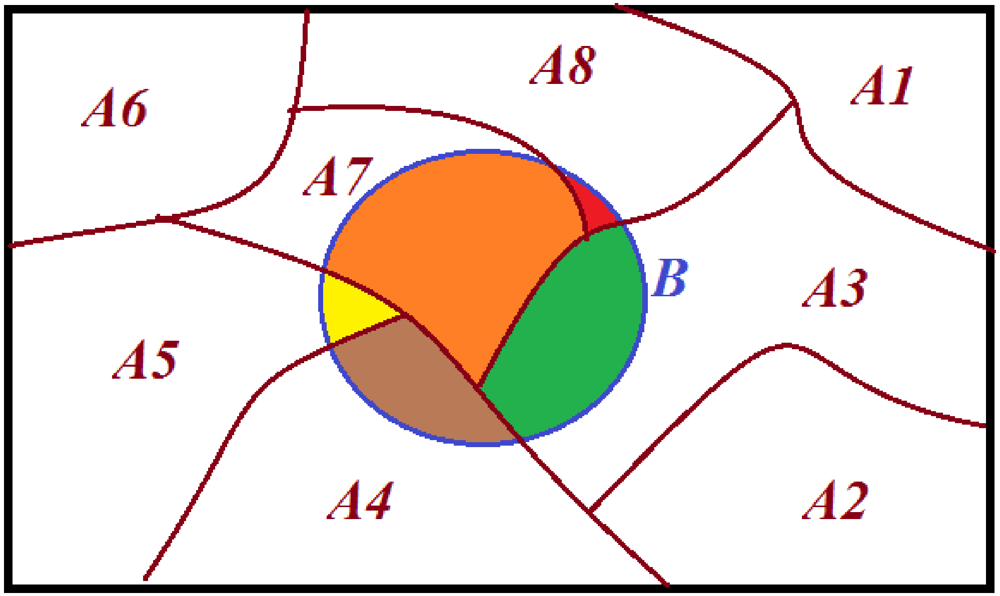
Probabilidad total

Teorema de probabilidad total:

Sea $A_1, A_2, \ldots$ una partición finita o numerable del espacio de estados S, es decir, $A_i \cap A_j = \emptyset$, para todo $i \neq j$ y $\bigcup_{i=1}^{\infty} A_i = S$ con $P(A_i) > 0$ para todo i. Entonces para cualquier $B \in \mathcal{F}$ se satisface:

$$P(B) = \sum_n P(B \cap A_n)$$

$$= \sum_n P(B|A_n)\, P(A_n)$$



Probabilidad total

Ejemplo (Blanco, 1.63):

El señor Rodríguez sabe que hay un 40% de probabilidad de que la empresa en la cual labora abra una sucursal en Montevideo. Si lo hace, la probabilidad de que él sea nombrado gerente de dicha sucursal es de un 80%. Si no lo hace, la probabilidad de que el Sr. Rodríguez sea nombrado gerente en otra sucursal es tan sólo 10%. Se desea calcular la probabilidad de que el Sr. Rodríguez sea nombrado gerente de una sucursal de su empresa.

Probabilidad total

Ejemplo (Blanco, 1.63):

El señor Rodríguez sabe que hay un 40% de probabilidad de que la empresa en la cual labora abra una sucursal en Montevideo. Si lo hace, la probabilidad de que él sea nombrado gerente de dicha sucursal es de un 80%. Si no lo hace, la probabilidad de que el Sr. Rodríguez sea nombrado gerente en otra sucursal es tan sólo 10%. Se desea calcular la probabilidad de que el Sr. Rodríguez sea nombrado gerente de una sucursal de su empresa.

Solución:

Sean: *B:="el Sr. Rodríguez es nombrado gerente"*

A:= "la empresa abre una sucursal en Montevideo"

Entonces por el teorema de probabilidad total:

$$\begin{aligned} P(B) &= P(B|A)P(A) + P(B|A^c)P(A^c) \\ &= 0.8 \times 0.4 + 0.10 \times 0.60 \\ &= 0.38 \end{aligned}$$

Regla de Bayes

Regla de Bayes:

Sea $A_1, A_2, \ldots$ una partición finita o numerable del espacio de estados S, con $P(A_i) > 0$ para todo i. Entonces se satisface para todo $B \in \mathcal{F}$ con $P(B) > 0$:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)}$$
$$= \frac{P(A_i)P(B|A_i)}{\sum_j P(B|A_j)\,P(A_j)}; \quad \text{para todo } i$$

UTADEO
UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO

Regla de Bayes

Regla de Bayes:

Sea $A_1, A_2, \ldots$ una partición finita o numerable del espacio de estados S, con $P(A_i) > 0$ para todo i. Entonces se satisface para todo $B \in \mathcal{F}$ con $P(B) > 0$:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)}$$
$$= \frac{P(A_i)P(B|A_i)}{\sum_j P(B|A_j)\,P(A_j)}; \quad \text{para todo } i$$

Interpretación:

Supóngase que los eventos $A_1, A_2, \ldots$ son todas las posibles causas, mutuamente excluyentes, de un evento B. Bajo el supuesto de que el evento B ha sido observado, la fórmula de Bayes permite conocer cuál de estas causas es la más probable de haber producido el evento B.

UTADEO
UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO

Regla de Bayes

Ejemplo (Blanco, 1.64):

El señor Rodríguez sabe que hay un 40% de probabilidad de que la empresa en la cual labora abra una sucursal en Montevideo. Si lo hace, la probabilidad de que él sea nombrado gerente de dicha sucursal es de un 80%. Si no lo hace, la probabilidad de que el Sr. Rodríguez sea nombrado gerente en otra sucursal es tan sólo 10%. Si se sabe que el Sr. Rodríguez fue nombrado gerente de una sucursal de su empresa ¿cuál es la probabilidad de que la empresa haya abierto una sucursal en Montevideo?.

Regla de Bayes

Ejemplo (Blanco, 1.64):

El señor Rodríguez sabe que hay un 40% de probabilidad de que la empresa en la cual labora abra una sucursal en Montevideo. Si lo hace, la probabilidad de que él sea nombrado gerente de dicha sucursal es de un 80%. Si no lo hace, la probabilidad de que el Sr. Rodríguez sea nombrado gerente en otra sucursal es tan sólo 10%. Si se sabe que el Sr. Rodríguez fue nombrado gerente de una sucursal de su empresa ¿cuál es la probabilidad de que la empresa haya abierto una sucursal en Montevideo?.

Solución:

Por la regla de Bayes:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} = \frac{0.4 \times 0.8}{0.38} = 0.84211$$

UTADEO
UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO

Actualización bayesiana

Elon Musk
@elonmusk

Replying to @cleantechnica

Technically, I tested positive, then negative twice, then positive again, so “Elon Musk Tests Negative for Covid” is an equally correct title. The “rapid antigen test” from BD seems to be about as useful as a flipping a coin.

Sections

The Washington Post
Democracy Dies in Darkness

Get one year for $29 US

Morning Mix

Elon Musk, who has cold symptoms, says his coronavirus tests are inconclusive: ‘Something extremely bogus is going on’

Actualización bayesiana

Elon Musk
@elonmusk

Replying to @cleantechnica

Technically, I tested positive, then negative twice, then positive again, so "Elon Musk Tests Negative for Covid" is an equally correct title. The "rapid antigen test" from BD seems to be about as useful as a flipping a coin.

Sections

The Washington Post
Democracy Dies in Darkness

Get one year for $29 US

Morning Mix

Elon Musk, who has cold symptoms, says his coronavirus tests are inconclusive: 'Something extremely bogus is going on'

T1: + T2: - T3: - T4: +

¿Cuál es la probabilidad de que Elon Musk esté contagiado?

(Idea tomada de: https://aprendendeconeli.com/)

Actualización bayesiana

¿Cuál es la probabilidad de que Elon Musk esté infectado dado que dio positivo en un test?

$$P(I|+)$$

Actualización bayesiana

¿Cuál es la probabilidad de que Elon Musk esté infectado dado que dio positivo en un test?

$$P(I|+)$$

Regla de Bayes

$$P(I|+) = \frac{P(+|I) \cdot P(I)}{P(+|I) \cdot P(I) \ + \ P(+|S) \cdot (1 - P(I))}$$

Sensibilidad: Probabilidad de que para un sujeto infectado se obtenga en la prueba un resultado positivo:

P(+ | I) ~ 84% = **0.84**

Especificidad: Probabilidad de que para un sujeto sano se obtenga en la prueba un resultado negativo: **P(- | S) ~ 99.5%** = 0.995

P(+ | S) = 1 - P(- | S) = **0.005**

Actualización bayesiana

$$P(I|+) = \frac{P(+|I)\cdot P(I)}{P(+|I)\cdot P(I) \; + \; P(+|S)\cdot(1-P(I))}$$

Probabilidad a priori $P(I)$: Nivel de “creencia” inicial de la probabilidad del evento de interés. En este caso es dependiente del sujeto, exposición, país. Como aproximación se toma igual a la ***prevalencia*** (proporción de la población general que está infectada):

Active Cases in the United States

US

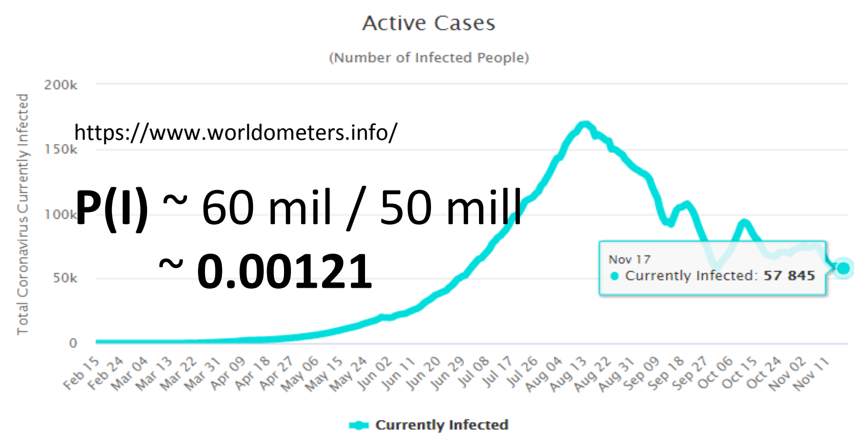
P(I) ~ 3,872 mill / 331 mill

~ **0.0117**

x2: **P(I) ~ 0.02**

Active Cases in Colombia

Colombia



P(I) ~ 60 mil / 50 mill

~ **0.00121**

x2: **P(I) ~ 0.002**

UTADEO
UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO

Actualización bayesiana

Elon Musk en USA:
P(I) = 0.02

$$\boldsymbol{P(I|+)} = \frac{P(+|I) \cdot P(I)}{P(+|I) \cdot P(I) \; + \; P(+|S) \cdot (1 - P(I))} = 0.774$$

$$\boldsymbol{P(I|+-)} = \frac{P(+|I) \cdot 0.774}{P(+|I) \cdot 0.774 \; + \; P(+|S) \cdot (1 - 0.774)} = 0.355$$

$$\boldsymbol{P(I|+--)} = 0.081$$

$$\boldsymbol{P(I|+--+) = 0.937}$$

Elon Musk en Colombia:
P(I) = 0.002

$$\boldsymbol{P(I|+)} = \frac{P(+|I) \cdot P(I)}{P(+|I) \cdot P(I) \; + \; P(+|S) \cdot (1 - P(I))} = 0.2519$$

$$\boldsymbol{P(I|+-)} = \frac{P(+|I) \cdot 0.2519}{P(+|I) \cdot 0.2519 \; + \; P(+|S) \cdot (1 - 0.2519)} = 0.0514$$

$$\boldsymbol{P(I|+--)} = 0.0086$$

$$\boldsymbol{P(I|+--+) = 0.594}$$

Regla de Bayes

Ejercicio 3: Regla de Bayes y pruebas

1. Escriba un código en R que permita calcular la probabilidad de que la persona **esté infectada** de COVID-19 dado que el **test da positivo**, P(I|+). Tome los valores de sensibilidad y especificidad del test dados en esta presentación. El código debe proporcionar una gráfica de esta probabilidad en función de la probabilidad a priori, P(I). Compare los valores de P(I|+) para diferentes países: Colombia, USA y España, tomando como P(I) la prevalencia (proporción de casos activos) en esos países.

2. Escriba un código en R para calcular la probabilidad de que la persona **no esté infectada** de COVID-19 dado que el **test da negativo** P(S|-). Proporcione la gráfica de P(S|-). Compare los resultados para los tres países.

3. Escriba un código en R que permita obtener la probabilidad de infección dados los resultados de dos tests: ++, +-, -+, --. Compare para los tres países.

Ayuda: Ver código *ReglaBayes.R* para el caso de un test de VIH

UTADEO
UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO