Introducción a Métodos Bayesianos

Víctor Morales Oñate
Sitio personal
ResearchGate
GitHub
LinkedIn

08 de mayo de 2019

Contents

Paquetes de esta sección

```
if(!require(ISLR)){install.packages("earth")}
if(!require(ISLR)){install.packages("caret")}
if(!require(ISLR)){install.packages("AmesHousing")}
```

Introducción

La estadística es el estudio de la incertidumbre. Pero, ¿Cómo le medimos? ¿Cómo tomamos decisiones en su presencia?

Una de las formas de lidiar con la incertidumbre de una forma más cuantificable es pensar en **probabilidades**

Pensemos en algunos ejmplos de probabilidad:

- Supongamos que estamos jugando con un dado justo, y preguntamos $P(X = 4)$
- También podríamos preguntar, ¿es un dado justo? Tiene sentido preguntar ¿Cuál es la probabilidad de que el dado sea justo? $P(justo)$
- Otro ejemplo es, ¿Cuál es la probabilidad de que llueva mañana? $P(llueva)$
- ¿Cuál es la probabilidad de que el universo de expadanda eternamente?

Paradigmas

Hay tres formas en las que podemos definir la probabilidad:

- Clásico: los resultados son igualmente probables, tienen igual probabilidad. Ejemplo: el lanzamiento de un dado justo. También podemos preguntar cosas relacionadas como ¿cuál es la probabilidad de que la suma sea 4 en el lanzamiento de dos dados?

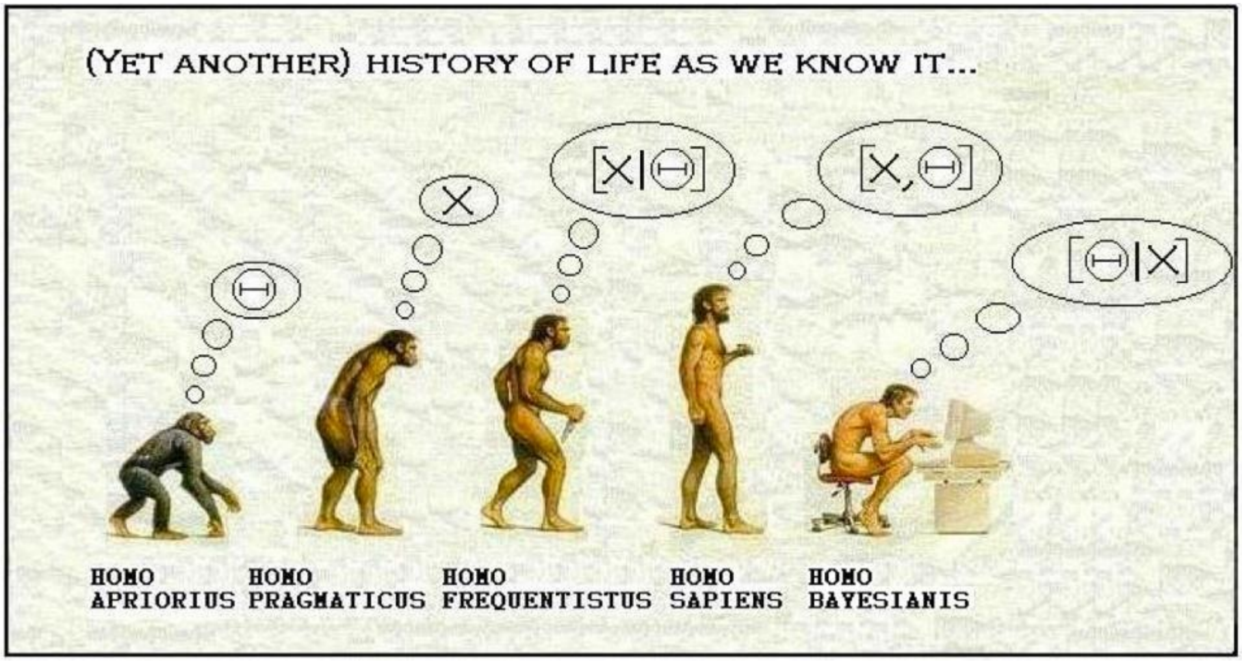

Figure 1:

- Frecuentista: Se asume una secuencia hipotética infinita de eventos. Luego vemos la frecuencia relativa de esa secuencia hipotética. En el caso del dado, podemos pensar en lanzar el dado infinitas veces. Entonces 1/6 de las veces aparecerá el 4. Coincidimos con el caso clásico.

Pero también podemos preguntar, $P(lluvia)$ mañana. En este caso debemos pensar en una secuencia infinita de *mañana* y ver qué fracción de esta secuencia infinita de *mañanas* ha llovido, lo cual es extraño de pensar.

O, en el caso de $P(justo)$, podemos lanzar el dado infinitas veces, pero eso no cambiará la probabilidad de que sea justo. Entonces, bajo el paradigma frecuentista, $P(justo) = \{0, 1\}$. No es una respuesta muy intuitiva.

El enfoque frecuentista trata de ser *objetivo* en cómo define probabilidades. Pero, como pueden ver, eso puede caer en cuestiones filosóficas profundas. Algunas veces, la *objetividad* es simplemente un lujo. Agunas veces tenemos interpretaciones que no son particularmente intuitivas.

- Bayesiano: Toma en cuenta la perspectiva personal. Tu probabilidad representa tu propia medida de incertidumbre, toma en cuenta lo que tu sabes de un problema en particular. Por lo tanto, lo que tu ves puede ser diferente de lo que alguien más piense.

Por ejemplo, $P(justo)$ puede ser diferente para cada persona.

El teorema de Bayes

La probabilidad condicional se da cuando consideramos dos eventos que se relacionan entre sí dado que sabemos que otro evento ha ocurrido:

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

Por ejemplo, consideremos una clase de 30 estudiantes. Supongamos que dentro de esta clase hay 9 mujeres. Tambien que tenemos 12 estudiantes de computación, de los cuales 4 son mujeres.

$$P(F) = \frac{9}{30} = \frac{3}{10}$$

$$P(C) = \frac{12}{30} = \frac{2}{5}$$

$$P(F \cap C) = \frac{4}{30} = \frac{2}{15}$$

Ahora podemos preguntar probabilidades marginales

$$P(F|C) = \frac{P(F \cap C)}{P(C)} = \frac{2/15}{2/5} = \frac{1}{3}$$

También podemos preguntar en la otra dirección

$$P(F|C^c) = \frac{P(F \cap C^c)}{P(C^c)} = \frac{5/30}{18/30} = \frac{5}{18}$$

Cuando asumimos **independencia**:

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

entonces

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

En nuestro ejemplo, vemos que

$$P(F|C) \neq P(F)$$

Por lo tanto, no son independientes.

Teorema de Bayes

Se usa este teorema para *revertir* la dirección de *condicionar*:

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

Con el ejemplo anterior, esto sería

$$P(C|F) = \frac{P(F|C)P(C)}{P(F|C)P(C) + P(F|C^c)P(C^c)}$$

$$P(C|F) = \frac{(1/3)(2/5)}{(1/3)(2/5) + (5/18)(3/5)} = 4/9$$

Ejemplo: VIH

Se sabe que una prueba de VIH tiene la siguiente precisión $P(+|VIH) = 0.977$, $P(+|noVIH) = 0.926$. Además un estudio muestra que $P(VIH) = 0.0026$.

Se desea saber $P(VIH|+)$:

$$P(VIH|+) = \frac{P(+|VIH)P(VIH)}{P(+|VIH)P(VIH) + P(+|noVIH)P(noVIH)}$$

Notemos que en este caso no tenemos los cálculos explícitamente. Debemos calcular algunos elementos.

$$= \frac{(0.977)(0.0026)}{(0.977)(0.0026) + (1 - 0.926)(1 - 0.0026)} = 0.033$$

Este resultado es sorprendete para la mayoría de personas. Este resultado es común para cuando se tiene enfermedades raras. El número de falsos positivos ($(1 - 0.926)(1 - 0.0026)$) supera grandemente a los verdaderos positivos.

Pese a que el test es muy preciso, obtenemos más falsos positivos que verdaderos positivos. Tendría más sentido hacer el test donde hay más prevalencia en la población en lugar de la población total donde es una enfermedad rara.

Taller

1. Si adivinas esta pregunta al azar, tienes una probabilidad de 0.25 de estar en lo correcto. ¿Qué paradigma probabilístico demuestra mejor este argumento?

- Clásico
- Frecuentista
- Bayesiano
- Ninguno

2. En una prueba de opción múltiple, no conoces la respuesta a una pregunta con tres alternativas. Sin embargo, una de las opciones contiene una palabra clave que el profesor usó frecuentemente durante la clase. En lugar de adivinar al azar, selecciona la opción que contiene la palabra clave, suponiendo que tiene una probabilidad mayor de 1/3 de ser correcta.

¿Qué paradigma probabilístico demuestra mejor este argumento?

- Clásico
- Frecuentista
- Bayesiano

3. En promedio, uno de cada tres estudiantes en su escuela participa en actividades extracurriculares. Tu concluyes que la probabilidad de que participe un estudiante seleccionado al azar de tu escuela es 1/3.

¿Qué paradigma probabilístico demuestra mejor este argumento?

- Clásico
- Frecuentista
- Bayesiano

Para las preguntas 4-6, considere el siguiente escenario:

> Tu amiga te ofrece una apuesta de que puede vencerte en un juego de ajedrez. Si tu ganas, ella te debe \$5, pero si ella gana, tu le debes \$3.

4. Supongamos que ella está 100% segura de que te vencerá. ¿Cuál es su retorno esperado para este juego?

5. Supongamos que solo tienes un 50% de confianza de que te vencerá (su probabilidad personal de ganar es $p = 0.5$). ¿Cuál es su (de ella) retorno esperado ahora?
6. Ahora, asumiendo que tu amiga solo acepte apuestas justas (retorno esperado de \$0), encuentre la probabilidad de que ella gane.
7. Digamos que te ofrecen este par de apuesta.

(i) si llueve o está nublado mañana, le pagas \$4, de lo contrario te paga \$6;

(ii) si hace sol, le pagas \$5, de lo contrario te paga \$5.

Supongamos que *lluvia*, *nublado* y *soleado* son los únicos eventos en consideración. Si realiza ambas apuestas simultáneamente, esto se denomina *libro holandés*, ya que tiene la garantía de ganar dinero. ¿Cuánto ganas sin importar el resultado?

8. Aparentemente tu amigo no entiende las leyes de la probabilidad. Vamos a examinar las apuestas que ofreció.

- Para que la apuesta (i) sea justa, su probabilidad de que llueva o esté cubierta debe ser de .6 (puede verificarlo calculando su rendimiento esperado y estableciéndolo en \$ 0).
- Para que la apuesta (ii) sea justa, su probabilidad de que sea soleada debe ser .5. Esto da como resultado un *libro holandés* porque las probabilidades de su amigo no son coherentes. No suman hasta 1. ¿Qué suman?