Estadística Bayesiana

La idea fundamental detrás de la Estadística Bayesiana es el Teorema de Bayes.

¿Cómo comenzar?

1. Probabilidades
2. Probabilidad Condicional
3. Teorema de Bayes
4. Estadística Bayesiana

1. Probabilidad

Es un número entre 0 y 1 (incluidos ambos) que representan un nivel de confianza frente a un determinado hecho o una predicción.

Ej: lanzar una moneda.

2. Probabilidad Condicional

Es una probabilidad basada en alguna información previa.

Ej: Probabilidad de tener un infarto.

Población de EEUU: 311.000.000
Sufren un infarto por año: 785.000 ciudadanos.

Seleccionados al alzar se podría decir que un ciudadano tiene aproximadamente el 0.25% la probabilidad de sufrir un infarto. O una probabilidad de 0.0025

Pero en realidad hay varias más variables que influyen en la probabilidad de tener un infarto, como por ejemplo: edad, colesterol, presión sanguínea, etc…

Cargando los factores de riesgo en una calculadora:
http://cvdrisk.nhlbi.nih.gov/calculator.asp

Va a arrojar resultados mayores y menores sobre el promedio. Por ej 0.33 o 0.23

Estadística Bayesiana

Notación:

p (A | B)

Se lee: La probabilidad de A, sabiendo que B es verdadero.

Cojoint Probability (Probabilidad de A y B) (ambos verdaderos)

p(A y B) = p(A) . p(B)

Por ejemplo: Probabilidad de que salga “cara” dos veces seguidas al tirar un moneda:
p(cara).p(cara) = 0.5 . 0.5 = 0.25

U = {(cara,cara),(cara,seca),(seca,cara),(seca,seca)} = 1 / 4 = 0.25

Probabilidad de que salga 1 seca y una cara, sin importar el orden = 1 / 2 = 0.5
P(cara) + P(seca) = 0.5

Esto funciona ya que ambos sucesos son INDEPENDIENTES

Independencia

Se dice que dos sucesos aleatorios son independientes cuando la probabilidad de cada uno de ellos no está influida porque el otro suceso ocurra o no.

Ej: Jugador encesta 80% de tiros libres:

P(encestar 3 tiros seguidos) = 0.8 x 0.8 x 0.8 = 0.512 = 51,2%

O visto de otra manera: 80% de encestar el primero.
Tiene el 80% de ese 80% de encestar el segundo: (80x80)/100 = 64%
Y de esos 64 el 80% va ser el tercer tiro: (80x64)/100 = 51,2%

Independencia

Se dice que dos sucesos aleatorios son independientes cuando la probabilidad de cada uno de ellos no está influida porque el otro suceso ocurra o no.

Ej: Jugador encesta 80% de tiros libres:

Pensemos ahora que nuestro equipo en un partido de la final de la NBA ese encuentra 1 punto abajo faltando 1 segundo para que termine el partido, y teniendo 3 tiros libres a favor.

-¿Cuál es la probabilidad de que al menos empate el juego?

P(al menos empatar) = P(al menos acertar un tiro libre) = [1 – P(de no acertar NINGUNO)]

P(de no meter NINGUN tiro libre) = 0.2^3 = 0,008 = 0.8%

Por lo tanto hay una probabilidad del 99.2% que al menos se empate el partido. Y 89,6% de ganarlo

Debatir sobre si estos resultados pertenecen realmente a un problema de sucesos independientes. Y relacionar esto con la importancia del análisis de datos más allá de los resultados estadísticos.

Dependencia

Ejemplo: De un mazo de cartas de poker, obtener dos reyes seguidos:

Total de cartas 52:
(1-10 + J + Q + K) 13 * 4 palos diferentes = 52

A = Obtener un King en la primera carta
B = Obtener un King en la segunda carta

P(A) = 4 / 52

P(B/A) = 3 / 51

P(A and B) = P(A) x P(B/A) = 12 / 2652 = 0,0045 = 0,45% aprox.

Teorema de Bayes:

¿Cómo cambia la forma de ver el mundo que nos rodea?

- No existen los blancos o negros, todo se transforma en escala de grises
- Cambia nuestras creencias en presencia de nuevos eventos

El teorema de Bayes es de enorme relevancia puesto que vincula la probabilidad de A dado B con la probabilidad de B dado A.

Problema de los Caramelos

Imaginemos dos bowls el primero con 30 caramelos de frutilla y 10 de limón, y el segundo con 20 de cada uno.

Ahora, sin ver, elegimos al azar un caramelo que resulta ser de frutilla.
¿Cuál es la probabilidad de que sea del Bowl 1?

Aquí se enlaza la probabilidad que sea de frutilla, con la probabilidad de que pertenezca a algún bowl.

Teorema de Bayes:

Deseamos la "P(Blowl_1 / frutilla)"

Por otro lado, tenemos que: P(frutilla / Blowl_1) = 3 / 4

Pero, tener en cuenta que P(A/B) es diferente a P(B/A)

Podemos decir que p(A and B) = p(B and A)

P(A and B) = p(A).p(B/A)
P(B and A) = p(B).p(A/B)

p(B).p(A/B) = p(A).P(B/A)

P(A/B) = [p(A).P(B/A)] / p(B)

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

Estadística Bayesiana

Teorema de Bayes:

Solución al Problema de los Caramelos:

p(B_1): Es la probabilidad de haber elegido el Bowl 1. Ya que lo elegimos al azar , y eran dos bowls podemos decir que la P(B_1) = 1 / 2

p(Frutilla si B_1) = 30/40 = 3 / 4

p(Frutilla) = 50 / 80 = 5 / 8

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

p(B_1 / Frutilla) = [1 / 2 * 3 / 4] / [5 / 8] = 0,6

¿Pueden pensar otra manera mas directa de arribar al mismo resultado?
Si es así, como lo enlazan con el Teorema de Bayes?