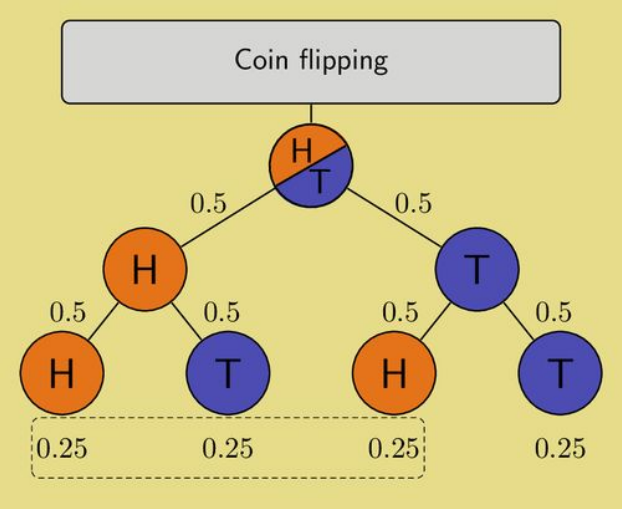
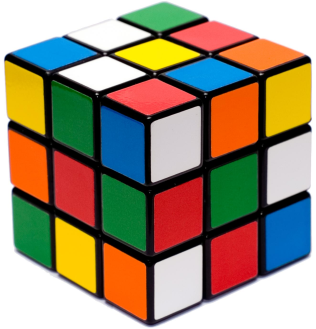
Teoria de Probabilidades

Introductorio a ML



¿cuàl es la probabilidad de ?

Se usan probabilidades de modo informal para expresar la información o la incertidumbre que se tiene acerca de observaciones de cantidades desconocidas



Términos

Experimento Aleatorio ($\mathcal{E}$) ↠ Resultados no predecibles sin una prueba

Espacio Muestral (**Ω**) : Conjunto de resultados posibles de un $\mathcal{E}$

Si lanzo dos veces una moneda □ 1 = {(C,C). (C,S), (S,C), (S,S)}

Eventos o sucesos: Subconjunto de un espacio Muestral

A: Si lanzo la moneda salga una cara *= {(C,C). (C,S), (S,C)} ... Podría haber otros conjuntos...¡SI!*

=> puedo aplicar operaciones entre conjuntos....Unión ∪ , Intersección ∩, Diferencia (-), Complemento A'

el uso de probabilidades para expresar la información se puede hacer de modo formal.

Estadística/Probabilidad Bayesiana

La información previa o externa que se posee en relación a los fenómenos que se tratan de modelizar juega un papel fundamental

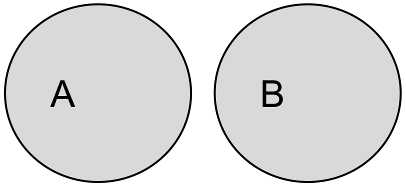
– Estimadores de los parámetros que tienen buenas propiedades estadísticas;

– Una descripción parsimoniosa (simple) de los datos observados;

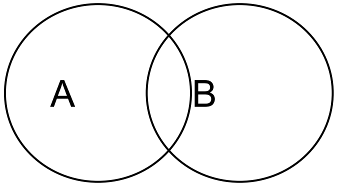
– Estimación de los datos missing y predicciones de futuras observaciones;

– Una metodología computacional potente para la estimación, selección y validación de modelo

Clases de sucesos probabilísticos

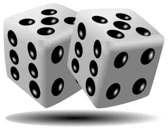


Mutuamente Excluyentes; No ocurren juntos $A \cap B = \emptyset$



Compatibles; Pueden suceder simultáneamente

- Independientes: La ocurrencia de uno no depende del otro
- Dependientes: La ocurrencia de uno depende de la la ocurrencia del otro



¡Fórmulas!

$$P(A) = \frac{\text{Número de resultados favorables al suceso A}}{\text{Número total de resultados posibles}} = \frac{n(A)}{n(\Omega)}$$

Entonces, la probabilidad de sacar Sello o Cara al lanzar una moneda es ½ o la probabilidad de sacar por ejemplo 4 al lanzar un dado es 1/6

mmm…Si tengo dos dados, ¿qué probabilidad tengo de sacar un par cuya suma sea múltiplo de 2? ..¿y si tengo 3 datos y si son muuuchos dados? ?? Los espacios muestrales se multiplican

Total de casos para el cálculo: 6 valores posibles un dado, 36 dos dados….infiriendo entonces. ***Total de casos = 6^n***

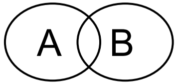
Si A es un suceso definido en el espacio muestral □, entonces:

0 ≤ P(A) ≤ 1 y P(A) + P(A') = 1 1 significa probabilidad máxima

El concepto básico en estadística bayesiana es el de probabilidad condicional: Para dos sucesos A y B: P(A|B) = P(A ∩ B) /P(B) ∝ P(A ∩ B)

Adición

Regla general de la adición de probabilidades para sucesos no excluyentes o intersecantes



Si A y B son dos eventos no mutuamente excluyentes (eventos intersecantes), de modo que ocurra A o B o ambos a la vez, entonces la probabilidad: $P(A \vee B) = P(A \cup A) = P(A) + P(B) - P(A \cap B)$

Dentro de una caja existen 10 fichas numeradas del 1 al 10. ¿Qué probabilidad existe de obtener en una sola extracción una ficha numerada con un número par o con un número primo?

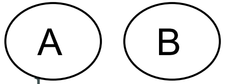
$$Espacio\ muestral = S = \{1,2,3,4,5,6,7,8,9,10\} \Longrightarrow n(S) = 10$$

A = número par = {2, 4, 6, 8, 10}
B = número primo = {2, 3, 5, 7}

$$Resultados\ favorables = A \cup B = \{2,4,6,8,10,3,5,7\} \Rightarrow n(E) = 8$$

$$P(A\ o\ B) = \frac{8}{10} = \frac{4}{5}$$

Regla general de la adición de probabilidades para sucesos excluyentes o no intersecantes



Si A y B son dos eventos no mutuamente excluyentes (eventos intersecantes), de modo que ocurra A o B o ambos a la vez, entonces la probabilidad: $P(A \vee B) = P(A \cup A) = P(A) + P(B)$ ya que la probabilidad de que ocurran los dos de forma común es cero $P(A \cap B) = \varnothing$

Multiplicación

Si A y B son dependientes   entonces: $P(A \cap B) = P(A) * P(B/A)$

La ocurrencia de un suceso marca la probabilidad de ocurrencia del segundo.

Si dentro de una caja tengo 4 caramelos de menta y 6 caramelos de miel, y al azahar saco uno tras otro dos caramelos. ¿qué probabilidad tengo de sacar un caramelo de miel en la segunda oportunidad? .....la probabilidad de extraer un caramelo de miel depende de qué haya sacado la primera vez....oo cuàl es probalidad de extraer primero un caramelo de menta y luego uno de miel......El espacio muestral se reduce en uno la segunda vez

$P(A \cap B) = 4/10 * 6/9 = 4/15$

Si A y B son independientes   entonces: $P(A \cap B) = P(A) * P(B)$

La ocurrencia de un suceso marca la probabilidad de ocurrencia del segundo.

Desde el punto de vista bayesiano, en la práctica, todas las probabilidades son condicionales porque casi siempre existe algún conocimiento previo o cierta experiencia acerca de los sucesos

Ejemplo

En el caso judicial en USA de Connecticut vs. Teal se alegó un caso de posible discriminación racial en base a un test para una promoción de estudiantes dada. Se observço que entre el grupo de personas que hacían un test, 48 eran negras (B) y 259 eran blancas (W). De este modo si se toma una persona al azar

$P(B) = 48/307 = 0{,}16$

$P(W) = 259/307 = 0{,}84$

De las personas negras 26 aprobaron el test (P) y el resto falló (F), mientras que en las personas blancas 206 aprobaron y el resto, no.

De este modo un total de 232 aprobaron. De modo que

$P(B|P) = 26/232 = 0{,}11$

$P(W|P) = 206/232 = 0{,}89$

Existe, por tanto, la tentación de asumir que existe una muy elevada discriminación racial... Sin embargo lo que se debería calcular es la distribución condicionada contraria: asumiendo que si se es negro (o blanco) cuál es la probabilidad de aprobar:

$P(P|B) = 26/48 = 0{,}54$

$P(P|W) = 206/259 = 0{,}80$

Ley de la probabilidad total

Para un suceso A y una partición B1, . . . , Bk,

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

Se puede aplicar el teorema a variables discretas:

$$f(x) = \sum_{y} f(x|Y = y)P(Y = y)$$

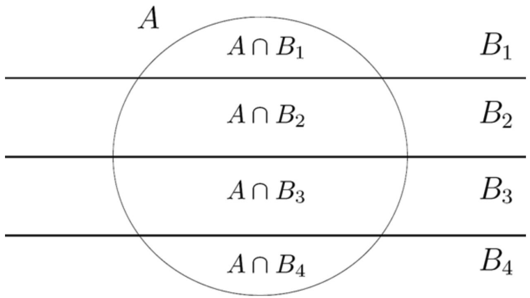
o a variables continuas:

$$f(x) = \int f(x|y)f(y)\,dy.$$

Teorema de Bayes

Se tiene que, para los sucesos A1, . . . , An y B,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)} \propto P(B|A_i)P(A_i)$$



En el caso que ***x*** sea contínua:

$$f(x|y) = \frac{f(y|x)f(x)}{f(y)} = \frac{f(y|x)f(x)}{\int_{\mathbb{R}} f(y|x)f(x)dx}$$

como el denominador f(y) es independiente de x, entonces se puede escribir el teorema en la forma de proporcionalidad ($\propto$):

$$f(x|y) \propto f(y|x)f(x)$$

Este resultado es útil para los cálculos porque implica que se pueden olvidar las constantes multiplicativas hasta el final de los cálculos en modelos complicados

Ejemplo

Los mellizos pueden ser de dos tipos: (M) monocigóticos (que proceden del mismo óvulo) o (D) dicigóticos (proceden de óvulos diferentes). Generalmente los monocigóticos se parecen mucho y son del mismo género, en tanto que los dicigóticos podrían ser bastante diferentes, incluso de géneros diferentes

Asumiendo que los dos géneros son igualmente probables e identificando para cada par de mellizos (mujer u hombre) como mm, hh, mh entonces:

$$P(mm|M) = P(hh|M) = \tfrac{1}{2} \qquad P(mh|M) = 0$$
$$P(mm|D) = P(hh|D) = \tfrac{1}{4} \qquad P(mh|D) = \tfrac{1}{2}$$

De lo que se deduce que

$$P(mm) = P(mm|M)\,P(M) + P(mm|D)\,P(D) = \frac{1}{2}P(M) + \frac{1}{4}(1 - P(M))$$

Por lo cual :

P(M) = **4 · P(mm) − 1**

Y se determina el modo de estimar la proporción de mellizos monocigóticos que existen en una población al observar la distribución de géneros con respecto al total de mellizos.

Distribución de probabilidad: Histograma, media, varianza y desviación estándar

Para la distribución de la probabilidad:

· Media: $\mu = \sum[x \cdot p(x)]$

· Varianza: $\sigma^2 = \sum[(x-\mu)^2 \cdot p(x)] = \sum[x^2 \cdot p(x)] - \mu^2$

· Desviación estándar: $\sigma = \sqrt{\sigma^2} = \sqrt{\sum[(x-\mu)^2 \cdot p(x)]} = \sqrt{\sum[(x^2 \cdot p(x)] - \mu^2}$

Regla de rango de la desviación estándar:

· Valor máximo usual: $= \mu + 2\sigma$

· Valor mínimo usual: $= \mu - 2\sigma$

Resultado	Probabilidad
1	0.05
2	0.10
3	0.30
4	0.33
5	0.15
6	0.07

$\Leftarrow$ Dado Trucado

Dado Confiable $\Rightarrow$

Resultado	Probabilidad
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Distribución Binomial

$P(x) = {}_{n}C_{x}P_{x}(1-p)^{n-x}$

P(x): Probabilidad de obtener *x* en cada *n* intentos

n:: Número de ocurrencias totales

x: Ocurrencias exitosas

p: probabilidad de éxito en cada prueba

Ej: Si lanzo una moneda 20 veces ¿cuál es la probabilidad de obtener cara 15 veces?

p= ½ (cada lanzamiento tiene una probabilidad de ½)

n=20

x=15

$\Rightarrow P(x) = {}_{20}C_{15} * P^{15} * (1-½)^{20-15}$

$$P(x) = \frac{20!}{(20-15)!} * \frac{1}{2}^{15} * \frac{1}{2}^{5}$$

P(x) = 0.015