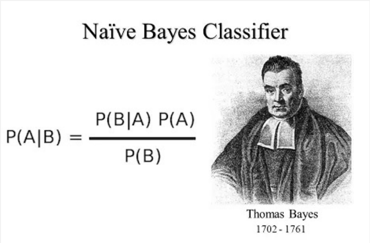


oldemar rodríguez

CONSULTOR en M1N3R14 D3 D4T0S

Aprendizaje Supervisado

Clasificación Bayesiana (Método Naïve (ingenuo) de Bayes)

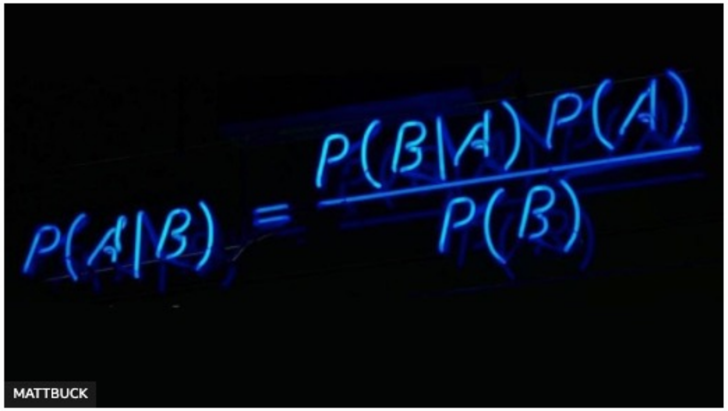
BBC NEWS | MUNDO

Noticias América Latina Internacional Medio ambiente Coronavirus Hay Festival Economía Centroamérica Cuenta BBC Extra

Qué es el teorema de Bayes, el potente método para generar conocimiento que nació cuando trataban de demostrar un milagro

Redacción
BBC News Mundo

11 diciembre 2021

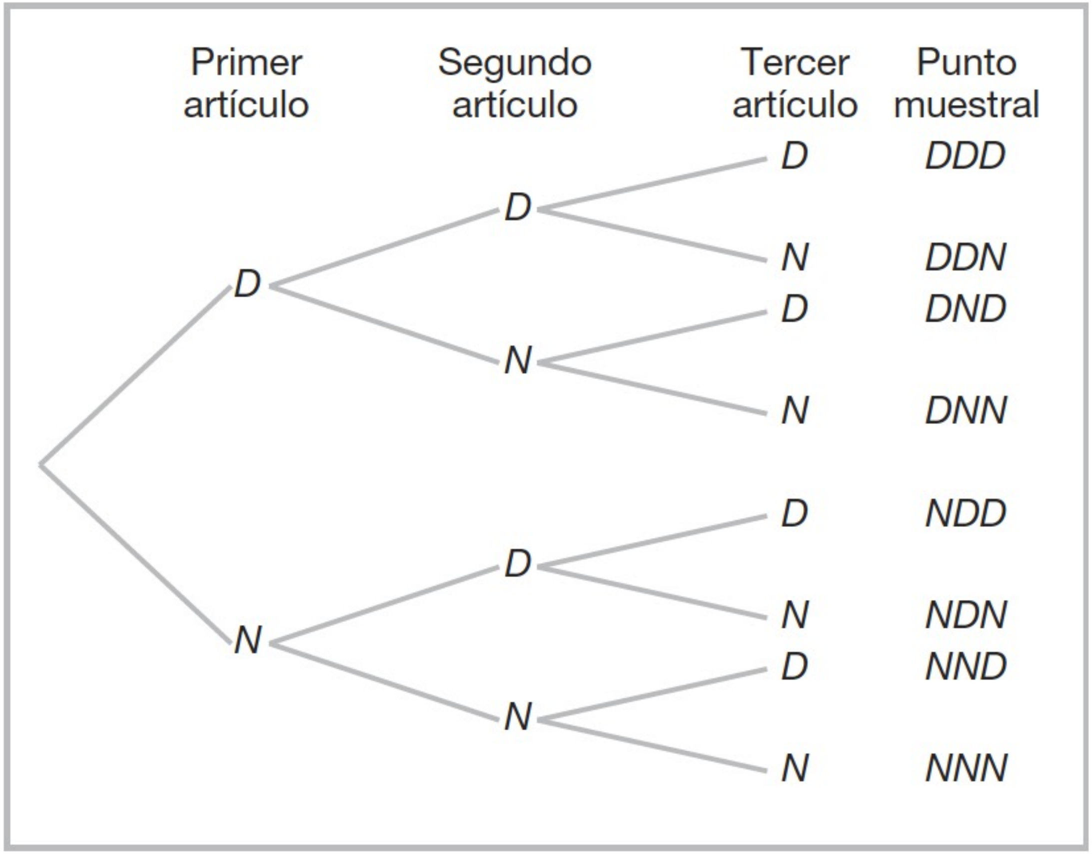


P(A|B) = (P(B|A)P(A))/P(B)

Conjunto Fundamental
Espacio Muestral

- ***Definición***: El resultado de un experimento aleatorio se llama ***evento***.
- La cuantificación de las posibilidades de tal evento se dé o suceda corresponde a la noción intuitiva de probabilidad.
- Para lograr esta cuantificación, es necesario describir de antemano, de manera muy precisa, todos los resultados posibles para los eventos. Este conjunto experimental se llama el *Conjunto Fundamental, Espacio Muestral* o *Universo* y se denota por **S**, muchos libros usan Ω para denotar el espacio muestral.
- Es decir, al conjunto de todos los resultados posibles de un experimento estadístico se le llama espacio muestral y se representa con el símbolo **S**.

- **Ejemplo**: Las seis caras numeradas de dado: S = {1, 2, 3, 4, 5, 6}.
- **Ejemplo**: Una bola se extrae de una urna que contiene una bola negra, dos blancas y cinco rojas entonces S = {negro, blanco, rojo}.
- **Ejemplo**: El espacio muestral S, de los resultados posibles cuando se lanza una moneda al aire, se puede escribir como S = {E, C}, en donde E y C corresponden a "escudo" y "corona", respectivamente.
- **Ejemplo**: Suponga que se seleccionan, de forma aleatoria, tres artículos de un proceso de fabricación. Cada artículo se inspecciona y se clasifica como defectuoso, D, o no defectuoso, N. Para listar los elementos del espacio muestral que brinde la mayor información, construimos el diagrama de árbol de la Figura de la siguiente filmina, de manera que las diversas trayectorias a lo largo de las ramas del árbol dan los distintos puntos muestrales. Al comenzar con la primera trayectoria, obtenemos el punto muestral DDD, que indica la posibilidad de que los tres artículos inspeccionados estén defectuosos. Conforme continuamos a lo largo de las demás trayectorias, vemos que el espacio muestral es S ={DDD, DDN, DND, DNN, NDD, NDN, NND, NNN}.


Primer artículo
Segundo artículo
Tercer artículo
Punto muestral
D
N
D
N
D
N
D
N
D
N
D
N
D
N
DDD
DDN
DND
DNN
NDD
NDN
NND
NNN

- **Definición**: Cada elemento $\omega \in S$ representa un **evento elemental**, y cualquier parte $A \subseteq S$, (o $A \in P(S)$) será un **evento**. A veces decimos que es el conjunto de posibles eventos y los eventos elementales son entonces los "singletons", es decir, los conjuntos reducidos a un solo elemento, de hecho, son estrictamente eventos, ya que pertenecen a $P(S)$, que no es el caso del punto ω.
- Es decir, un evento es un subconjunto de un espacio muestral.
- **Ejemplo**: Para un juego de datos se tiene que S = {1, 2, 3, 4, 5, 6}; el evento A = {1, 2} se traduce, representa simbólicamente el resultado "obtener un resultado menor o igual a 2".

- **Definición**: La intersección de dos eventos A y B es un evento, que se denota con el símbolo $A \cap B$, es el evento que contiene todos los elementos que son comunes a A y a B.
- **Definición**: Dos eventos A y B son mutuamente excluyentes o disjuntos si $A \cap B = \Phi$; es decir, si A y B no tienen elementos en común.
- **Definición**: La unión de dos eventos A y B es un evento, que se denota con el símbolo $A \cup B$, es el evento que contiene todos los elementos que pertenecen a A o a B, o a ambos.
- **Ejemplo**: Sea P el evento de que un empleado de una empresa petrolera seleccionado al azar fume cigarrillos. Sea Q el evento de que el empleado seleccionado ingiera bebidas alcohólicas. Entonces, el evento $P \cup Q$ es el conjunto de todos los empleados que beben o fuman, o que hacen ambas cosas.

Probabilidad de un evento

- La probabilidad de la ocurrencia de un evento que resulta de un experimento estadístico se evalúa utilizando un conjunto de números reales denominados pesos o probabilidades, que van de 0 a 1. Para todo punto en el espacio muestral asignamos una probabilidad tal que la suma de todas las probabilidades es 1.

- **Definición**:

La **probabilidad** de un evento A es la suma de los pesos de todos los puntos muestrales en A. Por lo tanto,

$$0 \leq P(A) \leq 1, \qquad P(\phi) = 0 \quad \text{y} \quad P(S) = 1.$$

Además, si $A_1, A_2, A_3, \ldots$ es una serie de eventos mutuamente excluyentes, entonces

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots.$$

- **Ejemplo**: Una moneda se lanza dos veces. ¿Cuál es la probabilidad de que ocurra al menos un Escudo (E)?
- **Solución**: El espacio muestral para este experimento es:

$$S = \{EE, EC, CE, CC\}$$

- Si la moneda está balanceada, cada uno de estos resultados tendrá las misma probabilidades de ocurrir. Por lo tanto, asignamos una probabilidad de ω a cada uno de los puntos muestrales. Entonces, $4\omega = 1$, es decir, $\omega = 1/4$. Si A representa el evento de que ocurra al menos un Escudo (E), entonces $A = \{EE, EC, CE\}$ y $P(A) = ¼ + ¼ + ¼ = ¾$.
- También se puede razonar como sigue: Del total de elementos en el espacio muestral S (que son 4) hay 3 elementos que son favorables, es decir, 3 en los que aparece (E), por lo tanto hay 3 favorables de 4 totales por los que la probailidad $P(A) = ¾$.

- **Ejemplo**: Se lanza un dado. Si E es el evento de que ocurra un número menor estricto que 4 en un solo lanzamiento del dado, calcule P(E).
- **Solución**: El espacio muestral es S = {1, 2, 3, 4, 5, 6}. Asignamos una probabilidad de w a cada número. Como la suma de las probabilidades debe ser 1, tenemos 6w = 1 o w = 1/6. Por consiguiente: como E = {1, 2, 3} se tiene que P(E) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2.
- Otra forma de verlo es que hay 3 casos favorables de 6 casos posibles, entonces: P(E) = 3/6 = 1/2.
- **Ejemplo**: En el mismo ejercicio anterior ¿Cuál es la probabilidad de que salga un número impar?
- **Ejemplo**: En el mismo ejercicio anterior ¿Cuál es la probabilidad de que salga un número par?

- **Regla General**: Si un experimento puede dar como resultado cualquiera de N diferentes resultados que tienen las mismas probabilidades de ocurrir, y si exactamente n de estos resultados corresponden al evento A, entonces la probabilidad del evento A es

  P (A) = n/N = casos favorables / casos totales.

- **Ejemplo**: A una clase de estadística para ingenieros asisten 25 estudiantes de ingeniería industrial, 10 de ingeniería mecánica, 10 de ingeniería eléctrica y 8 de ingeniería civil. Si el profesor elige al azar a un estudiante para que conteste una pregunta, ¿qué probabilidades hay de que el elegido sea a) estudiante de ingeniería industrial, b) estudiante de ingeniería civil o estudiante de ingeniería eléctrica?

- **Solución**: Las especialidades de los estudiantes de ingeniería industrial, mecánica, eléctrica y civil se denotan con I, M, E y C, respectivamente. El grupo está integrado por 53 estudiantes y todos tienen las mismas probabilidades de ser seleccionados.

a) Como 25 de los 53 individuos estudian ingeniería industrial, la probabilidad del evento I, es decir, la de elegir al azar a alguien que estudia ingeniería industrial, es $P(I) = 25/53$.

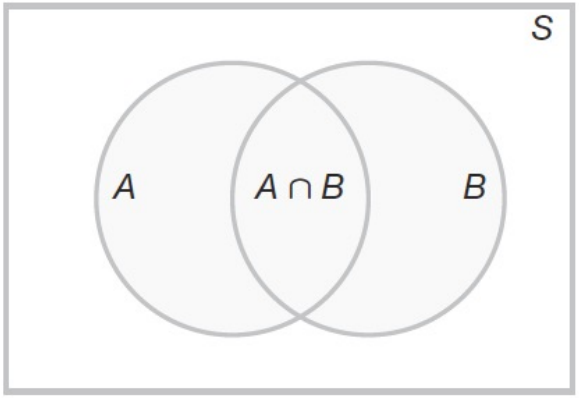
b) Como 18 de los 53 estudiantes son de las especialidades de ingeniería civil o eléctrica, se deduce que $P(C \cup E) = 18/53$.

Reglas Aditivas

- **Teorema**: Si A y B son dos eventos, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

- **Prueba**: Considere el diagrama de Venn de la figura al lado. $P(A \cup B)$ es la suma de las probabilidades de los puntos muestrales en $(A \cup B)$. Así, $P(A) + P(B)$ es la suma de todas las probabilidades en A más la suma de todas las probabilidades en B. Por lo tanto, sumamos dos veces las probabilidades en $A \cap B$. Como estas probabilidades se suman a $P(A \cap B)$, debemos restar esta probabilidad una vez para obtener la suma de las probabilidades en $A \cup B$.



Probabilidad condicional, independencia y regla del producto

- **Definición**: La probabilidad de que ocurra un evento B cuando se sabe que ya ocurrió algún evento A se llama probabilidad condicional y se denota con P(B|A). El símbolo P(B|A) por lo general se lee como "la probabilidad de que ocurra B, dado que ocurrió A", o simplemente, "la probabilidad de B, dado A".
- **Ejemplo**: Se lanza un dado ¿Cuál es la probabilidad de obtener un 3 dado que se sabe que el dado que salió es impar?
- **Solución**: B={3} y A={1,3,5} entonces P(B|A) = 1/3 (Método espacio muestral reducido).

Ejemplo (otra forma de verlo): Se lanza un dado y se tienen los siguientes eventos:

A: Se observa un número impar.

B: Se observa un 3.

P(B)=1/6 pues solo hay un caso favorable {3} y hay 6 casos posibles {1,2,3,4,5,6}

P(B|A)=1/3 pues solo hay un caso favorable {3} y hay únicamente 3 casos posibles {1,3,5}, como ya ocurrió A el dado debe ser impar por lo que el universo de posibilidades se reduce a los números impares.

Nótese que $A \cap B = \{3\}$ por lo que $P(A \cap B)=1/6$, además P(A)=3/6 pues hay 3 casos favorables {1,3,5} y 6 casos posibles {1,2,3,4,5,6}.

Entonces: $$\frac{P(A \cap B)}{P(A)} = \frac{1/6}{3/6} = \frac{1}{3} = P(B|A)$$ ESTO SIEMPRE SE CUMPLE

oldemar rodríguez
CONSULTOR en MINERIA DE DATOS

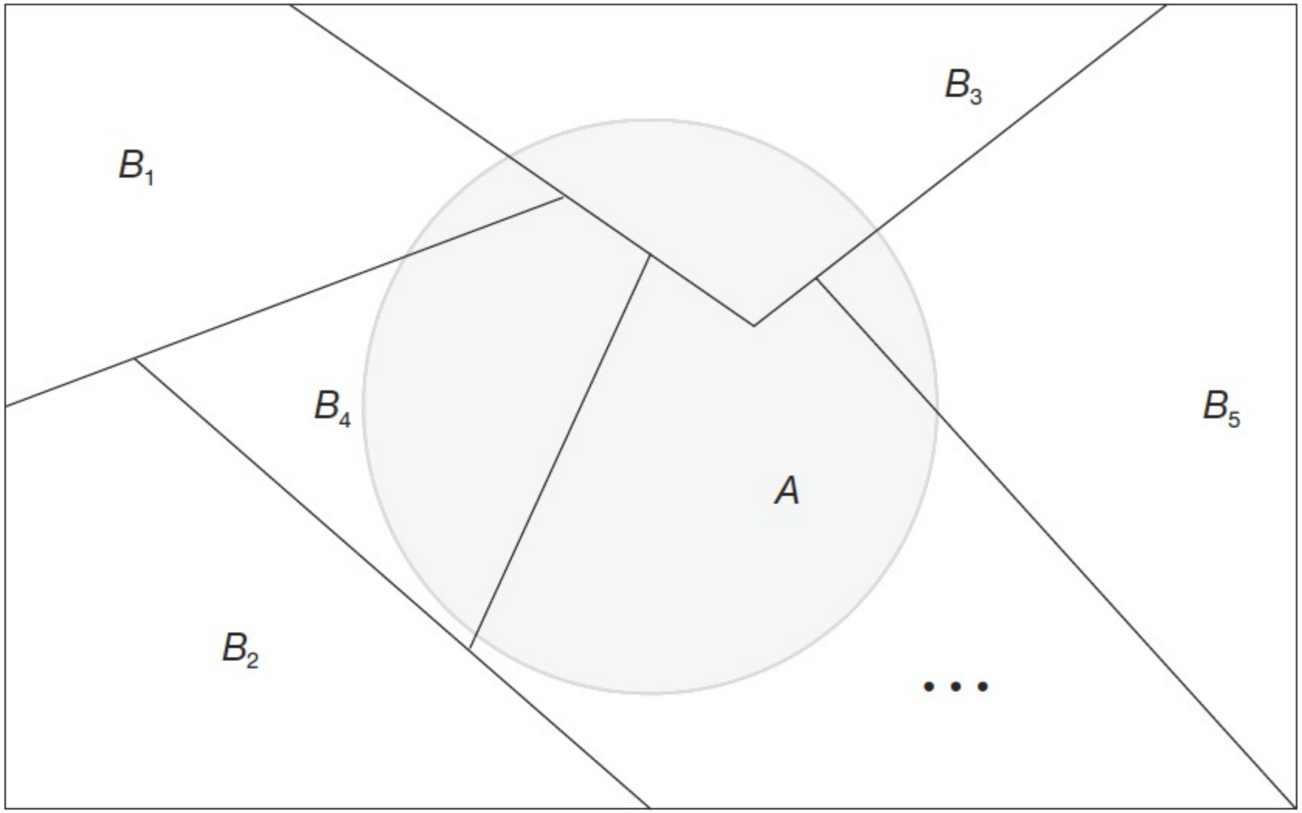
Probabilidad Total

- **Teorema de la Probabilidad Total**:

Si los eventos B_1, B_2,... B_k constituyen una partición del espacio muestral S, tal que $P(B_i) \neq 0$ para $i = 1, 2,..., k$, entonces, para cualquier evento A de S,

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

oldemar rodríguez
CONSULTOR en MINERÍA DE DATOS



Partición del espacio muestral *s*.

- **Prueba**:

Considere el diagrama de Venn . Se observa que el evento A es la unión de los eventos mutuamente excluyentes

$$B_1 \cap A, B_2 \cap A, \ldots, B_k \cap A;$$

es decir,

$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_k \cap A)$$

Por medio del corolario 2.2 del teorema 2.7 y el teorema 2.10 obtenemos

$$\begin{aligned} P(A) &= P[(B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_k \cap A)] \\ &= P(B_1 \cap A) + P(B_2 \cap A) + \cdots + P(B_k \cap A) \\ &= \sum_{i=1}^{k} P(B_i \cap A) \\ &= \sum_{i=1}^{k} P(B_i)P(A|B_i). \end{aligned}$$

$\blacksquare$

Teorema de Bayes

(**Regla de Bayes**) Si los eventos B_1, B_2,..., B_k constituyen una partición del espacio muestral S, donde $P(B_i) \neq 0$ para $i = 1, 2,...,k$, entonces, para cualquier evento A en S, tal que $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)} \quad \text{para} \quad r = 1, 2, \ldots, k.$$

- En 1763 se publicó el artículo "An essay towards solving a problem in the doctrine of chances", (Un ensayo hacia la solución de problemas en la disciplina del azar. T. del Ed.) escrito por el reverendo Thomas **Bayes** (1701-1761).

***Prueba*:** Mediante la definición de probabilidad condicional,

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

y después usando el teorema 2.13 en el denominador, tenemos

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

que completa la demostración.

Aprendizaje Supervisado
(Métodos Predictivos)
Clasificación Bayesiana
(Método Naïve Bayes)

Naïve Bayes Classifier

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

Thomas Bayes
1702 - 1761

oldemar rodríguez
CONSULTOR en MINERIA DE DATOS

Método Predictivo Naïve Bayes

- Se modela la probabilidad condicional de la variable a predecir Y, dadas las variables predictoras $X_1, \ldots, X_p$.
- Es decir, se modela $P\left(Y = k \middle| X_1 = x_1, \ldots, X_p = x_p\right)$.
- Es lo mismo escribir $P(Y = k|\boldsymbol{X} = \boldsymbol{x}) = P\left(Y = k \middle| X_1 = x_1, \ldots, X_p = x_p\right)$, donde $\boldsymbol{X} = \left(X_1, \ldots, X_p\right)$ y $\boldsymbol{x} = \left(x_1, \ldots, x_p\right)$.
- Asumimos que cada variable que define a un individuo corresponden a eventos o variables independientes.

oldemar rodríguez
CONSULTOR en MINERIA DE DATOS

Clasificación con Naïve Bayes

- Al usar el teorema de Bayes se tiene que

$$P(Y = k|\boldsymbol{X} = \boldsymbol{x}) = \frac{P(\boldsymbol{X} = \boldsymbol{x}|Y = k)P(Y = k)}{\sum_{i=1}^{K} P(\boldsymbol{X} = \boldsymbol{x}|Y = i)P(Y = i)}.$$

- Dado que asumimos independencia se tiene que

$$P(\boldsymbol{X} = \boldsymbol{x}|Y = i) = P(X_1 = x_1|Y = i) \cdots P\left(X_p = x_p\middle|Y = i\right) = \prod_{j=1}^{p} P\left(X_j = x_j\middle|Y = i\right).$$

- Para un individuo nuevo $\boldsymbol{x} = \left(x_1, \dots, x_p\right)$ se calcula $P(Y = k|\boldsymbol{X} = \boldsymbol{x})$ para cada clase k, y se predice $\boldsymbol{x}$ con la clase de mayor probabilidad.

oldemar rodríguez
CONSULTOR en MINERÍA DE DATOS

Ejemplo

- Suponga que tenemos la siguiente tabla de training con información de clientes, y queremos determinar si son buenos pagadores o no.

Id	Monto Crédito	Ingreso Neto	Monto Cuota	Grado Académico	Buen Pagador
1	2	4	1	4	Sí
2	2	3	1	4	Sí
3	4	1	4	2	No
4	1	4	1	4	Sí
5	3	3	3	2	No
6	3	4	1	4	Sí
7	4	2	3	2	No
8	4	1	3	2	No
9	3	4	1	3	Sí
10	1	3	2	4	Sí

- Llega un nuevo cliente y queremos determinar si es buen pagador o mal pagador.

Id	Monto Crédito	Ingreso Neto	Monto Cuota	Grado Académico	Buen Pagador
11	1	4	2	4	?

- Es decir, queremos obtener las siguientes probabilidades

$$P_{Sí} = P(BuenPagador = Sí|MontoCredito = 1, IngresoNeto = 4, MontoCuota = 2, GradoAcademico = 4)$$

$$P_{No} = P(BuenPagador = No|MontoCredito = 1, IngresoNeto = 4, MontoCuota = 2, GradoAcademico = 4)$$

- Si $P_{Sí}$ es mayor a P_{No} concluimos que el cliente sí es buen pagador, de lo contrario es mal pagador.

- Denote $P_{Sí} = P(BP = Sí|\boldsymbol{X} = \boldsymbol{x})$ y $P_{No} = P(BP = No|\boldsymbol{X} = \boldsymbol{x})$, donde $BP = BuenPagador$, $\boldsymbol{x} = (1, 4, 2, 4)$ y $\boldsymbol{X} = (MontoCredito, IngresoNeto, MontoCuota, GradoAcademico)$.

- Se tiene que

$$P_{Sí} = P(BP = Sí|\boldsymbol{X} = \boldsymbol{x})$$
$$= \frac{P(\boldsymbol{X} = \boldsymbol{x}|BP = Sí)P(BP = Sí)}{P(\boldsymbol{X} = \boldsymbol{x}|BP = Sí)P(BP = Sí) + P(\boldsymbol{X} = \boldsymbol{x}|BP = No)P(BP = No)}.$$

$$P_{No} = P(BP = No|\boldsymbol{X} = \boldsymbol{x})$$
$$= \frac{P(\boldsymbol{X} = \boldsymbol{x}|BP = No)P(BP = No)}{P(\boldsymbol{X} = \boldsymbol{x}|BP = Sí)P(BP = Sí) + P(\boldsymbol{X} = \boldsymbol{x}|BP = No)P(BP = No)}.$$

- Note que $P(BP = Sí) = \frac{6}{10} = 0.6$ y $P(BP = No) = \frac{4}{10} = 0.4$.

- Falta calcular $P(\boldsymbol{X} = \boldsymbol{x}|BP = Sí)$ y $P(\boldsymbol{X} = \boldsymbol{x}|BP = No)$.

- Note que el evento $\boldsymbol{X} = \boldsymbol{x}$ del nuevo individuo corresponde a 4 eventos independientes:

$$MontoCredito = 1,\ IngresoNeto = 4,$$
$$MontoCuota = 2,\ GradoAcademico = 4.$$

- Por tanto

$$P(\boldsymbol{X} = \boldsymbol{x} | BP = Sí) = P(MontoCredito = 1 | BP = Sí)\ P(IngresoNeto = 4 | BP = Sí)\ P(MontoCuota = 2 | BP = Sí)\ P(GradoAcademico = 4 | BP = Sí)$$
$$= \frac{2}{6}\frac{4}{6}\frac{1}{6}\frac{5}{6} = \frac{5}{162} \cong 0.03.$$

$$P(\boldsymbol{X} = \boldsymbol{x} | BP = No) = P(MontoCredito = 1 | BP = No)\ P(IngresoNeto = 4 | BP = No)\ P(MontoCuota = 2 | BP = No)\ P(GradoAcademico = 4 | BP = No)$$
$$= \frac{0}{4}\frac{0}{4}\frac{0}{4}\frac{0}{4} = 0.$$

Por propiedad de independiencia

- Entonces, al sustituir tenemos que

$$P_{Sí} = \frac{0.03 \times 0.6}{0.03 \times 0.6 + 0 \times 0.4} = 1 \quad \text{y}$$

$$P_{No} = \frac{0 \times 0.4}{0.03 \times 0.6 + 0 \times 0.4} = 0.$$

- Por lo tanto el individuo nuevo tiene una probabilidad máxima de ser un buen pagador, y se clasifica como $BuenPagador = Sí$.

Id	Monto Crédito	Ingreso Neto	Monto Cuota	Grado Académico	Buen Pagador
11	1	4	2	4	Sí

oldemar rodríguez
CONSULTOR en MINERÍA DE DATOS

Ejemplo de Clasificación Bayesiana

Supongamos que tenemos la siguiente tabla de datos con el ***Género*** y la ***Altura*** para 15 individuos, además se tiene una columna con la clasificación para cada individuo en P=***Pequeño***, M=***Mediano*** o A=***Alto***. La variable Género tiene dos modalidades F=Femenino y M=Masculino, mientras que la variables Altura tiene 6 modalidades 1 si Altura$\in$]0,1.6], 2 si Altura$\in$]1.6,1.7], 3 si Altura$\in$]1.7,1.8], 4 si Altura$\in$]1.8,1.9], 5 si Altura$\in$]1.9,2] y 1 si Altura$\in$]2,+∞].

Nombre	Género	Altura	Clase
Kristina	F	1	P
Jim	M	5	A
Maggi	F	4	M
Martha	F	4	M
Stephanie	F	2	P
Bob	M	4	M
Kathy	F	1	P
Dave	M	2	P
Worth	M	6	A
Steven	M	6	A
Debbie	F	3	M
Todd	M	5	M
Kim	F	4	M
Amy	F	3	M
Wynette	F	3	M

Ejemplo de Clasificación Bayesiana

- Supongamos que se tiene una nueva fila de la base de datos **t = (Adam, M, 5, ?).**
- El problema es: a partir de los datos históricos y usando Clasificación Bayesiana predecir si Adam corresponde a un individuo Pequeño, Mediano o Alto, es decir, saber si tiene mayor probabilidad de ser Pequeño, Mediano o Alto.
- Un análisis a la ligera diría que es alto pues Todd es M y 5 y fue clasificado como Alto. Sin embargo, esto no quiere decir que necesariamente sea Alto, pues por ejemplo podría ser que existan muchas personas Medianas lo cual aumentaría la probabilidad de ser Mediano.
- Lo que se hace en estos caso es calcular *P(Pequeño|t)*, *P(Mediano|t)* y *P(Alto|t)* para determinar cuál es mayor.

Ejemplo de Clasificación Bayesiana

$$P(Pequeño|t) = \frac{P(t|Peque\ ño) \cdot P(Peque\ ño)}{P(t|Peque\ ño) \cdot P(Peque\ ño) + P(t|Mediano\ ) \cdot P(Mediano\ ) + P(t|Alto\ ) \cdot P(Alto\ )}.$$

Tenemos que:

$$P(Pequeño) = \frac{4}{15}, P(Mediano) = \frac{8}{15}\ y\ P(Alto) = \frac{3}{15}$$

Como **t = (Adam, M, 5, ?)**, este es un evento que corresponde realmente a dos eventos independientes, ser M=Masculino y ser de Altura=5. Así:

$$P(t|Pequeño) = P(Masculino|Pequeño) \cdot P\big((Altura = 5)\big|Pequeño\big) = \frac{1}{4} \cdot \frac{0}{4} = 0.$$

$$P(t|Mediano) = P(Masculino|Mediano) \cdot P\big((Altura = 5)\big|Mediano\big) = \frac{2}{8} \cdot \frac{1}{8} = \frac{1}{32}.$$

$$P(t|Alto) = P(Masculino|Alto) \cdot P\big((Altura = 5)\big|Alto\big) = \frac{3}{3} \cdot \frac{1}{3} = \frac{1}{3}.$$

Ejemplo de Clasificación Bayesiana

Entonces:

$$P(Pequeño|t) = \frac{P(t|Pequeño)\cdot P(Pequeño)}{P(t|Pequeño)\cdot P(Pequeño) + P(t|Mediano)\cdot P(Mediano) + P(t|Alto)\cdot P(Alto)} = \frac{0\cdot\frac{4}{15}}{0\cdot\frac{4}{15}+\frac{1}{32}\cdot\frac{8}{15}+\frac{1}{3}\cdot\frac{3}{15}} = 0$$

$$P(Mediano|t) = \frac{P(t|Mediano)\cdot P(Mediano)}{P(t|Pequeño)\cdot P(Pequeño) + P(t|Mediano)\cdot P(Mediano) + P(t|Alto)\cdot P(Alto)} = \frac{\frac{1}{32}\cdot\frac{8}{15}}{0\cdot\frac{4}{15}+\frac{1}{32}\cdot\frac{8}{15}+\frac{1}{3}\cdot\frac{3}{15}} = 0{,}2$$

Ejemplo de Clasificación Bayesiana

$$P(Alto|t) = \frac{P(t|Alto)\cdot P(Alto)}{P(t|Pequeño)\cdot P(Pequeño) + P(t|Mediano)\cdot P(Mediano) + P(t|Alto)\cdot P(Alto)}$$

$$= \frac{\frac{1}{3}\cdot\frac{3}{15}}{0\cdot\frac{4}{15} + \frac{1}{32}\cdot\frac{8}{15} + \frac{1}{3}\cdot\frac{3}{15}} = 0{,}806$$

Por lo tanto Adam tiene mayor probabilidad de ser alto.

Gracias….