Teorema de Bayes

JFORES

PFC

Si un evento A puede ocurrir de n formas dentro de un espacio S con m elementos, la probabilidad de que ocurra A es:

- CF/TC = n/m

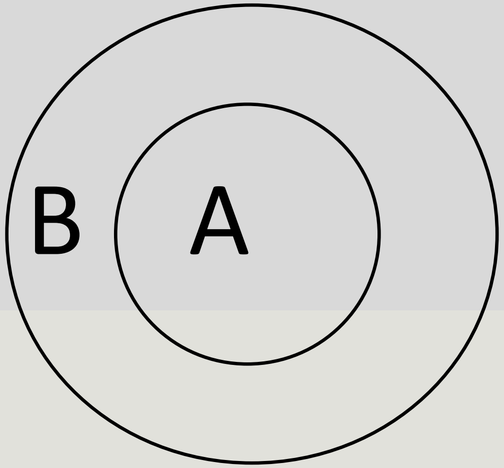
Si $A \subseteq B$, la probabilidad de seleccionar un elemento de A dado que está en B es:

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

Donde:

$|B|$ es la cardinalidad de B

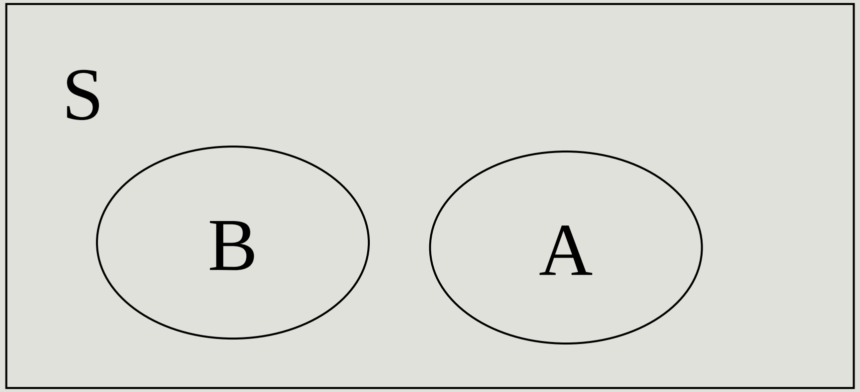
$|A|$ es la cardinalidad de A



P(A , B)

Si un evento *A* puede ocurrir de *m* formas distintas, y un evento *B* puede ocurrir de *n* formas distintas entonces: el número de formas en que *A* y *B* pueden ocurrir en secuencia es:

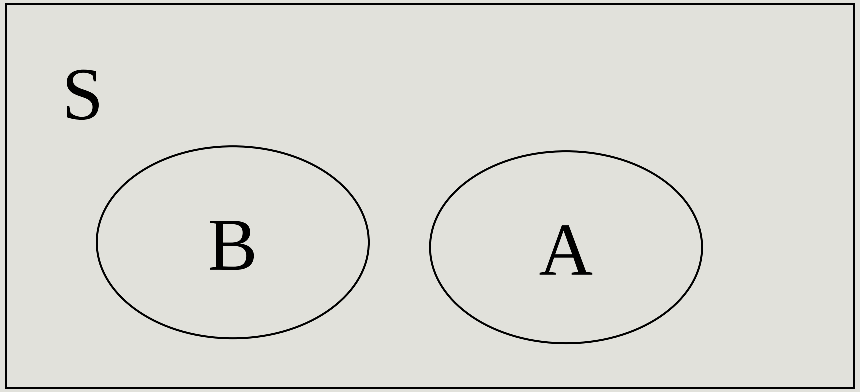
- $m \times n$.



P(A , B) = m*n

$P(A \cup B)$ si $A \cap B = \varnothing$

Dado el espacio de muestreo *S*, si P(A) es la probabilidad de que ocurra un evento de A y P(B) es la probabilidad de que ocurra un evento de B, la probabilidad de que ocurra un evento de A o de B es:

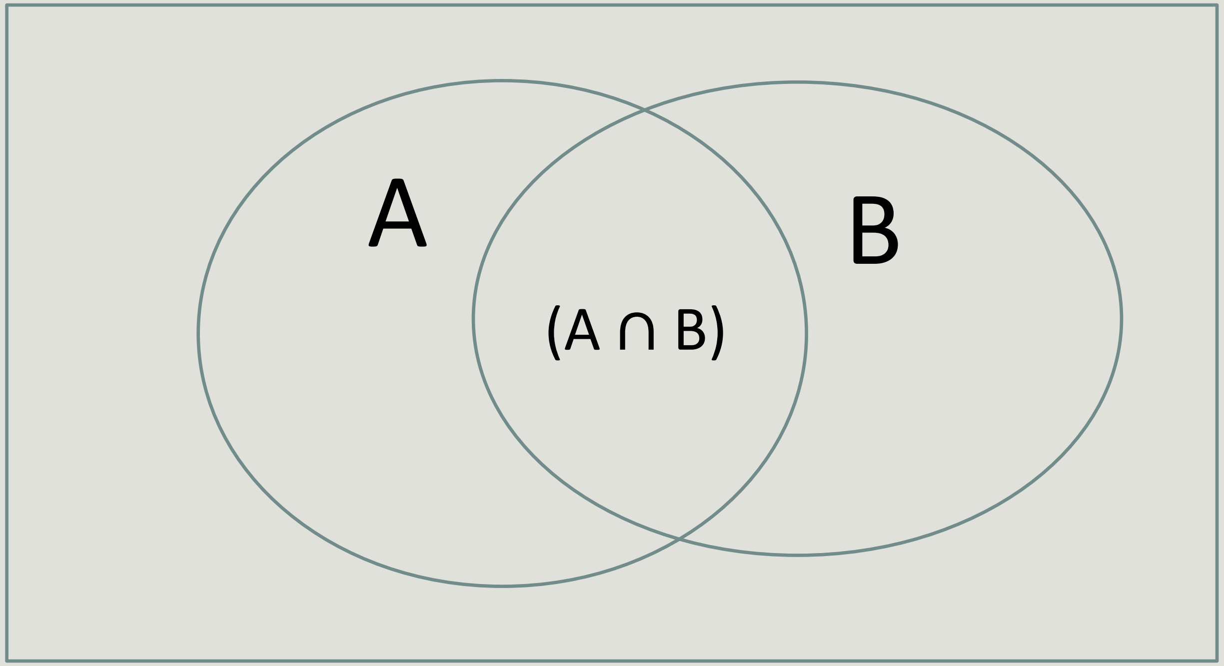


Para A y B $\subseteq$ S tal que $A \cap B = \varnothing$

$$P(A \cup B) = P(A) + P(B)$$

$P(A \cup B)$ si $A \cap B \neq \varnothing$

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



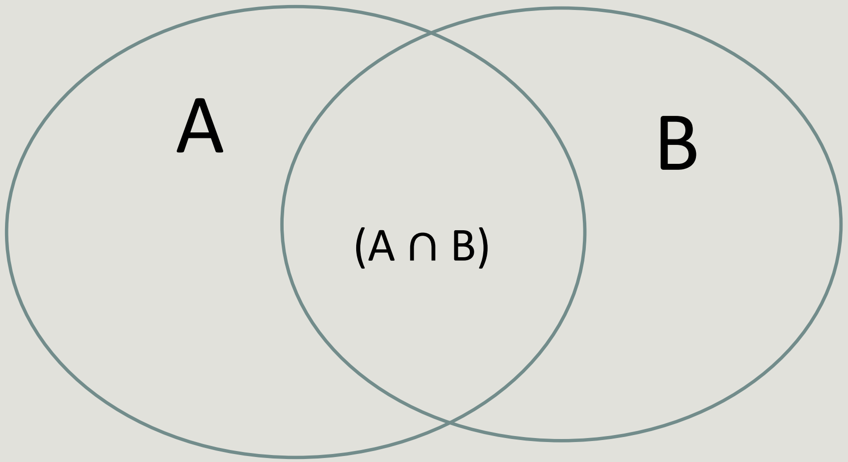
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Por que $A \cap B$ se esta tomando 2 veces

Probabilidad Condicional

Sea A y B eventos independiente tal que $P(B) \neq 0$. Entonces la probabilidad de que ocurra A dado que ocurrió B $P(A \mid B)$ es;

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



Ejemplo: Probabilidad de que ocurra un evento dado que ocurrió otro:

- Dado que el dado cayó impar, ¿Cuál es probabilidad de que sea un número primo?
- Dado que tiene fiebre, ¿Cuál es la probabilidad de que tenga gripe?

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

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

Ejemplo 1 Probabilidad condicional

Calcular la probabilidad de que al tirar un dado, salga un valor $<=2$ dado que el resultado ha sido < 4:

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

$=P(\{1,2\}|\{1,2,3\})$

$=P(\{1,2\} \cap \{1,2,3\})/P(\{1,2,3\})$

$=(2/6)/(3/6)$

$=2/3$

Ejemplo 2 Probabilidad condicional

Evento A: 75% de pacientes dio positivo a gripe común (GC) -> $P(A)=0.75$

Evento B: 68% de pacientes dio positivo coronavirus (CV)-> $P(B)=0.68$

Sabemos que el 85% de enfermos con GC habían dado positivo a CV -> $P(A|B)=0.85$

¿Si un enfermo X sabe que dio positivo en el test de GC, que probabilidad tiene de dar positivo en el test de CV?

$P(B|A)=?$

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

$=0.68*0.85$

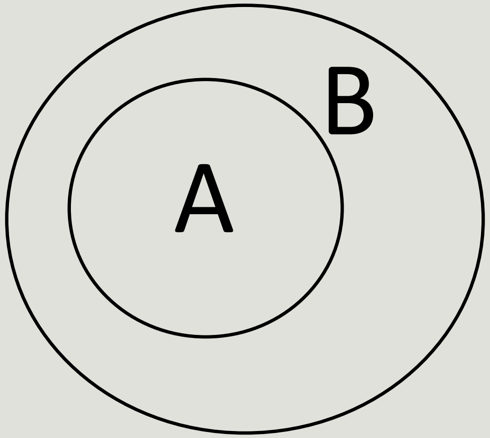
$=0.578$

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

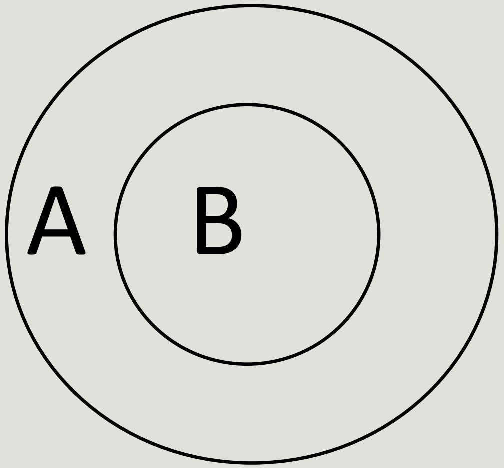
$=0.578/0.75 = 0.77$

R= 77%

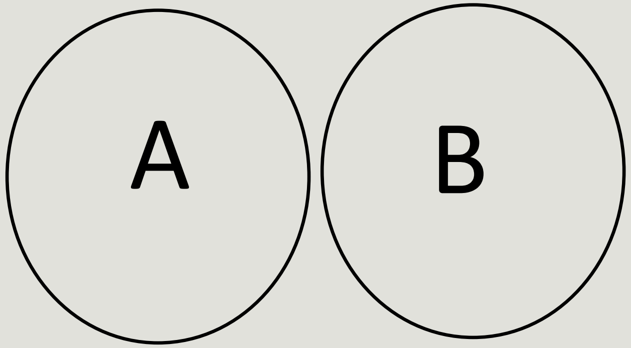
Casos probabilidad condicional



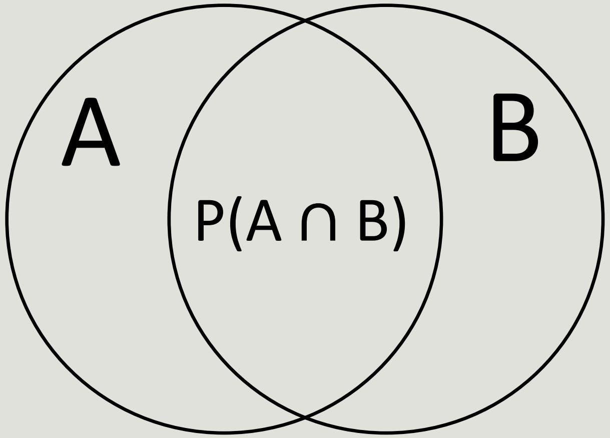
$P(A|B)$
$= P(A \cap B)/P(B)$
$= P(A)/P(B)$



$P(A|B)$
$= P(A \cap B)/P(B)$
$= P(B)/P(B)$
$= 1$



$P(A|B)$
$= P(A \cap B)/P(B)$
$= P(\emptyset)/P(B)$
$= 0$



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

Regla de Bayes

De la definición de probabilidad condicional se puede deducir:

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

Esto permite "invertir" las probabilidades, por ejemplo obtener la *P* de una enfermedad dado un síntoma, con conocimiento de la *P* de los síntomas dado que alguien tiene cierta enfermedad

Sabemos que:
P(A|B)=P(A ∩ B)/P(B)
P(A ∩ B)=P(A|B)*P(B)

De igual forma
P(B|A)=P(A ∩ B)/P(A)

Entonces:

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

Ejemplo

Si tenemos nueva información ¿Cómo modifica esto a la probabilidad a priori?

Ejemplo:

Se sabe que en México se dan 35 casos de cáncer por cada 100 mil mujeres

Entonces a priori sabemos que:

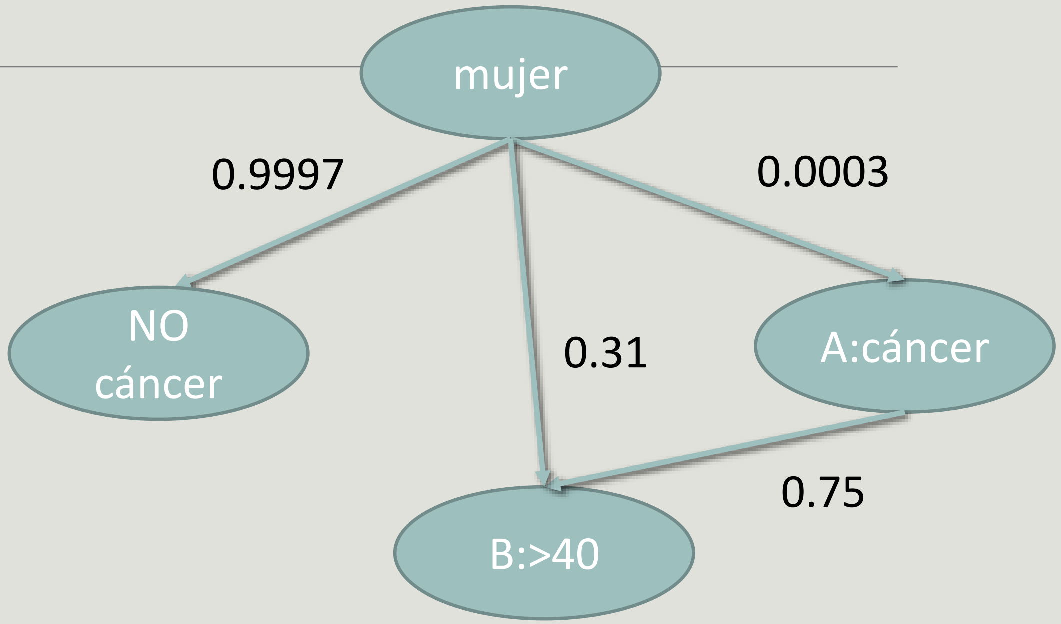
P(cs)=35/100000 = 0.00035= 0.035%

Hay 0.035% de que una mujer adquiera cáncer de seno,

Pero si tenemos la siguiente información nueva:

1. El 75% de los casos ocurren después de los 40 años.
2. 31% de las mujeres tienen mas de 40 años

¿Cuál es la probabilidad de adquirir cáncer de una mujer de 50 años?



P(A|B)=[P(A)*P(B|A)]/P(B)

P(A|B)=[0.0003 * 0.75]/0.31

0.0007258065

0.072

Probabilidad conjunta

Medida estadística que indica la probabilidad de que dos o más sucesos ocurran al mismo tiempo.

Si los eventos A y B son independientes, la probabilidad conjunta se calcula multiplicando sus probabilidades individuales:

$P(A \text{ y } B)=P(A)*P(B)$.

Si los eventos son dependientes, la ocurrencia de uno afecta la probabilidad del otro. la fórmula se ajusta para incluir la probabilidad condicional, que es la probabilidad de que ocurra el evento B dado que ya ha ocurrido el evento A. La fórmula se convierte en:

$P(A \text{ y } B)=P(A)*P(B|A)$

Teorema de bayes y probabilidad conjunta

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

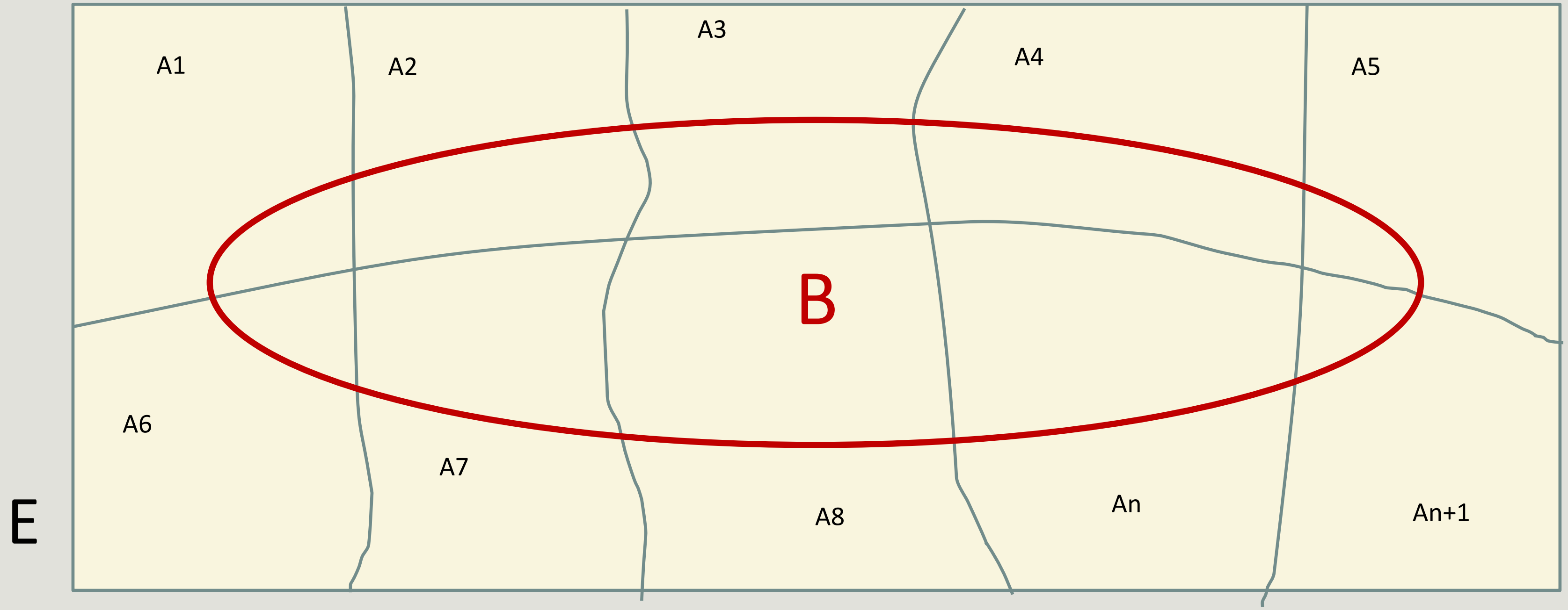
Aplicando la probabilidad conjunta: la probabilidad de que ocurra el evento B dado que ya ha ocurrido el evento A. Como $P(A \text{ y } B)=P(A)*P(B|A)$ tenemos

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

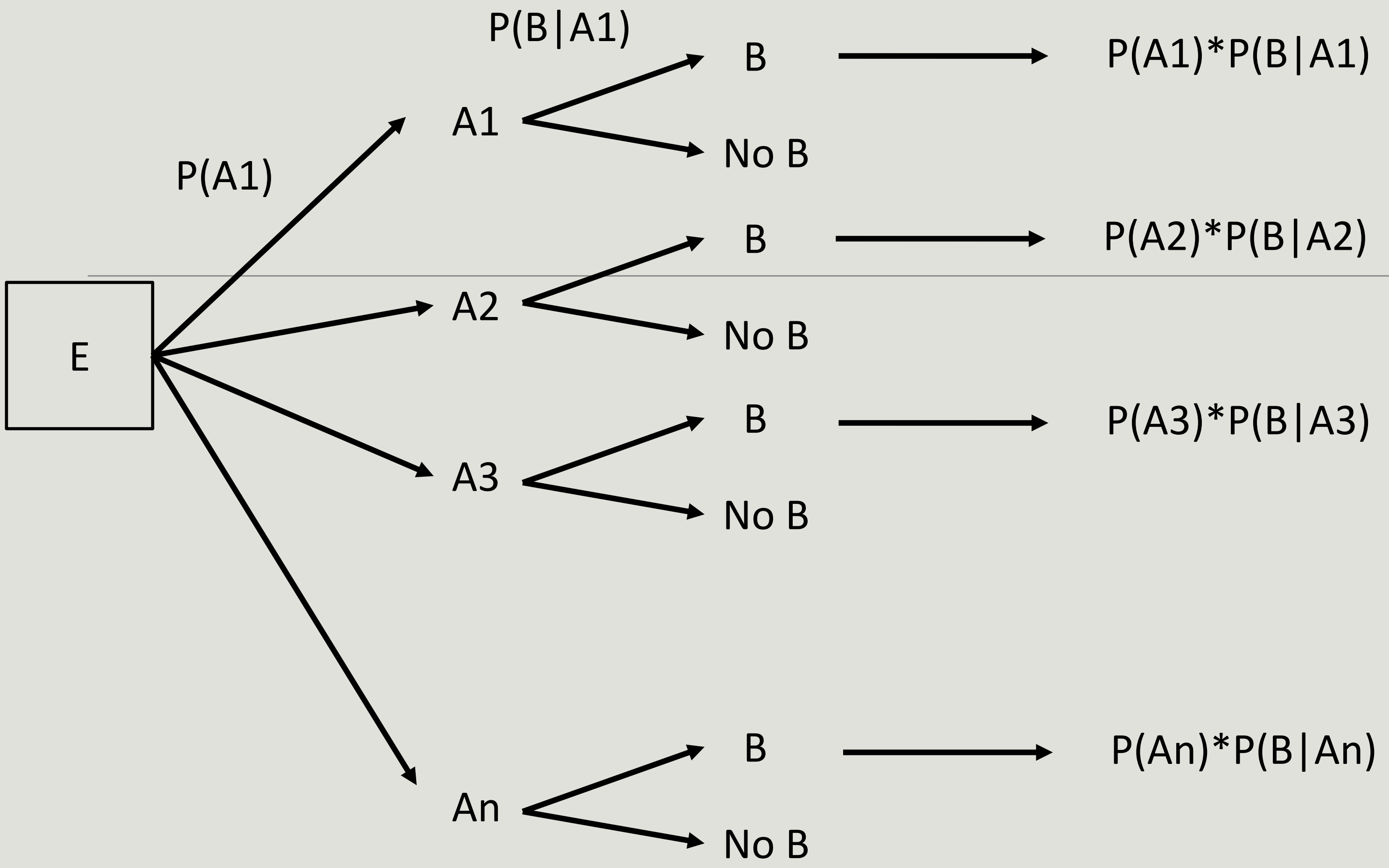
Los sucesos Ai .. An forman una partición del espacio muestral E si cumplen:

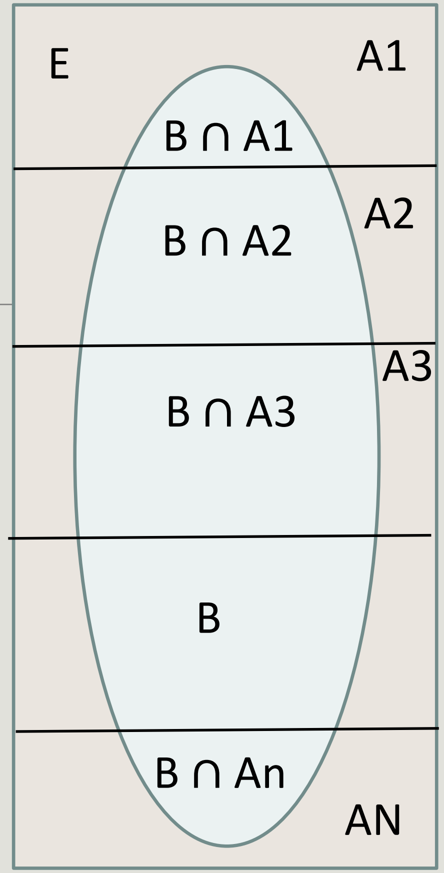
$Ai \cap Aj = \emptyset$

$A1 \cup A2 \ldots \cup An = E$



Espacio muestral

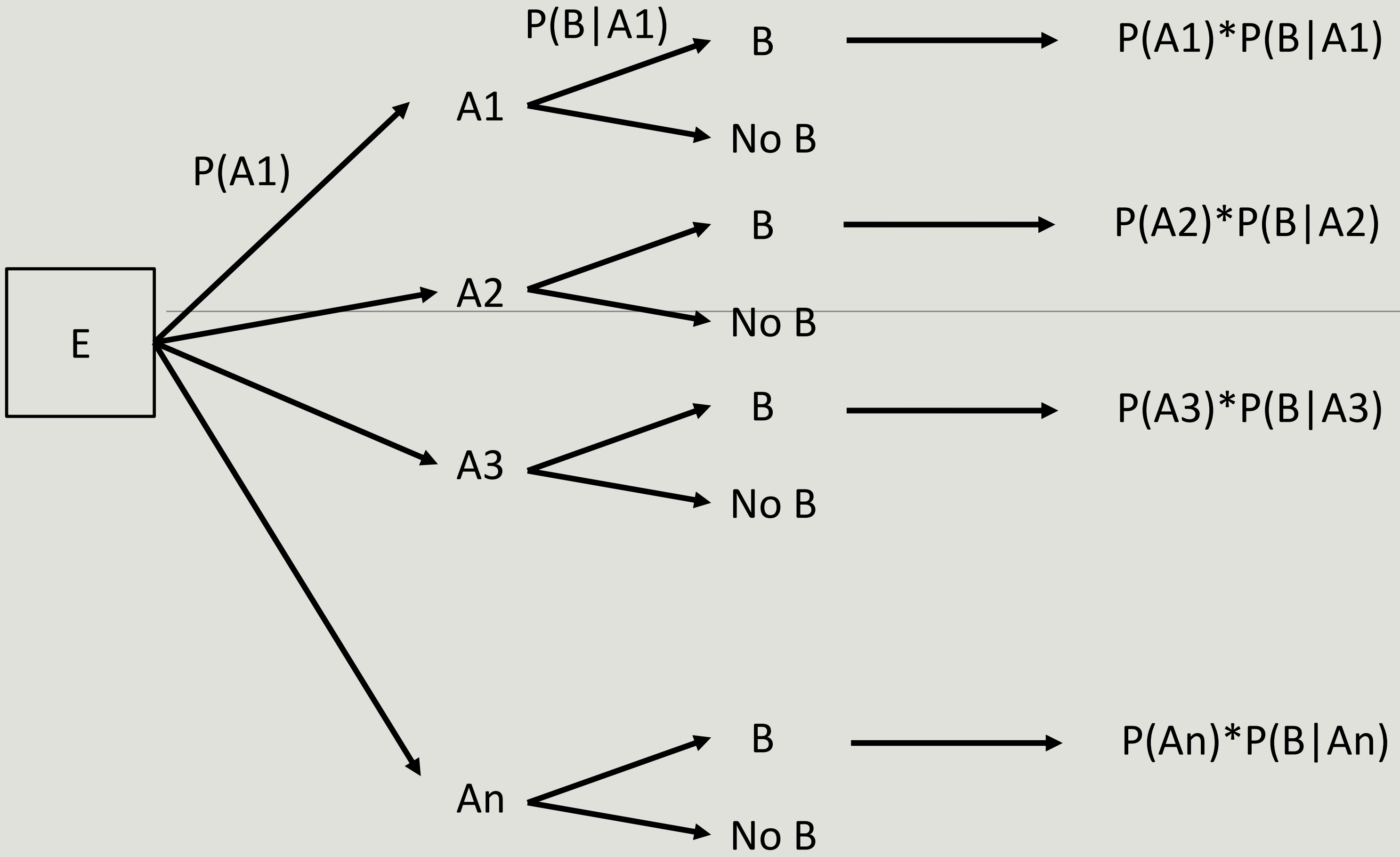


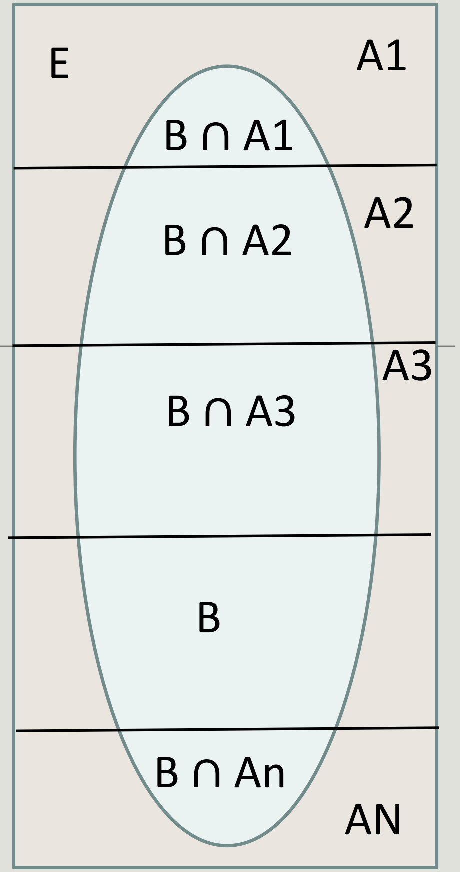


Teorema de la probabilidad total

P(B) = P(A1)*P(B|A1)+P(A2)*P(B|A2)+…+P(An)*P(B|An)

$= \sum(P[B|Ai] * P[Ai])$





Teorema de la probabilidad total

P(B)

= P(A1)*P(B|A1)+P(A2)*P(B|A2)+…+P(An)*P(B|An)

= $\sum(P[B|Ai] * P[Ai])$

$$\mathrm{P(Ai|B)} = \frac{P(Ai) * P(B|Ai)}{\sum(P[B|Ai] * P[Ai])}$$

Ejemplo bayes enfermedad datos

Variantes de una enfermedad:

Prevalencia   Presencia de síntomas

A1:…40%……………..2%

A2:…30%……………..5%

A3:…20%……………..10%

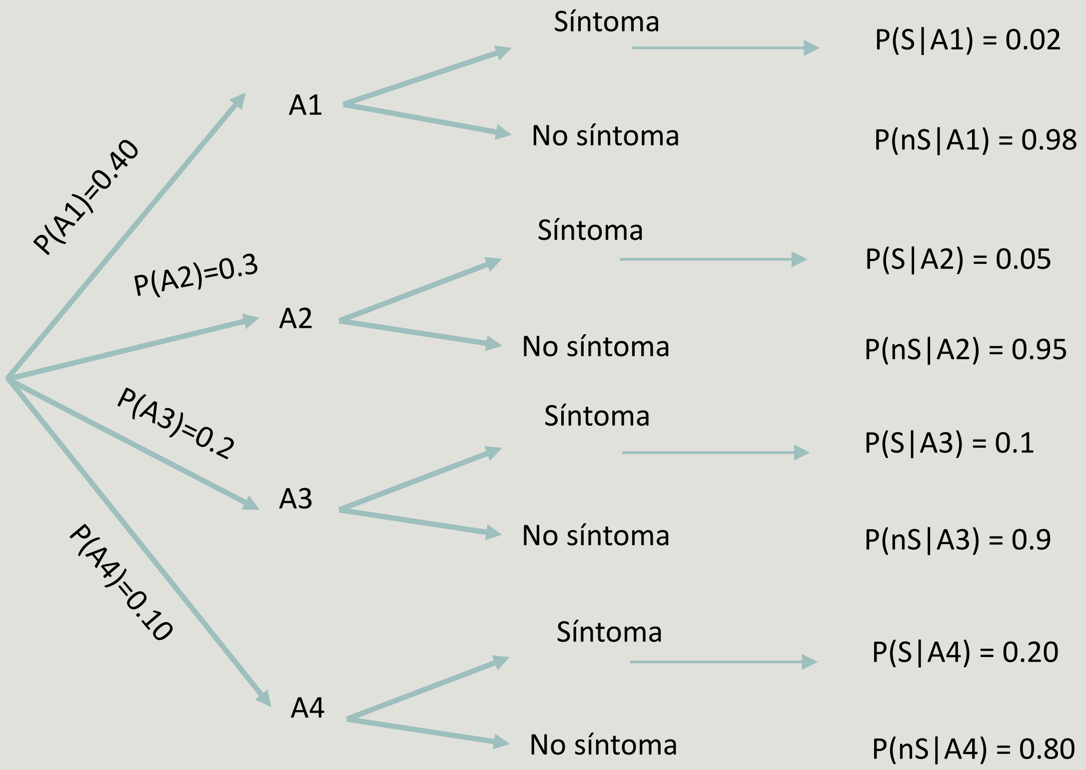
A4:…10%……………..20%

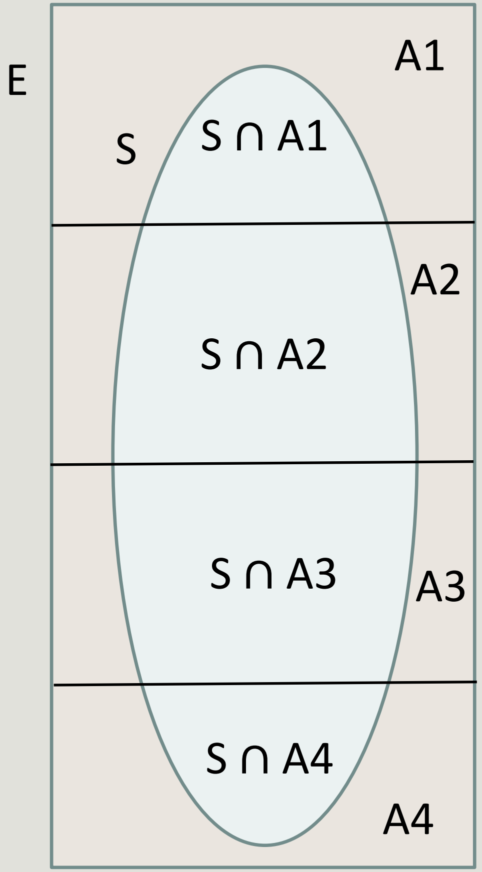
Nota: prevalencia es una **medida epidemiológica que indica la proporción de personas en una población que tienen una enfermedad o un determinado evento de salud en un momento específico o durante un periodo determinado.**

Si se observa la presencia del síntoma ¿que variante es mas probable?

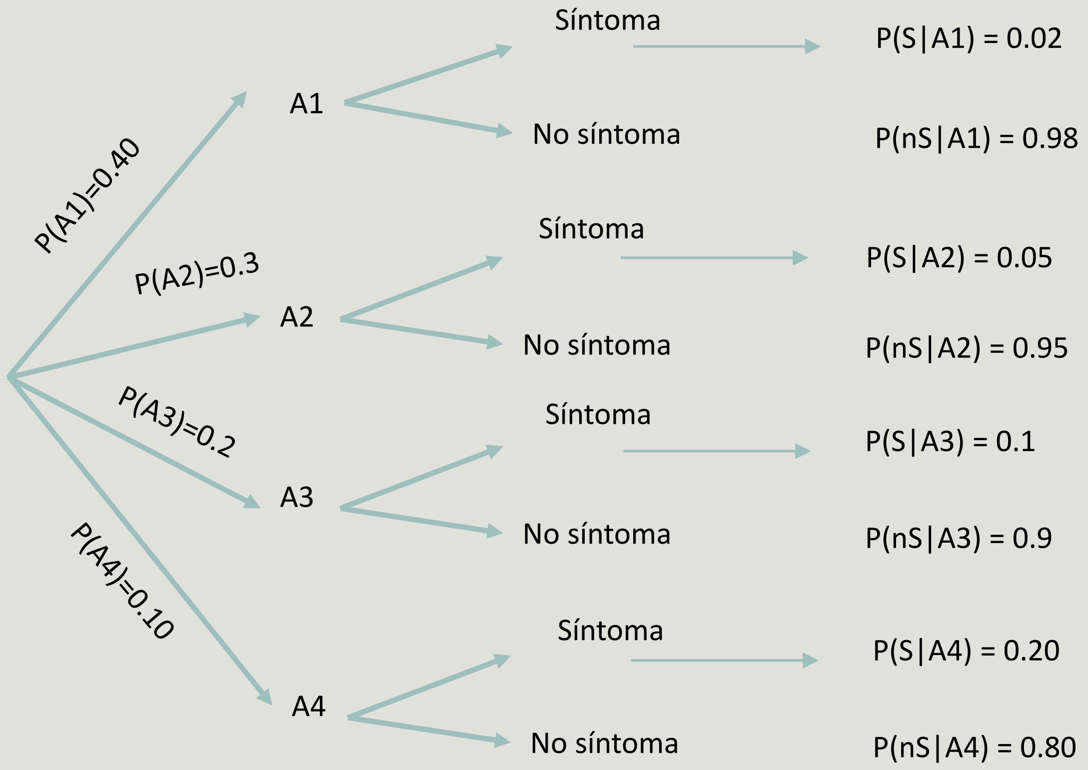
Análisis enfermedad

Se observa que Ai es una partición de espacio muestral





Solución ENFERMEDAD



Si se observa la presencia del síntoma ¿que variante es mas probable?

$P(A_1|S)= [P(A1) * P(S|A_1)]/ \sum(P(Ai) * P(S|Ai))$

$P(A_1|S)=(0.4 * 0.02)/(0.4*0.02+0.3*0.05+0.2*0.10+0.10*0.20)$

=0.008/0.063

=0.172

De la misma forma

$P(A_2|S)=(0.30*0.5)/(0.4*0.02+0.3*0.05+0.2*0.10+0.10*0.20)=0.238$

$P(A_3|S)=(0.20*0.10)/(0.4*0.02+0.3*0.05+0.2*0.10+0.10*0.20)=0.317$

$P(A_4|S)=(0.10*0.20)/(0.4*0.02+0.3*0.05+0.2*0.10+0.10*0.20)=0.317$

Regla de la Cadena

Sean $A1,A2,\ldots,An$ eventos dependientes. La probabilidad de que ocurran **todos** es:

$$P(A1\cap A2\cap\cdots\cap An)=P(A1)\cdot P(A2|A1)\cdot P(A3|A1\cap A2)\cdots P(An|A1\cap A2\cap\cdots\cap An-1)$$

Para 3 eventos:

$$P(A\cap B\cap C)=P(A)\cdot P(B|A)\cdot P(C|A\cap B)$$

Ejemplo

Se tiene una urna con 3 bolas rojas, 2 bolas verdes y 4 bolas blancas. Se extrae al azar, de una en una y sin reposición, tres bolas de esa urna. Calcular la probabilidad de que la secuencia extraída sea (roja, blanca, roja).

Total de bolas = 9

A extraer primear roja

B extraer segunda blanca

C extraer tercera roja

$P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B)$

P(A) = La primer es roja = 3/9 = 1/3 = 0.33

P(B|A) = La segunda blanca | la primera roja =4/8 = 1/2 = 0.5

$P(C|A \cap B)$ = 3ª roja|1ª roja y 2ª blanca = 2/7

$P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B)$

$= (1/3) \cdot (1/2) \cdot (2/7) = 2/42 = 1/21$

Respuesta final: 1/21