Módulo 4 – Inferencia Estadística

Probabilidad

Ciencia de Datos

Objetivos

- Explicar los principales conceptos de probabilidad asociados a un evento aleatorio.

- Realizar cálculos básicos de probabilidad.

Contenido

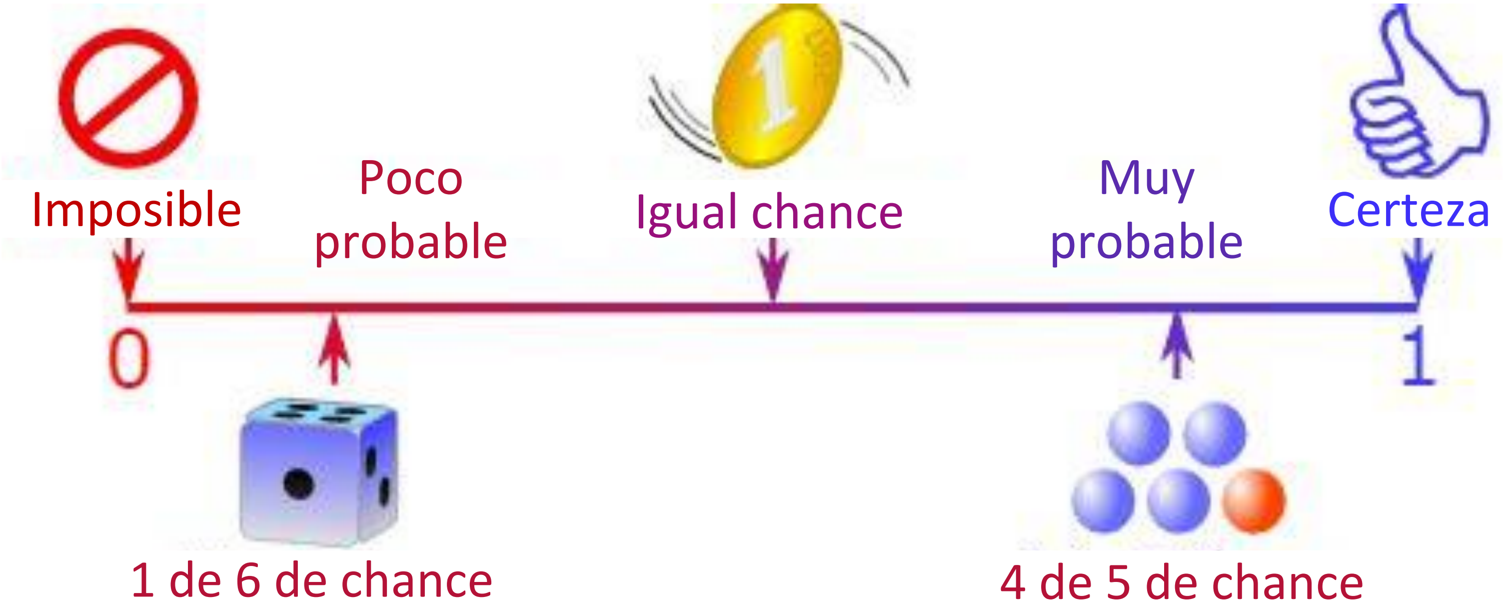
- Probabilidad.
- Suceso y Espacio Muestral.
- Reglas aditivas y multiplicativas.
- Probabilidad Condicionada.
- Árboles de Probabilidad.

¿Qué significa Probabilidad?

- La probabilidad es una medida numérica que se utiliza para cuantificar la posibilidad o chance de que un evento o suceso ocurra en una situación determinada. En otras palabras, es la medida de la certidumbre o incertidumbre de que algo suceda.

- Por ejemplo, si la probabilidad de que llueva mañana es 0.5, entonces esto significa que hay una chance del 50% de que llueva y una chance del 50% de que no llueva.

¿Qué significa Probabilidad?



¿Qué significa Probabilidad?

- La Teoría de Probabilidades, es una rama de las matemáticas que se ocupa del estudio de los fenómenos aleatorios y la cuantificación de la incertidumbre asociada a ellos. La teoría de probabilidades se utiliza para modelar y analizar situaciones en las que el resultado es incierto o aleatorio.

- La teoría de probabilidades tiene muchas aplicaciones en diversos campos, como la estadística, la física, la ingeniería, la economía, la informática, entre otros. Se utiliza para hacer predicciones y tomar decisiones en situaciones inciertas, como en la evaluación de riesgos, el diseño de experimentos, la optimización de procesos, entre otros.

Definición Probabilidad

Podemos determinar la probabilidad de ocurrencia de un acontecimiento o suceso A, y la denotaremos como P(A), si dividimos el número de casos favorables entre el número de casos igualmente posibles, es decir:

$$P(A) = \frac{Casos\ favorables}{Casos\ posibles}$$

Ejemplo de Probabilidad

- Por ejemplo, considere la siguiente ruleta con 16 zonas de distinto color: 7 rojas, 5 azules, 3 verdes y 1 amarilla.

- Cada color tiene una probabilidad distinta de salir tras hacer girar la ruleta:

- La probabilidad de que salga color rojo es 7/16.
- La probabilidad de que salga color azul es 5/16.
- La probabilidad de que salga color verde es 3/16.
- La probabilidad de que salga color amarillo es 1/16.

P(A) = (*Casos favorables*) / (*Casos posibles*)

Probabilidad

¿Cuál es la probabilidad de que al lanzar un dado salga un número mayor a 4?

¿Cuál es la probabilidad de que al lanzar un dado salga un número mayor a 5?

¿Cuál es la probabilidad de que al lanzar un dado salga número impar?

¿Cuál es la probabilidad de que al lanzar un dado salga número par?

¿Cuál es la probabilidad de que al lanzar un dado salga 1?

Paradoja del Cumpleaños

- Esta paradoja hace la siguiente pregunta: ¿cuántas personas se necesitan como mínimo para que sea más probable que al menos 2 de ellas cumplan años el mismo día?

- A pesar de lo que nos indica la intuición, si mantenemos el supuesto de que los años tienen 365 días, la paradoja establece que hacen falta 23 personas para que haya una probabilidad del 50% de que al menos 2 de ellas cumplan años el mismo día.
- Y resulta que, si en una fiesta hay más de 57 invitados, la probabilidad de que dos personas cumplan años el mismo día es del 99% .

Enfoques de Probabilidad

Para determinar la probabilidad, tenemos varios enfoques:

- Interpretación frecuentista .
- Interpretación clásica.

KIBERNUM

Interpretación Frecuentista

- En un experimento aleatorio, la frecuencia relativa asociada al suceso A denotada por $f_r(A)$, corresponde a la razón entre la frecuencia absoluta $f(A)$ y la cantidad veces que se realice el experimento.

$$\boldsymbol{f_r(A) = \frac{cantidad\ de\ veces\ que\ ocurre\ A}{Cantidad\ de\ veces\ que\ se\ realiza\ el\ experimento} = \frac{f(A)}{n}}$$

Una forma de calcular la probabilidad de un suceso es analizar la tendencia de la frecuencia relativa, al repetir un experimento infinitas veces.

Interpretación Clásica

Dos sucesos son equiprobables si tienen la misma probabilidad de ocurrencia.

Ej.: Lanzar una moneda equilibrada. Que salga sello tiene la misma probabilidad que salga cara.

$$P(A) = \frac{cantidad\ de\ casos\ favorables}{cantidad\ de\ casos\ posibles} = \frac{\#A}{\#\Omega}$$

Suceso y Espacio Muestral

Suceso y Espacio Muestral

Espacio Muestral es el conjunto formado por todos los posibles resultados que pueden darse en un experimento aleatorio. Cada uno de estos resultados es conocido como **evento o suceso elemental**. **Por ejemplo:**

S={1,2,3,4,5,6}

S={cara, sello}

Suceso y Espacio Muestral

Se denomina evento o suceso a cualquier subconjunto del espacio muestral formado por sucesos elementales.

Ejemplo: Lanzamiento de un dado.
Suceso A: Obtener un número par

- Un suceso **es seguro** cuando coincide con el espacio muestral del experimento.
- Un suceso **es imposible** si nunca ocurre, o, si no se presenta al realizar un experimento aleatorio.
- Dos sucesos son **mutuamente excluyentes** o disjuntos si no pueden suceder simultáneamente, es decir, si y sólo si su intersección es vacía.
- La cantidad de elementos que componen un suceso A se denomina **cardinalidad** de A y se denota $\#A$.
- Análogamente se denota **#Ω** es la cantidad de resultados posibles de un experimento aleatorio.

Ejemplos de Suceso y Espacio Muestral

Ejemplo 1

Se tienen 20 tarjetas numeradas correlativamente del 1 al 20 en una bolsa no transparente. Si se extrae una tarjeta:

1. ¿Cuál es la probabilidad de extraer un número par?
2. ¿Cuál es la probabilidad de extraer un múltiplo de 3?

Ejemplo 1

Espacio Muestral={1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20}
#Ω=20

¿Cuál es la probabilidad de extraer un número par?

$$\mathrm{A} = \{2,4,6,8,10,12,14,16,18,20\}$$
$$\#A = 10$$
$$P(A) = \frac{\#A}{\#\Omega} = \frac{10}{20} = \frac{1}{2}$$

¿Cuál es la probabilidad de extraer un múltiplo de 3?

$$\mathrm{B} = \{3,6,9,12,15,18\}$$
$$\#B = 6$$
$$P(B) = \frac{6}{20} = \frac{3}{10}$$

Ejemplo 2

Se encuestaron a alumnos de enseñanza media de un colegio particular sobre las preferencias de los deportes que practican, entregándose las respuestas en la siguiente tabla:

Deporte	Número de alumnos que lo práctica	f relativa
Futbol	120	0,48
Tenis	22	0,088
Handball	38	0,152
Voleibol	56	0,224
Otros	14	0,056

a) Si se escoge un alumno al azar. Determine la probabilidad que practique tenis.
b) ¿Cuál es la probabilidad que un alumno no juegue tenis?

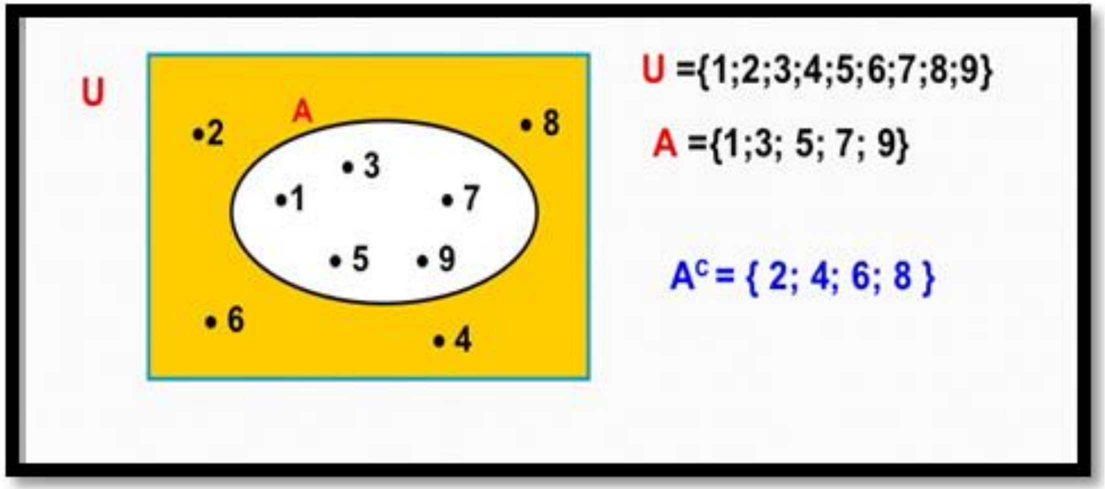
Propiedades de la Probabilidad

Conceptos básicos sobre conjuntos:

Conjunto: Es una colección de elementos. Se denotan con letras mayúsculas. El conjunto formado por todos los elementos. En probabilidad el universo es el espacio muestral(Ω).

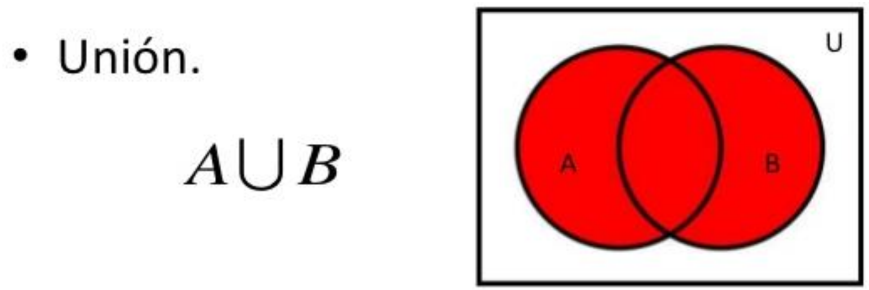
Cardinalidad de un conjunto: es la cantidad de elementos que tiene y se denota por $\#A$.

Complemento de A: es el conjunto formado por los elementos que no pertenecen a A.

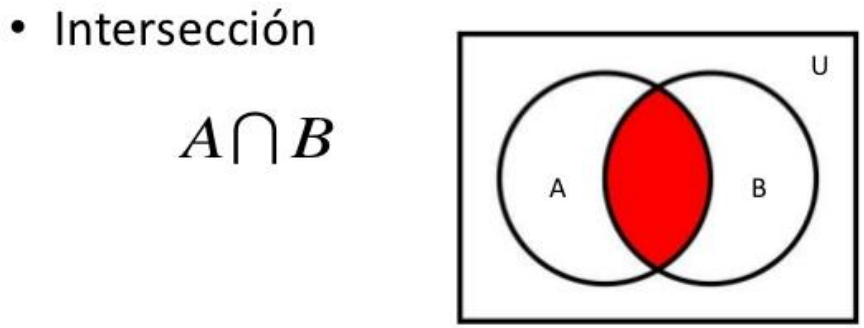


Propiedades de la Probabilidad

Unión de conjuntos: Corresponde a los conjuntos formados por los elementos que pertenecen a A o los elementos que pertenecen a B

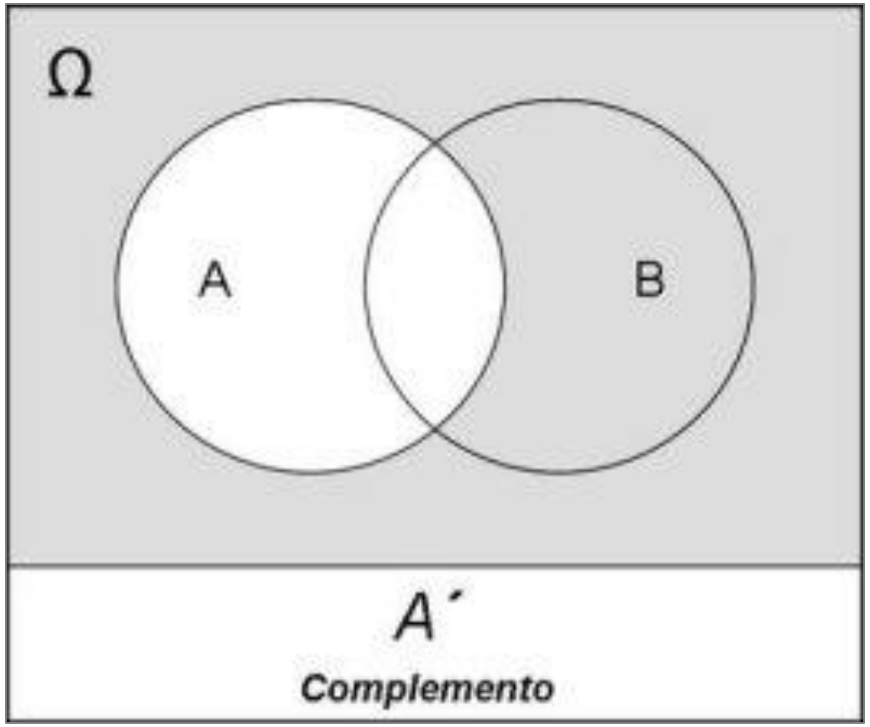


Intersección de conjuntos: Corresponden a los elementos que están simultáneamente en A y en B



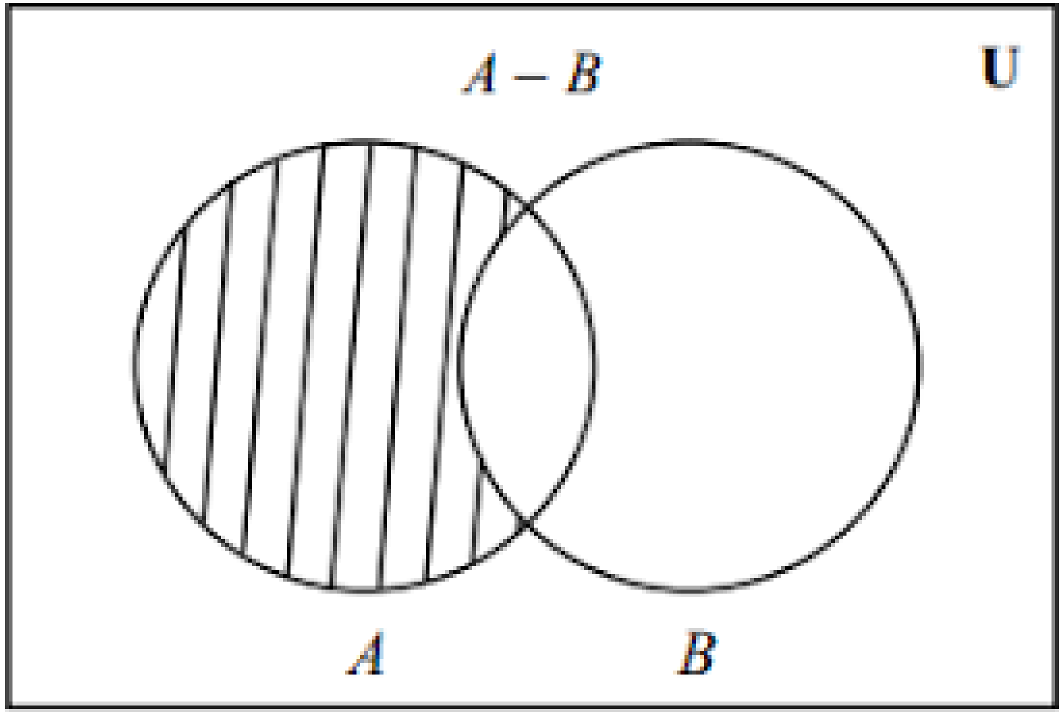
Propiedades de la Probabilidad

Complemento de A: Son todos los elementos que no están en A.



Propiedades de la Probabilidad

Diferencia entre A y B: Corresponde a los elementos que están en A y que no están en B.



Reglas Aditivas y Multiplicativas

Regla de la Adición

Imagina el experimento de lanzar un dado:

$$\Omega = \{1,2,3,4,5,6\}$$

Se definen dos sucesos:
A: Obtener un número menor que dos
A = { 1 }

B: Obtener un numero par
B = { 2, 4, 6}

Si ahora queremos obtener la probabilidad de A U B, entonces los casos favorables serán:

A U B = { 1, 2, 4, 6}

Es decir, al ser **disjuntos o mutuamente** excluyentes, se tiene:

P(A U B) = P(A) + P(B)

Regla de la Adición

Imagina el experimento de lanzar un dado:

$$\Omega = \{1,2,3,4,5,6\}$$

Se definen dos sucesos:
A: Obtener un numero par
A = { 2, 4, 6 }

B: Obtener un numero primo
B = { 2, 3, 5 }

Si ahora queremos obtener la probabilidad de A U B, entonces los casos favorables serán:

A U B = { **2**, 3, 4, 5, 6}

En este caso, hay elementos comunes entre A y B, por lo tanto:

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

Propiedades de la Probabilidad

En un experimento, dos sucesos son mutuamente excluyentes si la ocurrencia de uno excluye la ocurrencia de otro, por lo tanto, la intersección es vacía.

$$(A \cap B) = \emptyset \qquad\qquad P(A \cup B) = P(A) + P(B)$$

Si se tienen dos sucesos A y B cualesquiera de un espacio muestral E, la probabilidad de que ocurra $A \cup B$ se expresa como:

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

Regla Multiplicativa

La probabilidad de la intersección de dos eventos A y B se calcula como:

$$\mathbf{P}(\mathbf{A} \cap \boldsymbol{B}) = \mathbf{P}(\boldsymbol{B}) \cdot \mathbf{P}(\boldsymbol{A}/\boldsymbol{B})$$

Donde $P(A/B)$ corresponde a la probabilidad del evento A dado la ocurrencia de B. Se conoce como probabilidad condicional.

Dos eventos **son independientes** si $P(A \cap B) = P(A) \cdot P(B)$, o en forma equivalente, **dos eventos son independientes si la realización de uno no afecta la probabilidad del otro**, es decir $P(A/B) = P(A)$.

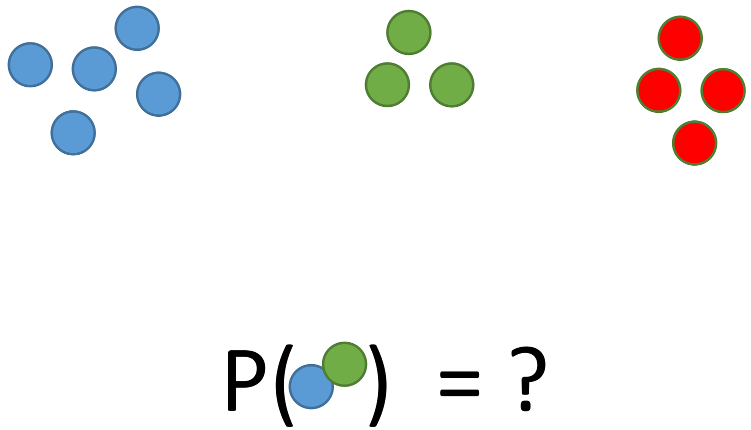
Esto se conoce como regla multiplicativa de las probabilidades, se utiliza para calcular la probabilidad que suceda un evento y otro a la vez.

Ejemplo

Se tiene una bolsa negra con 5 bolitas azules, tres bolitas verdes, y 4 rojas. Si extraemos dos bolas, ¿cuál es la probabilidad de extraer una bola azul y una bola verde?

Para resolver este problema, tenemos dos enfoques:

- **Con reposición**, es decir, se saca una bolita y se vuelve a poner dentro de la bolsa para la segunda extracción.
- **Sin reposición**, es decir, se saca una bolita y no se devuelve a la bolsa para realizar la segunda extracción.



Ejemplo (A)

A) Sin Reposición de bolas

Si definimos los eventos:
A: Extraer una bola azul
B: Extraer una segunda bola verde

Al ser con reposición, el evento B es totalmente independiente del evento A, lo que ocurra en A no influye en el resultado en B. Por lo tanto, la probabilidad que ocurran A y B es:

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

$\boldsymbol{P(A)}$ es igual a 5 casos favorables dividido en 12 bolitas en total.
$\boldsymbol{P(B)}$ es igual a tres casos favorables dividido en 12 bolitas en total.

Por lo tanto,

$$P(A \cap B) = P(A) \cdot P(B) = \frac{5}{12} \cdot \frac{3}{12} = \frac{5}{48}$$

Ejemplo (B)

B) Con Reposición de bolas.

Ahora no se devuelva la primera bolita extraída, por lo tanto, lo que sucede en la primera extracción afecta el número de casos posibles, ya que disminuye en una bolita los casos posibles para la segunda extracción. Por esto se llama probabilidad condicional, dado que el resultado de A afecta a B. Por lo tanto,

P(A)=5/12

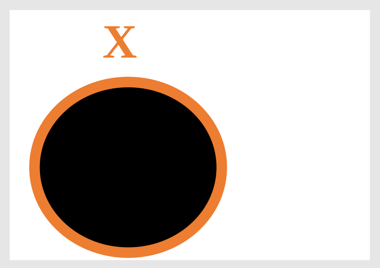
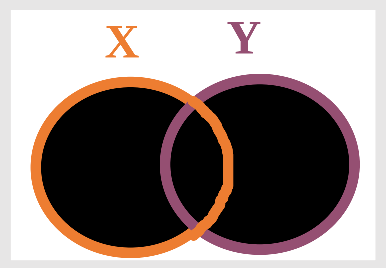
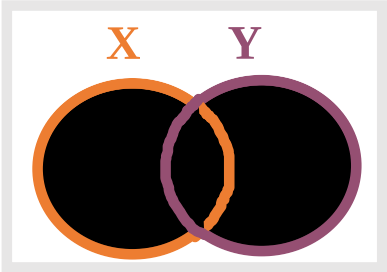
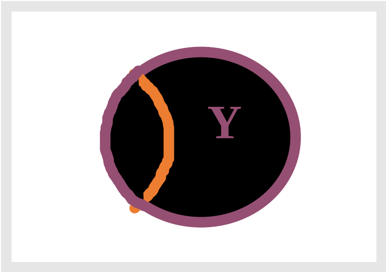
Y, como no devuelvo la bolita, ahora hay 11 casos posibles para la segunda extracción. Por esto:

P(B/A)= 3/11

En consecuencia,
$P(A \cap B)$**=5/12·3/11=15/132**

Probabilidad Condicional

Tipos de Probabilidades

Marginal	Unión	Conjunta	Condicional
$P(X)$ La probabilidad de que ocurra **X**	$P(X \cup Y)$ La probabilidad de que ocurra **X o Y**	$P(X \cap Y)$ La probabilidad de que ocurra **X e Y**	$P(X\|Y)$ La probabilidad de que ocurra **X sabiendo que ha ocurrido Y**
			

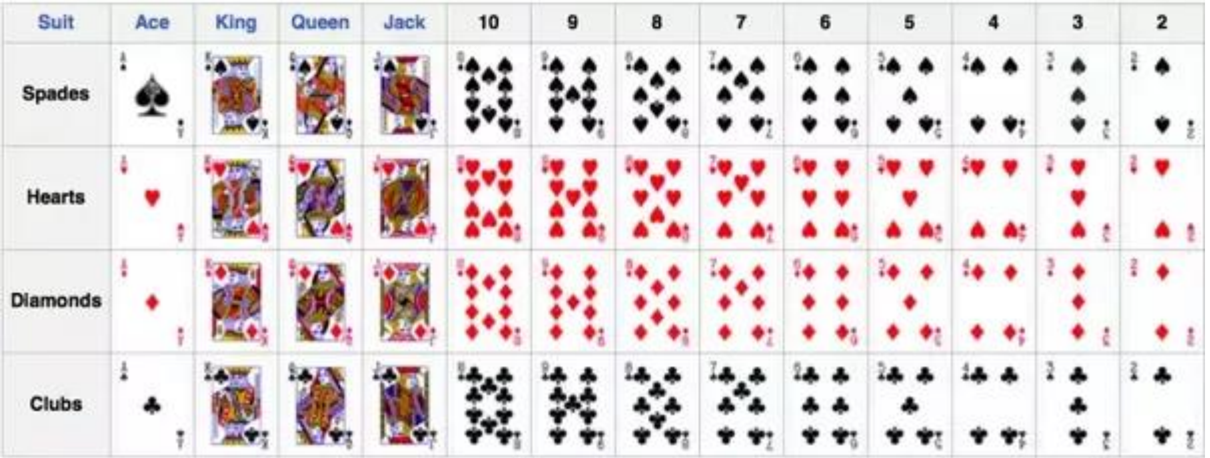
Probabilidad Condicional

Se llama probabilidad de A condicionada a, o probabilidad de un suceso A sabiendo que se ha producido un suceso B a:

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

Ejemplo 1

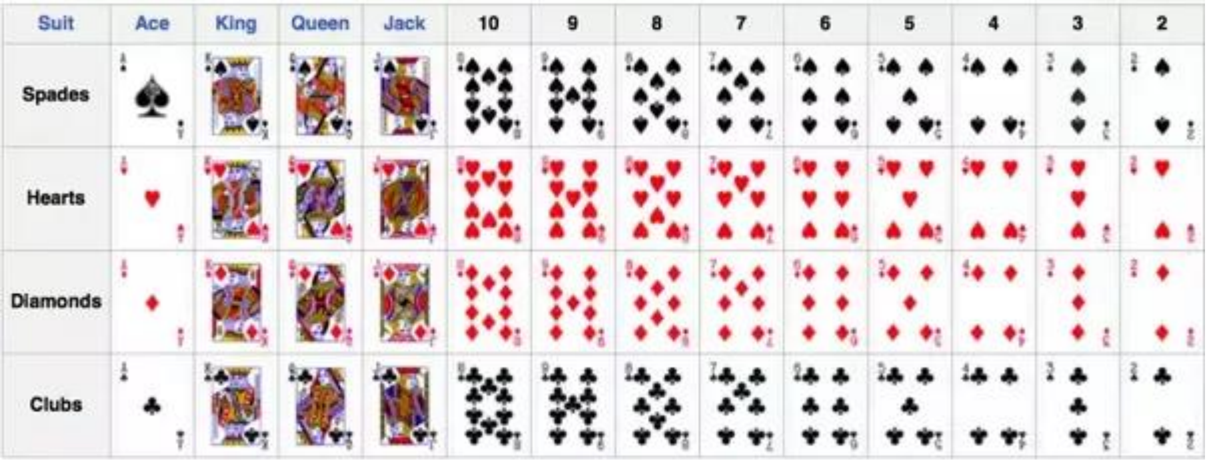
¿Cuál es la probabilidad que una carta escogida al azar sea un AS, sabiendo que es roja?



$$P(As|Rojo) = \frac{P(As \cap Rojo)}{P(Rojo)} = \frac{2/52}{26/52} = \frac{2}{26} = \frac{1}{13}$$

Ejemplo 2

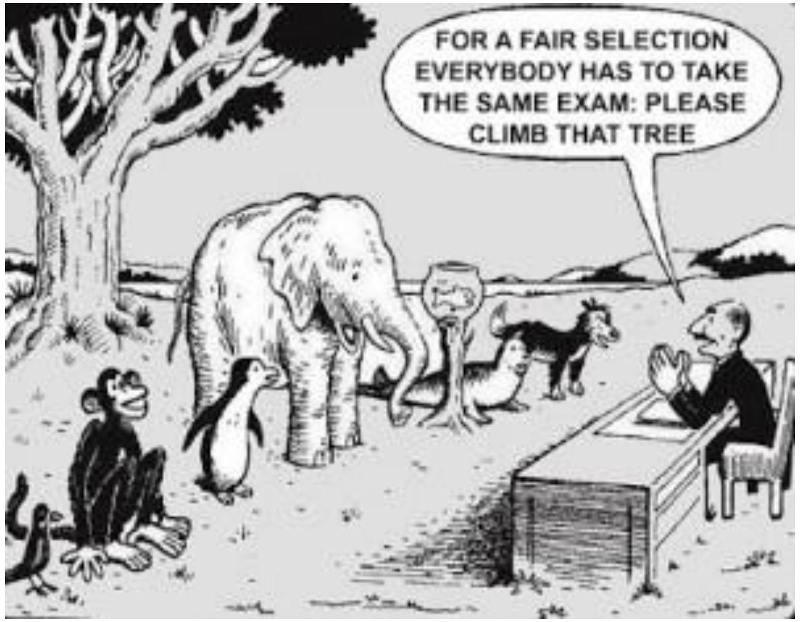
¿Qué resultado arrojaría si te piden el cálculo al revés? Es decir, la probabilidad que sea roja, sabiendo que es un AS.



$$P(Rojo|AS) = ?$$

Propiedades de la Probabilidad

la distinción entre $P(A|B)$ y $P(B|A)$ es muy importante y necesita reconocerse perfectamente. Veamos dos ejemplos que nos van a permitir enfatizar la diferencia entre ambas probabilidades. El primero es un ejemplo clásico pero muy clarificador : sea S el suceso *tengo dos brazos y dos piernas* y sea R el suceso *soy un mono*. Obviamente, $P(S|R) = 1$, mientras que $P(R|S) \neq 1$. La primera probabilidad es equivalente a afirmar que *si soy un mono entonces tengo dos brazos y dos piernas*, mientras que la segunda es equivalente a *si tengo dos brazos y dos piernas, no tengo porqué ser un mono*.



Ejemplo 2

Suponga que se estudia si el color del pelo está asociado al color de los ojos. Se analizaron 300 personas seleccionadas aleatoriamente con los siguientes resultados. Si se selecciona una de estas personas al azar, encuentre la probabilidad de que la persona tenga el pelo negro, dado que tiene los ojos café.

Color del pelo	Color de los ojos		
	Café	Azul	Otro
Negro	70	30	20
Rubio	20	110	50

Ejemplo 2

Primero asignamos letras a los eventos y calculamos los **totales** de la tabla:

	C	A	O	Total
N	70	30	20	**120**
R	20	110	50	**180**
Total	**90**	**140**	**70**	**300**

$$P\begin{pmatrix} \text{que tenga el pelo negro,} \\ \text{dado que tiene los ojos café} \end{pmatrix} = \frac{P(\text{que tenga pelo negro y ojos café})}{P(\text{que tenga los ojos café})}$$

$$P(N/C) = \frac{P(N \cap C)}{P(C)} = \frac{70/300}{90/300} = 0{,}78$$

Por lo tanto, hay un 78% de probabilidad de escoger a una persona que tenga el pelo negro, dado que tiene ojos café.

Ejemplo 3

Imagina que se lanza un dado honesto de seis caras. Se definen los siguientes eventos:

A: El puntaje obtenido es un número mayor que 2.
B: El puntaje obtenido es un número menor que 5.

Calcula la probabilidad de A dado que ha sucedido B.

Ejemplo 3

Si se describe el espacio muestral se tiene:

$\boldsymbol{\Omega} = \{\mathbf{1, 2, 3, 4, 5, 6}\}$
$A = \{ 3, 4, 5, 6 \}$
$B = \{ 1, 2, 3, 4 \}$

Recuerda que el evento A/B corresponde a los elementos de **A dado que ya se sabe que sucedió B.**

Por lo tanto, se tiene A/B={3,4}, ya que se ha ocurrido B, el espacio muestral se reduce a cuatro elementos, de los cuales dos pertenecen a A.

Por lo tanto,

$$P(A/B) = \frac{2}{4} = 0,5$$

Ejemplo 3

Aplicando fórmulas:

$$P(A) = \frac{4}{6} \qquad P(B) = \frac{4}{6} \qquad P(A \cap B) = \frac{2}{6}$$

Por lo tanto,

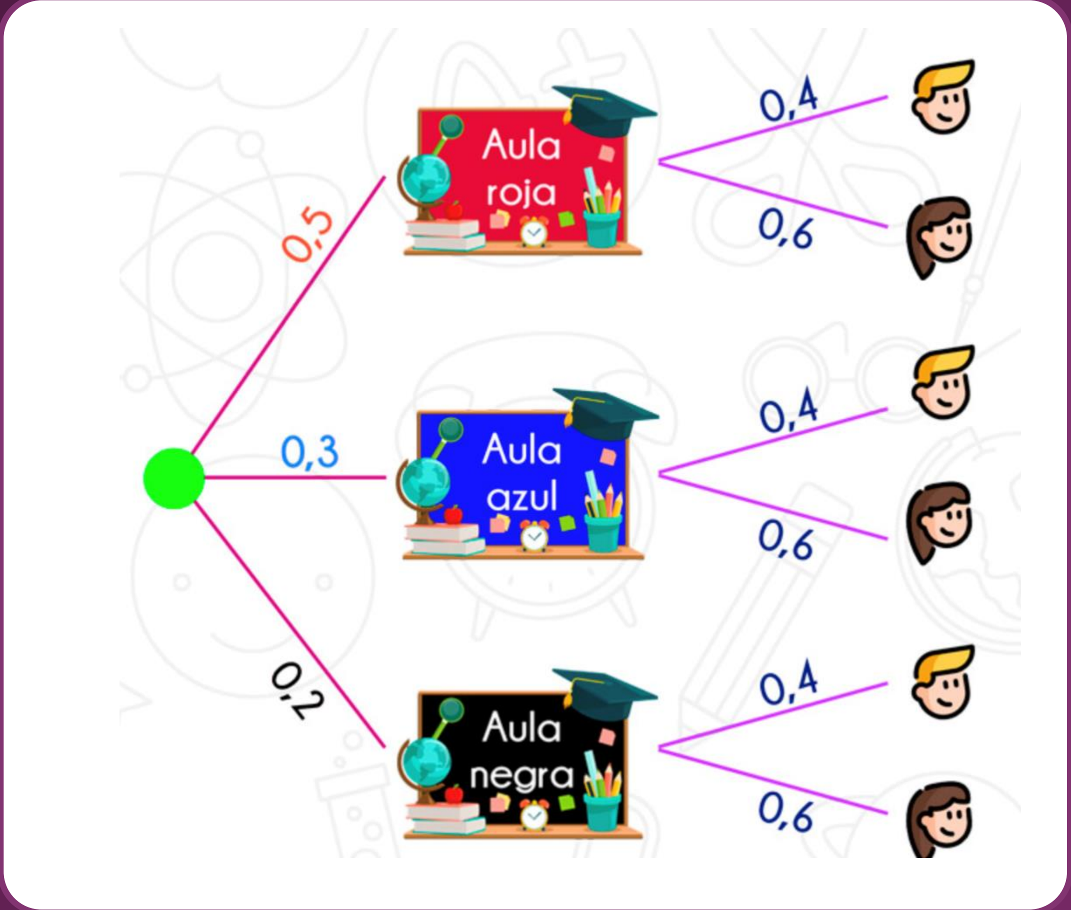
$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{2/6}{4/6} = 0{,}5$$

Con ambos métodos se llega al mismo resultado.

Árbol de Probabilidades

Árbol de Probabilidades

Los diagramas de árbol son muy útiles para el cálculo de probabilidades, sus ramas nos indican los distintos resultados al realizar experimentos sucesivos.



- En una academia hay 3 aulas: el aula roja, el aula azul y el aula negra.

- El aula roja tiene al 50 % de los estudiantes de la academia.
- El aula azul al 30 %.
- El aula negra al 20 %.
- Además, en cada aula hay un 40 % de hombres.

Diagrama de Árbol

Ahora calculamos la probabilidad de que si se selecciona un estudiante al azar, este sea un hombre del aula azul.

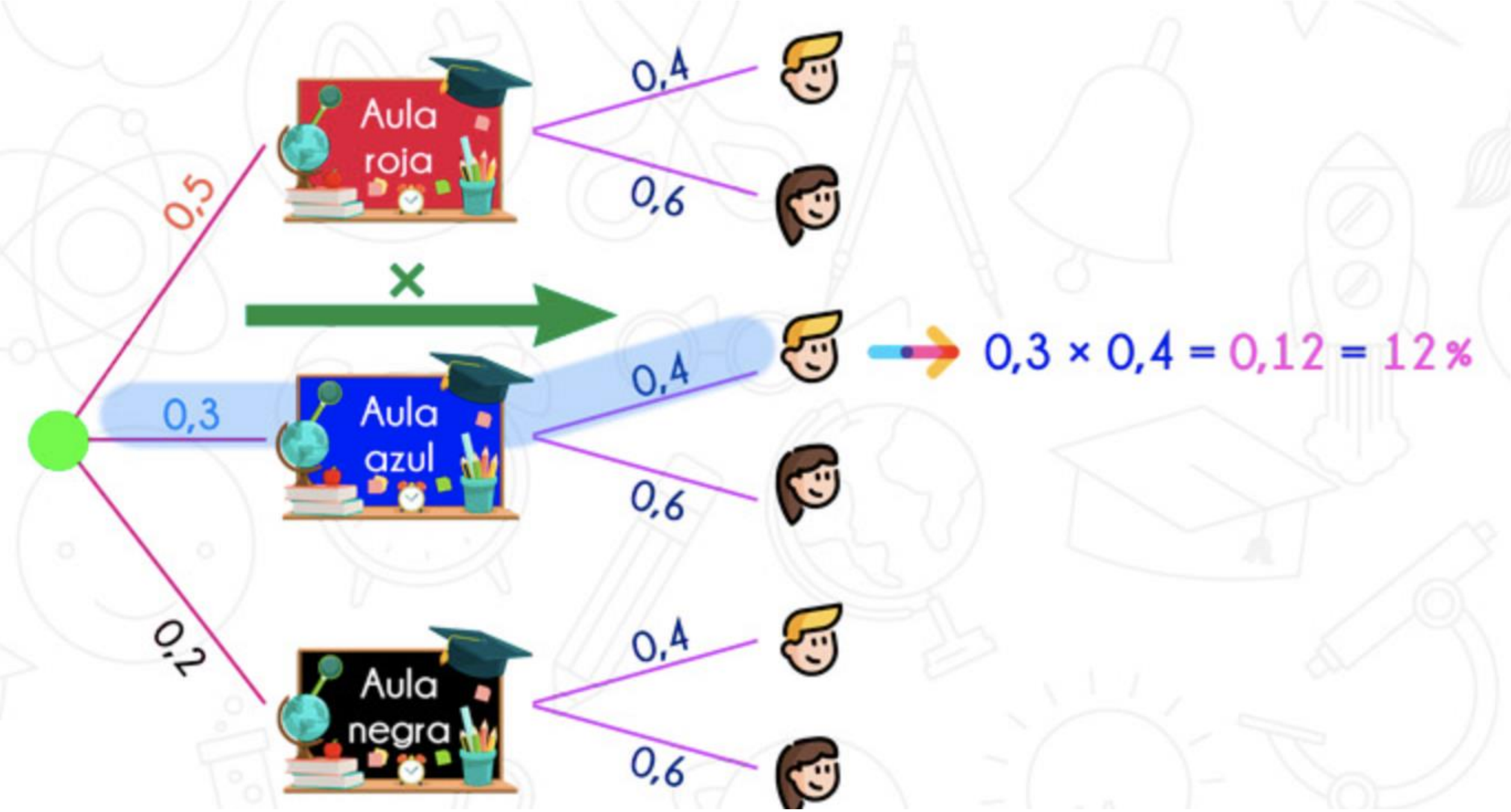
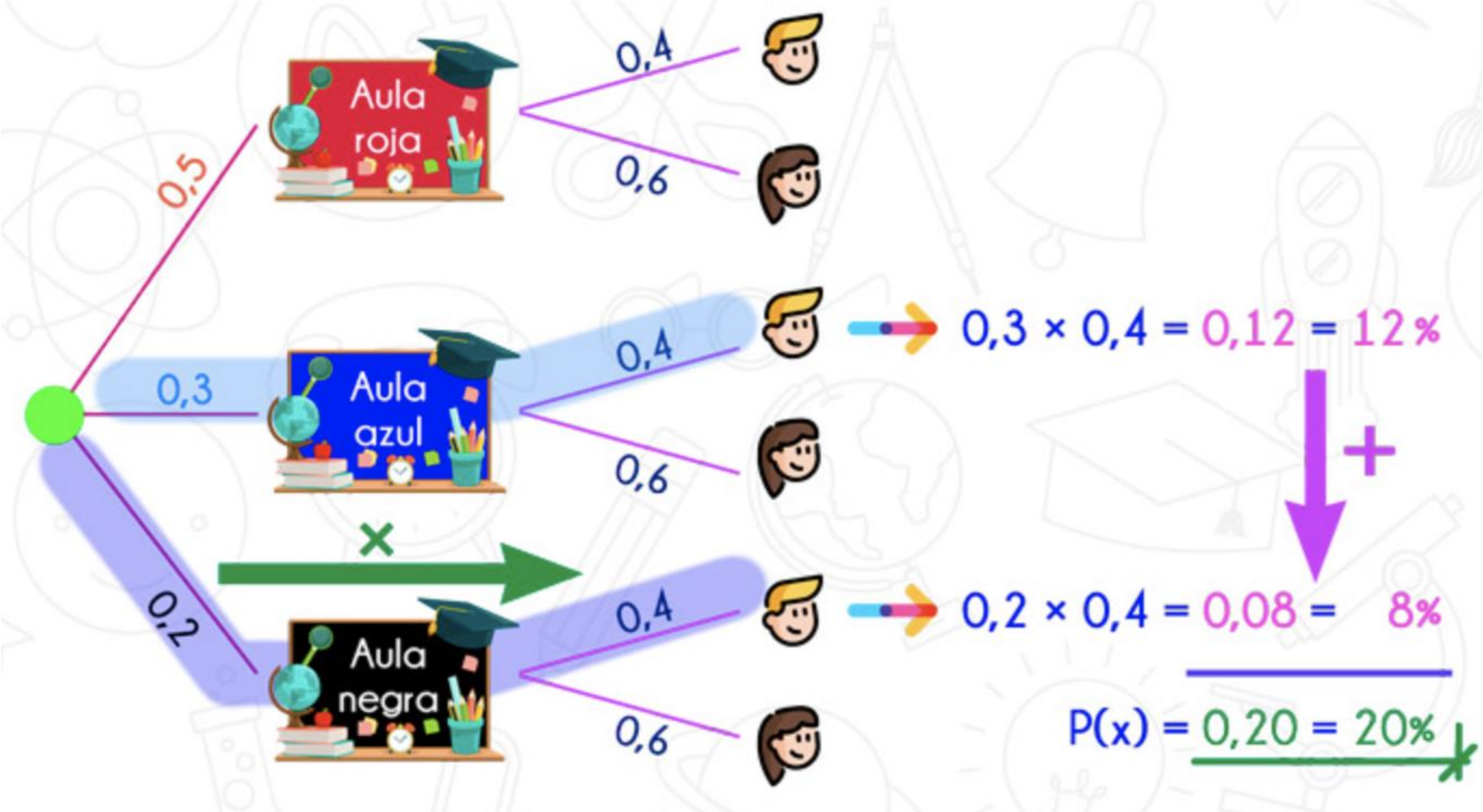


Diagrama de Árbol

Si se selecciona un estudiante al azar, **¿cuál es la probabilidad de que sea un estudiante hombre del aula negra o un hombre del aula azul?**



Recuerda que multiplicamos cuando avanzamos hacia la derecha y sumamos cuando avanzamos hacia abajo.

Dudas y consultas

Fin presentación