Inferencia Estadística

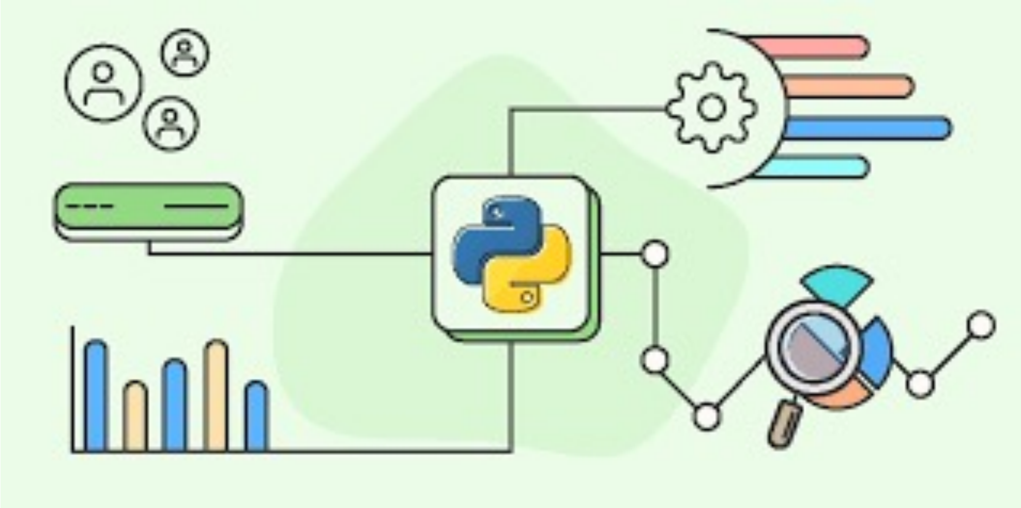
Probabilidad

Especialización en Ciencia de Datos

2021

Objetivos

- Utiliza los conceptos básicos de estadística Inferencial
- Explicar los principales conceptos de probabilidad asociados a un evento aleatorio.



Contenido:

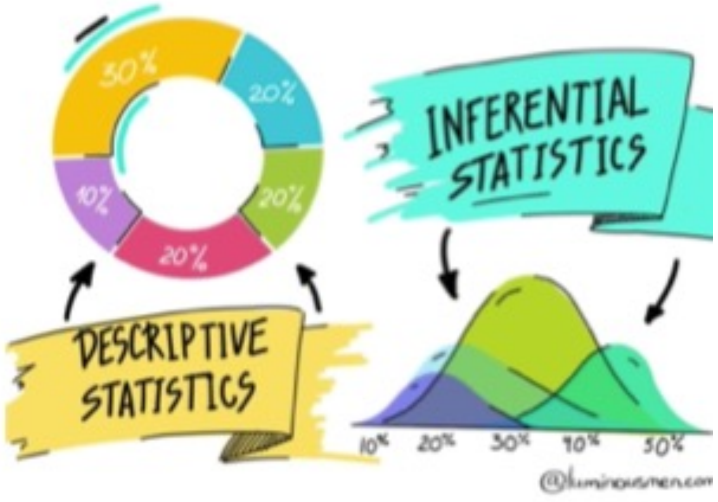
1. Inferencia Estadística
2. Probabilidad: definición
3. Árbol de Probabilidad
4. Árboles de decisión

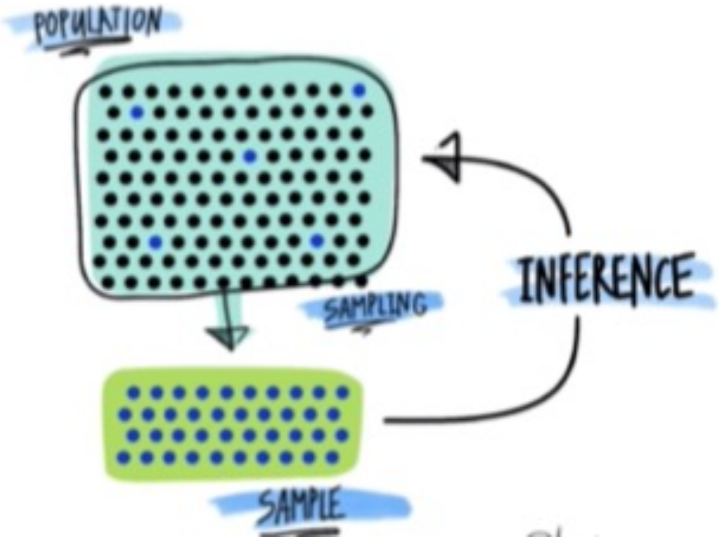
1. Inferencia Estadística

Estadística

- **Estadística Descriptiva:** provee información acerca de los datos que ya hemos obtenido.

- **Inferencia Estadística:** nos permite obtener conclusiones acerca de la población de la cual proviene la muestra de datos.
  - Estimación de parámetros
  - Cálculo de intervalos de confiabilidad
  - Testeo de hipótesis.





Se basan en el conocimiento teórico de la distribución de probabilidad del estadístico muestral que se utiliza como estimador de un parámetro.

Inferencia Estadística

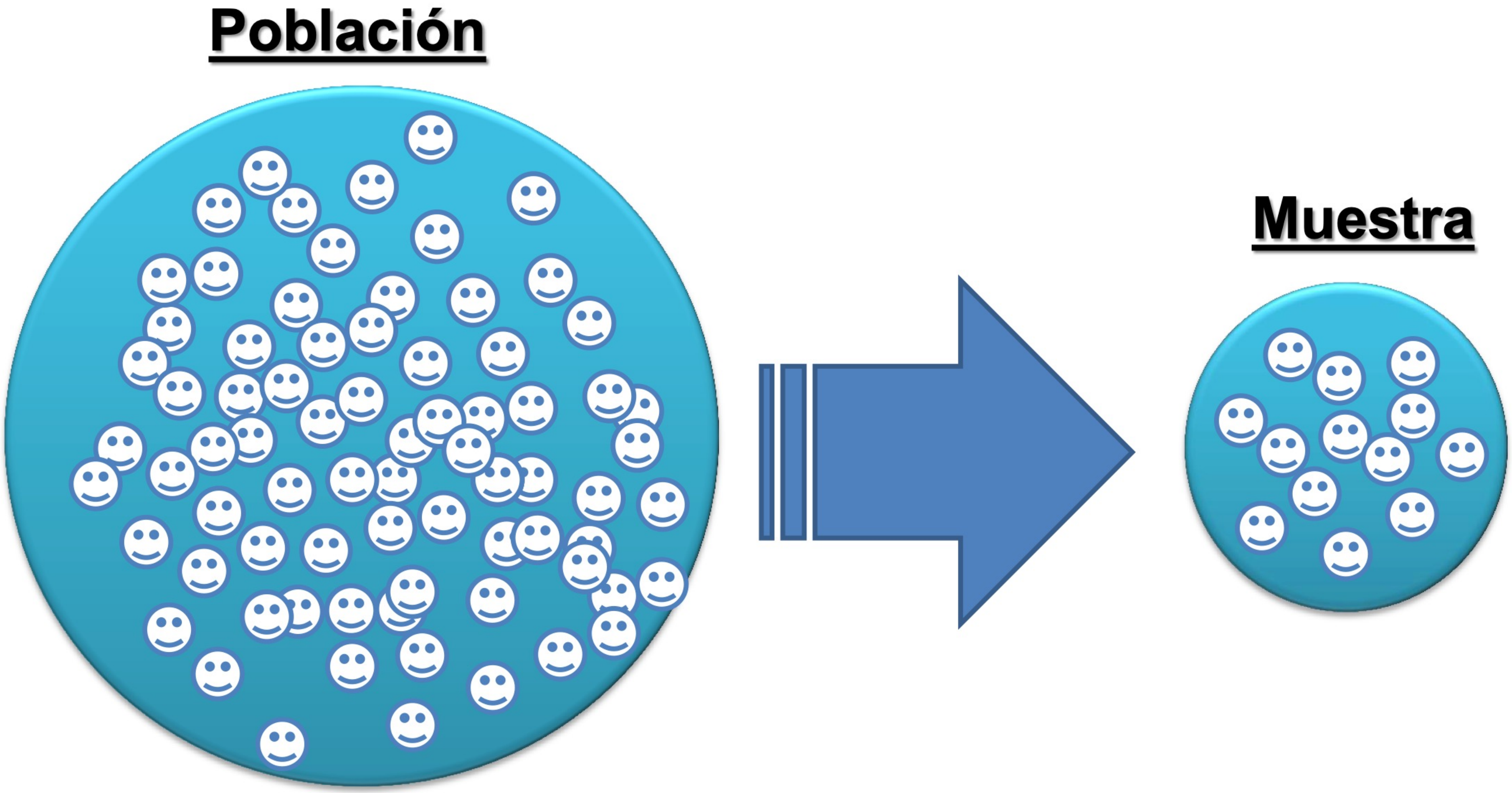
- ¿ Si tenemos un conjunto de mediciones, cuál será el valor promedio del siguiente conjunto de mediciones?

  - No podemos predecir el valor exacto, pero sí describir lo que esperamos en términos probabilísticos.
  - Dado un conjunto de datos, la inferencia estadística nos permite describir probabilísticamente que esperamos para muestras de datos adquiridos una y otra vez, incluyendo una medida de incerteza.

- ¿Cómo repetimos una y otra vez la adquisición de datos?
  - Podemos simularlos computacionalmente, generando "réplicas" o "permutaciones" de nuestras observaciones.
  - Resampleamos los datos una y otra vez, calculamos estadísticas de resumen, parámetros óptimos y sus respectivas distribuciones de probabilidad.

- **Bootstrapping** → uso de datos remuestreados para realizar inferencia estadística sobre una población.

Muestra poblacional



2. Probabilidad

Probabilidad

Se denota probabilidad de un suceso A como:

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

Para determinar esta probabilidad tenemos varios enfoques:

- Interpretación frecuentista.
- Interpretación clásica.

Probabilidad

- *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.

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

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.

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

Conceptos de Probabilidad

- *Experimento determinístico*

Es un experimento que se tiene certeza del resultado que se obtendrá al realizarlo varias veces bajo las mismas condiciones.

- *Experimento aleatorio*

Es un experimento que si se repite una cierta cantidad de veces, bajo las mismas condiciones, no se tiene certeza del resultado que se obtendrá, es decir, depende del azar.
Ej: lanzar un moneda

- *Espacio muestral y sucesos:*

El conjunto formado por todos los posibles resultados de un experimento. Cada uno de estos resultados es conocido como evento o suceso elemental

Probabilidad

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

Ejemplo: Lanzamiento de un dado.

Suceso A:Obtener un numero par

- Un suceso es seguro cuando coincide con el espacio muestral del experimento.
- Un suceso es imposible si nunca ocurre o sino se presenta al realizar un experimento aleatorio.
- Dos sucesos son mutuamente excluyentes o disjuntos si no pueden suceder simultáneamente, es decir, si y solo 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 $\#\Omega$ es la cantidad de resultados posibles de un experimento aleatorio.

Probabilidad

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

- ¿Cual es la probabilidad de extraer un numero par?
- ¿Cuál es la probabilidad de extraer un múltiplo de 3?

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

a) Casos favorables $A = \{2,4,6,8,10,12,14,16,18,20\}$ $\#A = 10$

$$P(A) = \frac{\#A}{\#\Omega}$$

$$P(A) = \frac{10}{20} = \frac{1}{2}$$

b) Casos favorables $B = \{3,6,9,12,15,18\}$ $\#B = 6$

$$P(B) = \frac{6}{20}$$

$$P(B) = \frac{3}{10}$$

Probabilidad

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	Numero 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) Cual es la probabilidad que un alumno no juegue tenis.

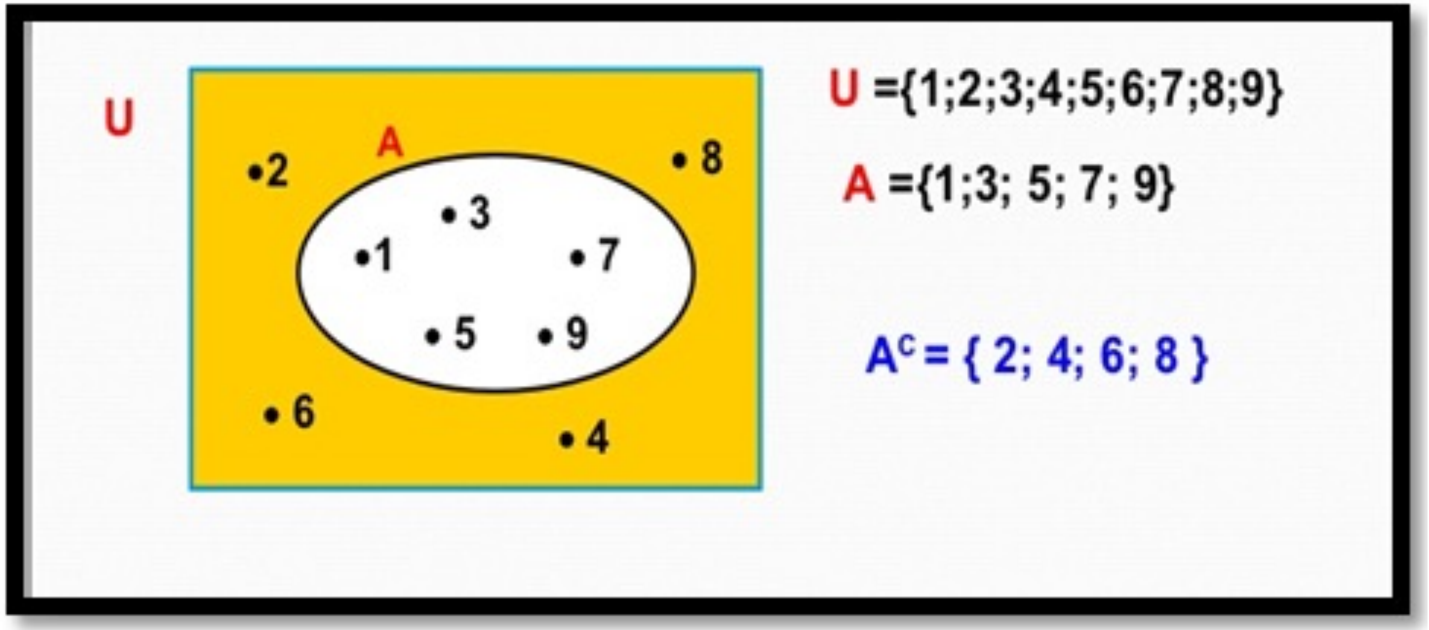
Propiedades de la Probabilidad

Conceptos básicos sobre conjuntos:

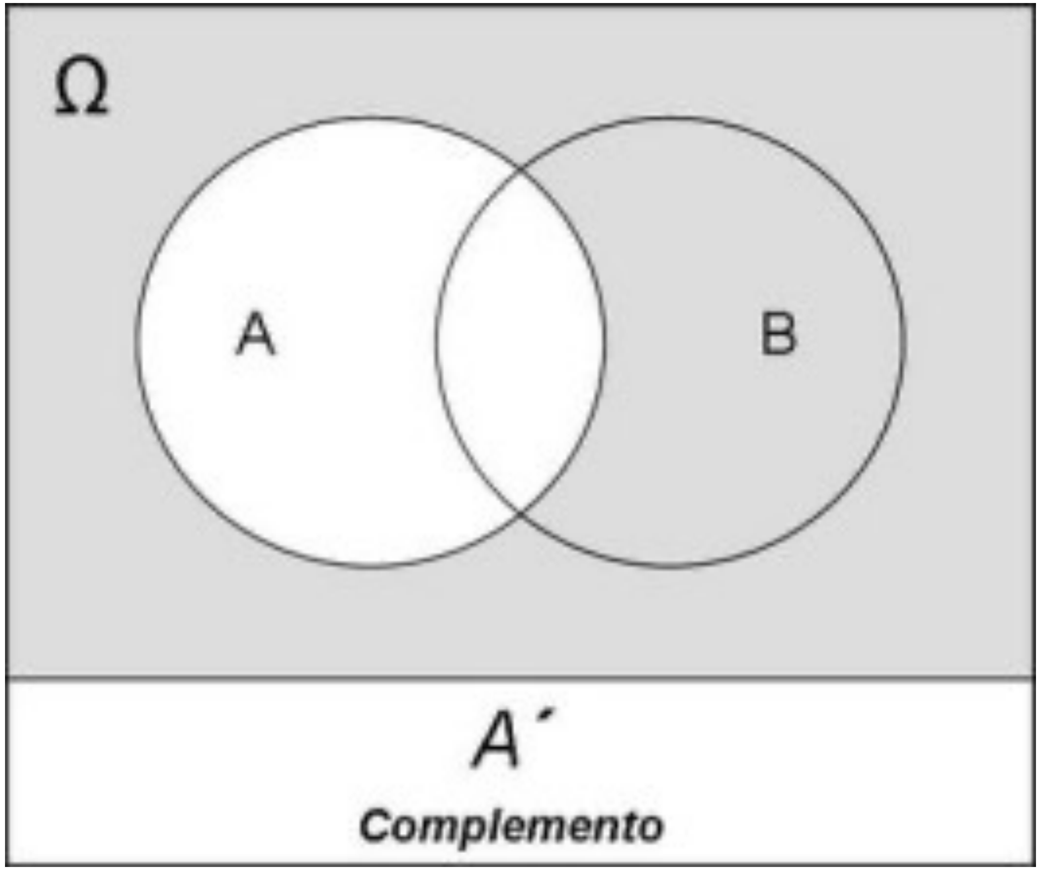
Conjunto: Es una colección de elementos. Se denotan con letras mayúsculas. El conjunto formado por todos los elementos. En probabilidad el universo es el espacio muestral(Ω).

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



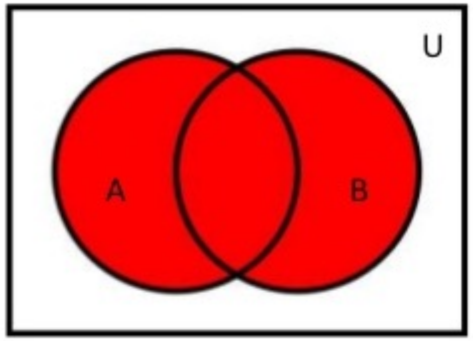
TODO LOS ELEMENTOS QUE NO ESTAN EN A

Propiedades de la Probabilidad

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

- Unión.

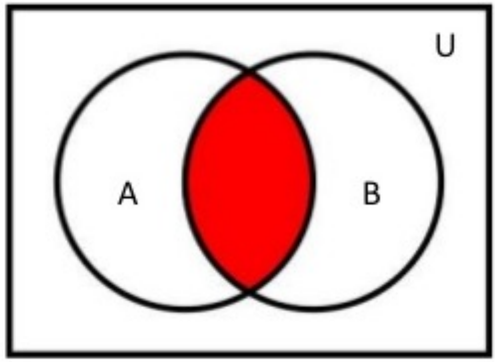
$$A \bigcup B$$



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

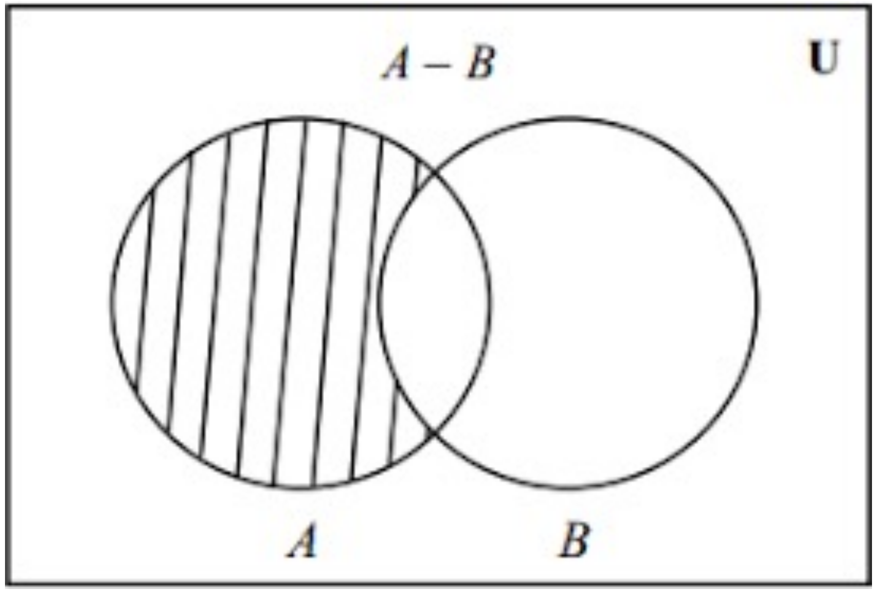
- Intersección

$$A \bigcap B$$



Propiedades de la Probabilidad

<u>Diferencia entre A y B</u>: Corresponde a los elementos que están en A y que no están en B.



REGLA DE LA ADICION
Imagina experimento lanzar un dado

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

Se definimos dos sucesos

A: Obtener un numero menor que dos

A=

B: Obtener un numero par

B=

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

Es decir al ser disjuntos o mutuamente excluyentes, se tiene

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

REGLA DE LA ADICION
Imagina experimento lanzar un dado

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

Se definimos dos sucesos

A: Obtener un numero par

A=

B: Obtener un numero primo

B=

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

TIENEN UN ELEMENTO QUE COMPARTEN, QUE TENGAN EN COMUN A Y B

Es decir , se tiene

P(A U B) = P(A) + P(B) – $P(A \cap B)$

RESUMEN

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.

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

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:

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

REGLA MULTIPLICATIVA DE LA PROBABILIDAD

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

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

donde $P(A/B)$ corresponde al 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 $PA\ /\ B) = P(A)$.*

ESTO SE CONOCE COMO LA REGLA MULTIPLICATIVAS DE LAS PROBABILIDAD

SE UTILIZA PARA CALCULAR PROBABILIDADES DE EVENTOS COMPUESTOS, ES DECIR 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 extraes dos bolas, cual es la probabilidad de extraer una bola azul y una bola verde

A) *CON REPOSICION DE BOLAS*.(*Es decir sacas una y la vuelves a poner dentro de la bolsa para la segunda extracción)*

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

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

$P(A)$ es igual a 5 casos favorables dividido en 12 bolitas en total

$P(B)$ es igual a tres casos favorables dividido en 12 bolitas en total

Por lo tanto

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

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

B) ***<u>SIN REPOSICION DE BOLITAS (</u>****Es decir se saca una bolita y no se devuelve a la bolsa para realizar la segunda extracción)*

Ahora no se devuelva la primera bolita extraída, por lo tanto lo que sucede en la primera extracción me afecta el numero 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.

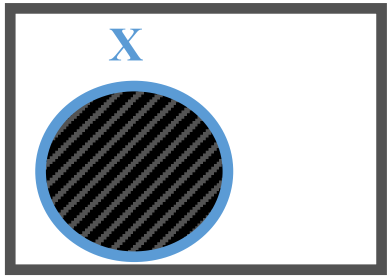
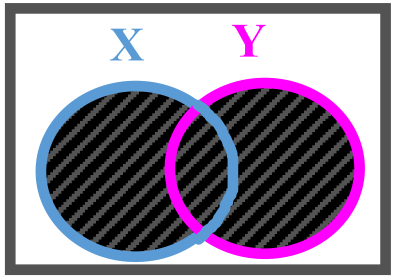
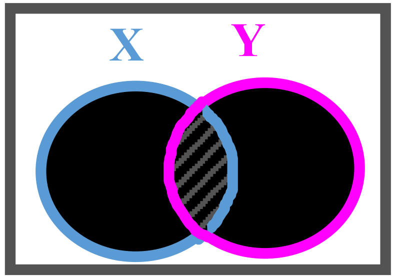
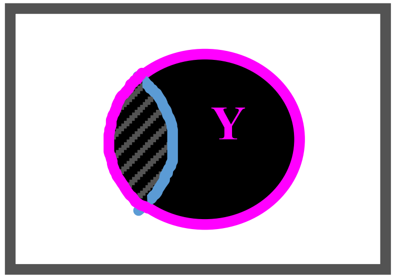
Por lo tanto

P(A)=$\frac{5}{12}$ y como no devuelvo la bolita, ahora tengo 11 casos posibles para la segunda extracción. Por esto

P(B/A)= $\frac{3}{11}$

$$P(A \cap B) = \frac{5}{12} \cdot \frac{3}{11} = \frac{15}{132}$$

Tipos de Probabilidades

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

Tipos de Probabilidades

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

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

Tipos de Probabilidades

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

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

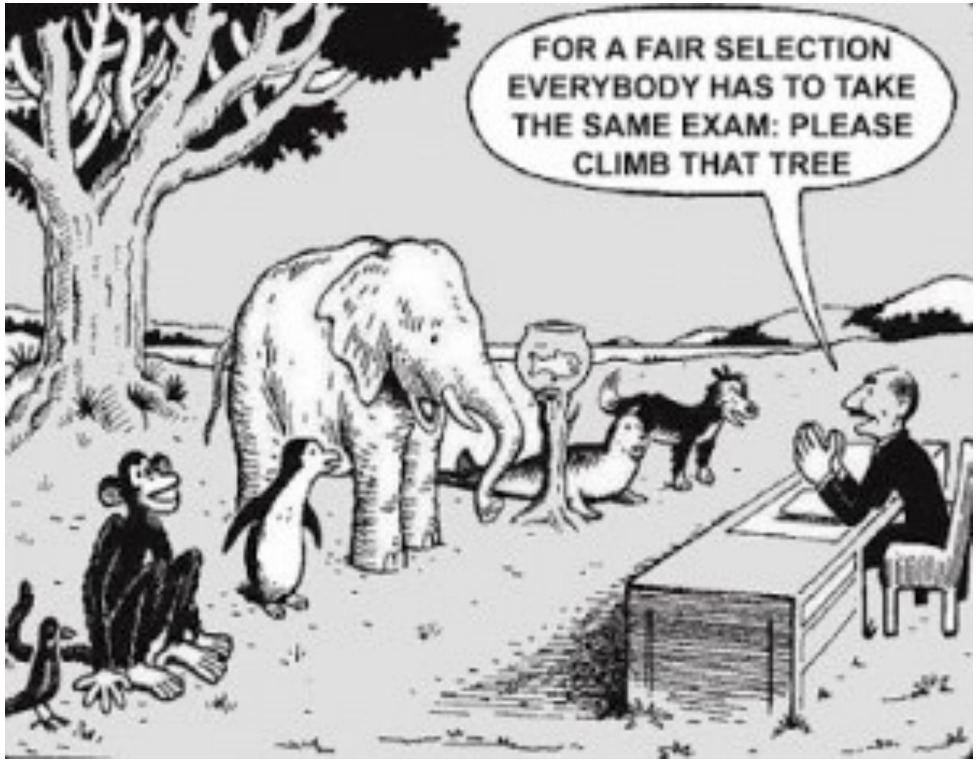
¿Cuál es la probabilidad de que una carta escogida al azar sea un as sabiendo que es roja?

Carta	Color		Total
	Rojo	Negro	
As	2	2	4
No-As	24	24	48
Total	26	26	52

Espacio restringido

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

¿Qué resultado te daría si te piden calcular al revés? La probabilidad de que sea roja, sabiendo que es un as es: (resuélvelo)



[redacted] 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 resuelto

1. 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:

Tabla de contingencia

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

a) 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.

Solución:

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**

La probabilidad condicional pedida es:

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

Matemáticamente se traduce en:

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

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

EJEMPLO 1

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

SOLUCION

Si describes el espacio muestral se tiene

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

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)=$\frac{2}{4} = \frac{1}{2}$

USANDO FORMULAS

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

Aplicas la formula P(A/B)=$\frac{P(A \cap B)}{P(B)}$=$\frac{2}{6} \div \frac{4}{6} = \frac{2}{4}$

CON AMBOS MÉTODOS LLEGAS A LO MISMO.

3. 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 % y el aula negra al 20 %. Además, en cada aula hay un 40 % de hombres.

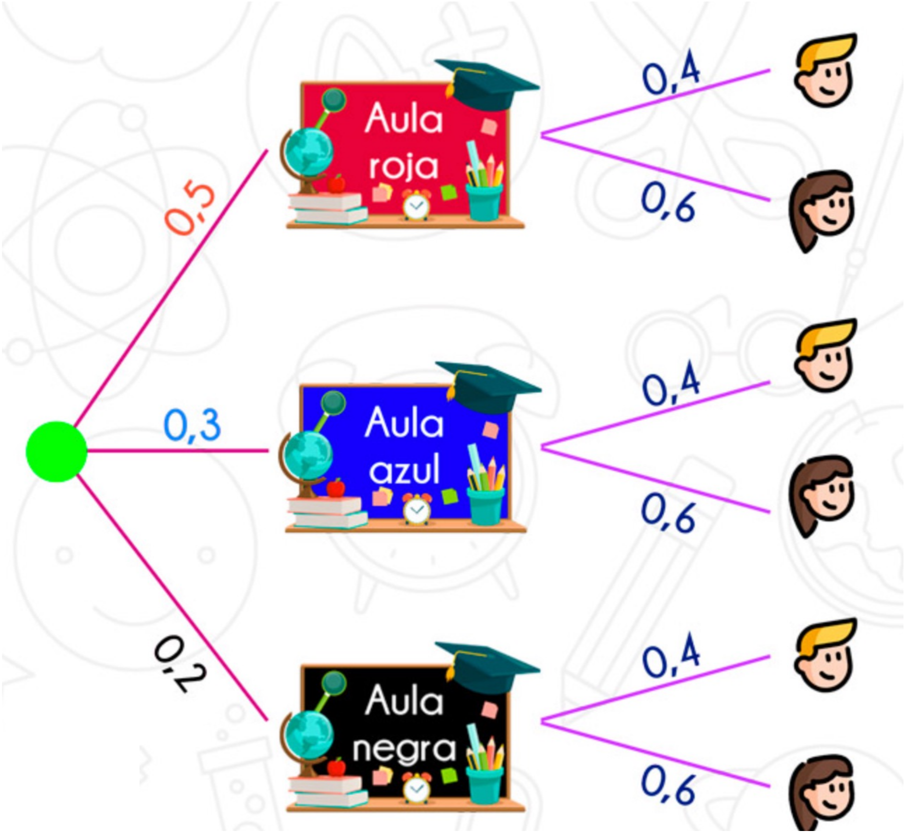


Diagrama de árbol

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

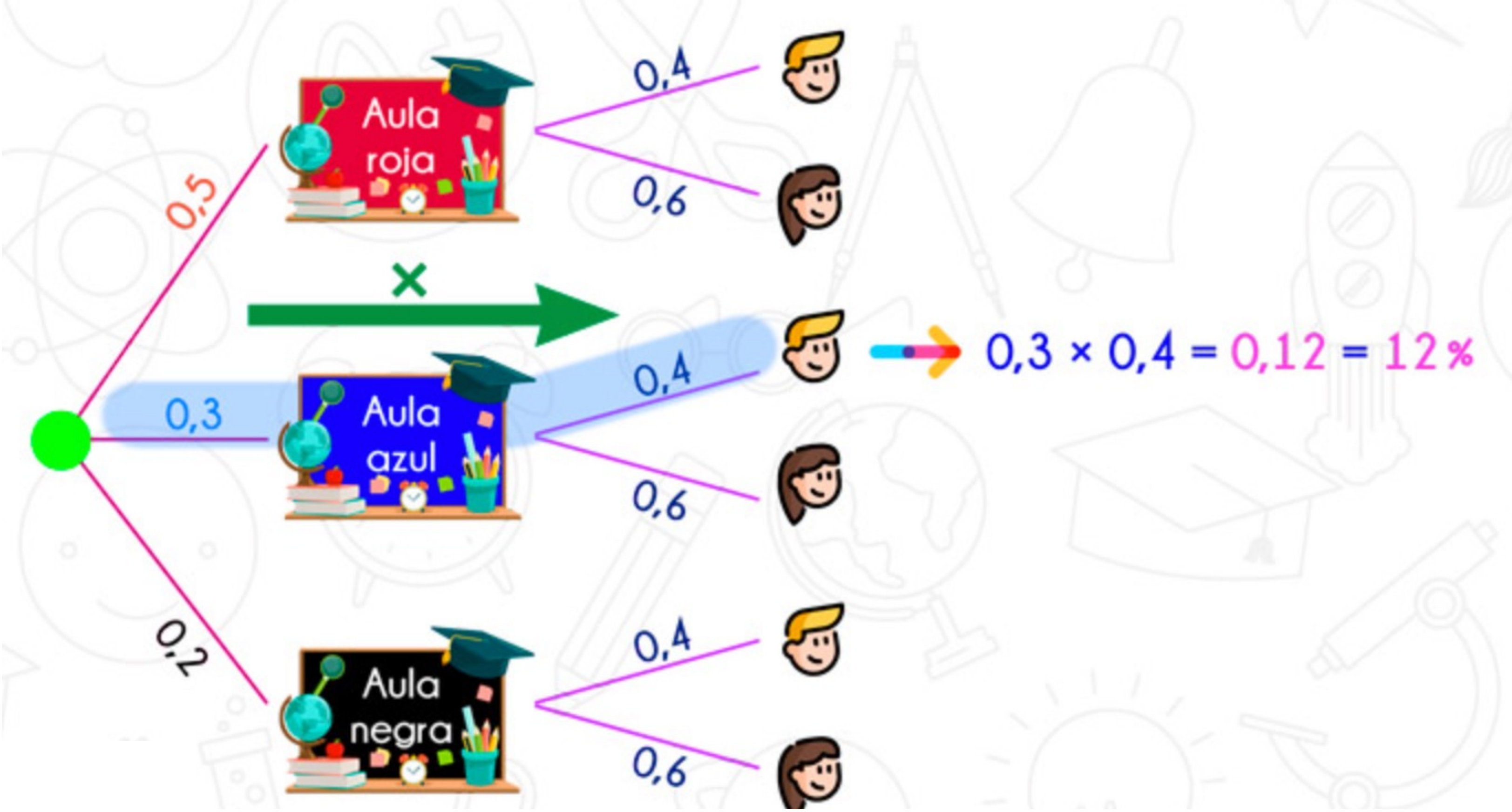
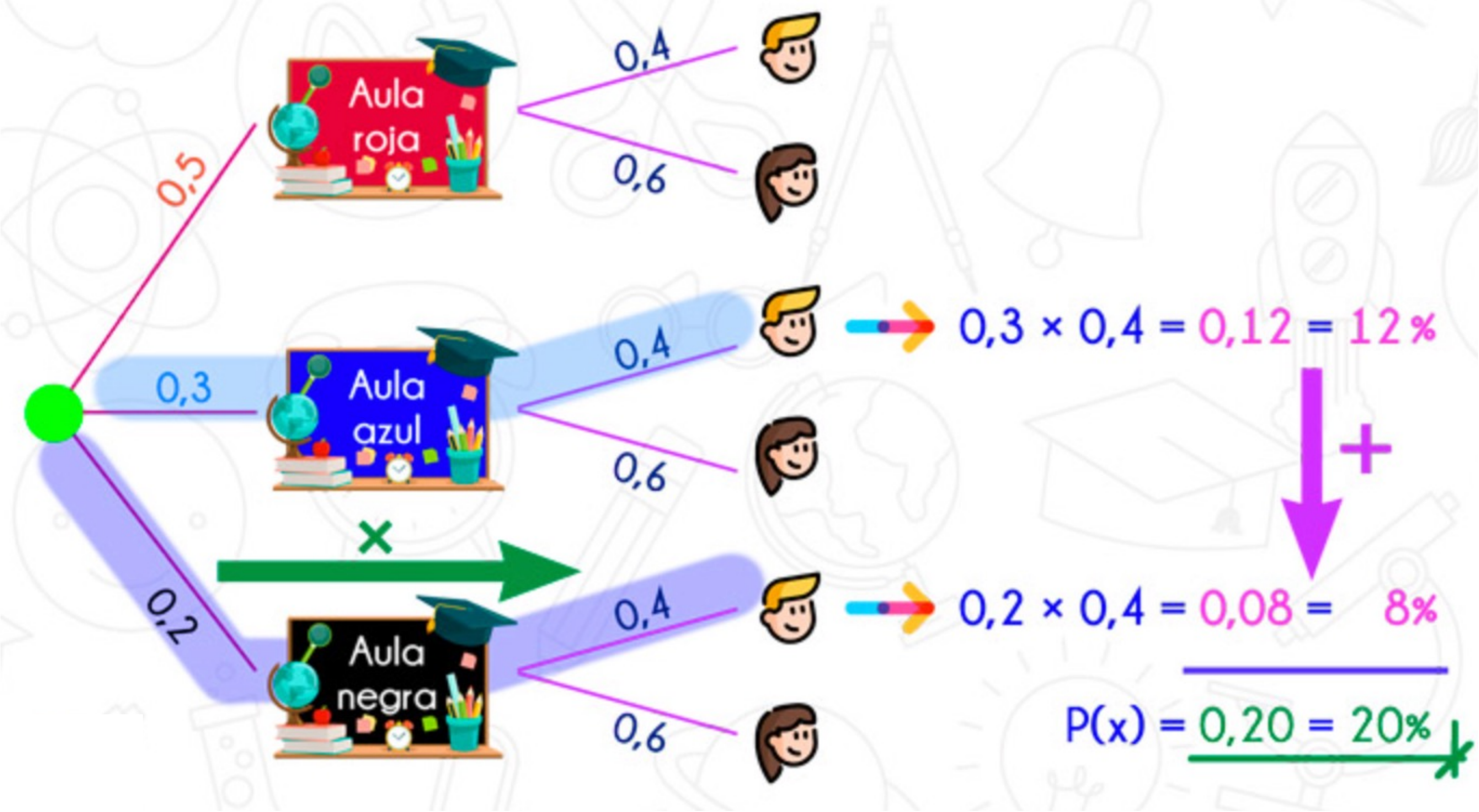


Diagrama de árbol

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.

A Practicar!!!!

KIBERNUM
Creando Juntos Excelencia

Software Engineering | IT Staffing | IT Academy | IT Consulting

Proyecto apoyado por CORFO

Gobierno de Chile