

Módulo: Fundamentos de probabilidad y estadística

DATA
ANALÍTICA



Agenda

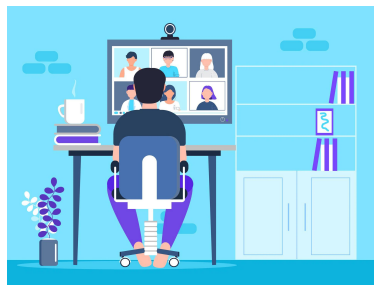
- 0. Reglas del Juego
- 1. Introducción a la estadística
- 2. Elementos de la estadística
- 3. Introducción a la probabilidad
- 4. Variables aleatorias
- 5. Distribuciones de probabilidad
 - 5.1. Distribuciones discretas
 - 5.2. Distribuciones continuas

Reglas del Juego

**Mantener el micrófono
apagado en caso de que
no vayamos a hablar.**



**Nos encantaría verte.
Ten tu cámara encendida y
conozcámonos
virtualmente.**



**Preguntar en caso que
tengan dudas.**



**Disfruta de este espacio.
Desconecta del resto y
participa.**



**Por cada clase tendremos
10 min o 15 min de receso.**





INTRODUCCIÓN A LA ESTADÍSTICA

¿Cómo podemos definir la estadística?

Ciencia aplicada que nos proporciona un conjunto de técnicas o procedimientos para **recopilar, organizar, analizar y presentar datos** con el fin de describirlos o de realizar generalizaciones válidas

“Es el **arte** de analizar los datos utilizando **técnicas matemáticas** para resolver problemas del mundo real”

Detección de correo spam

Pronóstico del tipo de cambio

Estimación de ingresos

Intención de voto en elecciones

Pronósticos de temperatura

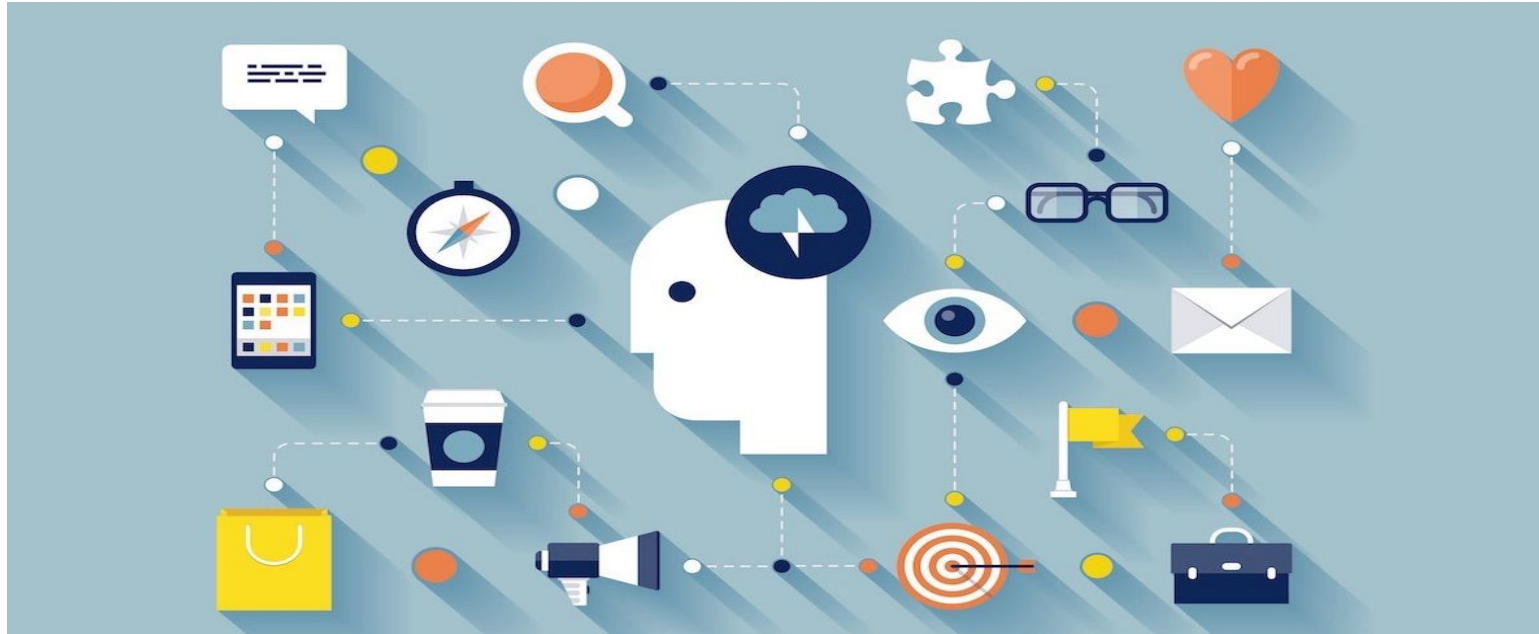
Detección de objetos

Detectar clientes propensos a mora

Pronósticos del ganador de un partido

Reducción de cantidad de variables

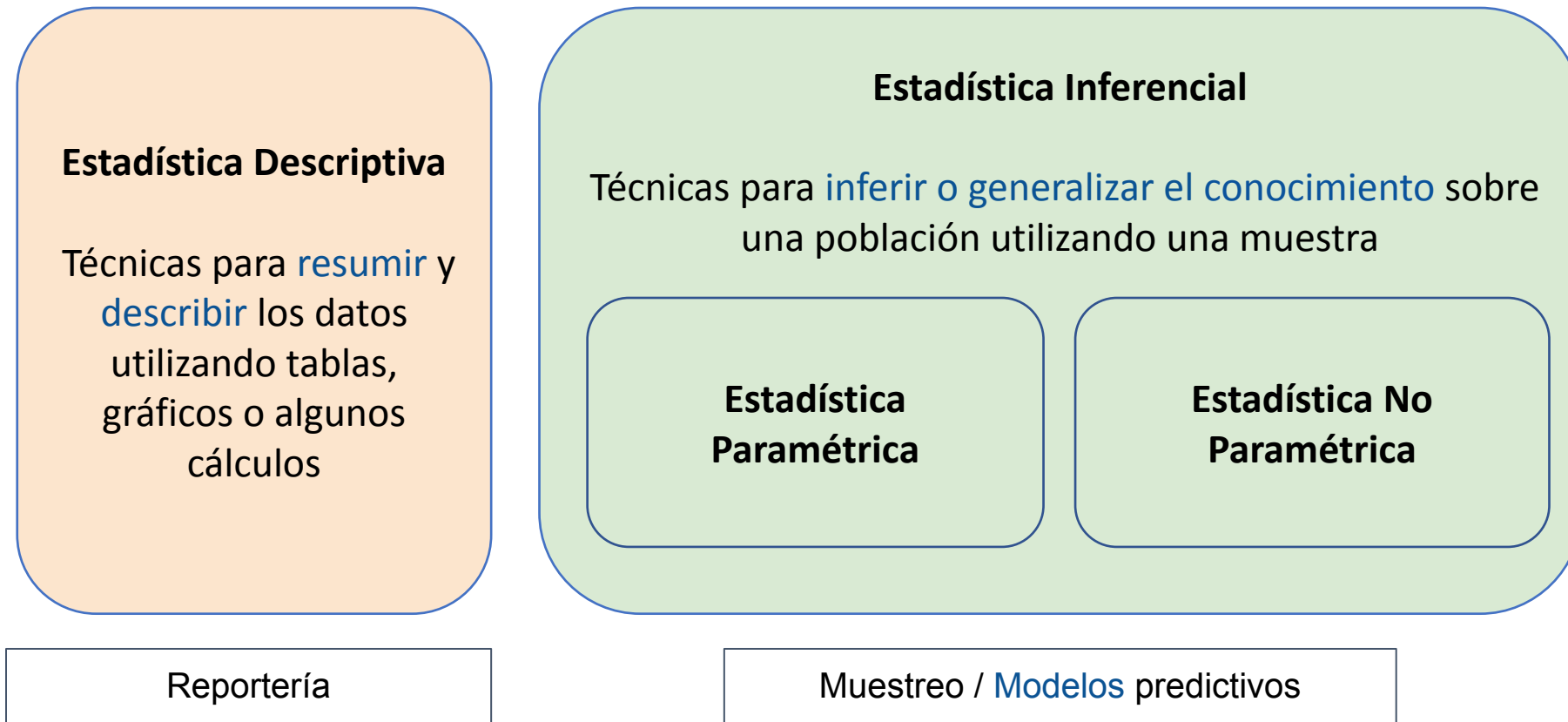
Para nuestra curiosidad...



[La Historia y aplicación de la Estadística - YouTube](#)

¿Cómo se clasifica o divide la estadística?

La estadística se clasifica en 2 grandes ramas:



Nota: Un modelo es una simplificación de la realidad mediante una representación matemática



ELEMENTOS DE LA ESTADÍSTICA

¿Cuáles son los elementos de la estadística?

- **Población:** Conjunto de elementos (personas, objetos, etc) sobre el cual se quiere obtener información observando o midiendo una o más características.

Empleados de una empresa del área de Data



Características o variables:

- Estado civil
- Grado instrucción
- Género
- Lugar donde vive
- Edad
- Ingresos
- Cantidad de hijos

OJO: Una población puede ser **finita** (existen N elementos) o puede ser **infinita** (en la práctica se refiere cuando el número de elemento es muy grande)

¿Cuáles son los elementos de la estadística?

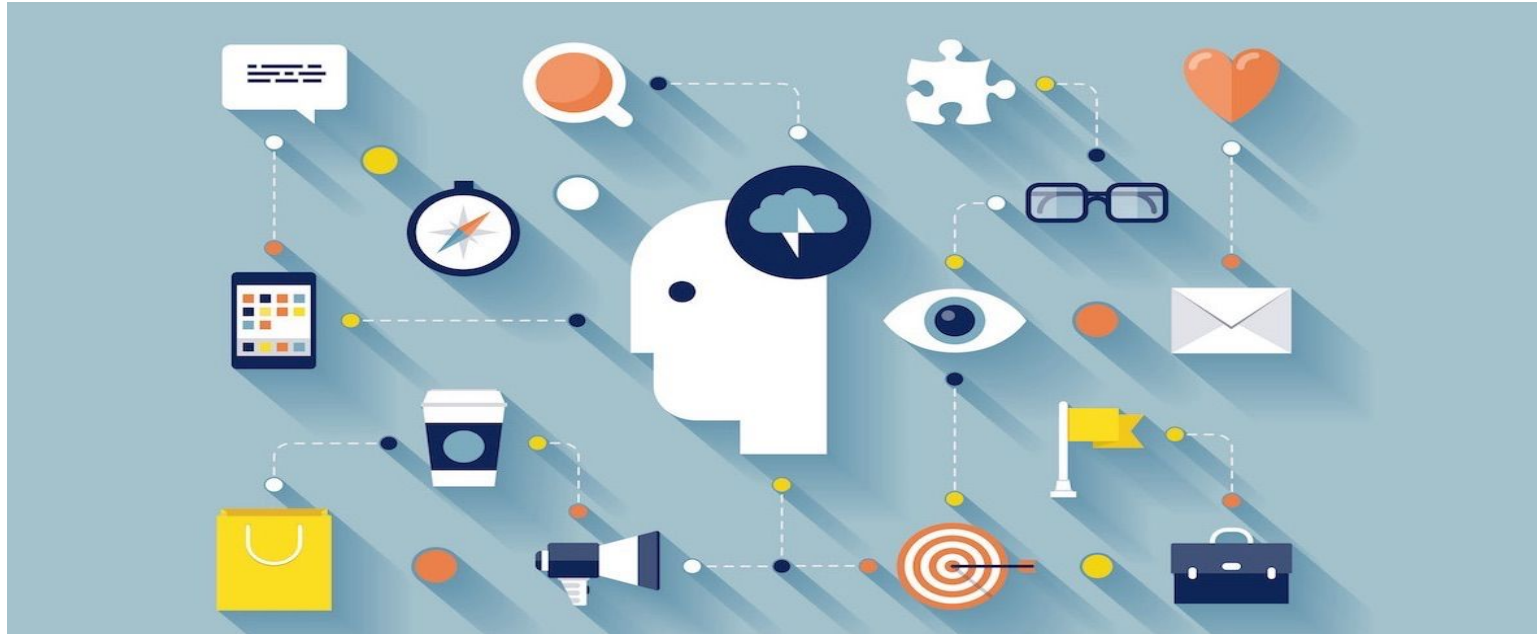
- **Muestra:** Es una parte o subconjunto de la población, el cual debe ser **representativa** y seleccionarse con **técnicas de muestreo**.

Empleados de una empresa del área de Data



OJO: Las técnicas de muestreo utilizan la **aleatoriedad**, esto significa que cada elemento de la población **tiene la misma posibilidad o probabilidad** de ser seleccionado. (**Muestreo probabilístico: Aleatorio simple**)

Para nuestra curiosidad...



[Técnicas de muestreo - YouTube](#)

¿Cuáles son los elementos de la estadística?

• **Parámetro**

Es una medida descriptiva que resume una característica de la **población**

Promedio poblacional (μ)

Varianza poblacional (σ^2)

Desconocido


• **Estadístico**

Es una medida descriptiva que resume una característica de la **muestra**

Promedio muestral (\bar{X})

Varianza muestral (s^2)

Conocido



Utilizamos los estadísticos de la muestra para estimar los parámetros de la población

Ejemplo de muestreo: Intención de voto en elecciones

Peruanos del padrón electoral



Parámetro: % de votos de cada candidato

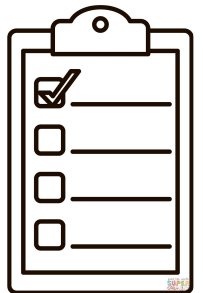
Muestra de votantes



Estadístico: % de votos de cada candidato

1er paso

2do paso



Ejemplo de muestreo: Modelos en Banca

Los bancos construyen distintos modelos estadísticos (admisión, **comportamental**, cobranzas, etc).

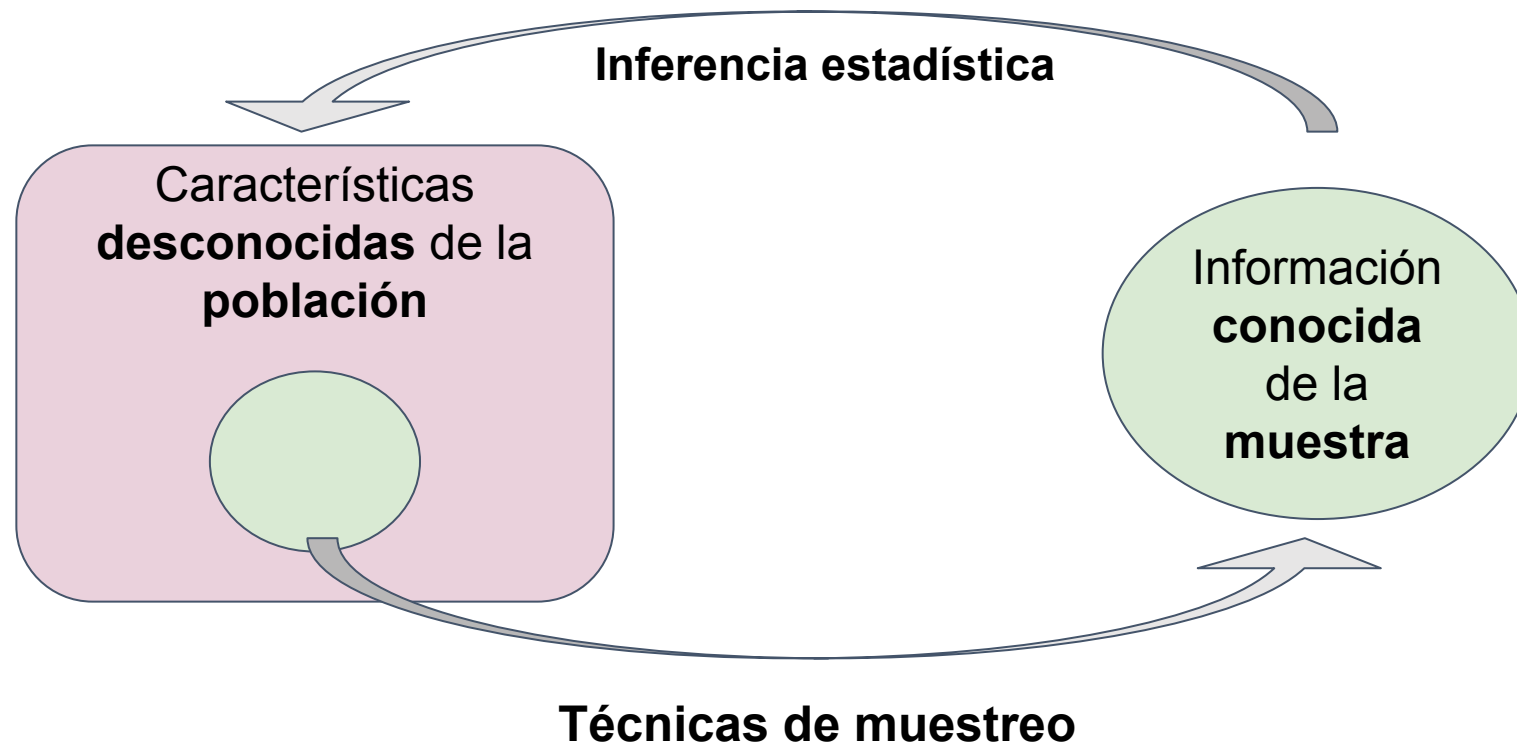
¿Cómo construyen un modelo de riesgo de crédito?

$$Y = f(X1, X2, X3, \dots)$$

- Y = Cliente cae en mora (deja de pagar su deuda)
- $X1, X2, X3, \dots$ = Características del cliente
- Población = Clientes del banco desde Enero 2016 hasta Febrero 2018
- Muestra = Aleatoriamente seleccionar el 70% u 80% de la población para la construcción del modelo analítico.

Para recordarlo siempre...

Ciclo básico de un análisis estadístico



Nota: Para realizar inferencia estadística se debe conocer sobre probabilidad

QUIZ TIME



INTRODUCCIÓN A LA PROBABILIDAD

¿Dónde encontramos el concepto de probabilidad?

¿Podré ganarme la Tinka algún día?

¿Las *chances* de que me quiera son altas?

¿Podré aprobar mi examen de admisión?

¿Cuán *probable* es que un cliente no pague su deuda?

¿Cuál es la *probabilidad* de que llueva mañana?

¿Cuál es la *probabilidad* de que Brasil gane el mundial?



Conceptos preliminares

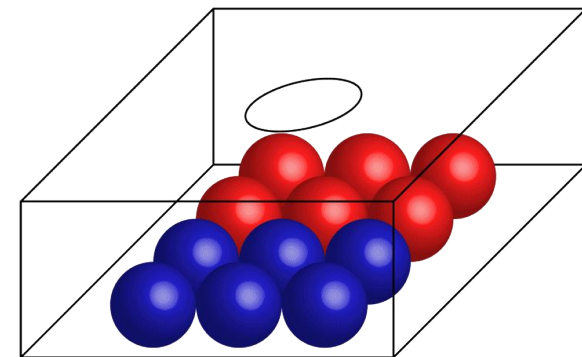
Para entender el estudio de la probabilidad existen *palabras claves* como: experimento aleatorio, espacio muestral, evento y probabilidad de un evento.

- **Experimento aleatorio:** Es un proceso que produce *diferentes resultados* a pesar de que se repita varias veces *bajo las mismas condiciones*.

Lanzamiento de un dado



Sacar una bola de una urna



Conceptos preliminares

- **Espacio muestral:** Son todos los posibles resultados de un experimento aleatorio.

Lanzamiento de un dado

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

Sacar una bola de una urna

$$S = \{\text{rojo}, \text{azul}\}$$

- **Evento o suceso:** Es un subconjunto del espacio muestral, para cada evento se define y se calcula su probabilidad.

Lanzamiento de un dado

$$A = \{\text{Obtener el número 2}\}$$

$$B = \{\text{Obtener un número par}\}$$

Sacar una bola de una urna

$$A = \{\text{Obtener una bola azul}\}$$

Conceptos preliminares

- **Probabilidad de un evento:** Indica la chance o posibilidad que el evento ocurra o suceda. Se denota por:

$$0 \leq P(A) \leq 1$$

- Cuando $P(A)$ se aproxima a 1 entonces se dice que el evento es muy o altamente probable
- Cuando $P(A)$ es igual a 1 entonces se dice que existe un certeza total o que siempre ocurre
- Cuando $P(A)$ se aproxima a 0 entonces se dice que el evento es poco probable
- Cuando $P(a)$ es igual a 0 entonces se dice que el evento es nulo o que nunca ocurre



Definiciones de probabilidad

- **Enfoque clásico:** La probabilidad se calcula como la división entre la cantidad de resultados favorables y la cantidad de elementos del espacio muestral (resultados posibles).

Lanzamiento de un dado

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

$A = \{\text{Obtener el número 2}\}$

$$P(A) = \frac{N(\{2\})}{N(S)} = \frac{1}{6} = 0.17$$

$N(\{2\}) = 1$ debido a que solo hay una cara con el número 2

$B = \{\text{Obtener un número par}\}$

$$P(B) = \frac{N(\{2, 4, 6\})}{N(S)} = \frac{3}{6} = 0.5$$

Definiciones de probabilidad

- **Enfoque de frecuencia relativa:** Repitiendo el experimento aleatorio muchísimas veces, la probabilidad se calcula como la división entre la **cantidad de ocurrencias del evento y la cantidad de repeticiones o simulaciones**.

Lanzamiento de un dado

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

$$A = \{\text{Obtener el número 2}\}$$



1. **Simular** el lanzamiento del dado unas 1000 veces y anotar los resultados.

2. **Contar** la cantidad de veces que ocurrió el número 2

3. **Dividir**

$$P(A) = \frac{174}{1000} = 0.174$$

Solución en excel



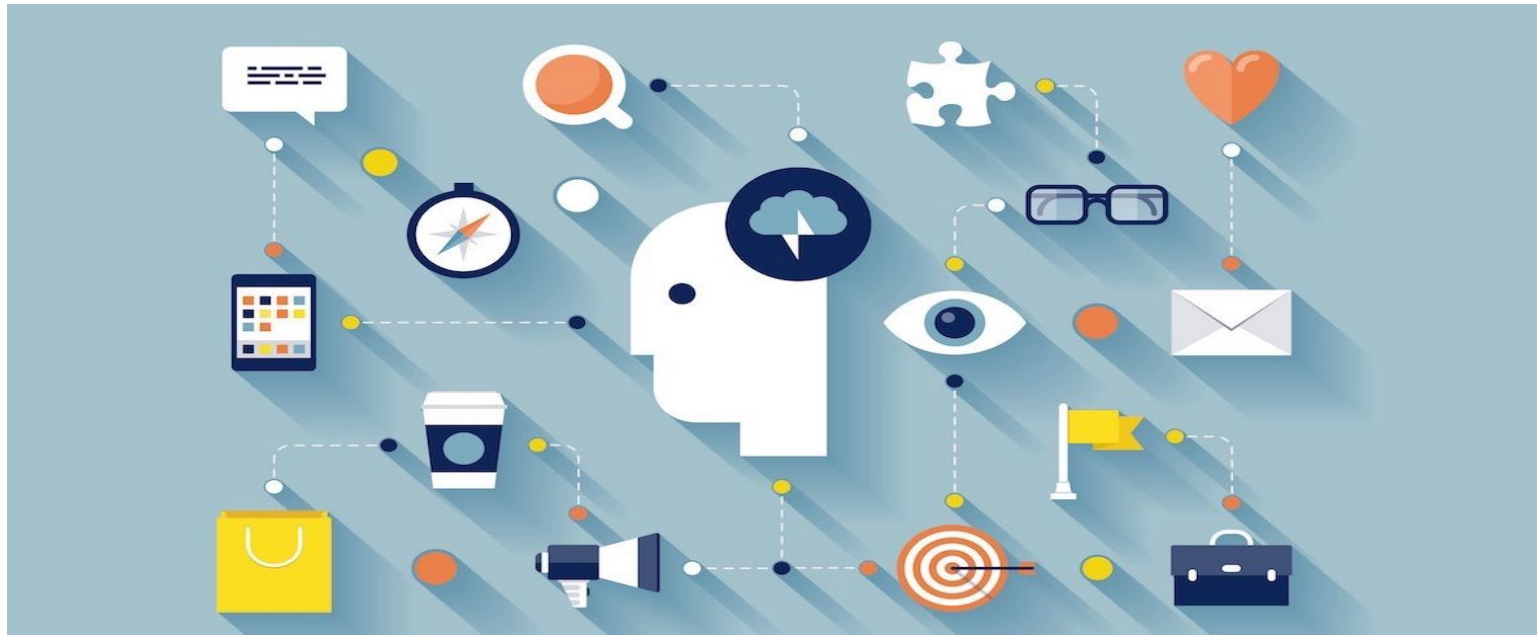
Para nuestra curiosidad...

Reto:

Calcular la probabilidad, según el *enfoque de frecuencia relativa*, de obtener un número par en el lanzamiento de un dado.



Para nuestra curiosidad...



[Experimento aleatorio, espacio muestral y evento o suceso - YouTube](#)

[Introducción a la Probabilidad Prof. Ronny Vallejos - YouTube](#)

QUIZ TIME



Challenge 6:

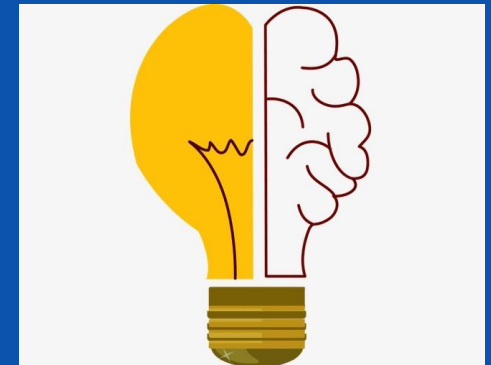
Resolver los ejercicios propuesto sobre la 1era parte de Introducción a la estadística y probabilidad.

Enviarlo por correo con el asunto: Challenge 6 – Fundamentos de Probabilidad y Estadística – [Apellidos y nombres]

Correo: team@dataanalitica.net

Referencia: [Probabilidades - Ejercicios Resueltos - Nivel 1A - YouTube](#)

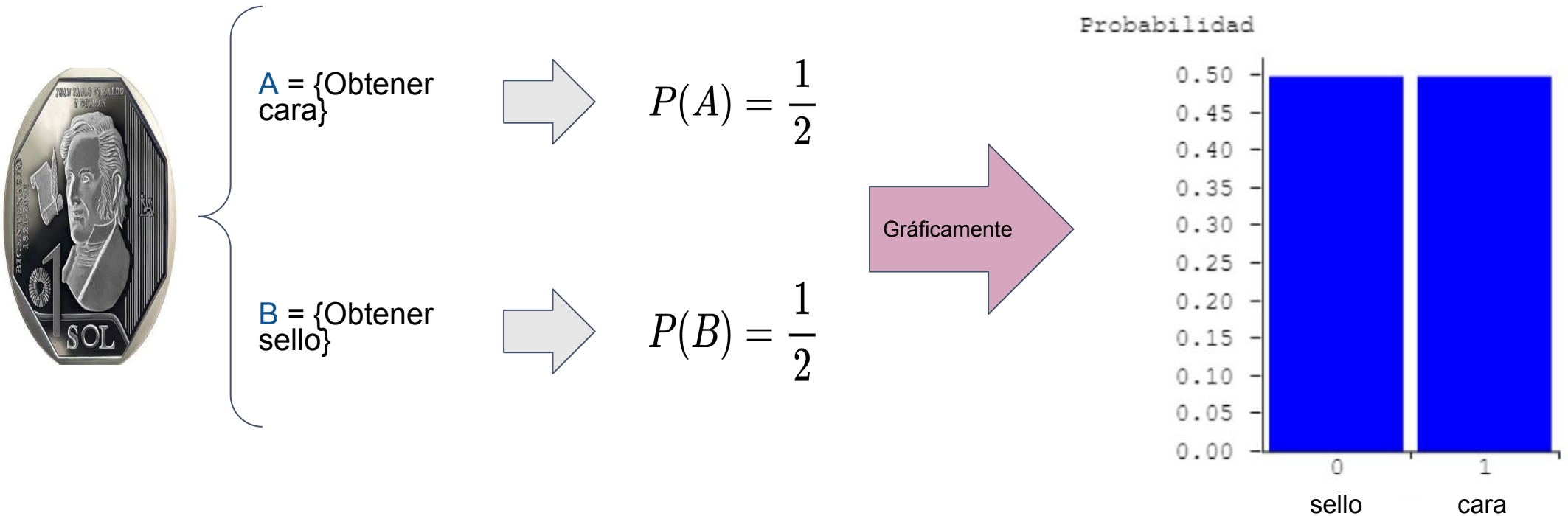
Fecha entrega: 01-marzo hasta las 10 pm



**En la próxima
clase...**

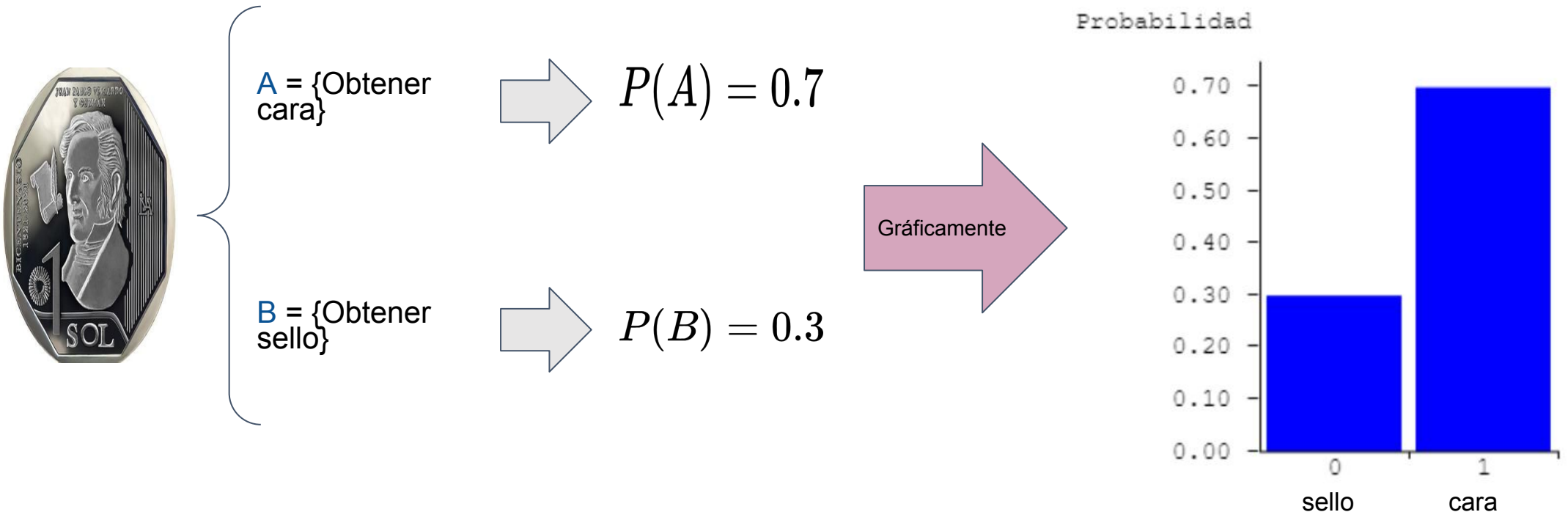
Caso: Lanzamiento de una moneda

¿Cuál es la probabilidad de obtener 1 cara en el lanzamiento de **una** moneda?



Caso: Lanzamiento de una moneda trucada

¿Cuál es la probabilidad de obtener 1 cara en el lanzamiento de **una** moneda trucada?

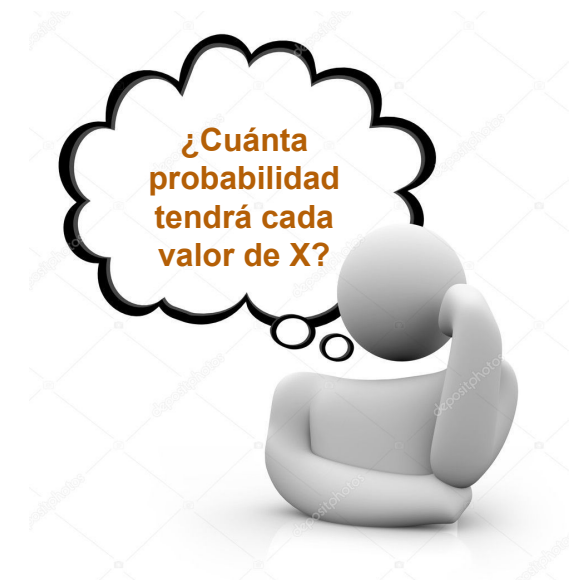
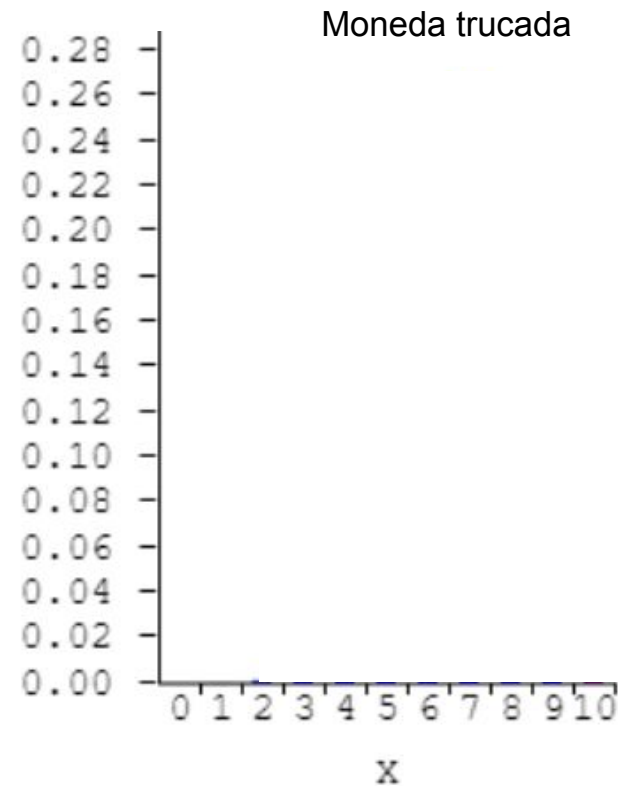
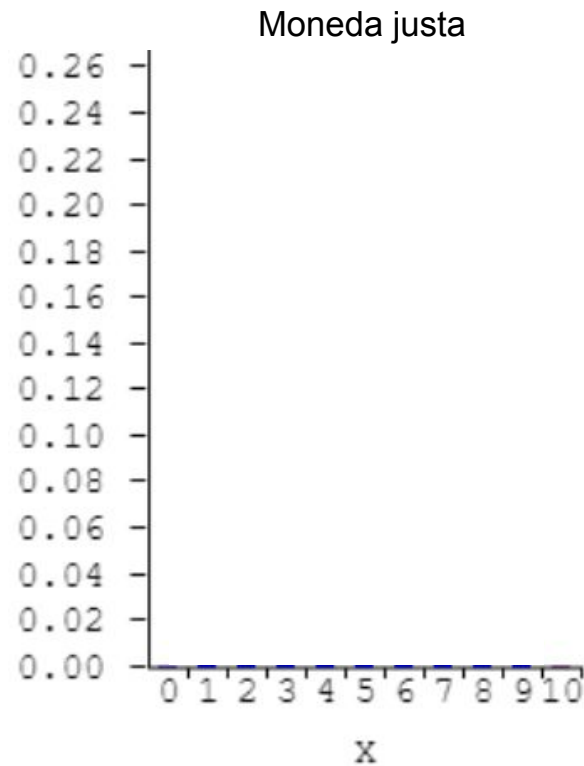


Sea la **variable aleatoria** X = # de caras en el lanzamiento de una moneda

“DISTRIBUCIÓN BERNOULLI”

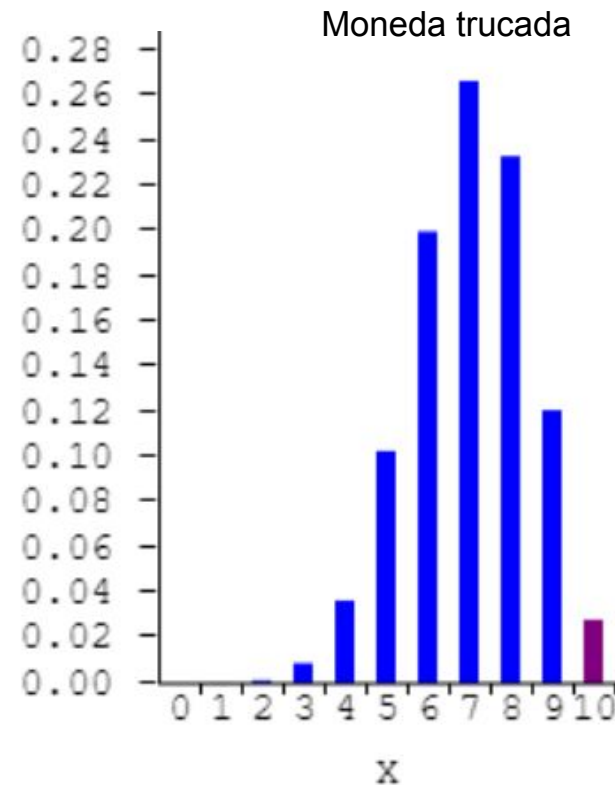
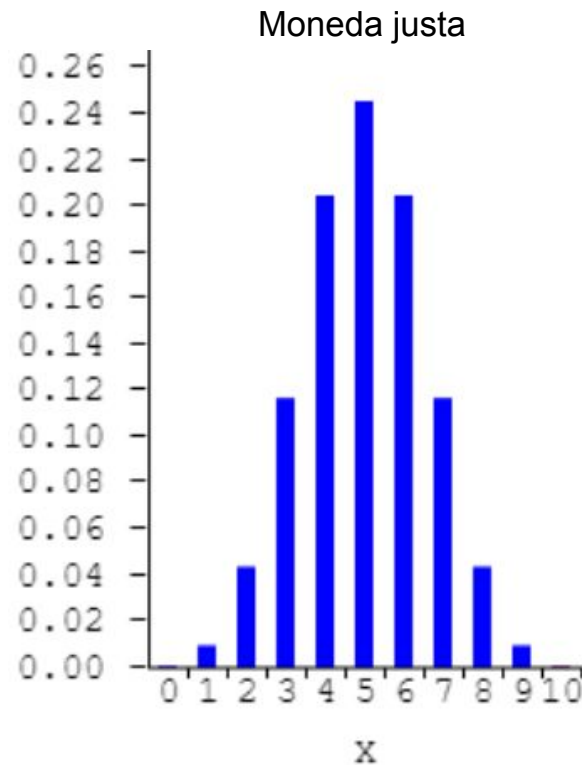
Caso: Lanzamiento de una moneda varias veces

¿Cuál es la probabilidad de obtener 4 caras en el lanzamiento de una moneda 10 veces?



Caso: Lanzamiento de una moneda varias veces

¿Cuál es la probabilidad de obtener 4 caras en el lanzamiento de una moneda 10 veces?



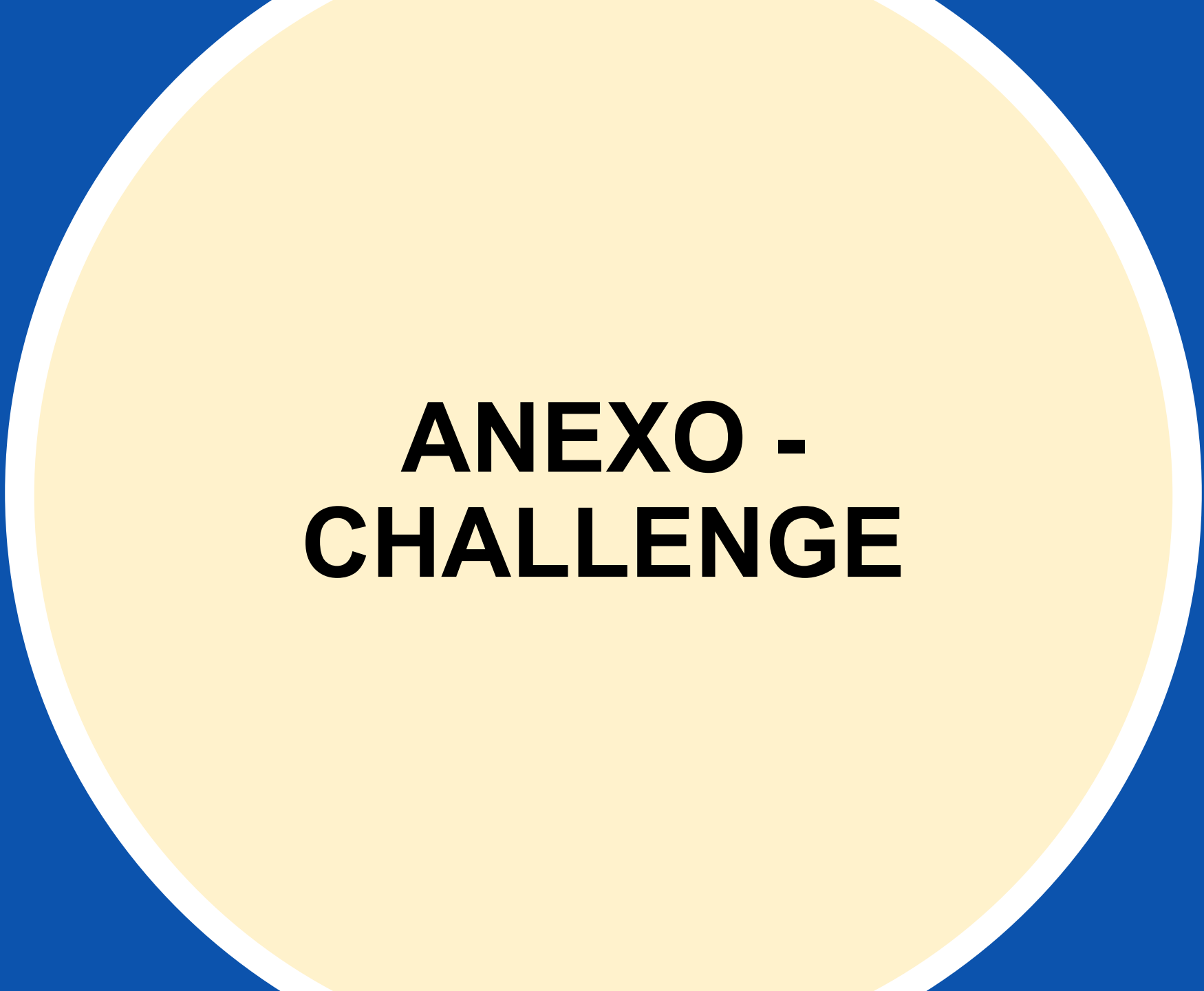
Sea la **variable aleatoria**
X = # de caras en el
lanzamiento de una
moneda 10 veces

**“ DISTRIBUCIÓN
BINOMIAL”**

GRACIAS

DATA
ANALÍTICA





ANEXO - CHALLENGE

1. Describe un caso donde apliques el ciclo básico de un análisis estadístico (Similar al caso de las elecciones o modelos en Banca). Sé lo más detallado posible (identificando los elementos de la estadística y/o utilizando gráficos de apoyo)
2. En una urna hay 4 bolas rojas, 3 bolas verdes y 2 bolas azules. Si se extrae una bola al azar, ¿cuál es la probabilidad de que sea verde?
3. Si una baraja de cartas contiene 52 cartas, ¿cuál es la probabilidad de sacar una carta de corazones al azar?
4. Si lanzamos dos dados justos, ¿cuál es la probabilidad de que la suma de los puntos sea 7?
(Sugerencia: Enlista los posibles resultados de lanzar dos dados para obtener tu espacio muestral y cuenta cuántos resultados suman 7)
5. Si se lanzan dos monedas (suponer que son monedas no trucadas), ¿cuál es la probabilidad de obtener al menos una cara? (Sugerencia: Enlista los posibles resultados de lanzar dos monedas: Cara-cara, cara-sello,etc)
6. Una urna contiene 5 bolas rojas y 3 bolas verdes. Si se extraen dos bolas al azar sin reemplazo, ¿cuál es la probabilidad de que ambas sean rojas? (Sugerencia: Calcula la probabilidad de que la primera bola sea roja y multiplicarla con la probabilidad de que la segunda bola sea roja, recuerda que para la segunda selección hay menos bolas)