Python - Probabilidad

Conceptos

Rogelio Ferreira Escutia

Profesor / Investigador
Tecnológico Nacional de México
Campus Morelia

Probabilidad

- **Sentido común:**
  - **¿Va a llover?**

Importancia

Importancia

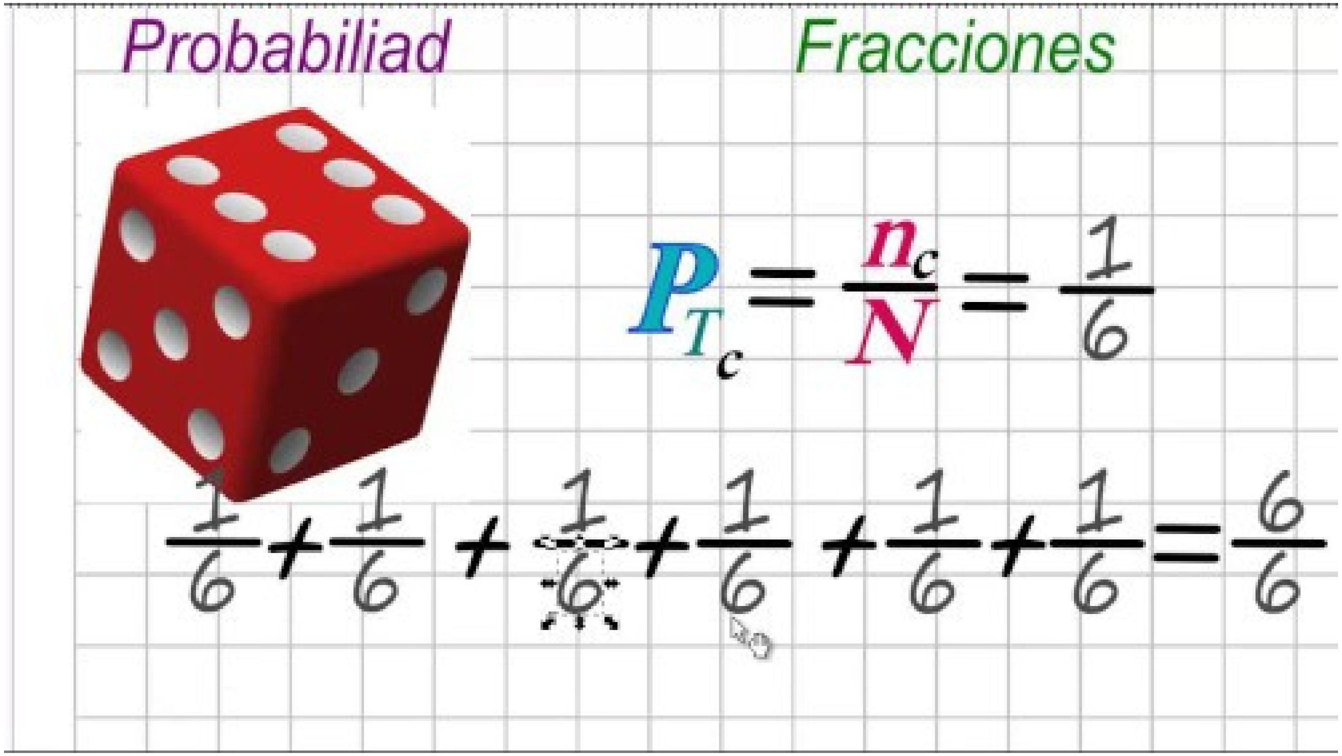
- **Otra área importante para el manejo y análisis de datos es la “Probabilidad”, ya que el solo hecho de llevar una estadística sobre ciertos hechos nos llevar a calcular de manera inmediata un conjunto de valores probables para cada hecho analizado.**

- **Los datos que provienen de eventos aleatorios que se suscitan en muchas áreas nos lleva al cálculo de las posibilidades de que ciertos eventos ocurran en un cierto tiempo.**

Conceptos

Probabilidad

- **La “Teoría de la Probabilidad” es una rama de la matemática que se encarga de estudiar los fenómenos aleatorios y es una medida para determinar si un evento ocurrirá o nó.**



Propiedades básicas (1)

- **La probabilidad se expresa como un número que está entre 0 y 1:**

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

- **Si tenemos la certeza de que el evento B ocurrirá, entonces:**

$$P(B) = 1$$

Propiedades básicas (2)

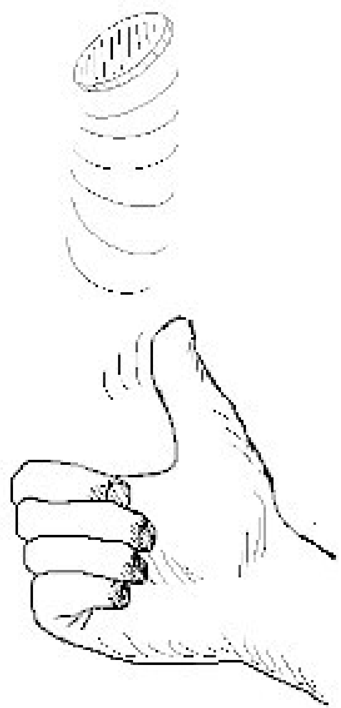
- **Si A y B son eventos mutuamente excluyentes:**

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

Espacio de muestreo

- **Es el conjunto de todos los posibles resultados de un experimento aleatorio.**

- **En el caso de una moneda (en México), sería águila ó sol.**



$$S = \{\text{águila}, \text{sol}\}$$

Lanzar una Moneda (1)

- **Lanzar 10 veces una moneda e imprimir el resultado:**

```python
#   Biblioteca para números aleatorios enteros
from random import randint
#   Se definen valores para la moneda
#       aguila = 1    sol = 0
#   Se inicializa la moneda
aguila = 0
sol = 0
print("\n")
for lanzamiento in range(1, 11):
    valor = randint(0, 1)
    if (valor==1):
        moneda = "aguila"
        aguila = aguila + 1
    else:
        moneda = "sol"
        sol = sol + 1
    print("Lanzamiento ", lanzamiento, " valor ", moneda)
print("\nResultado Final: Aguilas=",aguila," Soles=",sol,"\n")
```

Lanzar una Moneda (1)

- **Salida:**

```
Lanzamiento  1  valor  sol
Lanzamiento  2  valor  aguila
Lanzamiento  3  valor  aguila
Lanzamiento  4  valor  aguila
Lanzamiento  5  valor  sol
Lanzamiento  6  valor  aguila
Lanzamiento  7  valor  sol
Lanzamiento  8  valor  sol
Lanzamiento  9  valor  aguila
Lanzamiento  10  valor  aguila

Resultado Final: Aguilas= 6  Soles= 4
```

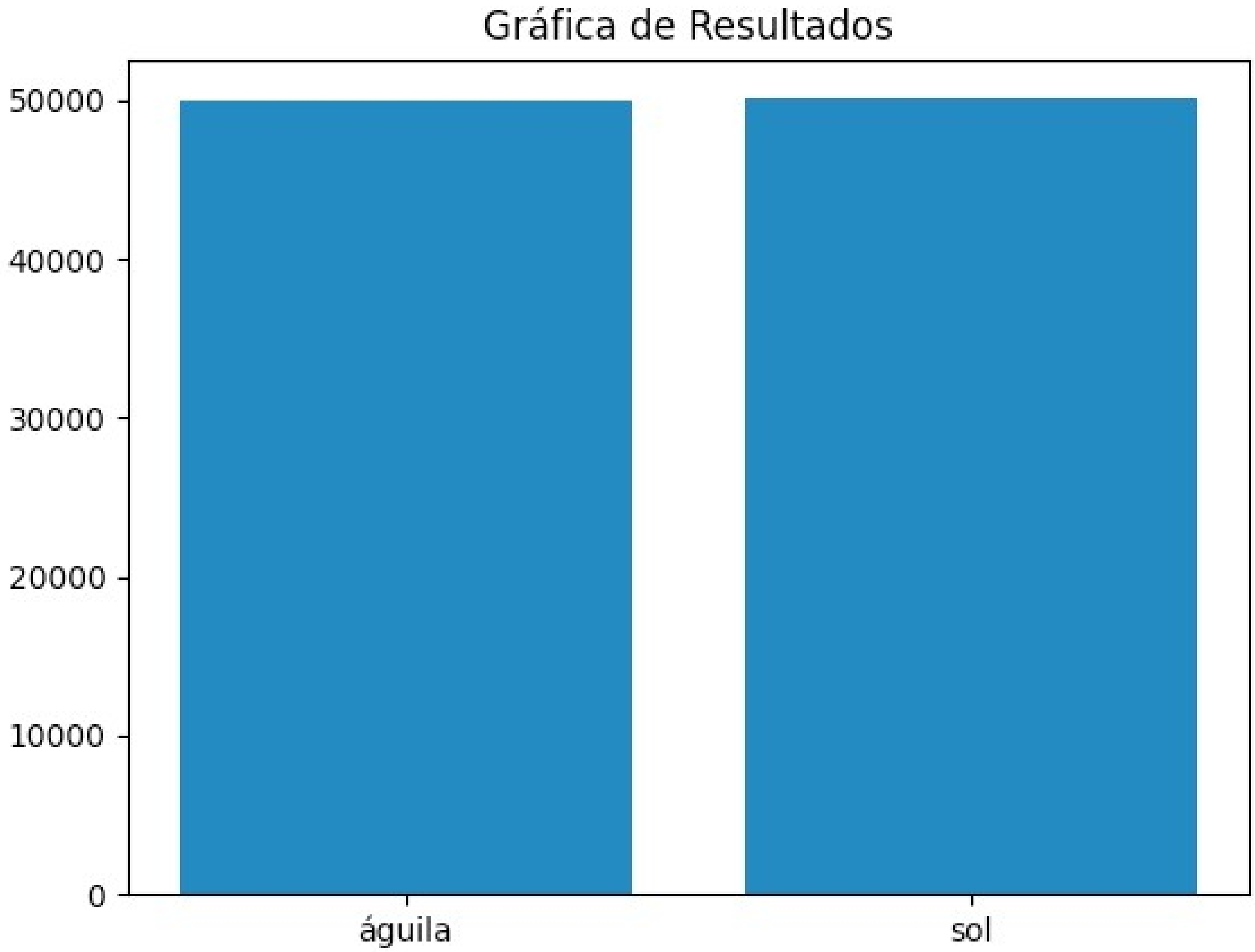
Ley de los Grandes números (1)

- **Si lanzamos una moneda muchas veces, la probabilidad de que caiga águila ó sol, tenderá a ser de 0.5 (½). A esto se le conoce como la “Ley de los grandes números”.**

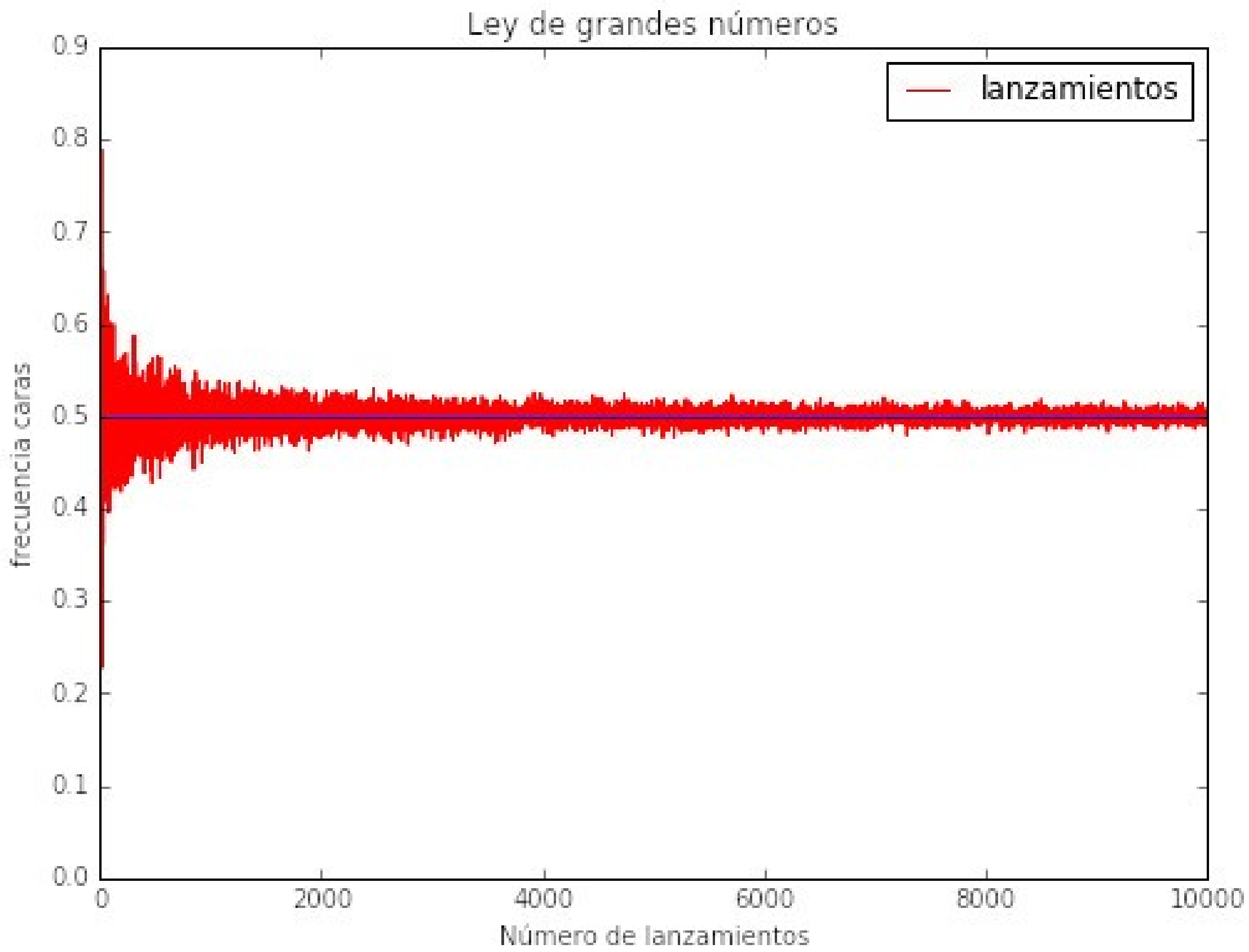
```python
from random import randint          # Biblioteca para números aleatorios enteros
import matplotlib.pyplot as plt     # Biblioteca para graficar
aguila = 0                          # Se inicializan los contadores
sol = 0
for lanzamiento in range(1, 100001):
    valor = randint(0, 1)           # valores para la moneda: aguila=1, sol=0
    if (valor==1):
        aguila = aguila + 1
    else:
        sol = sol + 1
print("\nResultado Final: Aguilas=",aguila," Soles=",sol,"\n")
etiquetas = ["águila", "sol"]       # Lista con las etiquetas
resultados = [aguila, sol]          # Lista con los resultados
plt.bar(etiquetas, resultados)      # Creamos una gráfica de barras
plt.title("Gráfica de Resultados")  # Agregamos un título a la gráfica
plt.show()                          # Mostrar la gráfica en pantalla
```

Ley de los Grandes números (2)

- **Salida:**



Ley de los Grandes números (3)



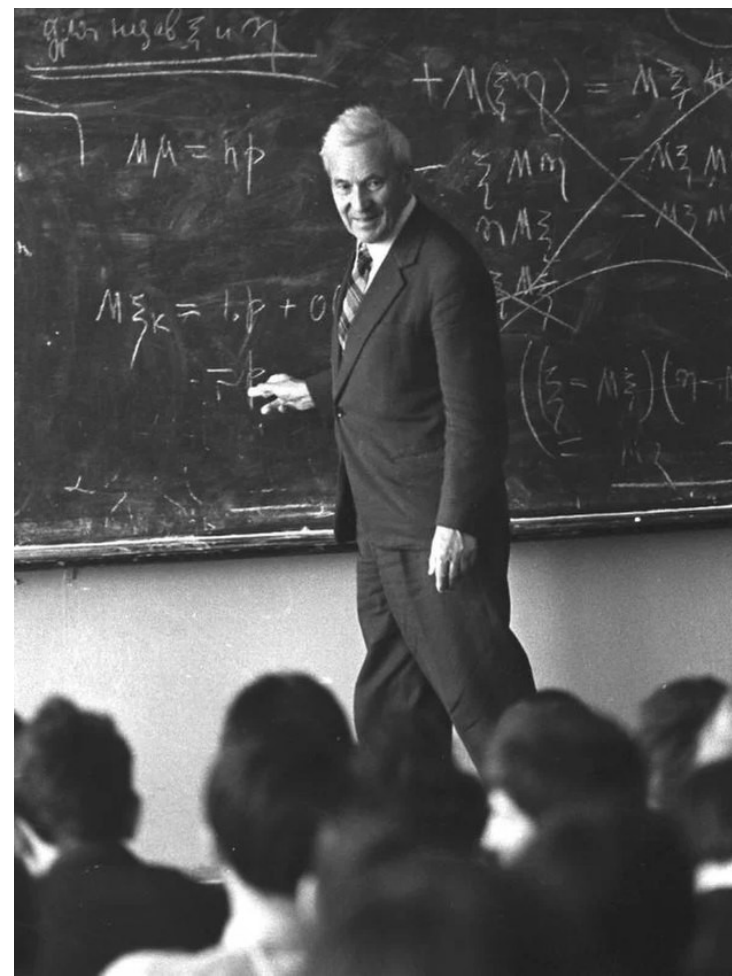
"Introducción a la teoría de probabilidad con Python", Raul E. Lopez Briega,
https://relopezbriega.github.io/blog/2016/11/26/introduccion-a-la-teoria-de-probabilidad-con-python/, mayo 2021

Espacio de Probabilidad

Definición

- **En 1933 Andréi Kolmogorov introdujo el concepto de "Espacio de Probabilidad".**

- **Un espacio de Probabilidad está constituido por 3 partes:**
  - **Espacio de Muestreo.**
  - **Espacio del Evento.**
  - **Ley de Probabilidad.**



1) Espacio de Muestreo

- **La primera parte de un "Espacio de Probabilidad" es el "Espacio de Muestreo", el cual es el conjunto de todas las posibles salidas de un experimentoy está representada por el símbolo griego "Ω".**

- **El espacio de muestreo puede ser de 2 maneras:**
  - **Discreto.**
  - **Continuo.**

1) Espacio de Muestreo

- **Algunos ejemplos de un espacio de muestreo discreto son:**
  - **Lanzar una moneda: Ω = { águila, sol }**
  - **Lanzar un dado: Ω={1,2,3,4,5,6}**
  - **Números enteros par:Ω={2,4,6,8,…}**

- **Algunos ejemplos de un espacio de muestreo contínuo son:**
  - **Ángulo de fase de un voltaje: Ω = { Θ | 0<=Θ<=2π }**
  - **Unahora: Ω={t|0<=t<=60minutos}**

2) Espacio del Evento

- **“Espacio del Evento” es toda la colección de eventos posibles y que está representado por el símbolo "Ƒ" y es un subconjunto de “Ω”.**

- **Un ejemplo es:**
  - **Experimento: Lanzar un dado.**
  - **Espacio de muestreo: Ω={1,2,3,4,5,6}**
- **2 Eventos**
  - **Evento1={Números par}={2,4,6}**
  - **Evento2={Mayores de 4}={5,6}**

3) Espacio de Probabilidad

- **Espacio de Probabilidad" es el "Ley de Probabilidad", la cual es la probabilidad de cada uno de los eventos.**

- **Si calculamos las probabilidades para los eventos anteriores nos queda:**

  – **P(Evento1)=3/6**

  – **P(Evento2)=2/6**

Probabilidad Condicional

Probabilidad condicional (1)

- **Dentro del área de ciencia de datos, es muy importante el descubrir las relaciones entre 2 eventos (por ejemplo A y B), y ver si es posible que un evento pueda causar o alterar al otro (que A cause a B), lo que indica una relación entre los eventos, lo que denominamos "Probabilidad Condicional".**

- **Esta es muy importante, ya que este relación puede desencadenar otros eventos, por ejemplo el evento B cause a C, por lo tanto podemos encontrar que si A implica a B, y B implica a C, podemos encontrar una relación entre A y C.**

Probabilidad condicional (2)

- **Nos indica la probabilidad de un evento (A) dado que ocurrió otro evento (B):**

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

Independencia

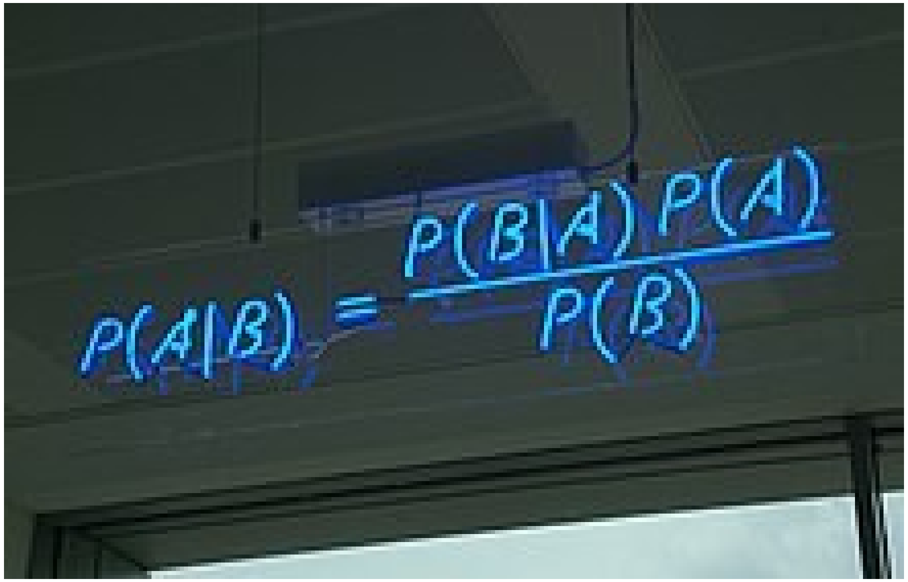
- **Cuando tenemos 2 eventos, en donde ninguno de los 2 depende del otro, entonces decimos que son “independientes”, y matemáticamente lo podemos expresar como:**

$$P[A \cap B] = P[A]P[B]$$

Teorema de Bayes

Teorema de Bayes

- **El teorema de Bayes, en la teoría de la probabilidad, es una proposición planteada por el matemático inglés Thomas Bayes (1702-1761)1 y publicada póstumamente en 1763,2 que expresa la probabilidad condicional de un evento aleatorio A A dado B B en términos de la distribución de probabilidad condicional del evento B B dado A A y la distribución de probabilidad marginal de solo A A.**



Ejemplo (1)

- **¿Cuál es la probabilidad de tener cancer, dado que tomo café?**

- **Las probabilidades a priori son:**

$$P(\text{Coffee}) = .65$$

$$P(\text{Cancer}) = .005$$

$$P(\text{Coffee}|\text{Cancer}) = .85$$

"Essential Math for Data Science", Thomas Nield, O'Reilly 2022

Ejemplo (2)

- **Sustituyendo ecuaciones:**

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

$$P\left(\mathrm{Cancer}|\mathrm{Coffee}\right) = \frac{P(\mathrm{Coffee}|\mathrm{Cancer})^{*}P(\mathrm{Coffee})}{P(\mathrm{Cancer})}$$

$$P\left(\mathrm{Cancer}|\mathrm{Coffee}\right) = \frac{.85^{*}.005}{.65} = .0065$$

“Essential Math for Data Science”, Thomas Nield, O’Reilly 2022

Ejemplo (3)

- **Código:**

```
p_coffee_drinker = .65
p_cancer = .005
p_coffee_drinker_given_cancer = .85

p_cancer_given_coffee_drinker = p_coffee_drinker_given_cancer *
    p_cancer / p_coffee_drinker

# prints 0.006538461538461539
print(p_cancer_given_coffee_drinker)
```

“Essential Math for Data Science”, Thomas Nield, O’Reilly 2022

Variables Aleatorias

Variables Aleatorias (1)

- **Las variables aleatorias de un evento, como lanzar un dado, se pueden mapear eventos a números, para posteriormente convertirse en parte de un dataset el cual pueda ser analizado en un proceso de ciencia de datos.**

- **Por ejemplo, se tienen 4 cartas (♠ ♣ ♥ ♦), para las cuales les construimos su espacio de probabilidad y nos queda como sigue:**

  – **Cartas: Ω={♠,♣,♥,♦}**

Variables Aleatorias (2)

- **La probabilidad de sacar cada de una de las cartas queda de la siguiente manera:**
  - **P[{♠}]=1⁄4**
  - **P[{♣}]=1⁄4**
  - **P[{♥}]=1⁄4**
  - **P[{♦}]=1⁄4**

Rogelio Ferreira Escutia

Profesor / Investigador
Tecnológico Nacional de México
Campus Morelia

rogelio.fe@morelia.tecnm.mx

rogeplus@gmail.com

xumarhu.net

@rogeplus

https://www.youtube.com/channel/UC0on88n3LwTKxJb8T09sGjg

rogelioferreiraescutia