Curso de

Programación Dinámica y Estocástica con Python

David Aroesti

Objetivos

- Aprender cuándo utilizar Programación Dinámica y sus beneficios
- Entender la diferencia entre programas deterministas y estocásticos
- Aprender a utilizar Programación Estocástica
- Aprender a crear simulaciones computacionales válidas

Introducción a la Programación Dinámica

[El nombre] Programación Dinámica se escogió para esconder a patrocinadores gubernamentales el hecho que en realidad estaba haciendo Matemáticas. La frase Programación Dinámica es algo que ningún congresista puede oponerse.

Richard Bellman

Programación Dinámica

- **Subestructura Óptima**. Una solución global óptima se puede encontrar al combinar soluciones óptimas de subproblemas locales.

- **Problemas empalmados**. Una solución óptima que involucra resolver el mismo problema en varias ocasiones.

Memoization

- La *Memorización* es una técnica para guardar cómputos previos y evitar realizarlos nuevamente.

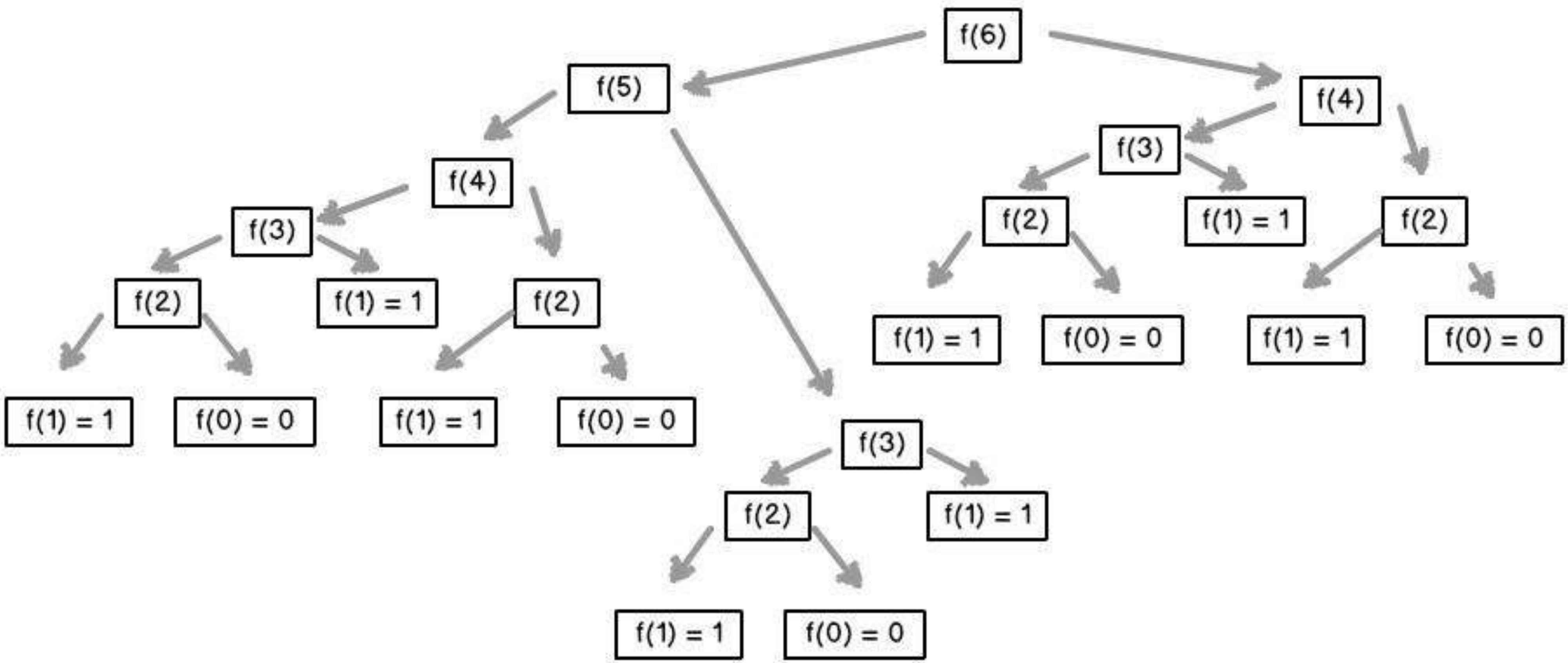
- Normalmente se utiliza un diccionario, donde las consultas se pueden hacer en O(1).

- Intercambia tiempo por espacio.

Números de Fibonacci

$$F_n = F_{n\text{-}1} + F_{n\text{-}2}$$

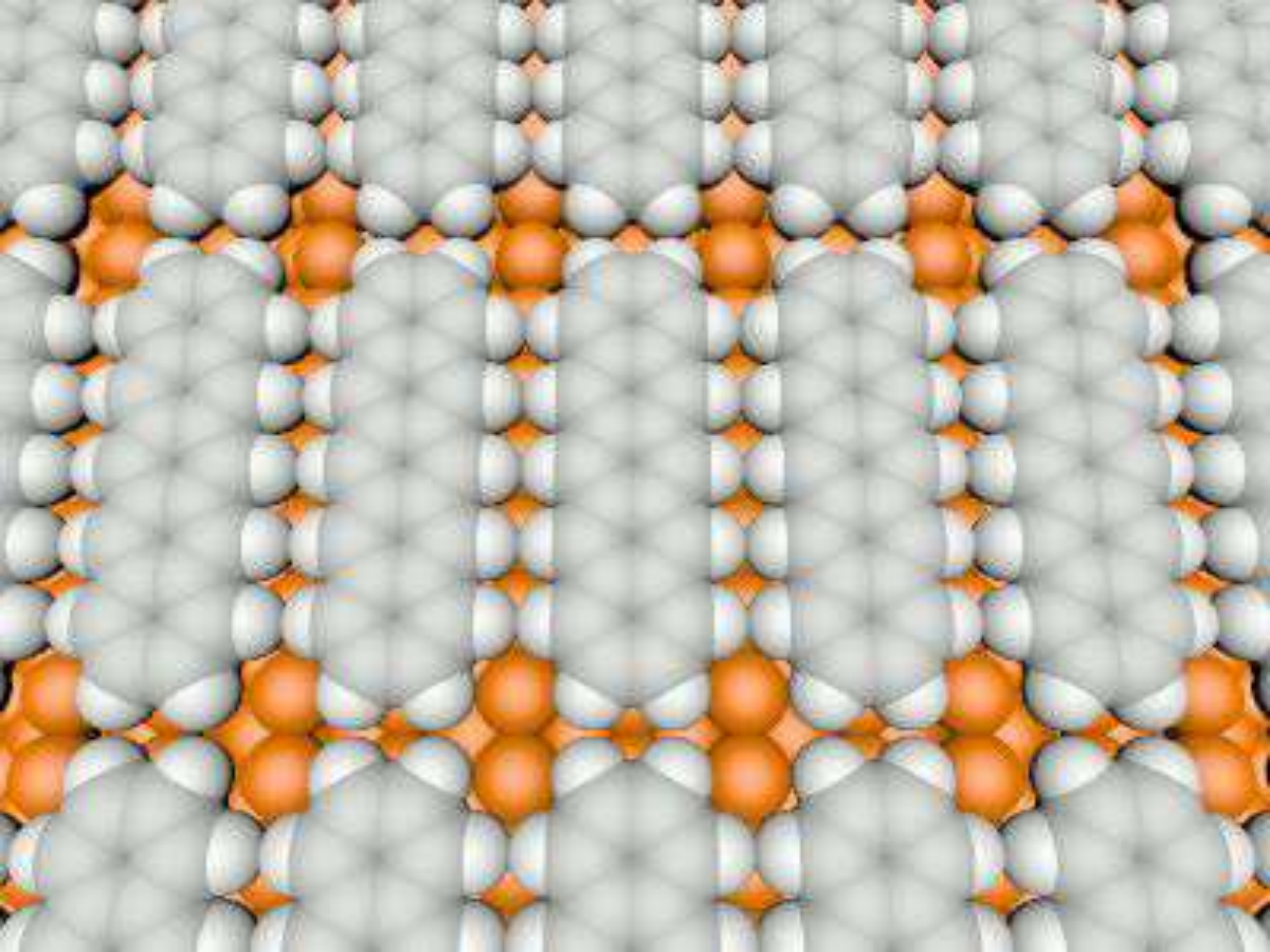
Números de Fibonacci



¿Qué son los caminos aleatorios?

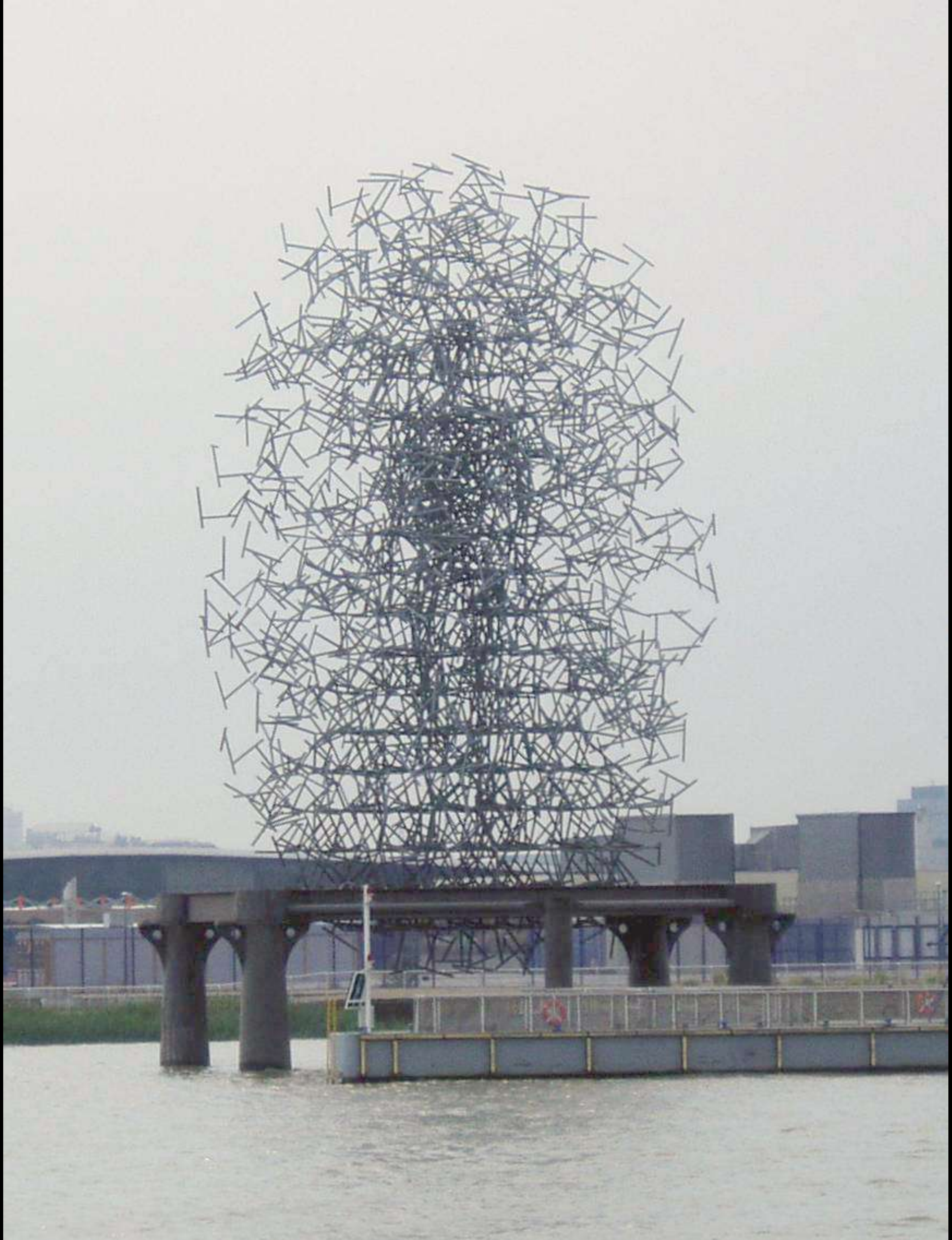
Caminos aleatorios

- Es un tipo de simulación que elige aleatoriamente una decisión dentro de un conjunto de decisiones válidas
- Se utiliza en muchos campos del conocimiento cuando los sistemas no son deterministas e incluyen elementos de aleatoriedad

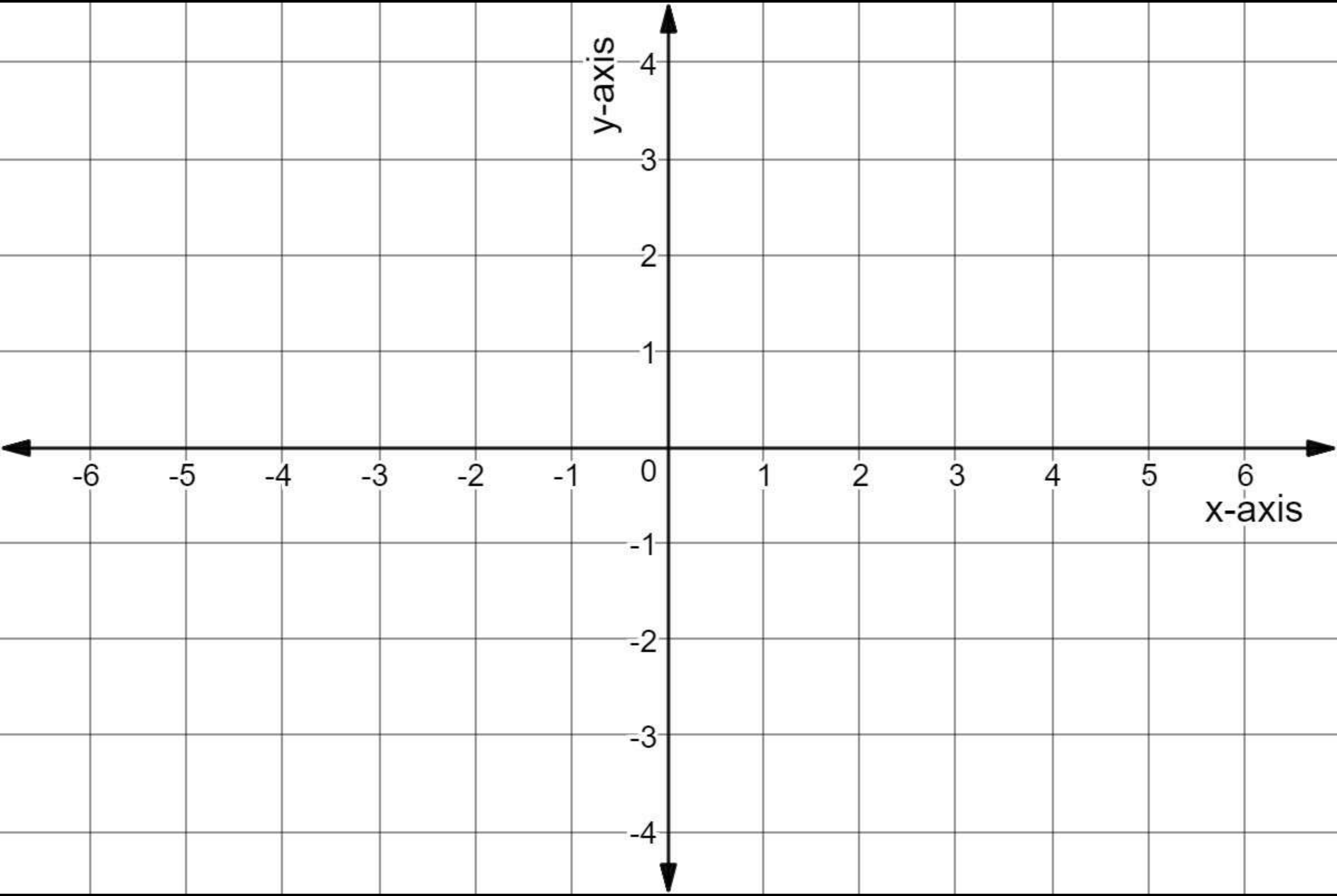


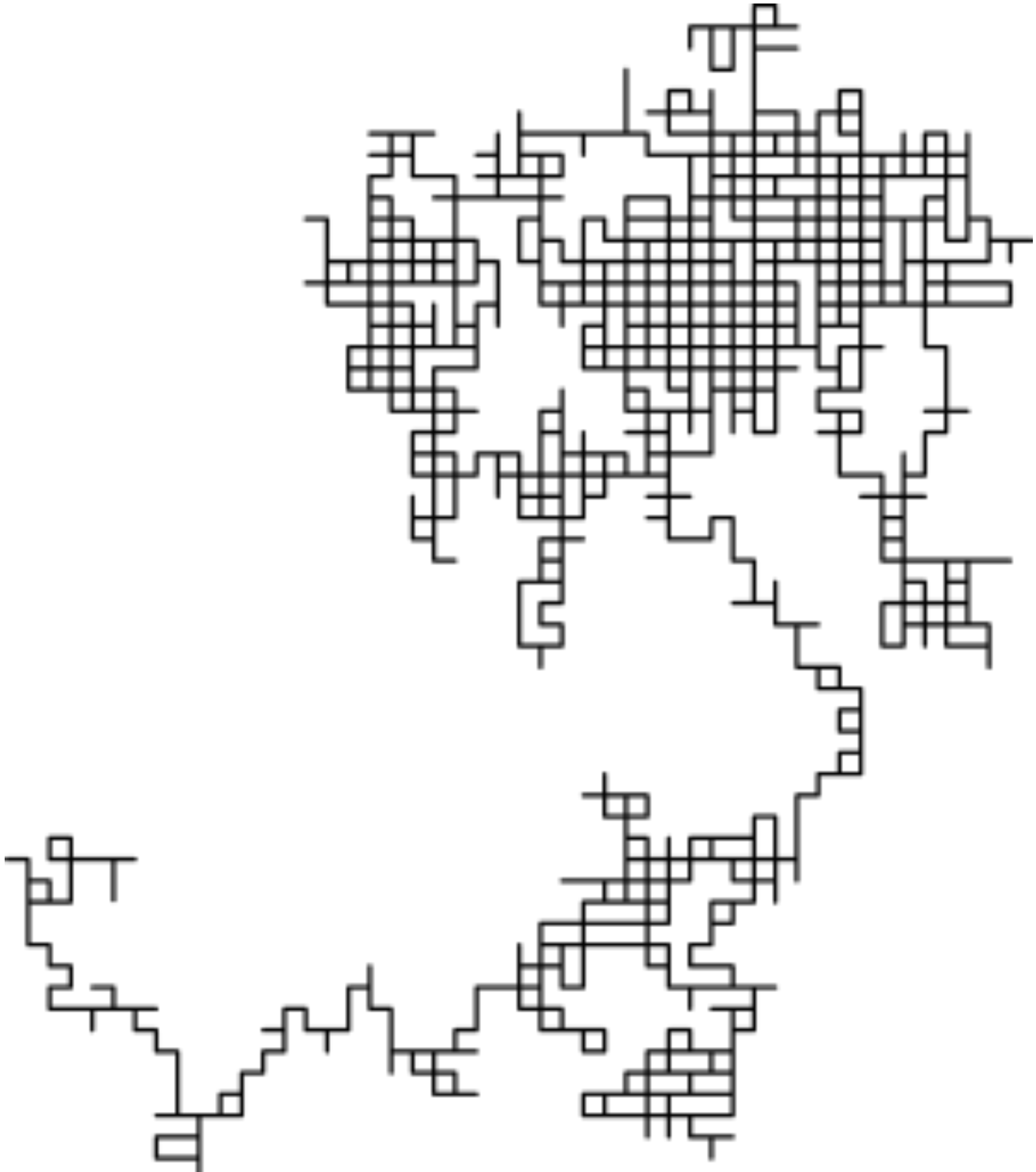

esa
Simulation of galaxy merger

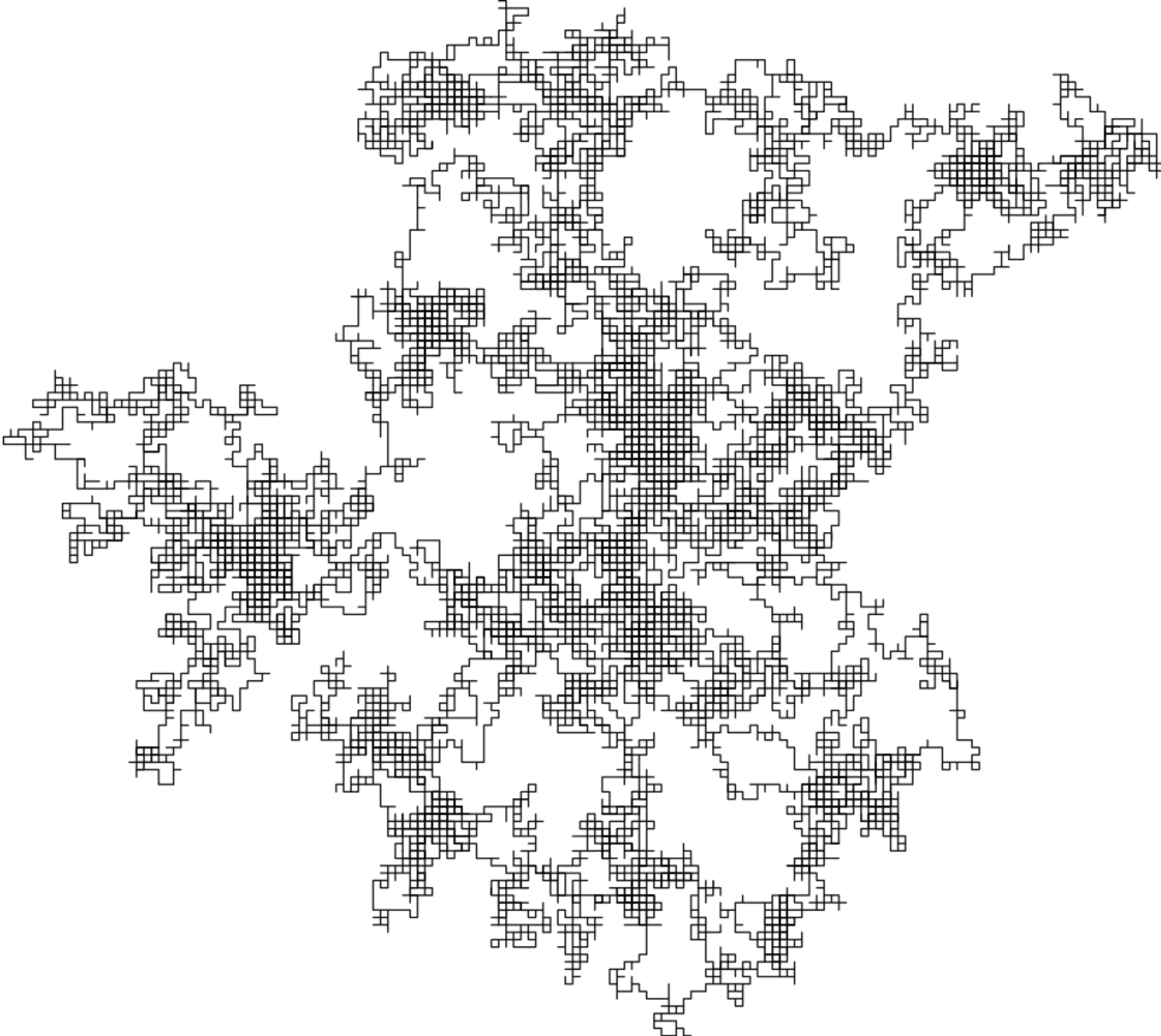

10.01
10.01
0.52%
0.47%
0.65%
MakeAGIF.com

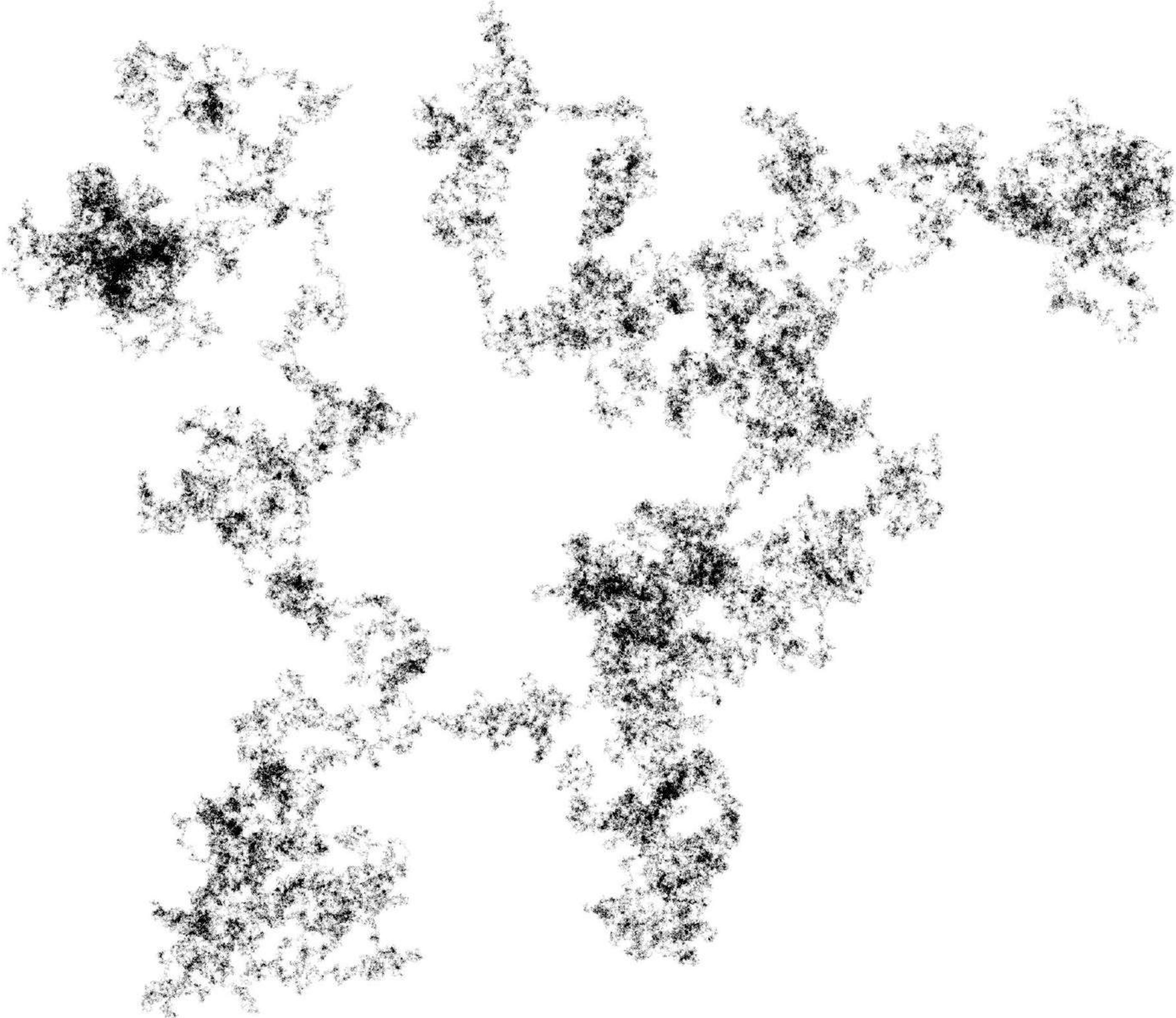


Camino de Borrachos


y-axis
4
3
2
1
0
-1
-2
-3
-4
-6
-5
-4
-3
-2
-1
1
2
3
4
5
6
x-axis







Introducción a la Programación Estocástica

Programación Estocástica

- Un programa es determinístico si cuando se corre con el mismo *input* produce el mismo *output*

- Los programas determinísticos son muy importantes, pero existen problemas que no pueden resolverse de esa manera

- La programación estocástica permite introducir aleatoriedad a nuestros programas para crear simulaciones que permiten resolver otro tipo de problemas

Programación Estocástica

- Los programas estocásticos se aprovechan de que las distribuciones probabilísticas de un problema se conocen o pueden ser estimadas


TO MAIN ST

Cálculo de Probabilidades

Probabilidades

- La probabilidad es una medida de la certidumbre asociada a un evento o suceso futuro y suele expresarse como un número entre 0 y 1

- Una probabilidad de 0 significa que un suceso jamás sucederá

- Una probabilidad de 1 significa que un suceso está garantizado de suceder en el futuro

Probabilidades

- Al hablar de probabilidad preguntamos qué fracción de todos los posibles eventos tiene la propiedad que buscamos

- Por eso es importante poder calcular todas las posibilidades de un evento para entender su probabilidad

- La probabilidad de que un evento suceda y de que no suceda es siempre 1

Probabilidades

- P(A) + P(~A) = 1
  - Ley del complemento

- P(A y B) = P(A) * P(B)
  - Ley multiplicativa

- P(A o B) = P(A) + P(B) (mutuamente exclusivos)
- P(A o B) = P(A) + P(B) - P(A y B) (no exclusivos)
  - Ley aditiva

Inferencia Estadística

Inferencia Estadística

- Con las simulaciones podemos calcular las probabilidades de eventos complejos sabiendo las probabilidades de eventos simples
- ¿Qué pasa cuando no sabemos las probabilidades de los eventos simples?
- Las técnicas de la Inferencia Estadística nos permiten inferir/concluir las propiedades de una población a partir de una muestra **aleatoria**.

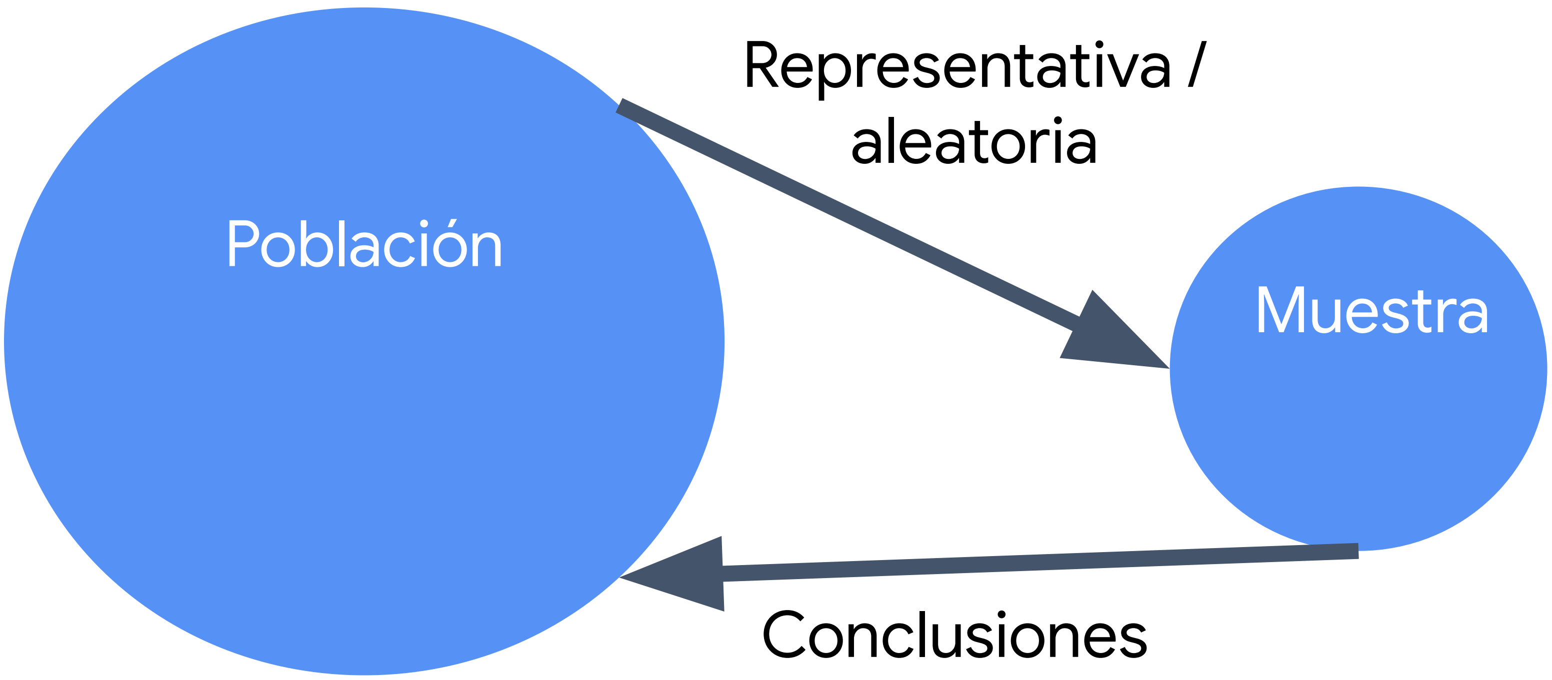
“

El principio guía de la Inferencia Estadística es que una muestra aleatoria tiende a exhibir las mismas propiedades que la población de la cual fue extraída.

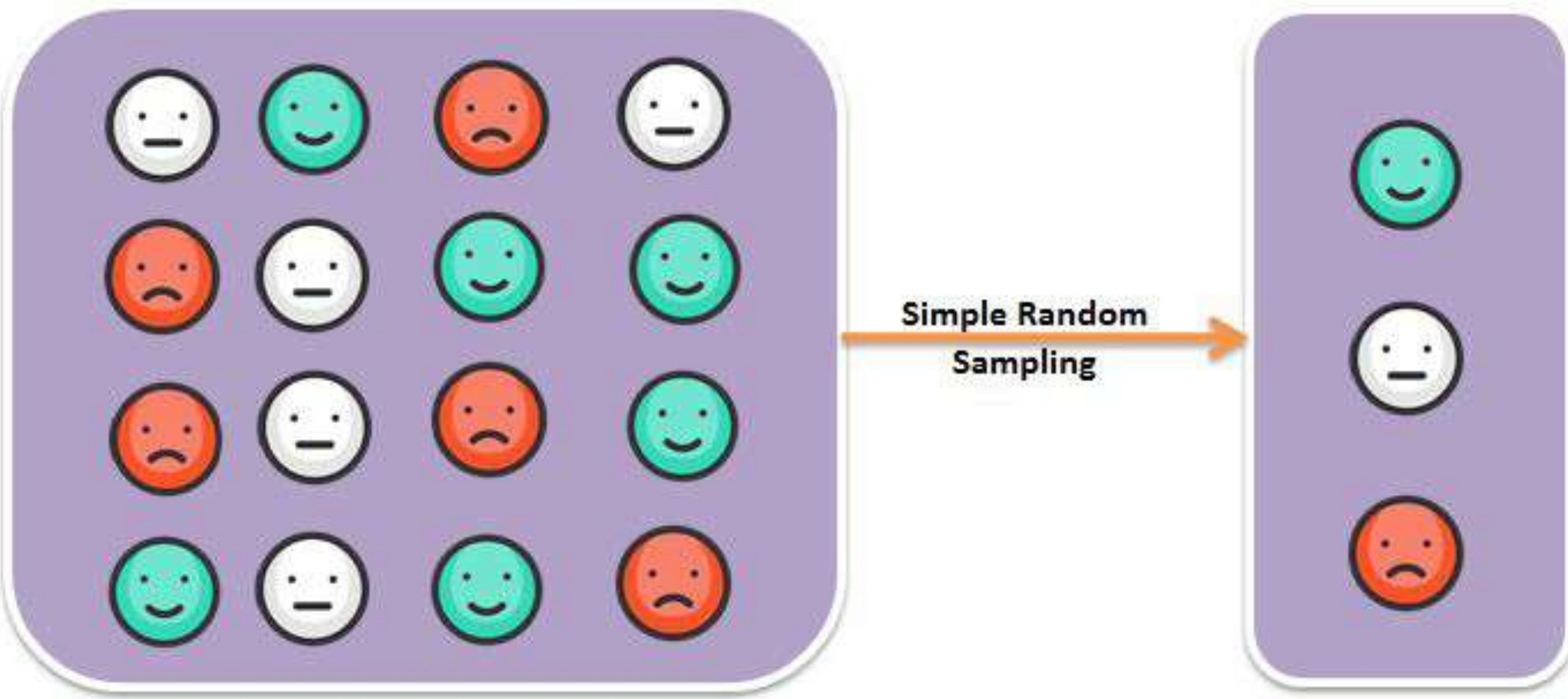
”

John Guttag

Proceso



Representativa / aleatoria
Población
Muestra
Conclusiones


Simple Random
Sampling

Ley de los grandes números

- En pruebas independientes repetidas con la misma probabilidad p de un resultado, la fracción de desviaciones de p converge a cero conforme las cantidad de pruebas se acerca al infinito.

$$\mathrm{P}\left(\lim_{n\to\infty} \overline{X}_n = \mu\right) = 1,$$

Falacia del apostador

- La falacia del apostador señala que después de un evento extremo, ocurrirán eventos menos extremos para nivelar la media

- *La regresión a la media* señala que después de un evento aleatorio extremo, el siguiente evento probablemente será menos extremo.

Media

Media

- Es una medida de tendencia central
- Comúnmente es conocida como el promedio
- La media de una población se denota con el símbolo μ. La media de una muestra se denota con $\bar{x}$

Media

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Media

$$\bar{x} = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Varianza y Desviación Estándar

Varianza

- La varianza mide qué tan propagados se encuentran un conjunto de valores aleatorios de su media.

- Mientras que la media nos da una idea de dónde se encuentran los valores, la varianza nos dice que tan dispersos se encuentran.

- La varianza siempre debe entenderse con respecto a la media.

Varianza

$$\mathrm{Var}(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

Desviación estándar

- La desviación estándar es la raíz cuadrada de la varianza.

- Nos permite entender, también, la propagación y se debe entender siempre relacionado a la media.

- La ventaja sobre la varianza es que la desviación estándar está en las mismas unidades que la media.

Desviación estándar

$$\sigma = \sqrt{\sum_{i=1}^{N} p_i (x_i - \mu)^2},$$

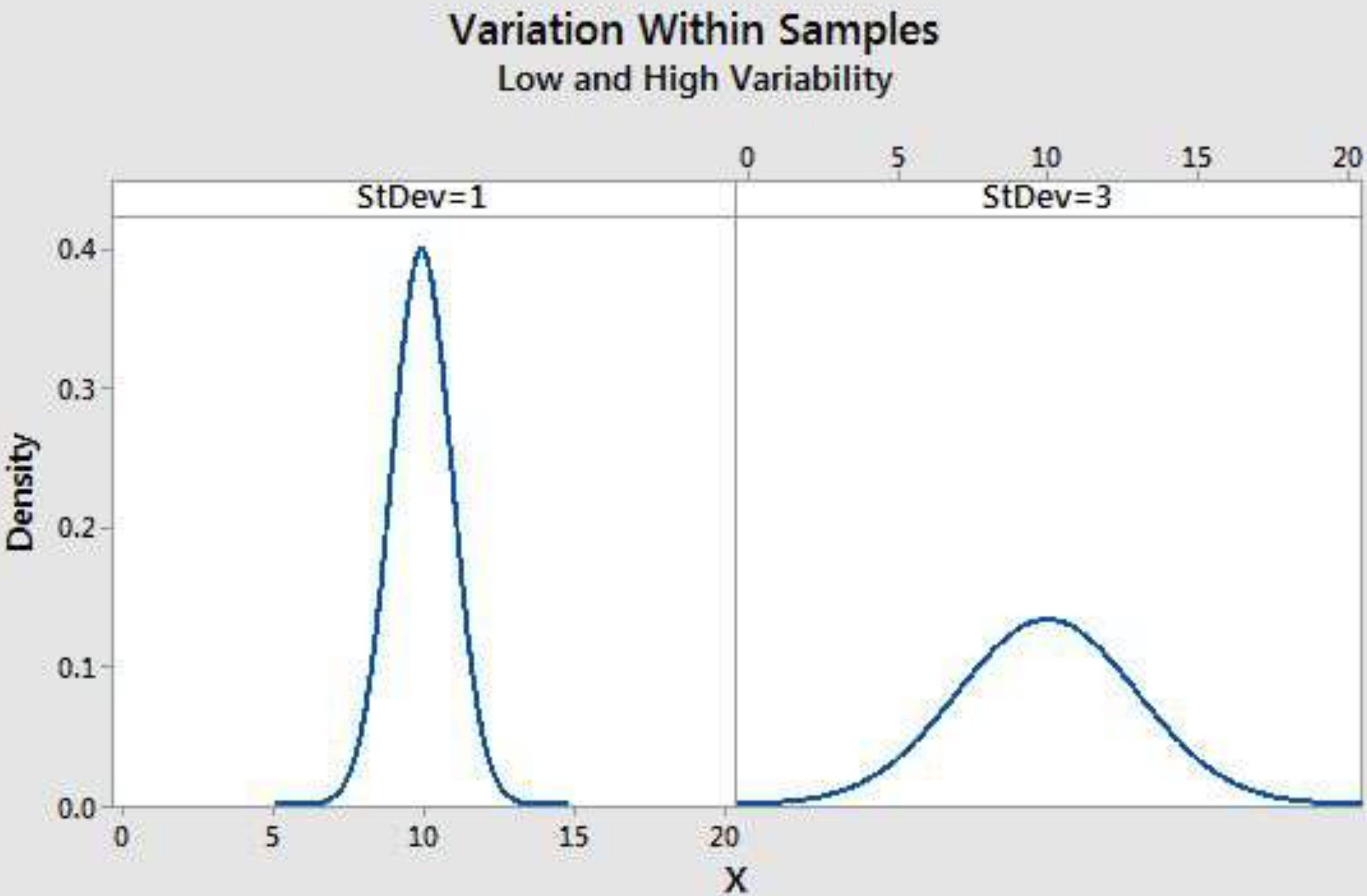
Distribución normal

Distribución normal

- Es una de las distribuciones más recurrentes en cualquier ámbito
- Se define completamente por su media y su desviación estándar
- Permite calcular intervalos de confianza con la regla empírica

Distribución normal

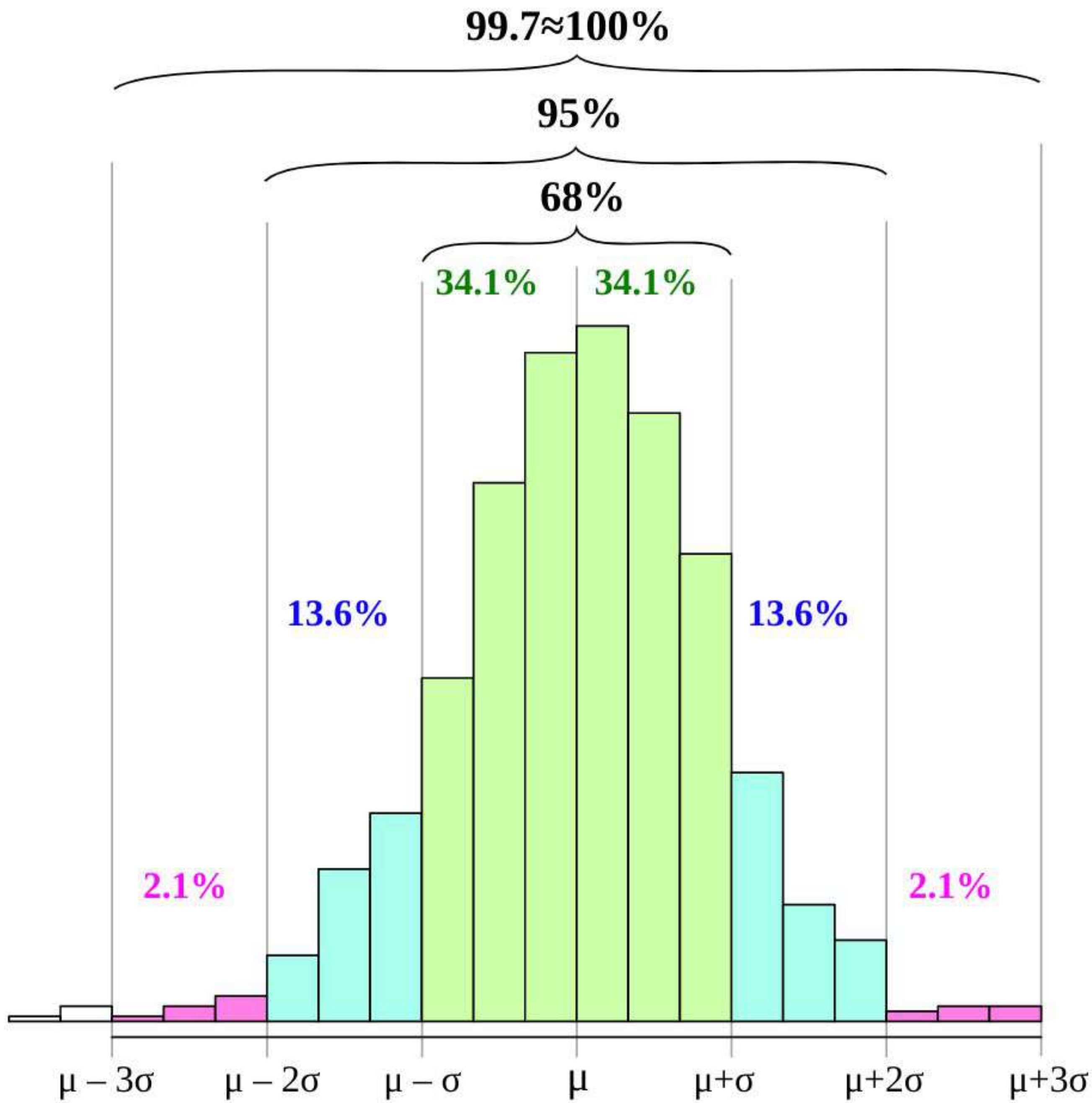
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$


Variation Within Samples
Low and High Variability
StDev=1
StDev=3
Density
X
0.0
0.1
0.2
0.3
0.4
0
5
10
15
20

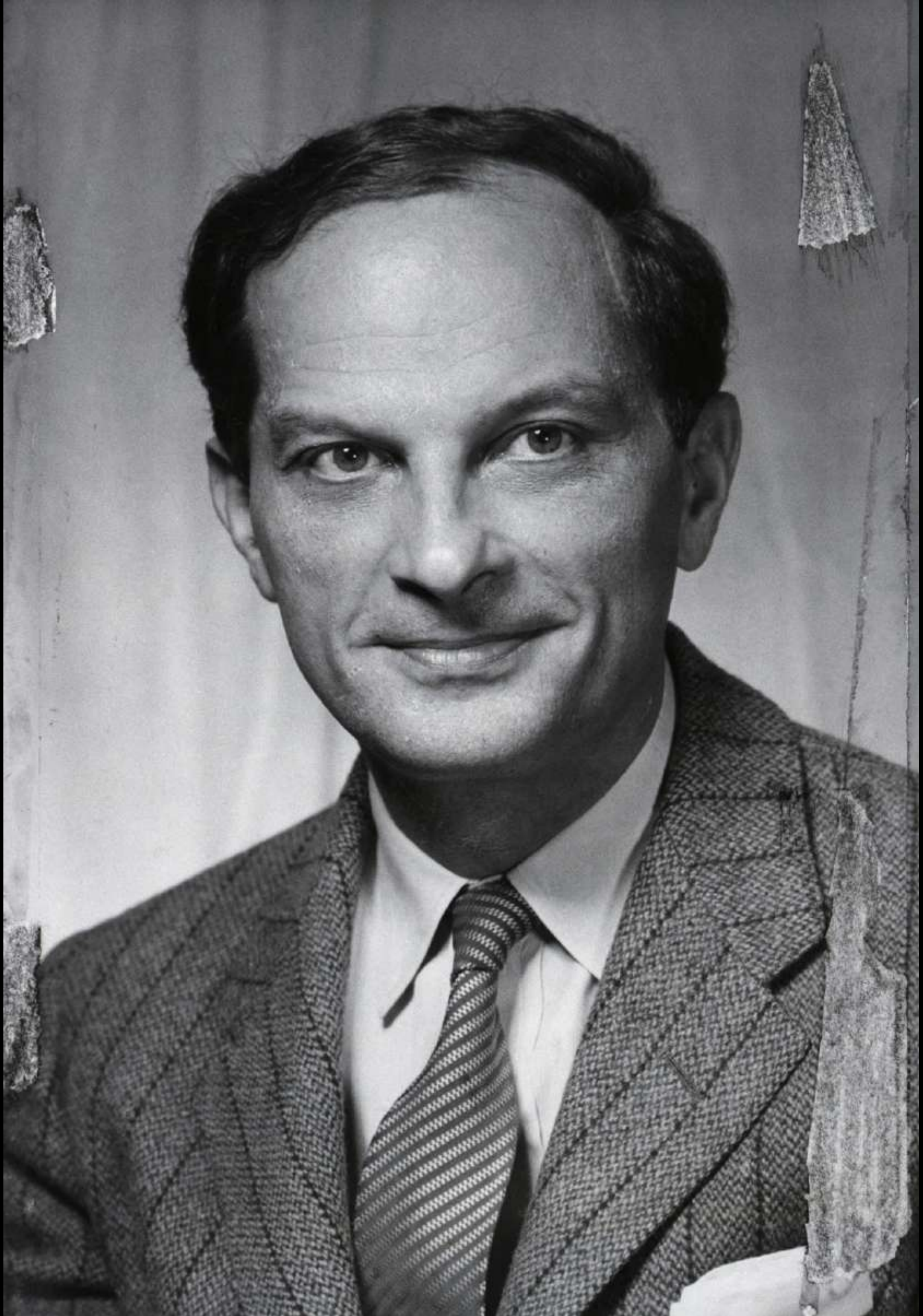
Regla empírica

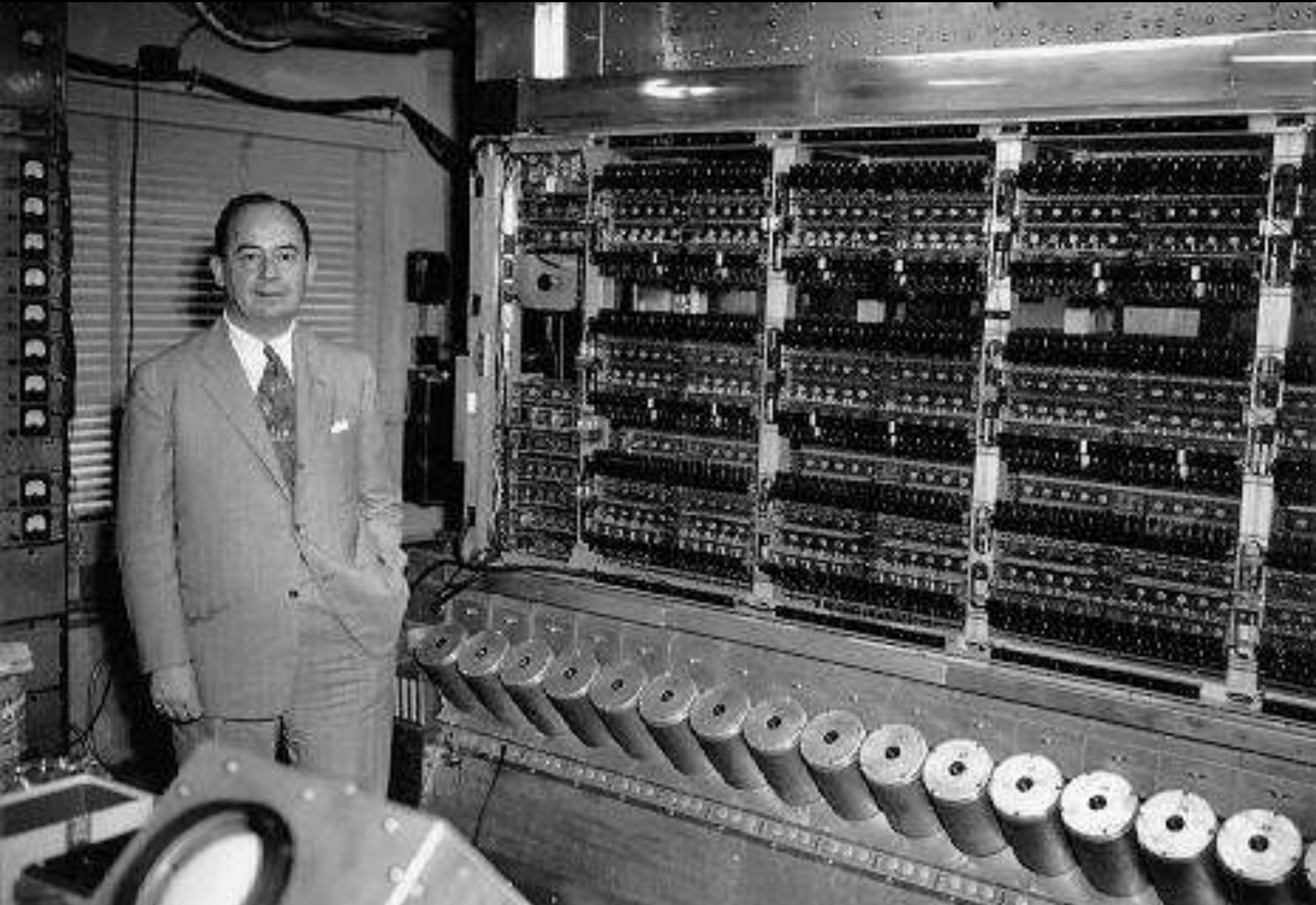
- También conocida como la regla 68-95-99.7
- Señala cuál es la dispersión de los datos en una distribución normal a uno, dos y tres sigmas
- Permite calcular probabilidades con la densidad de la distribución normal

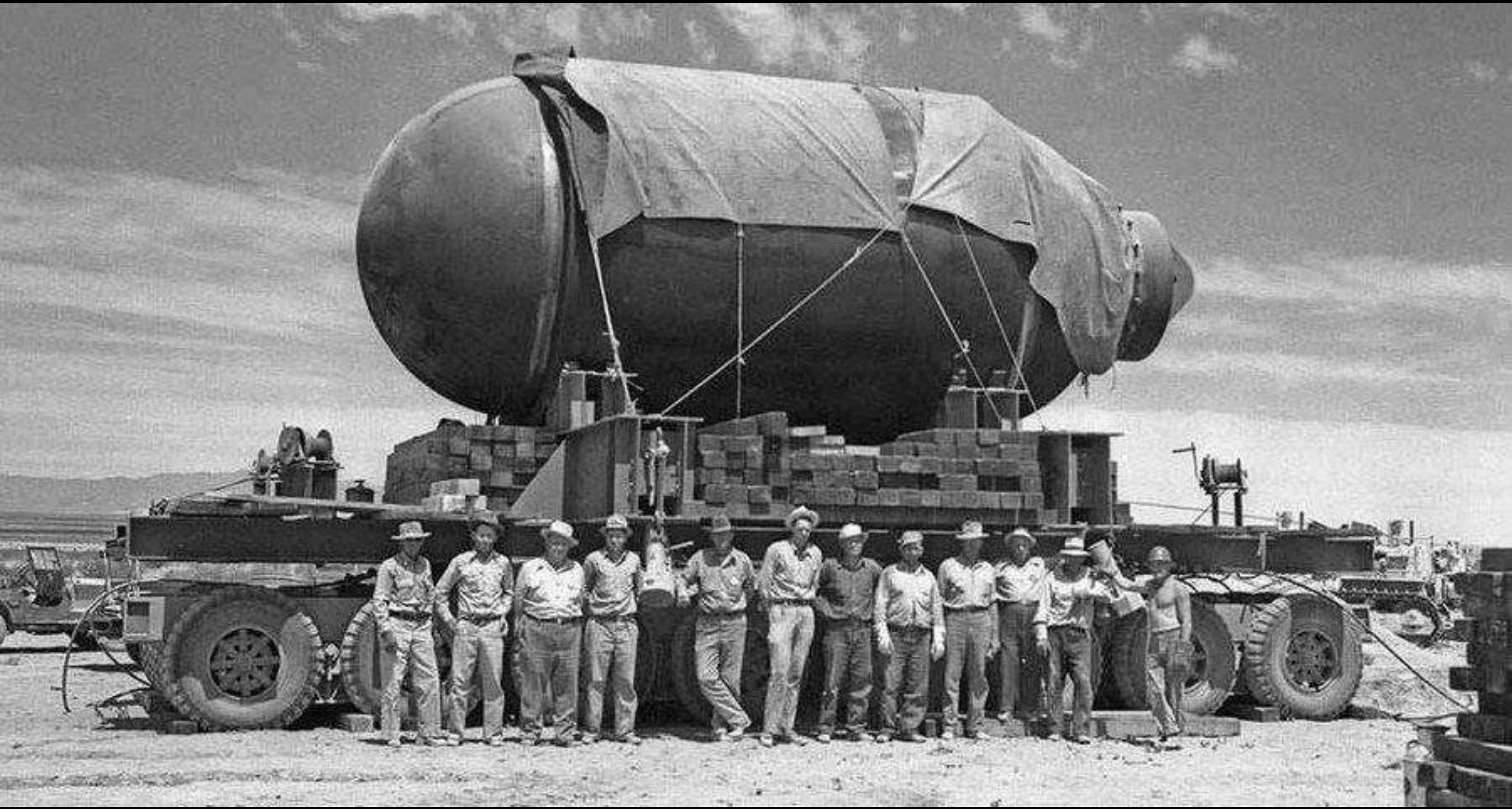
$$\Pr(\mu - 1\sigma \leq X \leq \mu + 1\sigma) \approx 0.6827$$
$$\Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.9545$$
$$\Pr(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.9973$$

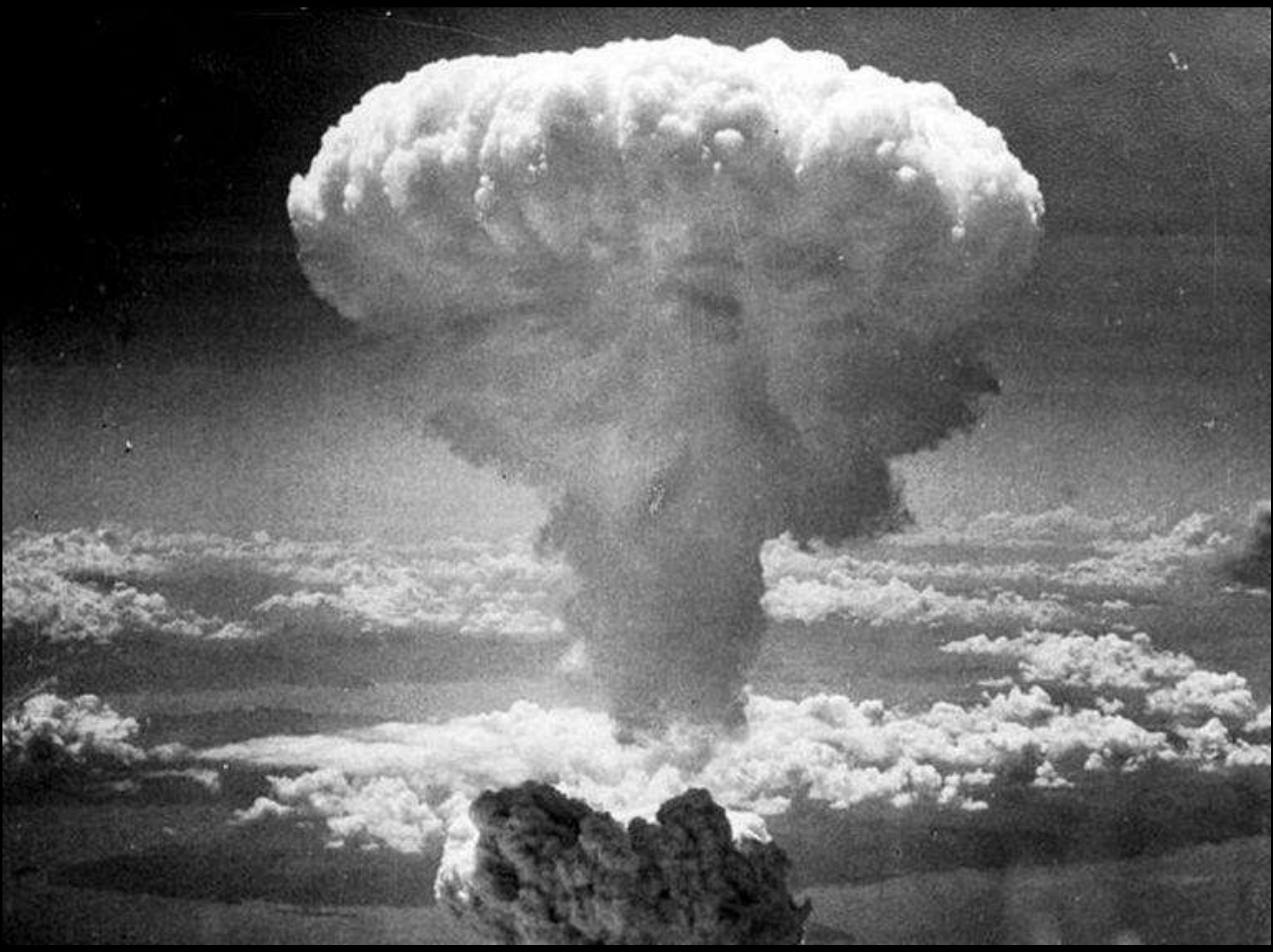

99.7≈100%
95%
68%
34.1%
34.1%
13.6%
13.6%
2.1%
2.1%
μ – 3σ
μ – 2σ
μ – σ
μ
μ+σ
μ+2σ
μ+3σ

¿Qué son las simulaciones de Montecarlo?







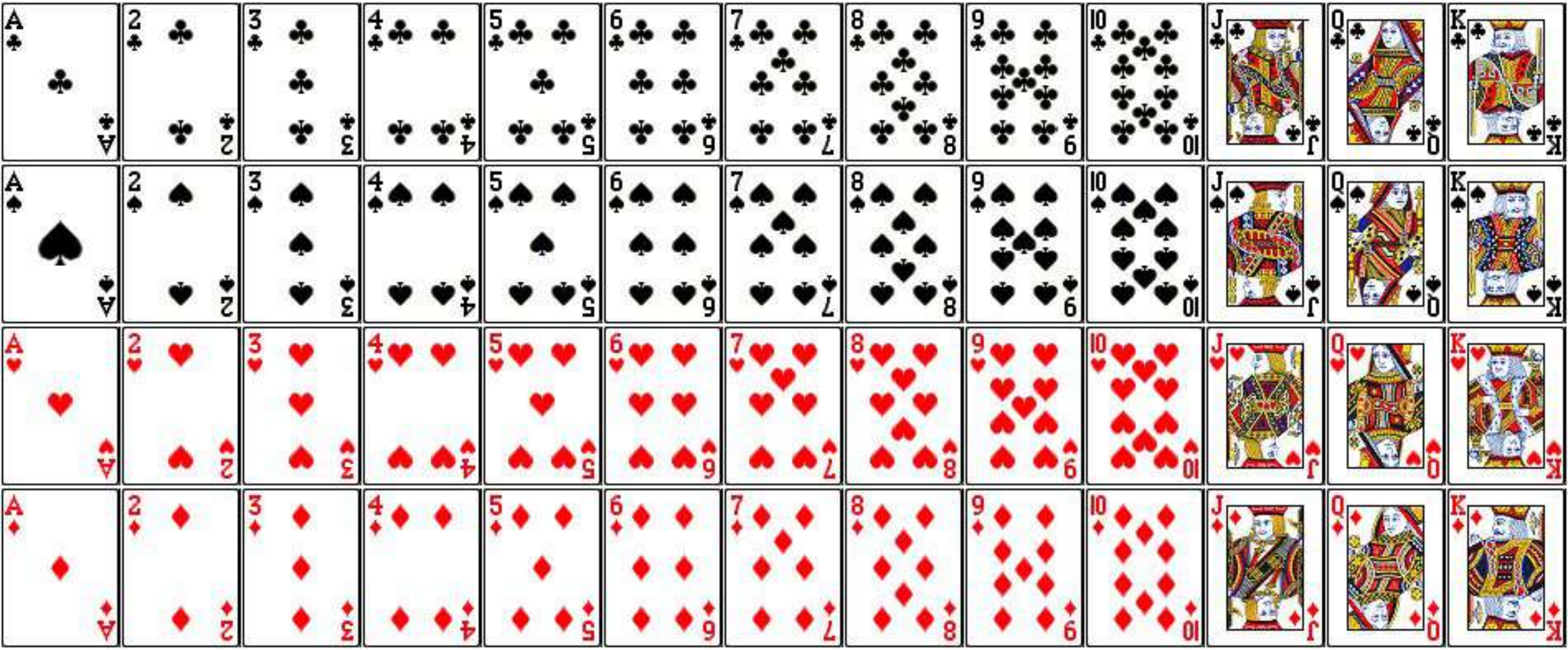


Simulaciones de Montecarlo

- Permite crear simulaciones para predecir el resultado de un problema
- Permite convertir problemas determinísticos en problemas estocásticos
- Es utilizado en una gran diversidad de áreas, desde la ingeniería hasta la biología y el derecho.

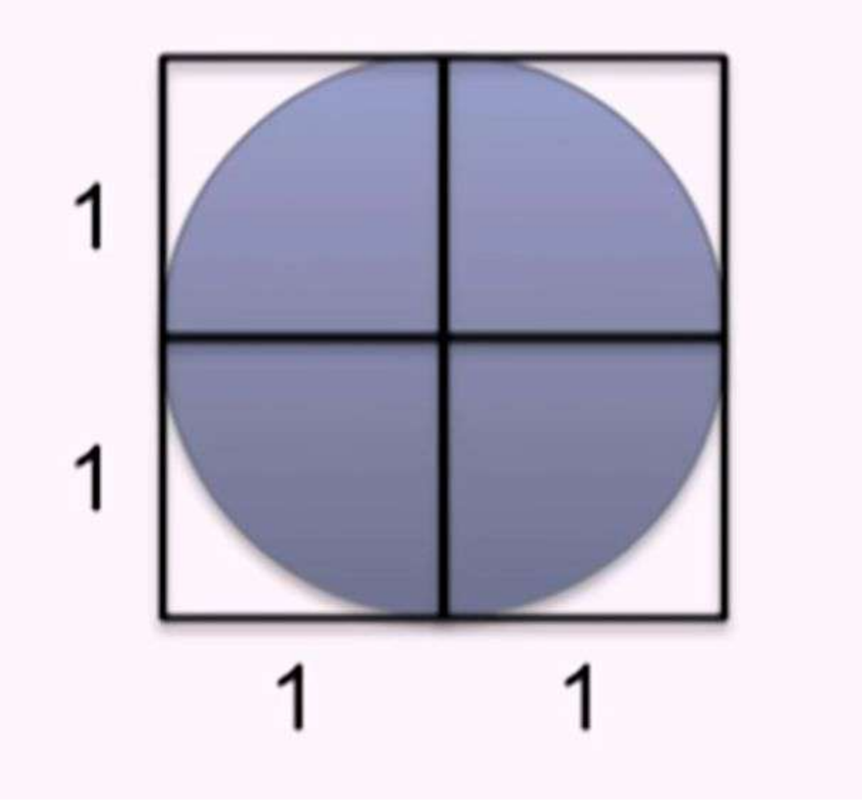
Simulación de Barajas

Simulación de barajas



Cálculo de Pi

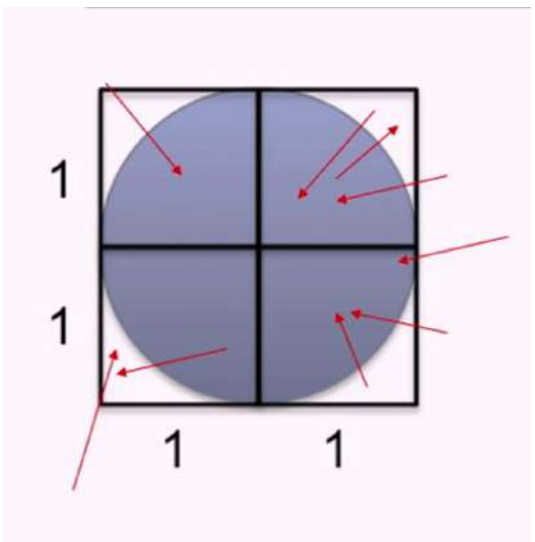
Cálculo de Pi



Acuad = b * h
Acirc = Πr**2

Acuad = 2 * 2 = 4
Acirc = Π

Cálculo de Pi



$$\frac{\text{agujas en circ}}{\text{agujas en cuad}} = \frac{\text{área circ}}{\text{área cuad}}$$

$$\text{área circ} = \frac{\text{área cuad} * \text{agujas circ}}{\text{agujas cuad}}$$

$$\text{área circ} = \frac{4 * \text{agujas circ}}{\text{agujas cuad}}$$

Muestreo

Muestreo

- Hay ocasiones en la que no tenemos acceso a toda la población que queremos explorar
- Uno de los grandes descubrimientos de la estadística es que las muestras aleatorias tienden a mostrar las mismas propiedades de la población objetivo
- El tipo de muestreo que hemos hecho hasta ahora es muestreo probabilístico

Muestreo

- En un muestreo aleatorio cualquier miembro de la población tiene la misma probabilidad de ser escogido

- En un muestreo estratificado tomamos en consideración las características de la población para partirla en subgrupos y luego tomamos muestras de cada subgrupo.
    - Incrementa la probabilidad de que el muestreo sea representativo de la población

Muestreo estratificado

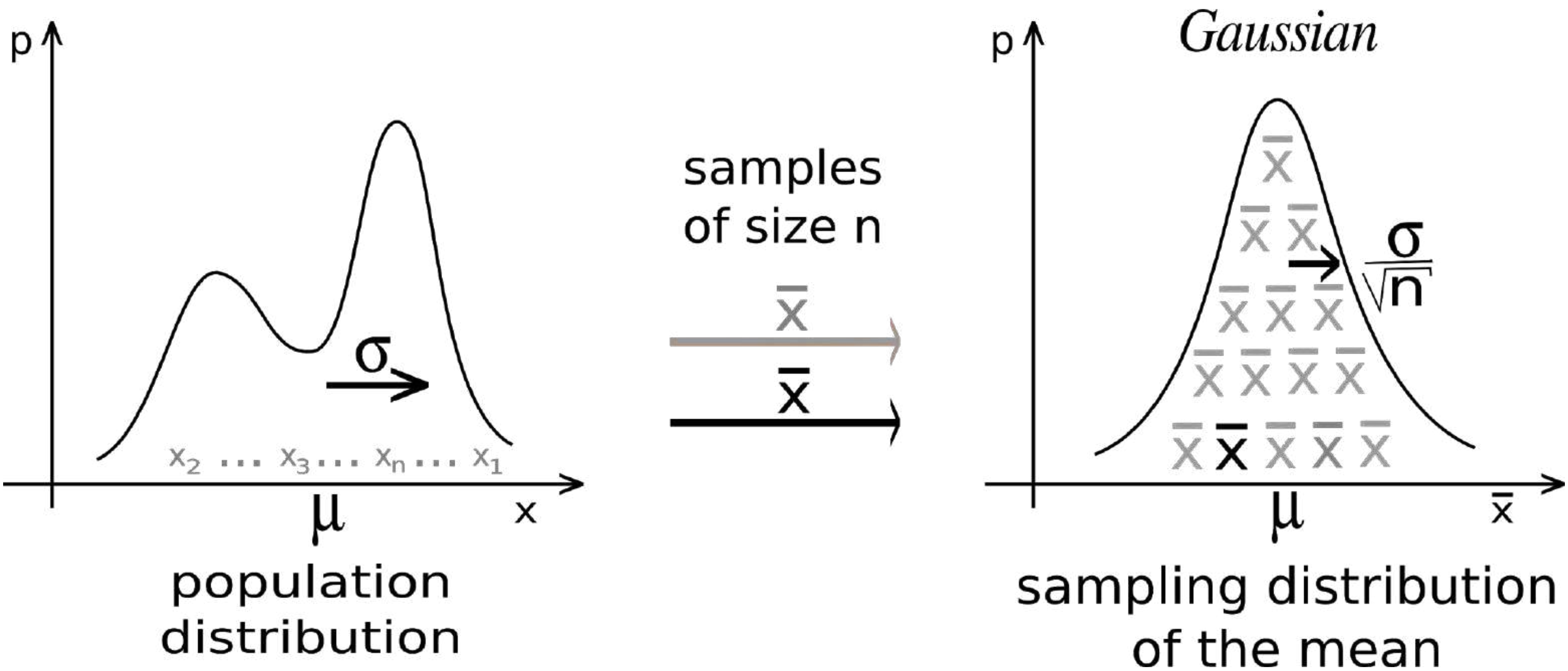
Teorema del límite central

Teorema del límite central

- Es uno de los teoremas más importantes de la estadística
- Establece que muestras aleatorias de cualquier distribución van a tener una distribución normal
- Permite entender cualquier distribución como la distribución normal de sus medias y eso nos permite aplicar todo lo que sabemos de distribuciones normales

Teorema del límite central

- Mientras más muestras obtengamos, mayor será la similitud con la distribución normal
- Mientras la muestra sea de mayor tamaño, la desviación estándar será menor

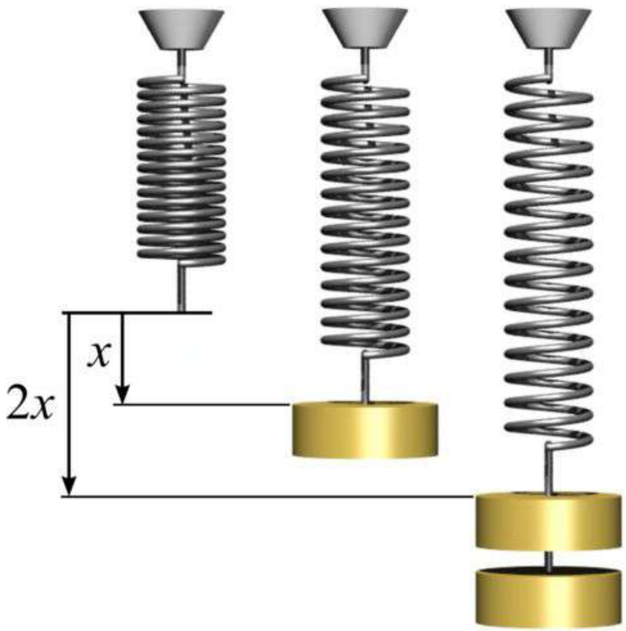

Gaussian
p
σ
samples
of size n
$\bar{x}$
$\bar{x}$
$\frac{\sigma}{\sqrt{n}}$
x_2 ... x_3 ... x_n ... x_1
μ
x
$\bar{x}$
population
distribution
sampling distribution
of the mean

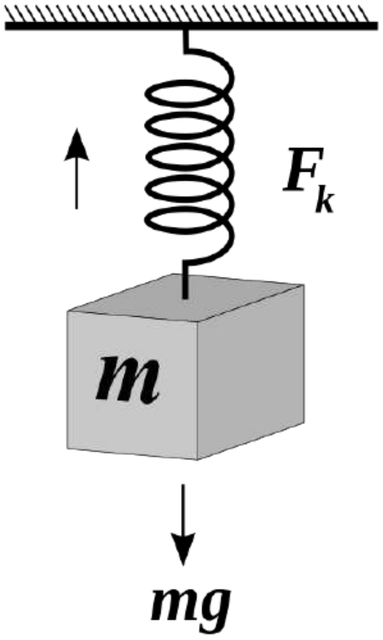
¿Cómo trabajar con datos experimentales?

Datos experimentales

- Es la aplicación del método científico
- Es necesario comenzar con una teoría o hipótesis sobre el resultado al que se quiere llegar
- Basado en la hipótesis se debe crear un experimento para validar o falsear la hipótesis
- Se valida o falsea una hipótesis midiendo la diferencia entre las mediciones experimentales y aquellas mediciones predichas por la hipótesis

Ley de la elasticidad de Hooke





$$F = -k * \delta$$

Regresión Lineal

Regresión lineal

- Permite aproximar una función a un conjunto de datos obtenidos de manera experimental
- No necesariamente permite aproximar funciones lineales, sino que sus variantes permiten aproximar cualquier función polinómica

Conclusiones

Conclusiones

- La programación dinámica permite optimizar problemas que tienen subestructura óptima y subproblemas empalmados
- Las computadoras pueden resolver problemas determinísticos y estocásticos
- Podemos generar simulaciones computacionales para responder preguntas del mundo real
- La inferencia estadística nos permite tener confianza de que nuestras simulaciones arrojan resultados válidos