Investigación Operativa
Simulación
Clase 01

Investigación Operativa UTN FRBA
Curso: I4051
Docente: Martin Palazzo

Simulación

Primer Parcial

1. ***Simulación***
2. Procesos estocásticos: Cadenas de Markov
3. Filas de Espera
4. Grafos y Redes de Proyectos

Segundo Parcial

1. Programación Lineal
2. Algoritmo del Simplex, Solución Dual y Primal
3. Transporte y Asignación
4. Inventarios

Procéso Estocástico

Es aquel caracterizado por:

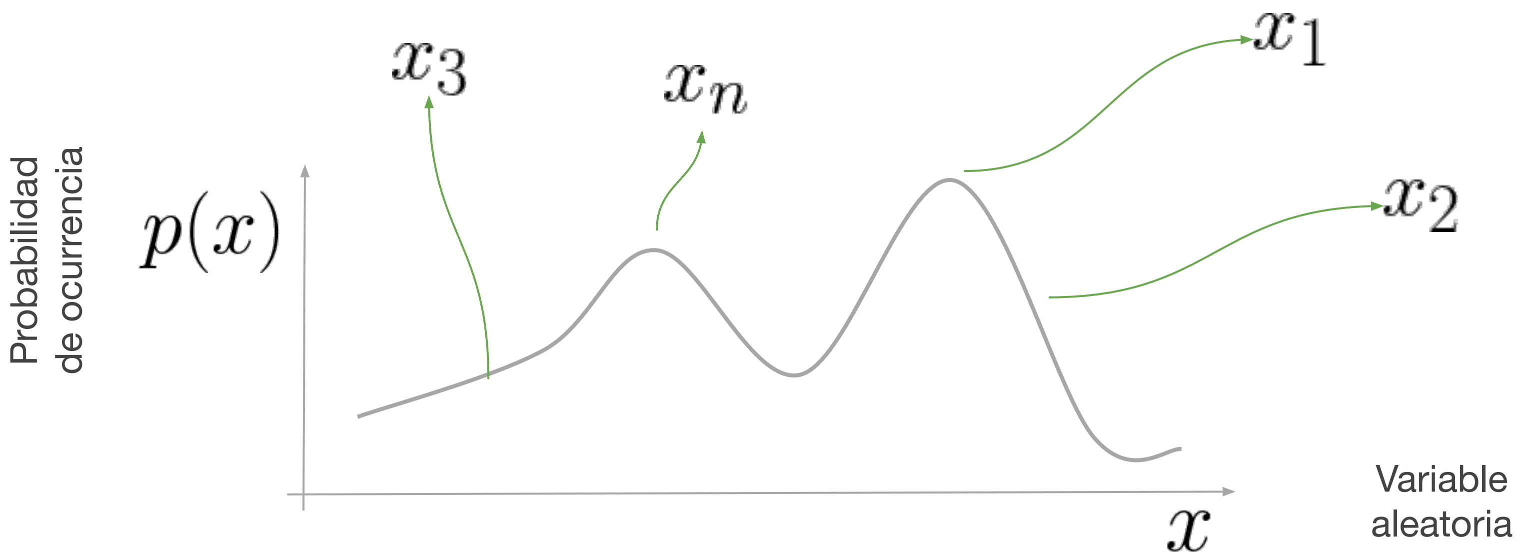
- una o más variables aleatorias (VA)
- Las VAs evolucionan en el tiempo bajo cierta distribución de probabilidad.

Objetivo de la simulación

El objetivo que queremos lograr es poder simular un proceso aleatorio (una variable aleatoria) con una **función de densidad de probabilidad** determinada.

Existen muchas funciones de densidad de probabilidad. Algunas populares como la Normal, la Poisson, la Exponencial y otras arbitrarias determinadas por el problema.

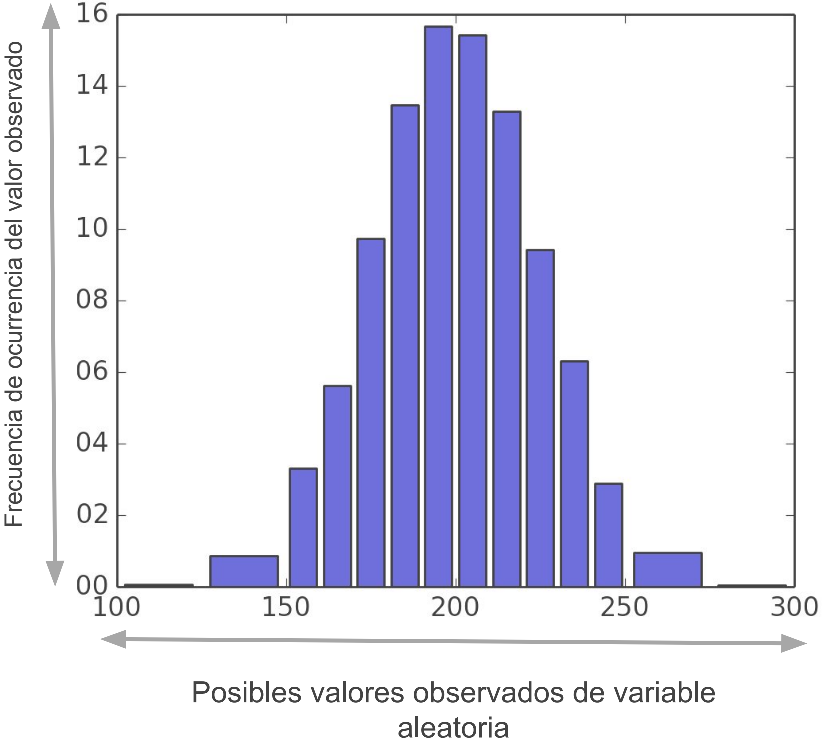
Muestreo desde una función de densidad de probabilidad



Suponiendo que **conocemos** la función de densidad de probabilidad de una variable aleatoria, vamos a **muestrear** multiples veces dicha funcion y obtener distintos valores de la variable aleatoria a simular. En el caso contrario si solo tenemos los datos y no conocemos la funcion de densidad que los genero se abordaran estrategias de maxima verosimilitud o metodos de estimacion no parametrica de la densidad [2]

[1] https://en.wikipedia.org/wiki/Maximum_likelihood_estimation
[2] https://en.wikipedia.org/wiki/Kernel_density_estimation .

Histograma de frecuencias



El histograma representa la frecuencia relativa de aparición de un valor de la variable aleatoria mediante la altura de las barras.

En el eje X tendremos los distintos valores que puede tomar una variable aleatoria a observar. En vez de contar valores únicos contamos todos los valores que caigan en un rango, es decir, la primer barra por ejemplo cuenta la cantidad de veces que la VA tomo los valores entre 100 y 125. La segunda barra cuenta la cantidad de veces que la VA tomo valores entre 125 y 150, etc.

Entonces al tomar muchas muestras (muestrear, samplear) una variable aleatoria podemos empíricamente entender cómo se distribuye los valores que la VA puede tomar. Entonces podemos decir que con un histograma podemos aproximar empíricamente la distribución de probabilidad.

Histograma de frecuencias

Cantidad de muestras por bin/caja

$$n_k = \sum \delta(x_{(kj)})$$

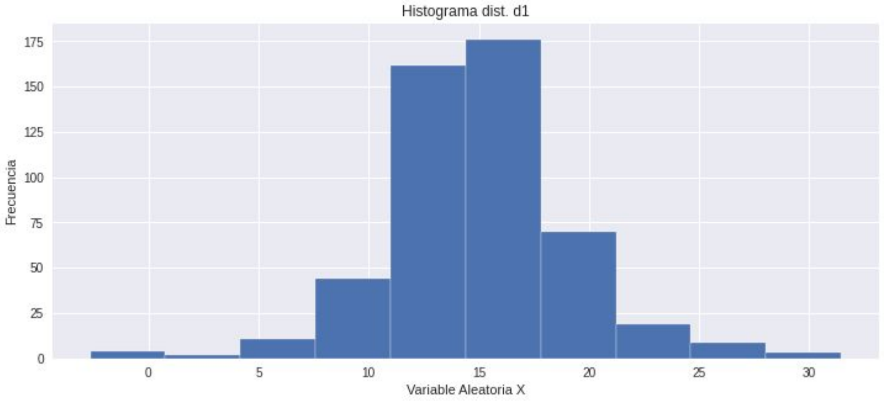
Funcion delta (contador)

$$\delta(x_{(ij)}) = 1$$

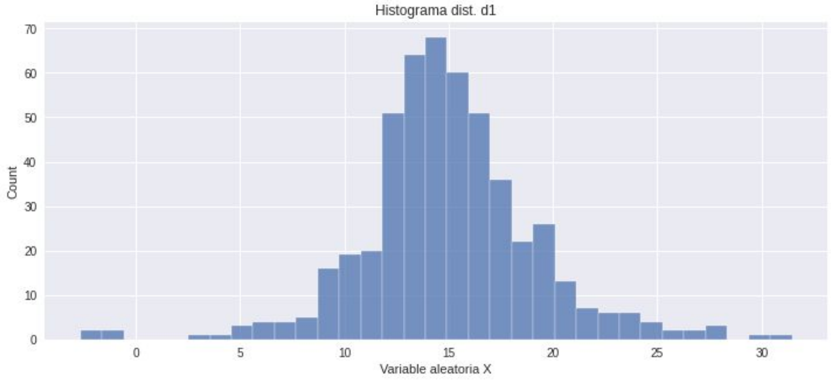
Muestras totales en los K bins

$$n = \sum_{i=1}^{k} \sum_{j=0}^{n_k} \delta(x_{(kj)})$$

Histograma de frecuencias

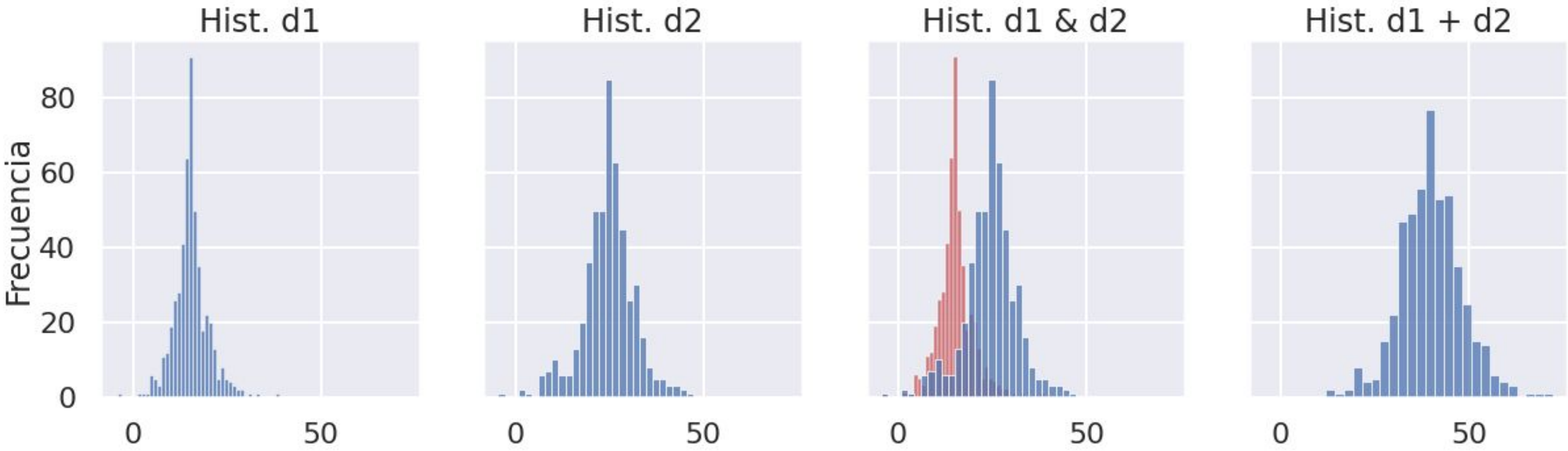


Histograma con bins = 10



Histograma con bins = 40

Histogramas



Histograma sobre 500 muestras de una distribución d1 normal mu = 15, sigma = 3

Histograma sobre 500 muestras de una distribución d2 normal mu = 25, sigma = 5

Los dos histogramas en simultáneo.

Histograma realizado sobre la suma de las 2 muestras obtenidas de las distribuciones d1 y d2.

Elementos para definir un sistema de simulación

1. Definir el estado del sistema.
2. Identificar los estados posibles del sistema que pueden ocurrir.
3. Identificar los eventos posibles que cambian el estado del sistema.
4. Contar con un reloj de simulación, localizado en alguna dirección del programa de simulación, que registrará el paso del tiempo (simulado).
5. Un método para generar los eventos de manera aleatoria de los distintos tipos.
6. Una fórmula para identificar las transiciones de los estados que generan los diferentes tipos de eventos.

Procesos estocásticos: categorización según memoria

Proceso estocástico

Proceso aleatorio puro

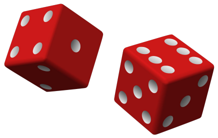
Proceso de markov

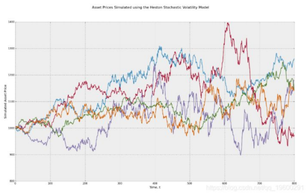
Procesos con memoria

El estado del sistema en cada tiempo t es independiente de cualquier estado pasado. La probabilidad de estado se determinará por una variable aleatoria.

El estado del sistema en cada tiempo t es determinado directamente por el estado del sistema en el tiempo t-1 . Es un proceso con memoria = 1 .

El estado del sistema en cada tiempo t es determinado por los estados que tomo el sistema en un período de tiempo $\Delta t = m$ con m> 1 .





Procesos estocásticos: categorización según variable aleatoria

Proceso estocástico

Variable aleatoria discreta

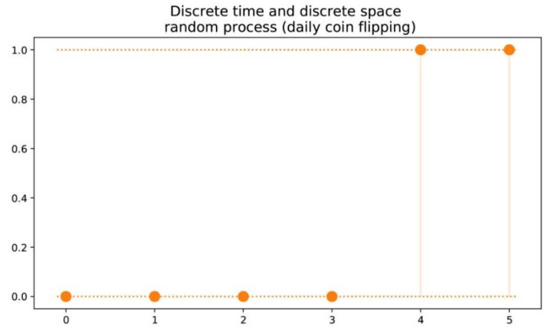
El estado del sistema toma valores enteros. Se puede discretizar el sistema en n estados.

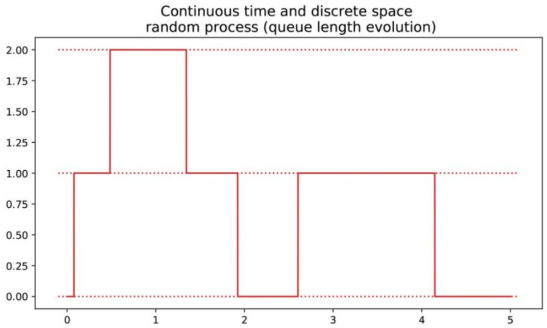
Variable aleatoria contínua

El estado del sistema toma valores reales continuos. El sistema no se discretiza en estados discretos.

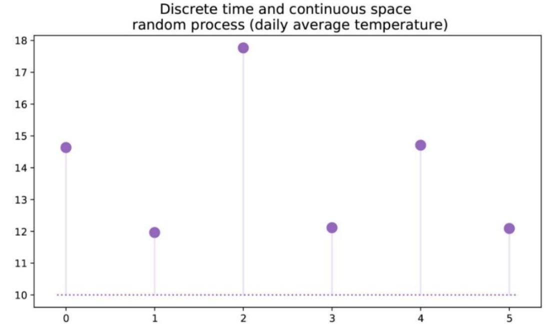
Procesos estocásticos: categorías por V.A. y parámetro

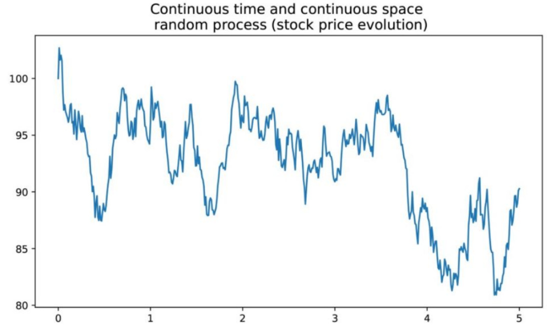
Variable aleatoria discreta





Variable aleatoria contínua





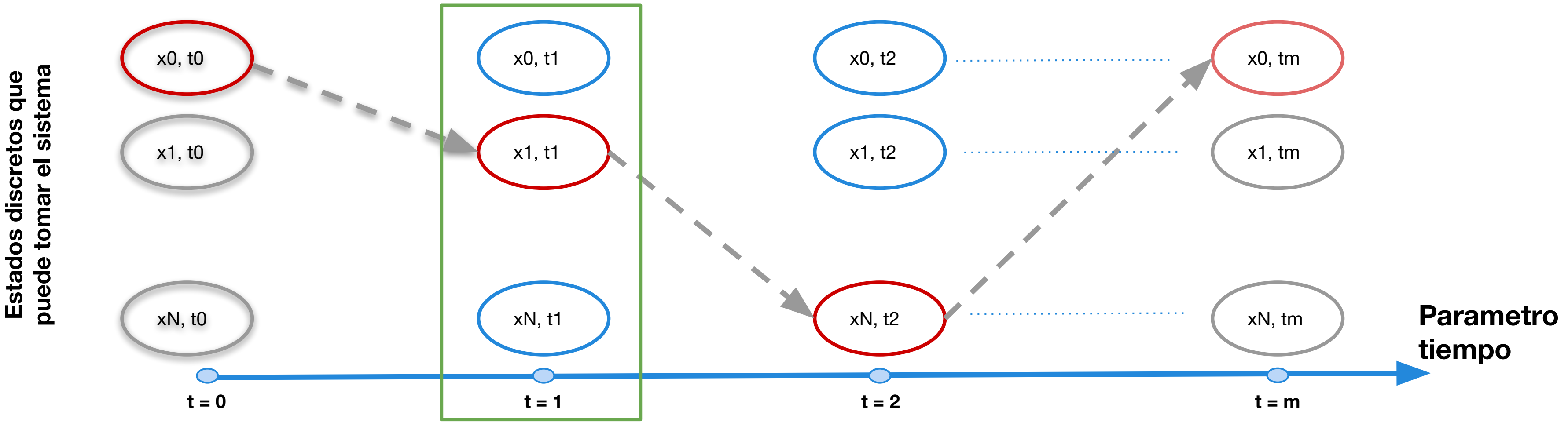
Parámetro tiempo discreto

Parámetro tiempo contínuo

*image credits: https://towardsdatascience.com/brief-introduction-to-markov-chains-2c8cab9c98ab

Procesos estocásticos

- El **estado** de un sistema es un vector de números o símbolos que describen completamente el estado de situación del sistema en cada momento



Vamos a definir un sistema caracterizado por N posibles estados (discretos o continuos). En cada instante "t" el sistema tomará uno de los N posibles estados.

Procesos estocásticos

Proba estados

t = 0

$$P(x=0 \mid t=0)$$
$$P(x=1 \mid t=0)$$
$$P(x=2 \mid t=0)$$
$$\dots$$
$$P(x=i \mid t=0)$$
$$\dots$$
$$P(x=n \mid t=0)$$

t = 1

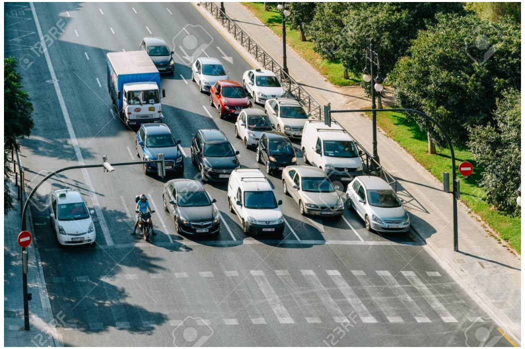
$$P(x=0 \mid t=1)$$
$$P(x=1 \mid t=1)$$
$$P(x=2 \mid t=1)$$
$$\dots$$
$$P(x=i \mid t=1)$$
$$\dots$$
$$P(x=n \mid t=1)$$

t = 2

$$P(x=0 \mid t=2)$$
$$P(x=1 \mid t=2)$$
$$P(x=2 \mid t=2)$$
$$\dots$$
$$P(x=i \mid t=2)$$
$$\dots$$
$$P(x=n \mid t=2)$$

El sistema se modela compuesto por **variables aleatorias x = f(t)** que tomarán distintos valores a medida que evoluciona el parámetro (tiempo). Para cada estado posible del sistema, dado un instante de tiempo "t" existirá una probabilidad asociada **P(x | t)**.

Procesos estocásticos



Proba estados

t = 0

$$P(x = 0 \mid t = 0)$$
$$P(x = 1 \mid t = 0)$$
$$P(x = 2 \mid t = 0)$$
$$\dots$$
$$P(x = i \mid t = 0)$$
$$\dots$$
$$P(x = n \mid t = 0)$$

t = 1

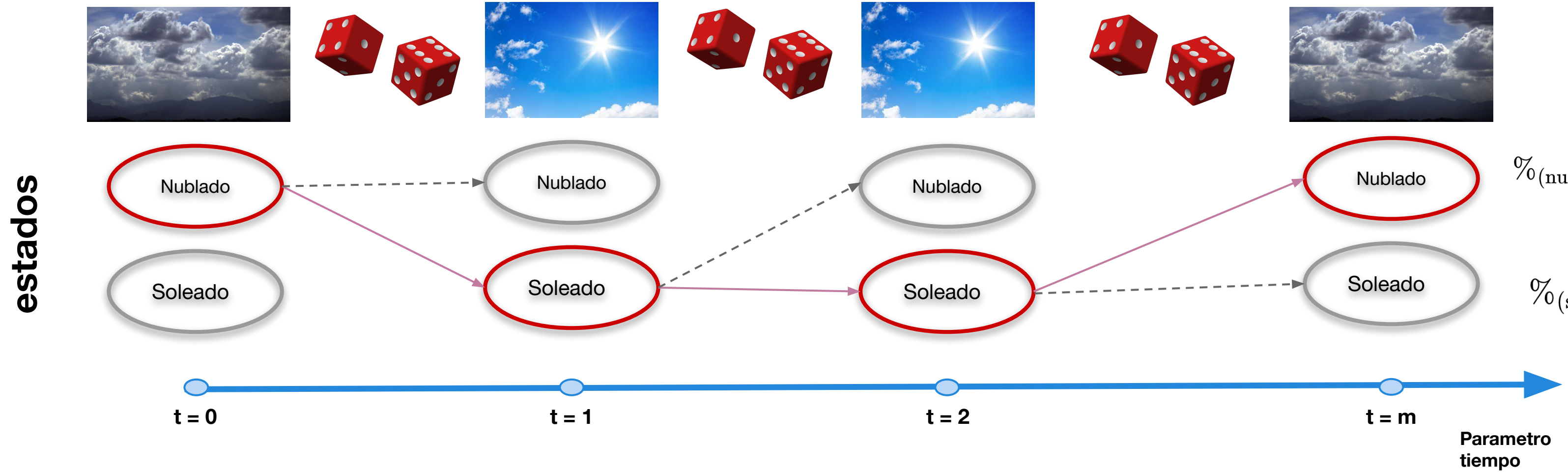
$$P(x = 0 \mid t = 1)$$
$$P(x = 1 \mid t = 1)$$
$$P(x = 2 \mid t = 1)$$
$$\dots$$
$$P(x = i \mid t = 1)$$
$$\dots$$
$$P(x = n \mid t = 1)$$

t = 2

$$P(x = 0 \mid t = 2)$$
$$P(x = 1 \mid t = 2)$$
$$P(x = 2 \mid t = 2)$$
$$\dots$$
$$P(x = i \mid t = 2)$$
$$\dots$$
$$P(x = n \mid t = 2)$$

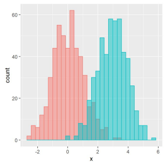
Estado X = Cantidad de vehículos esperando en el semáforo.

Procesos estocásticos: obteniendo la densidad empíricamente



$$\%_{(\text{nubes})} = \sum_{i=1}^{m} \frac{(x_i = \text{nubes})}{m}$$

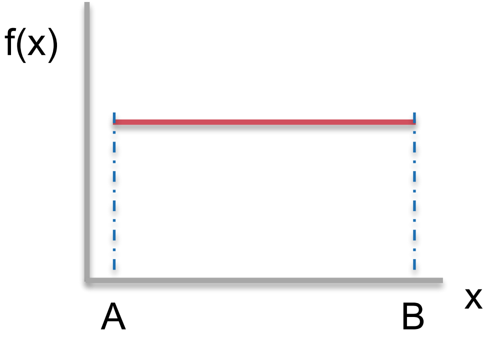
$$\%_{(\text{sol})} = \sum_{i=1}^{m} \frac{(x_i = \text{sol})}{m}$$



Supongamos que observamos un sistema con N = 2 estados discretos posibles: soleado y nublado. Suponemos que cada dia el clima depende de una distribución de probabilidad que no conocemos y es independiente del clima de los días anteriores. Para “aprender” la distribución de probabilidad de los estados existentes lo que podemos hacer es observar M veces el sistema, contar cuantas veces ocurrió cada estado, armar un histograma de frecuencias y asi estimar la posible distribución de probabilidad que gobierna dichos estados en el sistema.

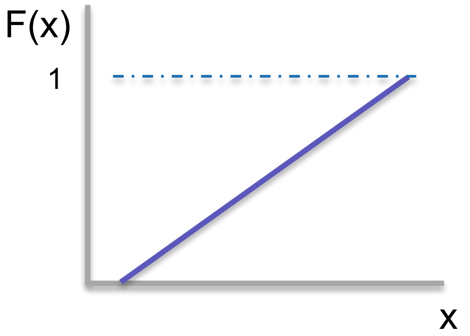
distribución uniforme

Distribución de densidad de probabilidad.



$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if} \quad a \leq x \leq b \\ 0 & \text{else} \end{cases}$$

Distribución de probabilidad acumulada.



$$F(x) = \begin{cases} 0 & \text{if} \quad x < a \\ \frac{x-a}{b-a} & \text{if} \ a \leq x \leq b \\ 1 & \text{if} \quad x > b \end{cases}$$

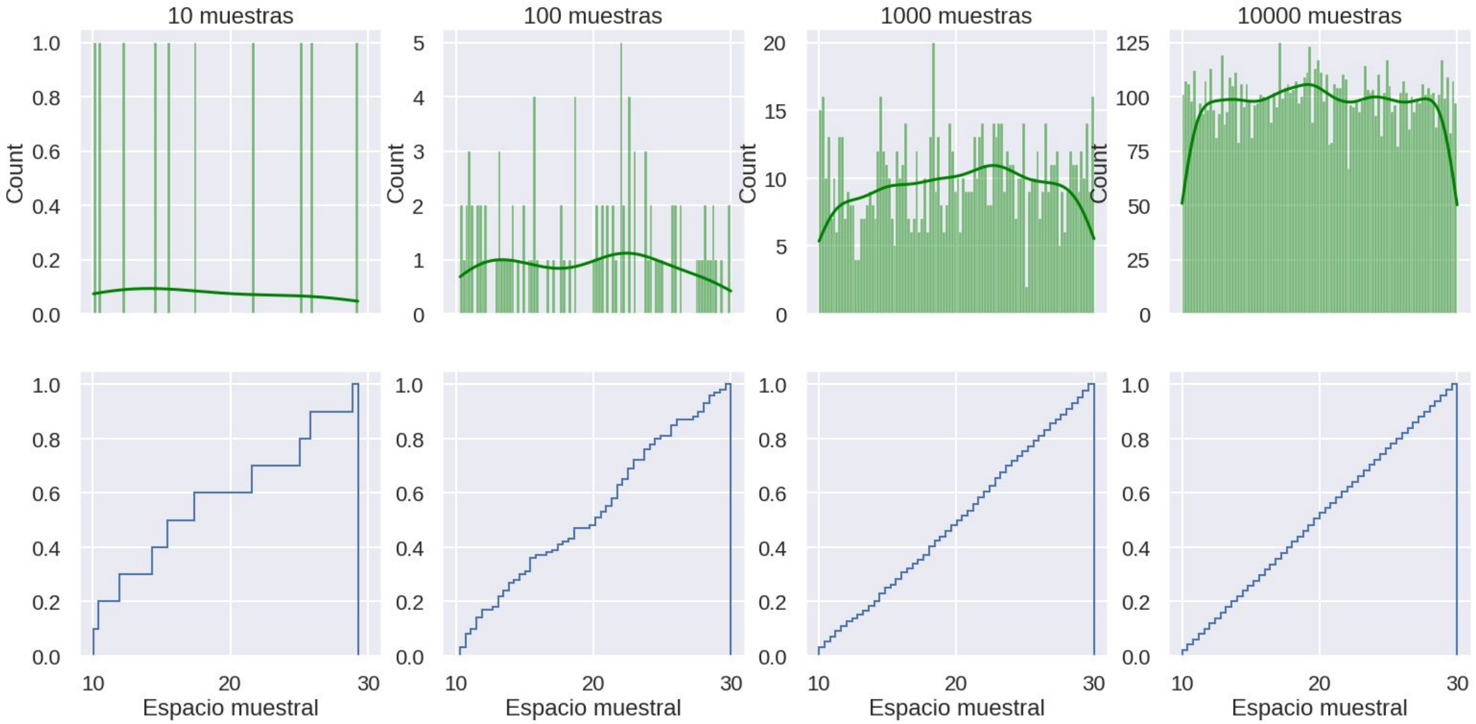
Rango de valores posibles.

La distribución de probabilidad uniforme asigna la misma probabilidad de ocurrencia a cada valor dentro del rango que puede generar una variable aleatoria.

muestreo desde distribución uniforme

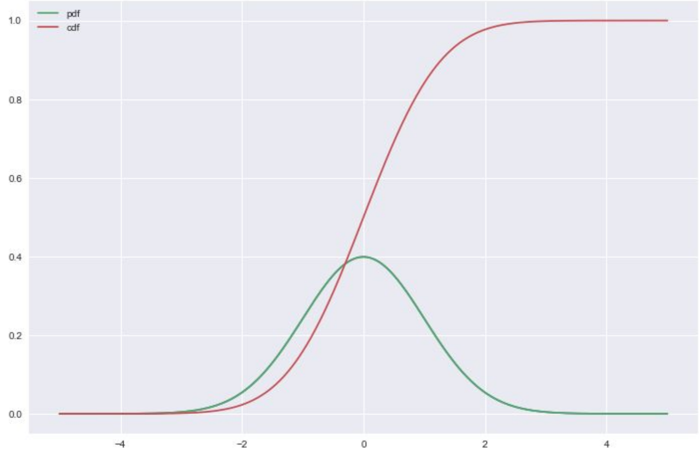
Probabilidad
Empirica
De ocurrencia

Probabilidad
Acumulada de
Ocurrencia



distribución gaussiana - normal

Distribución de densidad (verde) y acumulada (roja) de probabilidad.



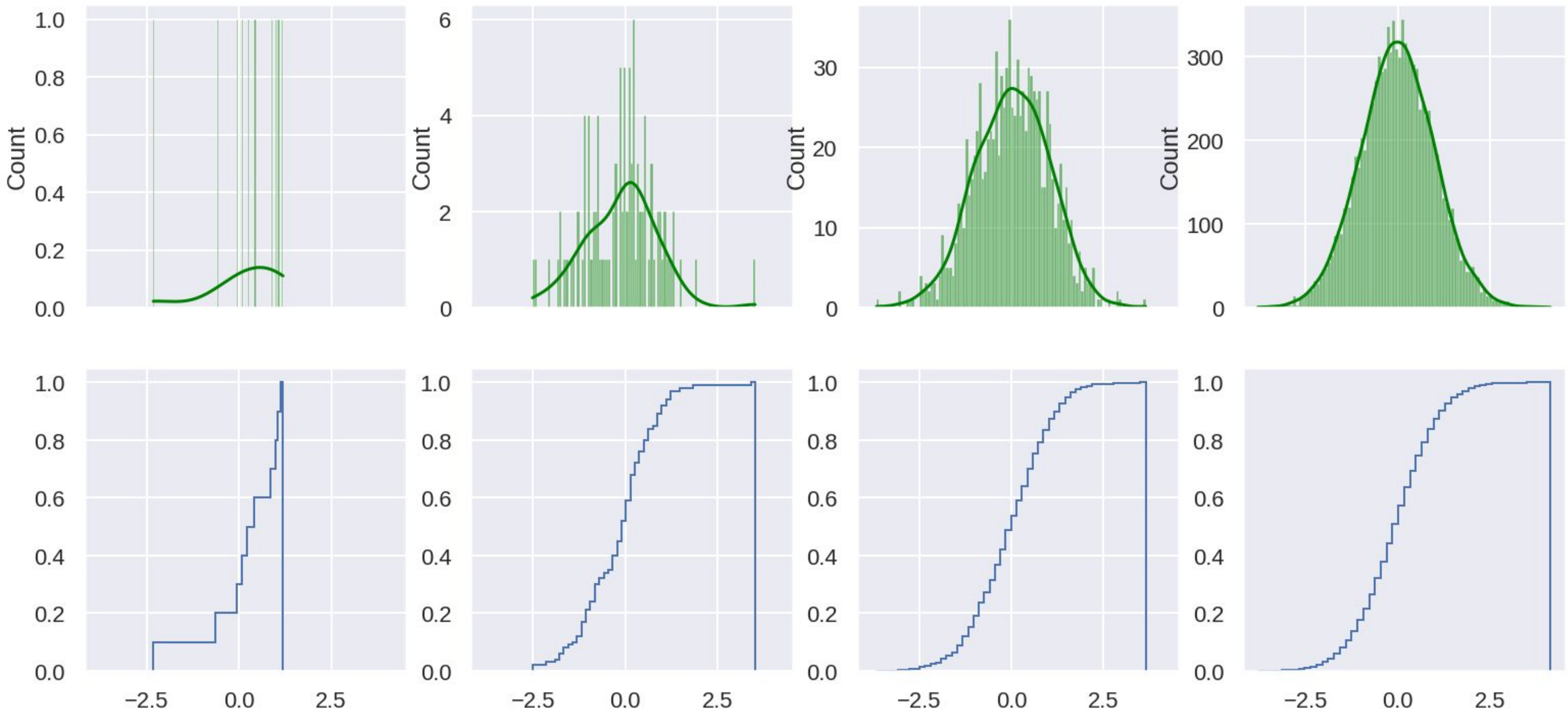
Rango de valores posibles de la VA.

Distribución simétrica de VA continua. El parámetro mu define la esperanza y el sigma el desvío standard.

Suele utilizarse para modelar procesos reales en ciencias naturales, sociales, etc.

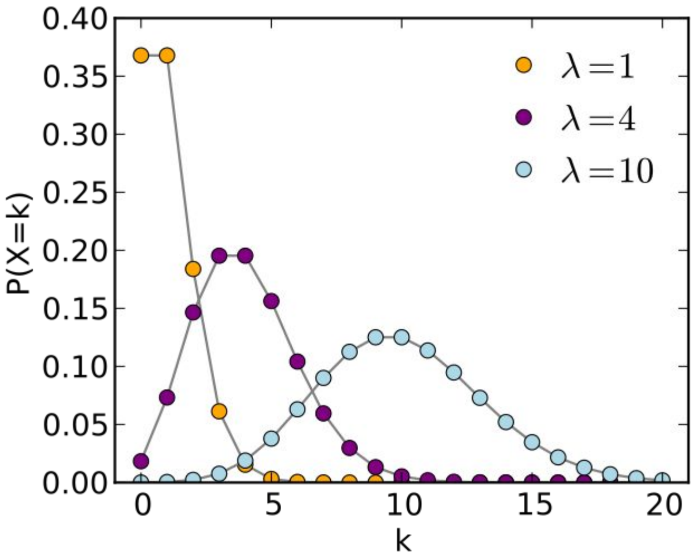
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

Muestreo de una distribución normal



distribución poisson

Distribución de densidad para distintos valores del parámetro lambda.



Rango de valores posibles de la VA.

Distribución simétrica de VA discreta. El parámetro lambda define la frecuencia de ocurrencia de eventos media.

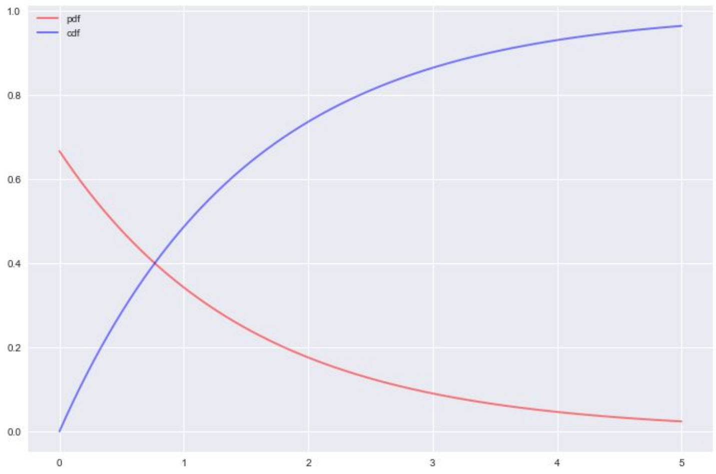
Se utiliza para modelar la **cantidad de eventos discretos en un intervalo de tiempo fijo.**

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

distribución exponencial

Distribución de densidad (rojo).

Acumulada (azul).



Rango de valores posibles de la VA.

Distribución simétrica de VA continua. El parámetro lambda define la frecuencia de ocurrencia de eventos media (es el mismo lambda que la poisson).

Se utiliza para modelar el **tiempo que transcurre entre dos eventos de un proceso de Poisson.**

$$f(x) = \lambda \mathrm{e}^{-\lambda x}$$

Caso simulación con 2 dados

simulación: ejercicio 2 dados

Simular un sistema con **2 dados** equilibrados, donde el resultado de cada jugada es la suma del resultado de cada dado. En función del resultado de cada jugada, determinar si conviene apostar de manera recurrente en un sistema de juego donde se gana 1 bitcoin cuando la suma de los 2 dados supera el valor de 6, 7 u 8, según cada caso y se pierde un bitcoin si sale la opción opuesta. Determinar rango, espacio muestral, distribución de probabilidad utilizada en cada dado y en conjunto.

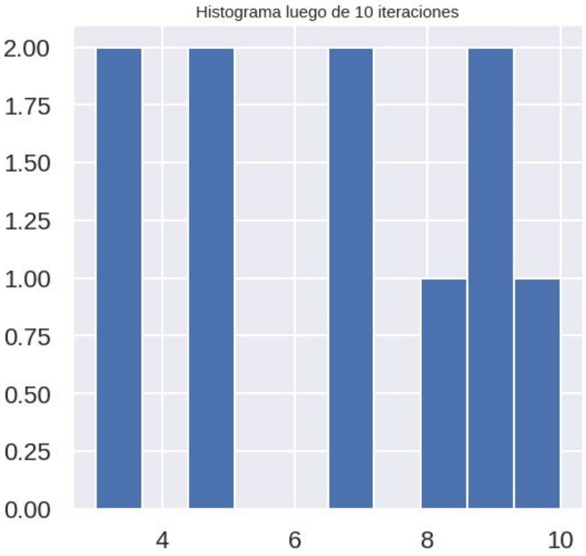
simulación: ejercicio 2 dados

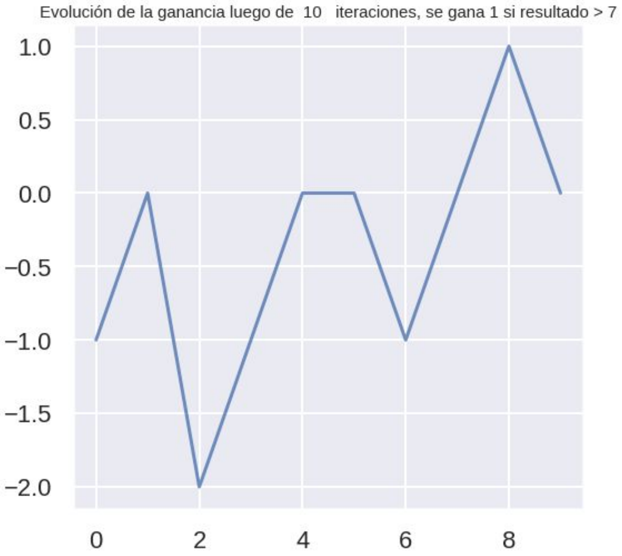
Distribución de probabilidad utilizada: uniforme

Espacio muestral:
{11,12,13,14,15,16,21,22,23,24,25,26,31,32,33,34,35,36,41,42,43,44,45,46,51,52,53,54,55,56,61,62,63,64,65,66}

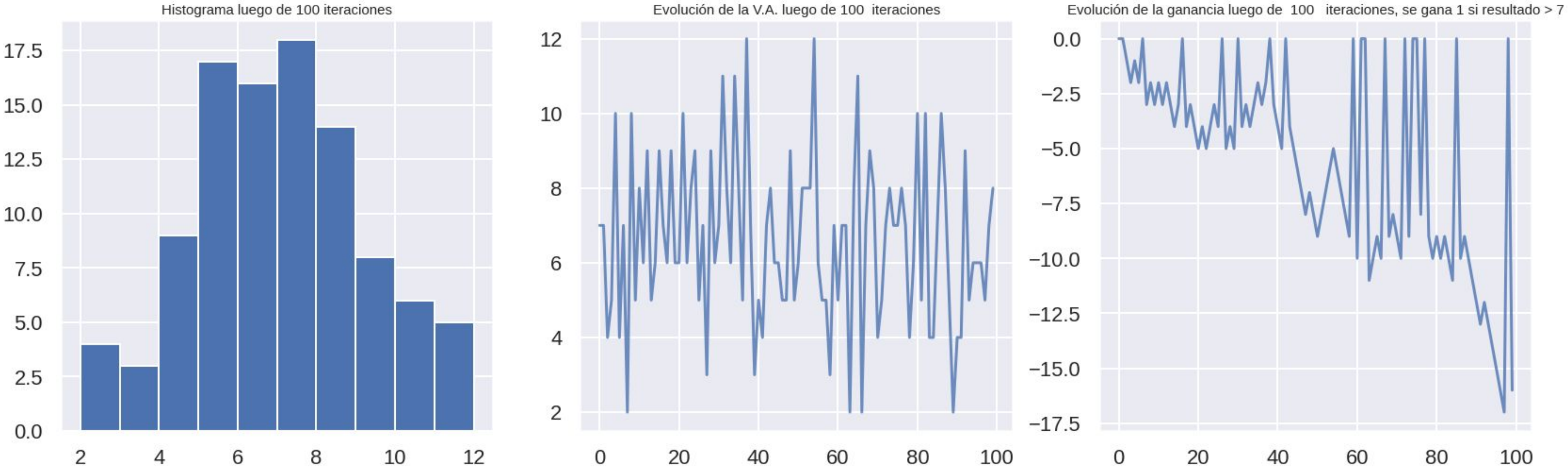
Rango: {2,3,4,5,6,7,8,9,10,11,12}

n = 10 muestras

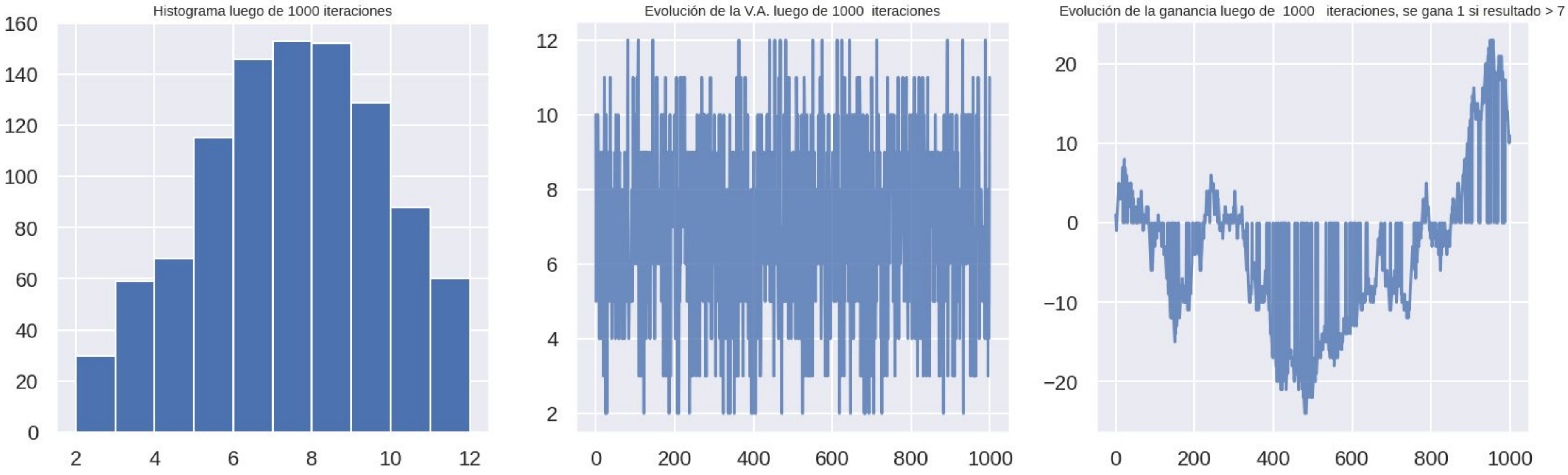




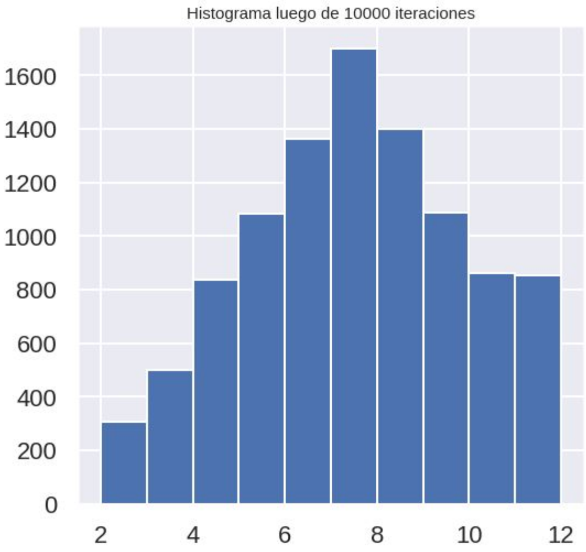
n = 100 muestras

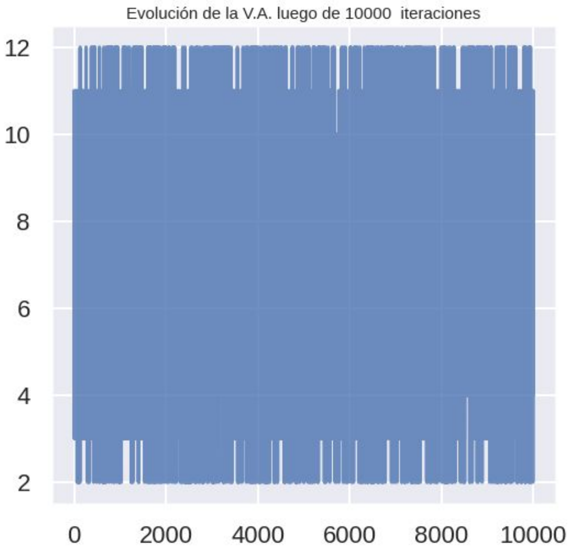


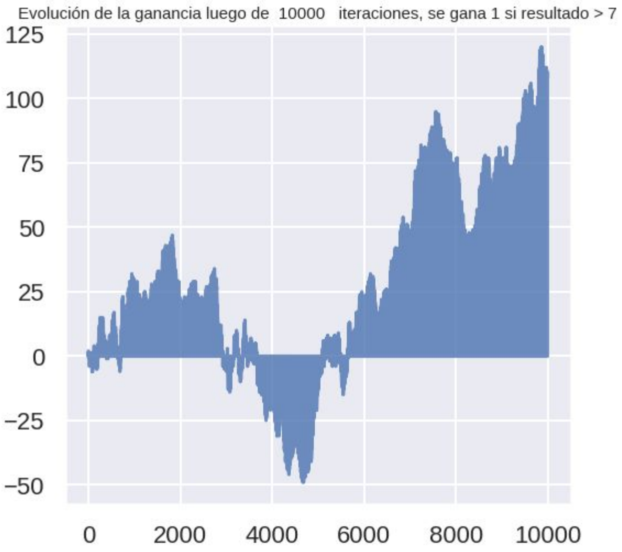
n = 1000 muestras



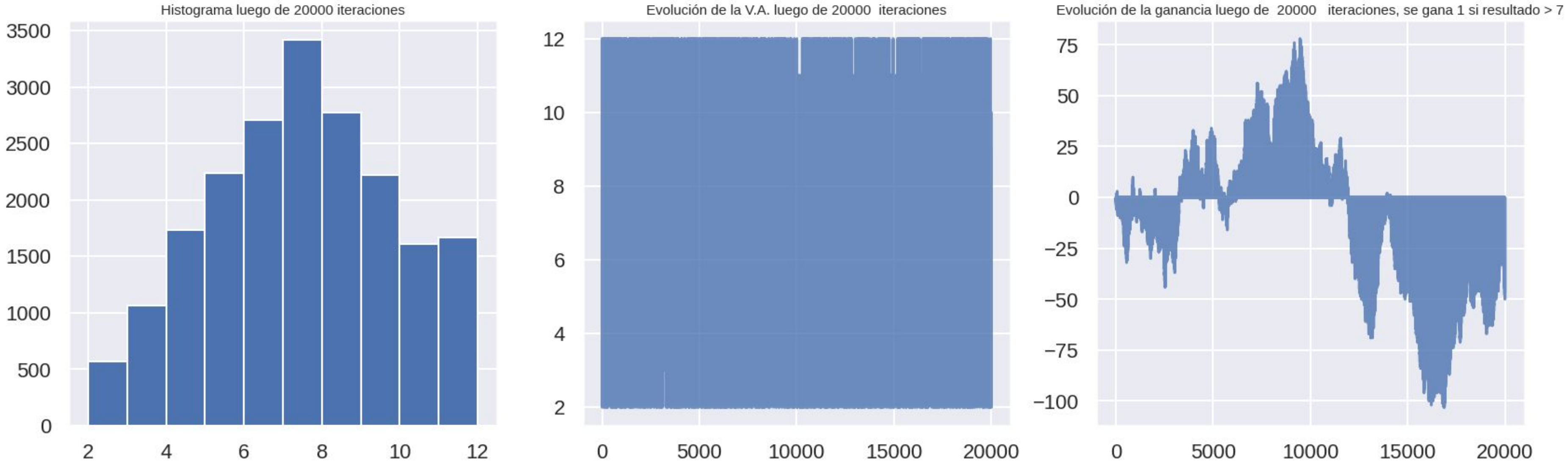
n = 10000 muestras







n = 20000 muestras



Método de la transformada inversa

Simulacion

Método de la transformada inversa

Objetivo: Generación de observaciones aleatorias a partir de una distribución de probabilidad determinada

Partiendo de una sucesión de números aleatorios uniformes queremos generar una sucesión de observaciones aleatorias que sigan una distribución de probabilidad determinada (p. 890 Hillier).

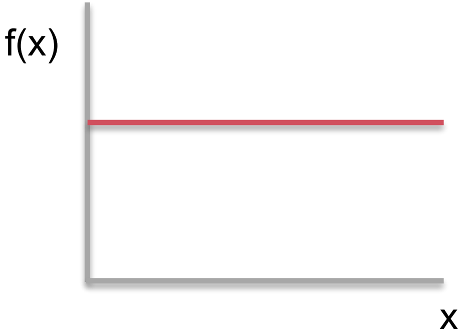
1. Generaremos números aleatorios uniformes (ejemplo: un dado de N caras) entre 0 y 1 representados con ‘U’
2. Establecer **F(x) = u** siendo F(x) la distribución acumulada de probabilidad.
3. Despejar x que será la observación simulada

método de la transformada inversa: distribución exponencial

Función de probabilidad uniforme (dado de n caras)

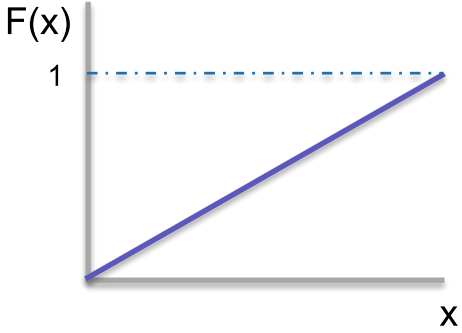
Función de probabilidad Exponencial

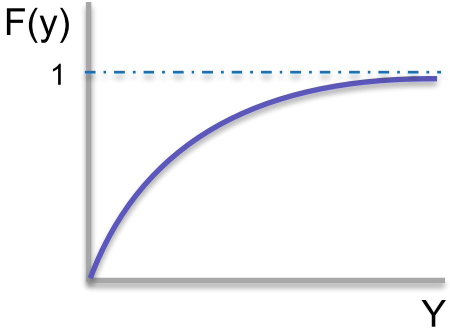
Función de densidad



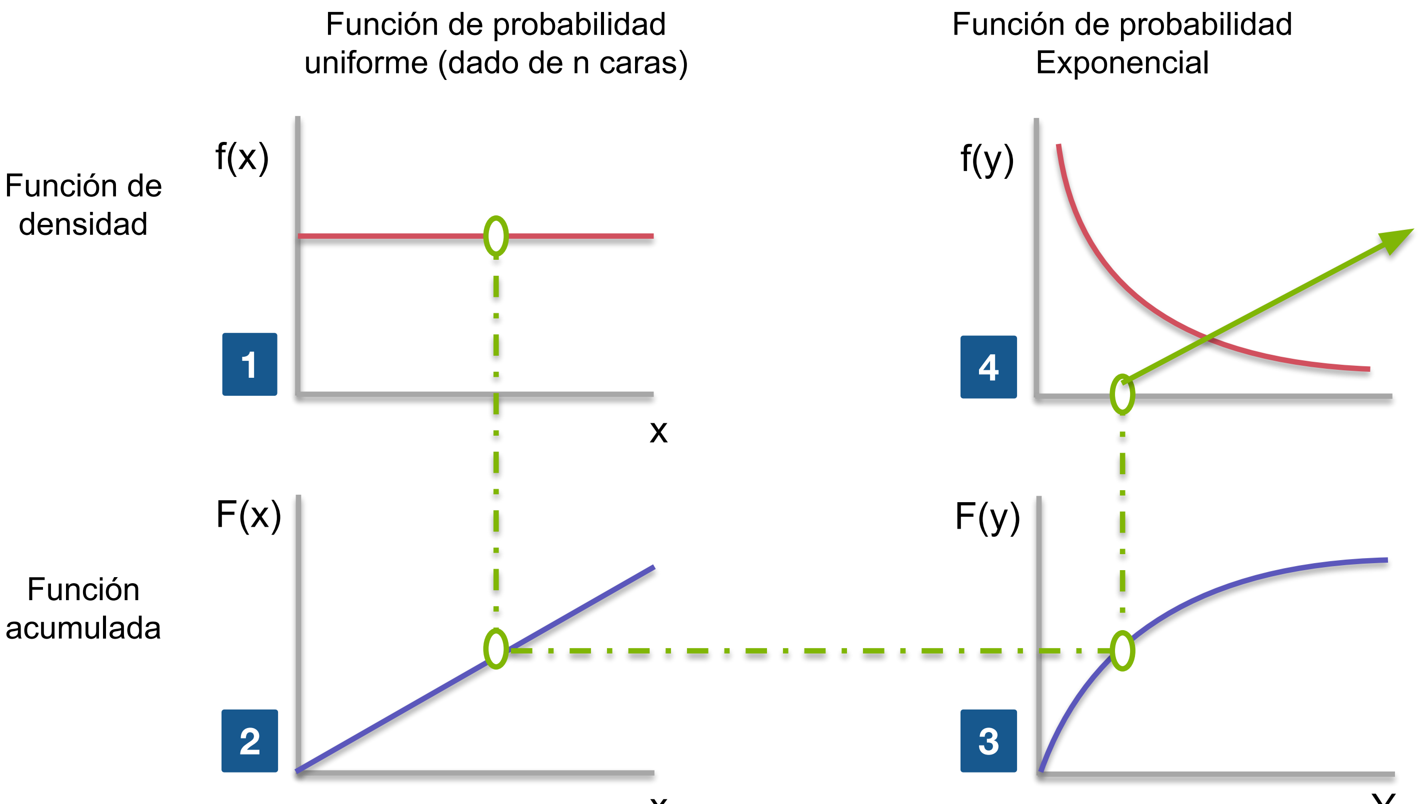
Distribución de probabilidad a simular

Función acumulada



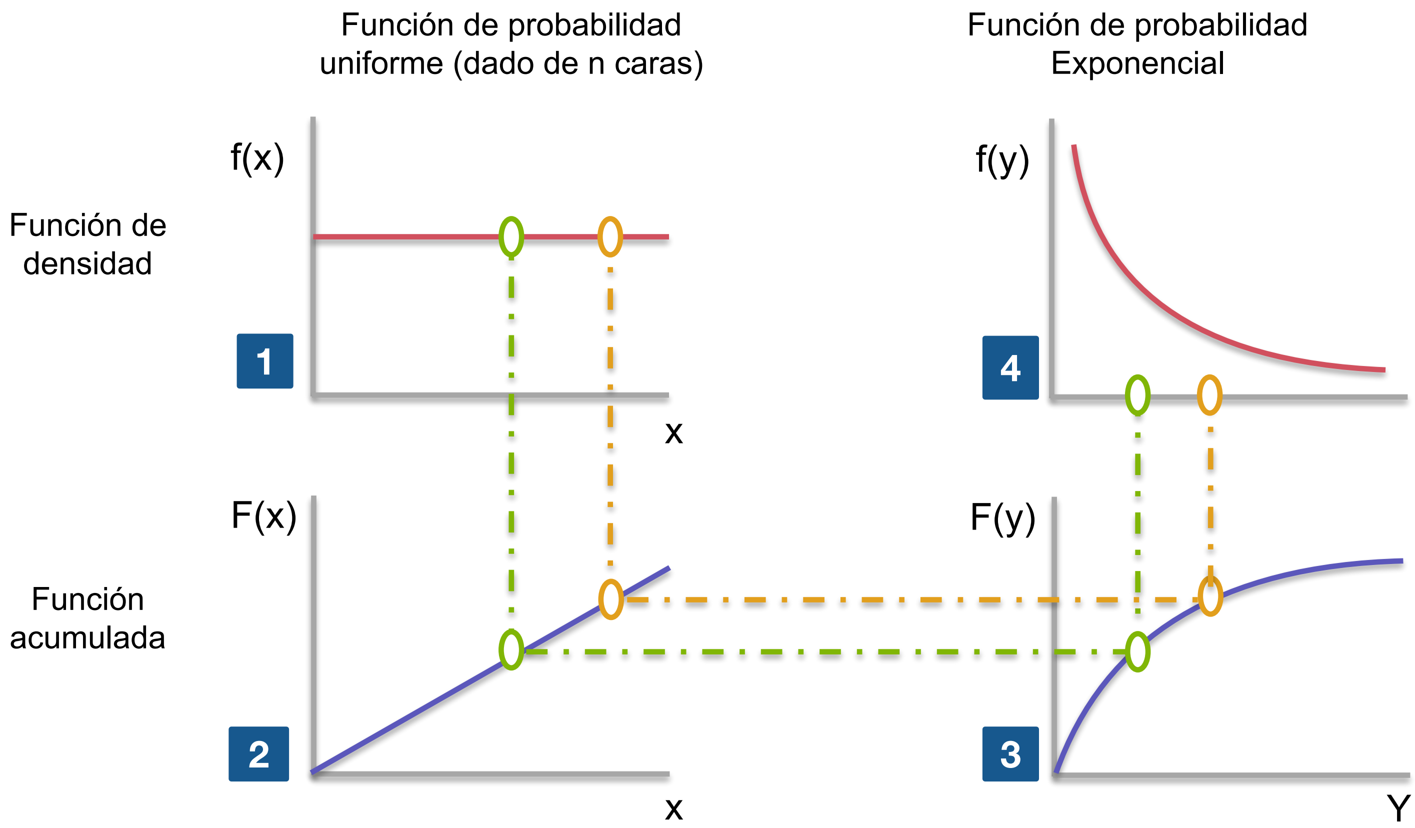


método de la transformada inversa: distribución exponencial



A partir de una distribución uniforme obtenemos el valor aleatorio generado bajo una distribución exponencial. Esto puede aplicarse a cualquier tipo de distribución.

método de la transformada inversa: distribución exponencial



Iterando múltiples veces podemos simular valores aleatorios de una determinada distribución de probabilidad (en este caso exponencial).

Distribución exponencial: tiempo entre arribos

Utilizamos la función de densidad de probabilidad **exponencial**, donde x representa el tiempo entre eventos con una **media [1/lambda].**

$$f(x) = \lambda \mathrm{e}^{-\lambda x}$$

Función de densidad acumulada exponencial

$$F(x) = \int_0^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}$$

Expresión para calcular el tiempo entre arribos siendo U una v.a. Uniforme entre 0 y 1.

$$u \sim U(0,1)$$

$$\boxed{x = -\left(\frac{1}{\lambda}\right) \ln(1-u)}$$

simulación: método de la transformada inversa

Ejemplo de simulación partiendo de una distribución de probabilidad específica

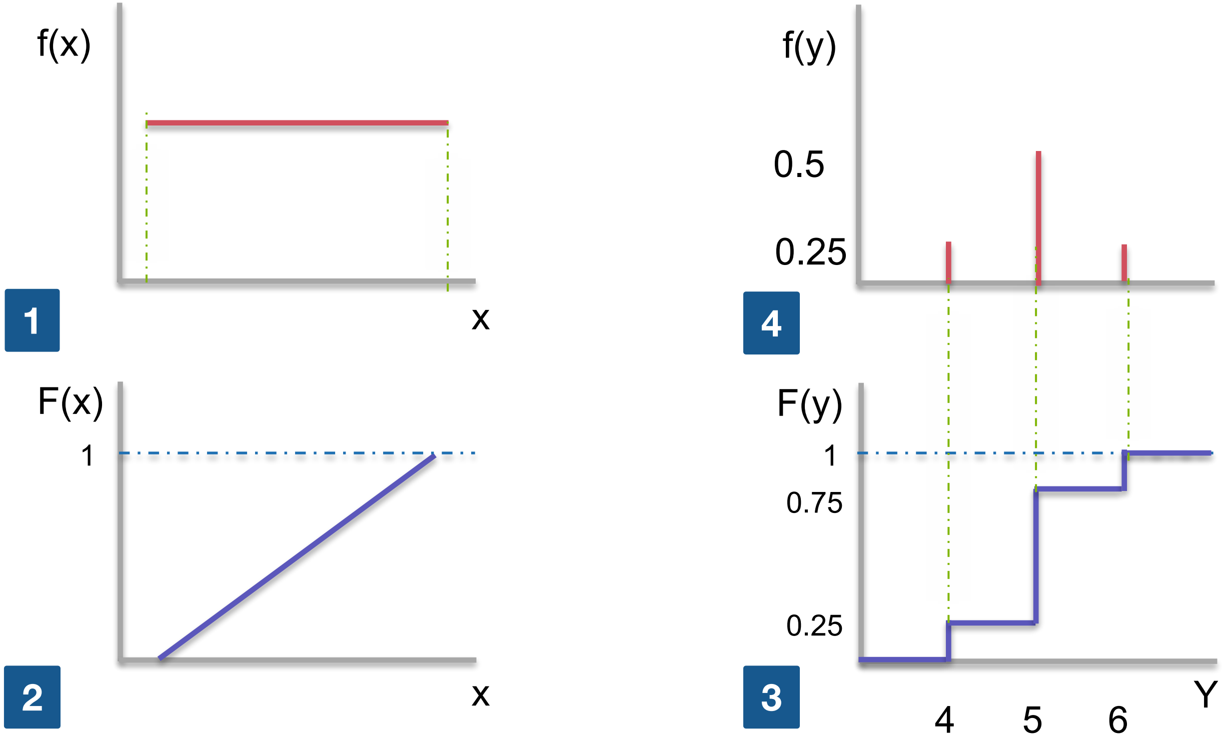
El tiempo de falla en una máquina es cada 4, 5 o 6 días con probabilidades de 0.25, 0.5 y 0.25. Quisiéramos simular las fallas de la máquina. Para eso definimos la distribución de densidad y su acumulada respectiva.

Número de días	Probabilidad	Acumulada
4	0.25	0.25
5	0.5	0.75
6	0.25	1

simulación: método de la transformada inversa

Distribución de probabilidad uniforme.

Function acumulada de distribución uniforme.



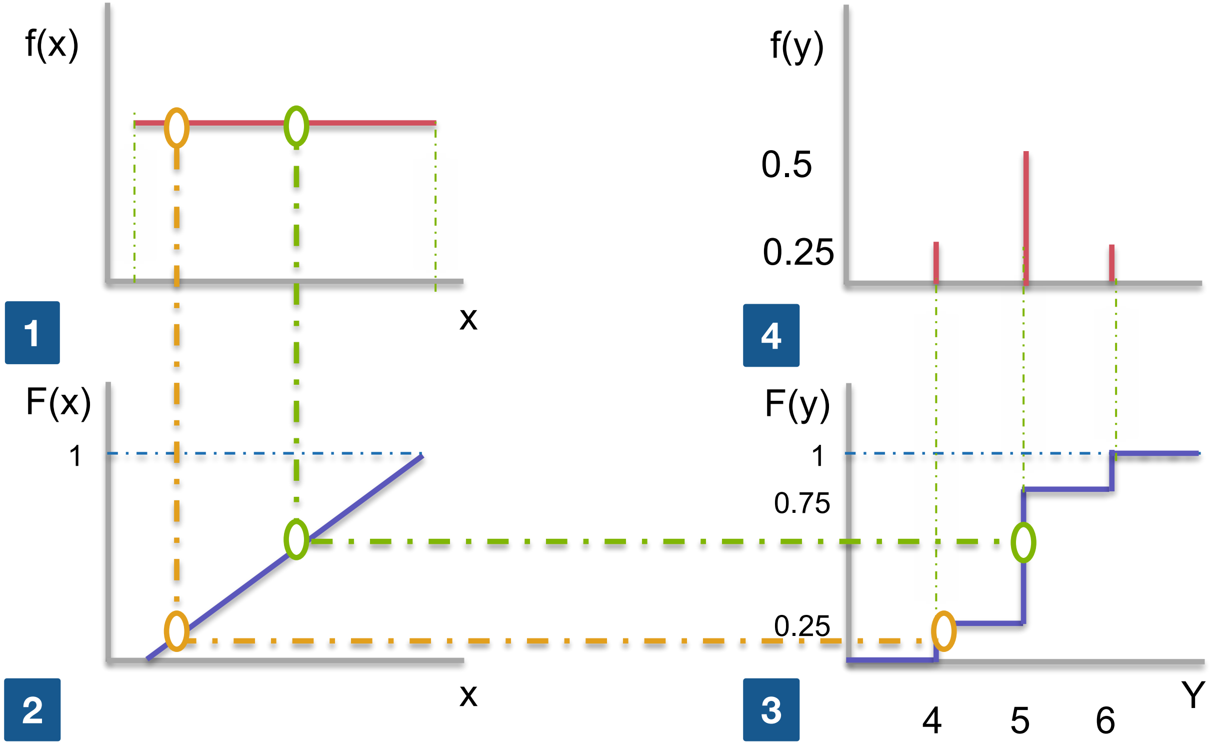
Distribución de probabilidad especial (en partes)

Function acumulada de distribución especial.

simulación: método de la transformada inversa

Distribución de probabilidad uniforme.

Function acumulada de distribución uniforme.



Distribución de probabilidad especial (en partes)

Function acumulada de distribución especial.