Modelado y Simulación Numérica

Daniel Pérez Palau

Tema 6. Generación de variables aleatorias

Calendario

	Semana	Tema	Refuerzo	Laboratorio	Actividad
09/11/2020					
16/11/2020	1	S0 + T1			
23/11/2020	2	T2			
30/11/2020	3	T3		L1	
07/12/2020	4	T4			
14/12/2020	5	T5			L1
21/12/2020	--	Semana de repaso	R-L1		
28/12/2020	--	Semana de repaso			
04/01/2021	6	T6 + repaso			
11/01/2021	7	T6			
18/01/2021	8	T7			
25/01/2021	9	T7			AG
01/02/2021	10	T8			
08/02/2021	11	T9		L2	
15/02/2021	12	T10	R-AG1		
22/02/2021	13	T11			L2
01/03/2021	14	Sesión examen	R-L2		
08/03/2021	15	Repaso (sesión doble)			
15/03/2021	16	Semana			

Próximas sesiones
T6-> (04/01 17:00CET)

Contenidos

- Tema 1. Conceptos generales de modelado matemático y simulación
- Tema 2. Modelado matemático de sistemas físicos
- Tema 3. Sistemas físicos y sus modelos
- Tema 4.Simulación
- Tema 5. Generación de números aleatorios
- Tema 6. Generación de variables aleatorias
- Tema 7. Medidas estadísticas
- Tema 8. Simulación de Monte Carlo
- Tema 9. Conceptos y elementos de simulación con eventos
- Tema 10. Modelado y simulación de sistemas de eventos discretos
- Tema 11. *Software* para modelado matemático y simulación

Objetivos

- Conocer métodos para generar variables aleatorias de forma eficiente.

Introducción

- Una **variable aleatoria (v.a.)** se define como una función que asigna a cada posible resultado de un experimento un número real.

- No conocer a priori el valor que va a tomar una variable no significa que sea imposible llegar a saber muchas cosas importantes sobre ella. **La distribución de una variable aleatoria determina qué valores puede tomar y con qué probabilidad.**

- **El conjunto de valores que puede tomar recibe el nombre de soporte.**

- En función del soporte podemos clasificarlas como:
  - **Distribuciones discretas**: tienen soporte finito o numerable.
  - **Distribuciones continuas**: Tienen soporte no numerable.

Variables aleatorias continuas

- En este tema nos centraremos en la generación de v.a. continuas.
  - Las discretas las veremos en los temas 9 y 10.

- Una v.a. continua se caracteriza por:
  - y la **función de densidad:** Es una función real de variable real no negativa que toma valores positivos sólo en el soporte de la variable (fuera del soporte valdrá cero). Se verifica que,

$$P(a < X < b) = \int_a^b f(x)dx\,, \qquad \int_{S_X} f(x)dx = 1$$

  - la **función de distribución F(x)** de la v.a.: mide la probabilidad acumulada desde $-\infty$ hasta el valor x.

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

Se verifica $f(x) = \frac{dF(x)}{dx}$

Métodos de generación: generalidades

- Estas variables presentarán una distribución estadística específica, cuyos parámetros serán conocidos o podremos estimarlos.

- Asumiremos que la distribución está completamente caracterizada.

- Seleccionaremos el método de generación atendiendo a diferentes factores, como por ejemplo:
  - precisión,
  - eficiencia de cálculo,
  - complejidad,
  - robustez,
  - facilidad de implementación,
  - …

Métodos de generación: bibliotecas

Existen bibliotecas que disponen de las v.a.

- Excel (rand(), norminv(rand(),mu,sig),…)
- Matlab (rand, normrnd(mu,sig),…)

Si no se dispone de una rutina de generación, será necesario programarla con el fin de incluirla en el modelo que se esté construyendo.

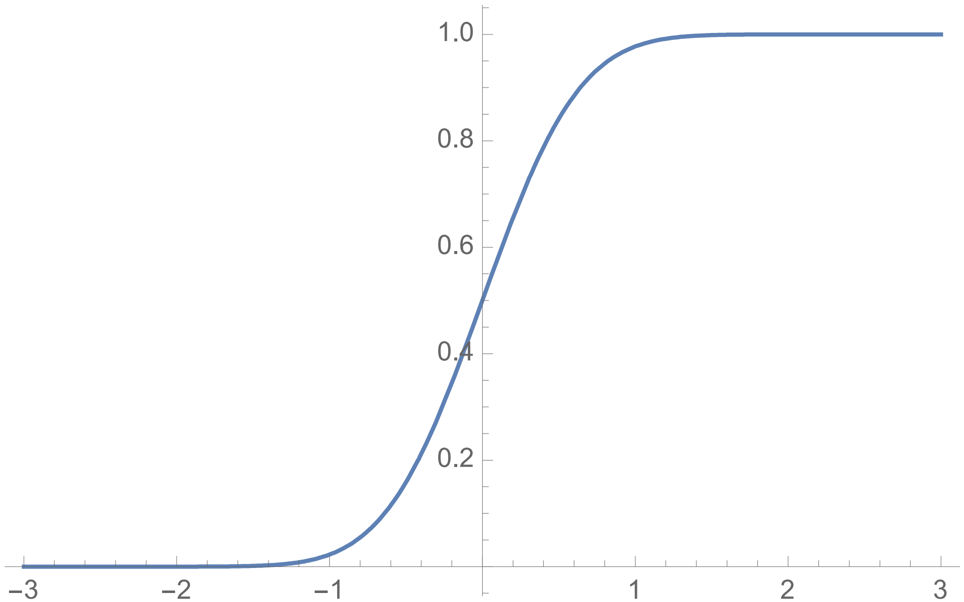
Veremos los siguientes métodos de generación de variables aleatorios:

- Transformada inversa
- Composición
- Aceptación rechazo
- Convolución

El método de la transformada inversa

Dada una v.a. X, con una función de distribución conocida $F(x)$, si disponemos de un número aleatorio uniforme p entre 0 y 1, el objetivo será obtener el valor de x que verifique que $F(x) = p$. Para ello recurriremos a despejar x en la expresión anterior, obteniendo,

$$x = F^{-1}(p)$$



El método de la transformada inversa

Este método se usa generalmente para distribuciones sencillas en las que F^{-1}es fácilmente calculable.

Ejemplos:

- La distribución uniforme(a,b)
- La distribución exponencial
- La distribución Weibull.

Computacionalmente, es el método de generación de v.a. más sencillo, pero no siempre es el más eficiente.

Ejemplo 1 (uniforme)

Continuemos por la distribución uniforme, que tiene la función de densidad,

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{otro caso} \end{cases}$$

Y función de distribución,

$$F(x) = \int_{\infty}^{x} f(t)dt = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

Ejemplo 1 (uniforme)

- Para generar la variable aleatoria, seguiremos los siguientes pasos:
  1. Obtener el valor de $F(x)$. En el caso de la uniforme,
  $$F(x) = \frac{x-a}{b-a}, \qquad a \leq x \leq b$$
  2. Fijar $F(x) = R$, siendo R un número aleatorio con distribución uniforme en el intervalo $(0,1)$. (ver tema anterior)
  3. A partir de la expresión $\frac{x-a}{b-a} = R$ despejamos x:
  $$x = a + R(b-a)$$
  Este expresión es el generador para la distribución uniforme
  4. Generamos números aleatorios uniformemente distribuidos en $[0,1], R_1, R_2, R_3, \ldots R_i$ y sustituirlos en la expresión anterior con el fin de obtener una v.a. distribuida uniformemente:
  $$X_i = a + R_i(b-a)$$

Ejemplo 2 (exponencial)

Continuemos por la distribución exponencial, que tiene la función de densidad,

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0 \quad , & \mathrm{x} < 0 \end{cases}, \qquad \lambda > 0$$

Y función de distribución,

$$F(x) = \int_{\infty}^{x} f(t)dt = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0 \quad , & \mathrm{x} < 0 \end{cases}$$

λ se interpreta como el número medio de ocurrencias por unidad de tiempo (nº coches en un semáforo/dia, nº de correos/hora,…)

Ejemplo 2 (exponencial)

Para generar la variable aleatoria, seguiremos los siguientes pasos:

1. Obtener el valor de $F(x)$. En el caso de la exponencial,
$$F(x) = 1 - e^{-\lambda x}, \qquad x \geq 0$$
2. Fijar $F(x) = R$, siendo R un número aleatorio con distribución uniforme en el intervalo $(0, 1)$. (ver tema anterior)

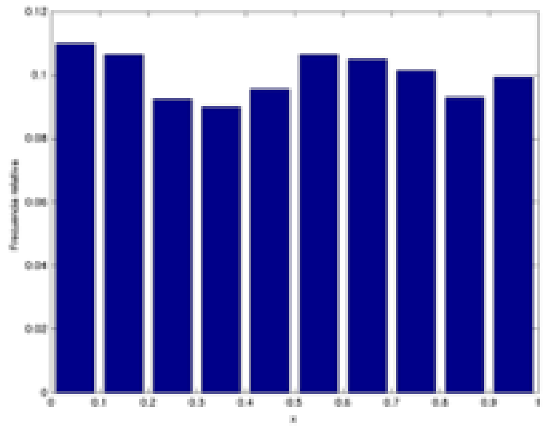
3. A partir de la expresión $1 - e^{-\lambda x} = R$ despejamos x:
$$x = -\frac{\ln(1-R)}{\lambda}$$

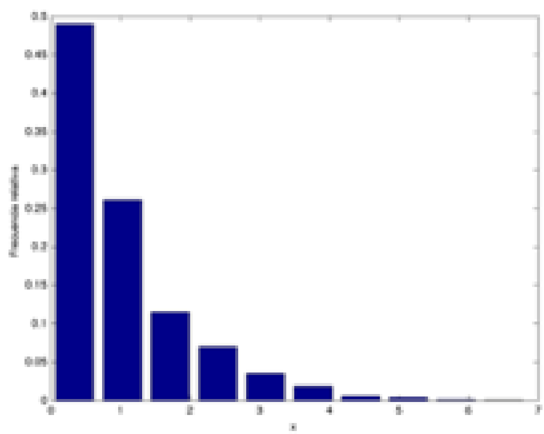
Este expresión es el generador para la distribución exponencial

4. Generamos números aleatorios uniformemente distribuidos en $[0,1], R_1, R_2, R_3, \dots R_i$ y sustituirlos en la expresión anterior con el fin de obtener una v.a. distribuida uniformemente:
$$X_i = -\frac{\ln(1-R)}{\lambda}$$

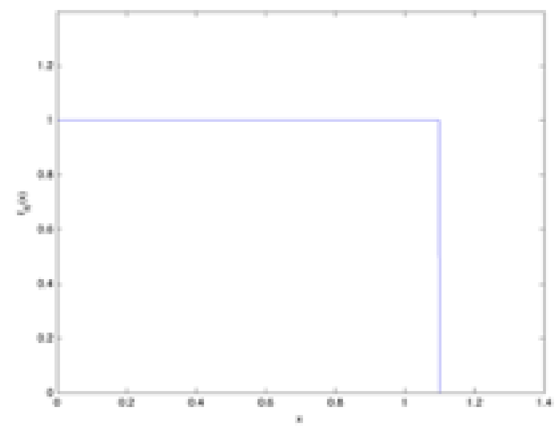
Ejemplo 2 (exponencial)

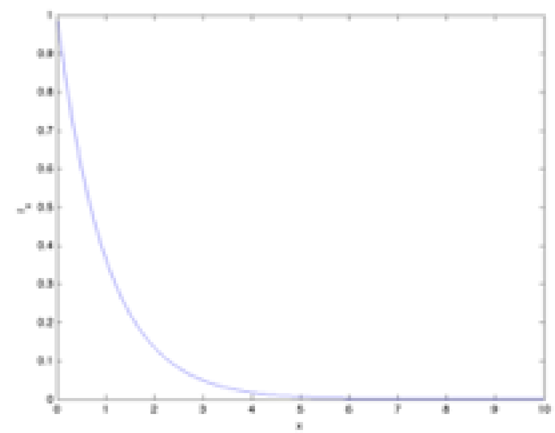
Histogramas:





Funciones de densidad





Ejemplo 3 (Weibull)

Continuemos por la distribución exponencial, que tiene la función de densidad,

$$f(x) = \begin{cases} \dfrac{\beta}{\alpha^{\beta}} x^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, & x \geq 0 \\ 0 \qquad\qquad , & \mathrm{x} < 0 \end{cases}$$

Y función de distribución,

$$F(x) = \int_{\infty}^{x} f(t)dt = \begin{cases} 1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, & x \geq 0 \\ 0 \qquad , & \mathrm{x} < 0 \end{cases}$$

Representa la tasa de fallo en máquinas o componentes electrónicos.

Ejemplo 3 (Weibull)

Para generar la variable aleatoria, seguiremos los siguientes pasos:

1. Obtener el valor de $F(x)$. En el caso de la Weibull,

$$F(x) = 1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, \qquad x \geq 0$$

2. Fijar $F(x) = R$, siendo R un número aleatorio con distribución uniforme en el intervalo $(0, 1)$. (ver tema anterior)

3. A partir de la expresión $1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}} = R$ despejamos x:

$$x = \alpha(-\ln(1 - R))^{\frac{1}{\beta}}$$

Esta expresión es el generador para la distribución Weibull.

4. Generamos números aleatorios uniformemente distribuidos en $[0,1], R_1, R_2, R_3, \ldots R_i$ y sustituirlos en la expresión anterior con el fin de obtener una v.a. distribuida uniformemente:

$$X_i = \alpha(-\ln(1 - R_i))^{\frac{1}{\beta}}$$

Métodos de Composición

Cuando la función de distribución no se puede invertir pero se puede definir como suma de otras funciones de distribución de la siguiente manera.

$$F(x) = \sum_{i=1}^{n} p_i F_i(x)$$

Con $p_i \geq 0$ y $\sum_{i=1}^{n} p_i = 1$.

Si nos proporcionan la función de densidad, entonces tiene que verificar:

$$f(x) = \sum_{i=1}^{n} p_i f_i(x)$$

Con p_i verificando las mismas propiedades y f_i funciones de densidad.

Atención: para que una función de densidad esté bien definida, la integral sobre su soporte debe ser 1 y siempre positiva.

Métodos de Composición

Una vez hemos identificado los pesos p_i y las funciones de distribución F_i invertimos cada distribución por separado para generar la v.a. de forma que obtenemos n generadores g_i.

Para decidir el valor de la variable aleatoria tendremos que generar dos números aleatorios: R_1 y R_2.

Construiremos los valores $P_i = \sum_{j=1}^{i} p_i$

El valor de la v.a. resultante es $x_i = g_k(R_2)$ con k el menor valor que satisface $R_1 \leq P_k$

Ejemplo (Método Composición)

Utilizar el método de composición para generar una variable aleatoria que tenga como función de densidad:

$$f(x) = \frac{1}{6} + (x-1)^2, \qquad 0 \leq x \leq 2$$

Para ello encontramos las funciones de densidad que componen el sumatorio:

$$a\int_0^2 1dx = ax\Big|_0^2 = 2a = 1 \rightarrow f_1(x) = \frac{1}{2}$$

$$b\int_0^2 (x-1)^2 dx = b\frac{(x-1)^3}{3}\Big|_0^2 = b\left(\frac{1}{3}+\frac{1}{3}\right) = \frac{2b}{3} = 1 \rightarrow f_2(x) = \frac{3}{2}(x-1)^2$$

Ejemplo (Método Composición)

Utilizar el método de composición para generar una variable aleatoria que tenga como función de densidad:

$$f(x) = \frac{1}{6} + (x-1)^2, \qquad 0 \leq x \leq 2$$

Tenemos que imponer:

$$f(x) = \frac{1}{6} + (x-1)^2 = \mathrm{p}_1 \frac{1}{2} + p_2 \frac{3}{2}(x-1)^2$$

Por lo que nos quedarán los pesos:

$$p_1 = \frac{1}{3}, \qquad p_2 = \frac{2}{3}$$

Estos pesos verifican: $p_1 + p_2 = 1$ y son positivos.

Ejemplo (Método Composición)

$$f(x) = \frac{1}{6} + (x-1)^2 = \frac{1}{3}\frac{1}{2} + \frac{2}{3}\frac{3}{2}(x-1)^2$$

La función de distribución resultante será:

$$F(x) = \frac{2}{6}\frac{x}{2} + \frac{4}{6}\frac{3}{2}\frac{(x-1)^3}{3}$$

Invertimos las funciones de distribución:

$$F_1(x) = \frac{x}{2} = R_2 \rightarrow x = g_1(R_2) = 2R_2$$

$$F_2(x) = \frac{3}{2}\frac{(\mathrm{x}-1)^3}{3} + \frac{1}{2} = R_2 \rightarrow x = g_2(R_2) = (2R_2 - 1)^{\frac{1}{3}} + 1$$

Ejemplo (Método Composición)

$$F_1(x) = \frac{x}{2} = R_2 \rightarrow x = g_1(R_2) = 2R_2$$

$$F_2(x) = \frac{\frac{3}{2}(\mathrm{x}-1)^3}{2 \cdot 3} = R_2 \rightarrow x = g_2(2R_2)^{\frac{1}{3}} + 1$$

Generamos los valores P_i:

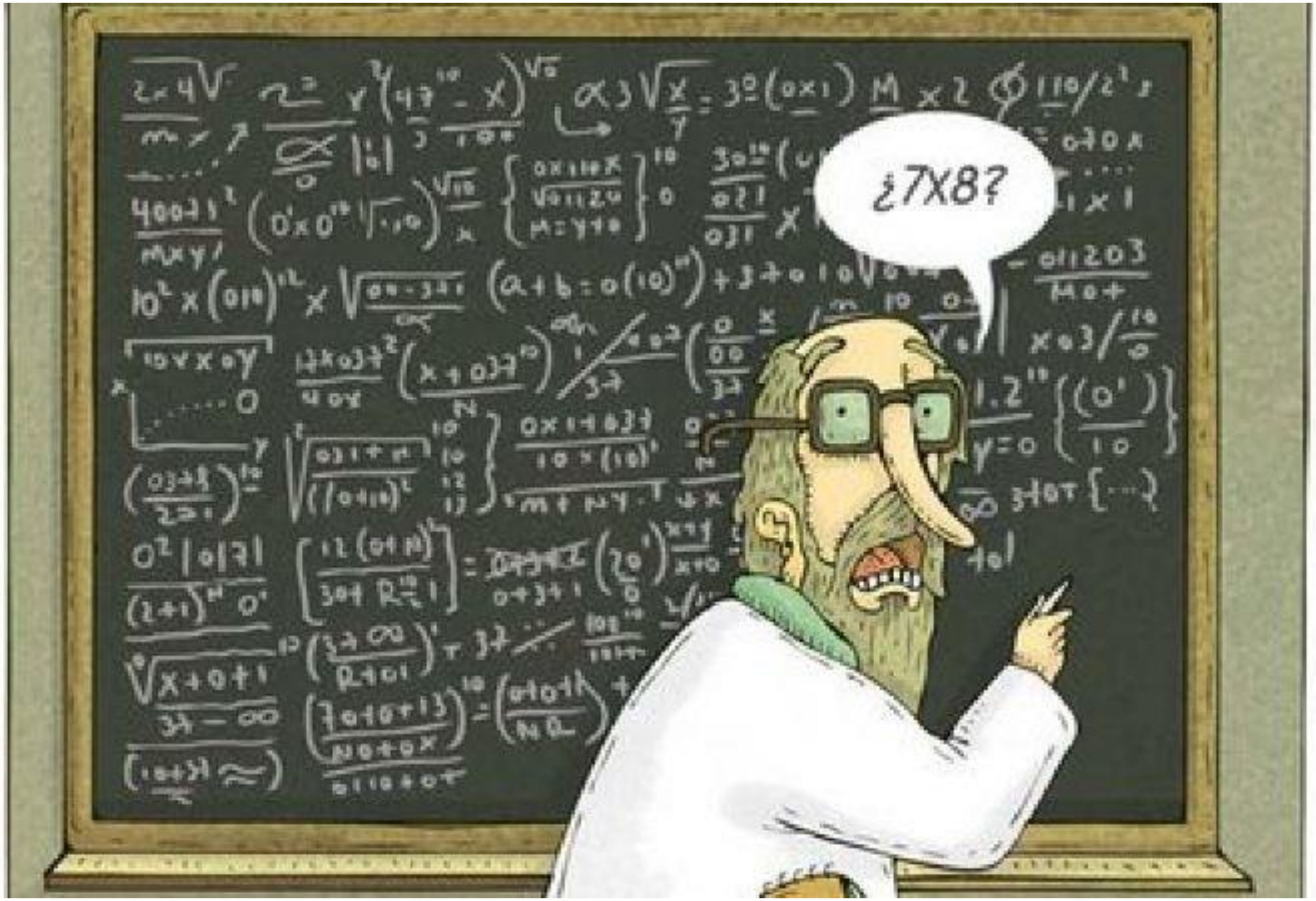
$$P_1 = p_1 = \frac{1}{3}, \qquad P_2 = p_1 + p_2 = \frac{1}{3} + \frac{2}{3} = 1$$

Generamos los valores aleatorios R_1 y R_2.

si $R_1 < P_1 \rightarrow \mathrm{X_i} = \mathrm{g_1(R_2)}$

si $(\mathrm{P_1} <) R_1 < P_2 \rightarrow \mathrm{X_i} = \mathrm{g_2(R_2)}$

¿Dudas?



Calendario

	Semana	Tema	Refuerzo	Laboratorio	Actividad
09/11/2020					
16/11/2020	1	S0 + T1			
23/11/2020	2	T2			
30/11/2020	3	T3		L1	
07/12/2020	4	T4			
14/12/2020	5	T5			L1
21/12/2020	--	Semana de repaso	R-L1		
28/12/2020	--	Semana de repaso			
04/01/2021	6	T6 + repaso			
11/01/2021	7	T6			
18/01/2021	8	T7			
25/01/2021	9	T7			AG
01/02/2021	10	T8			
08/02/2021	11	T9		L2	
15/02/2021	12	T10	R-AG1		
22/02/2021	13	T11			L2
01/03/2021	14	Sesión examen	R-L2		
08/03/2021	15	Repaso (sesión doble)			
15/03/2021	16	Semana			

Próximas sesiones
T6b-> 14/01 18:00 CET
T7a-> 22/01 17:00 CET

Método de Aceptación-Rechazo

Hay dos versiones de este método. Nosotros analizaremos la más sencilla, ya que es la más generalizada.

Consideremos una v.a. X con soporte en el intervalo finito $[a, b]$.

En este intervalo la función de densidad de X es conocida, está **acotada** y es **no nula**

La función de densidad es suficientemente compleja como para no poder utilizar el método de la transformada inversa.

Método de Aceptación-Rechazo

El procedimiento de generación de una v.a. X con función de densidad $f(x)$ es el siguiente:

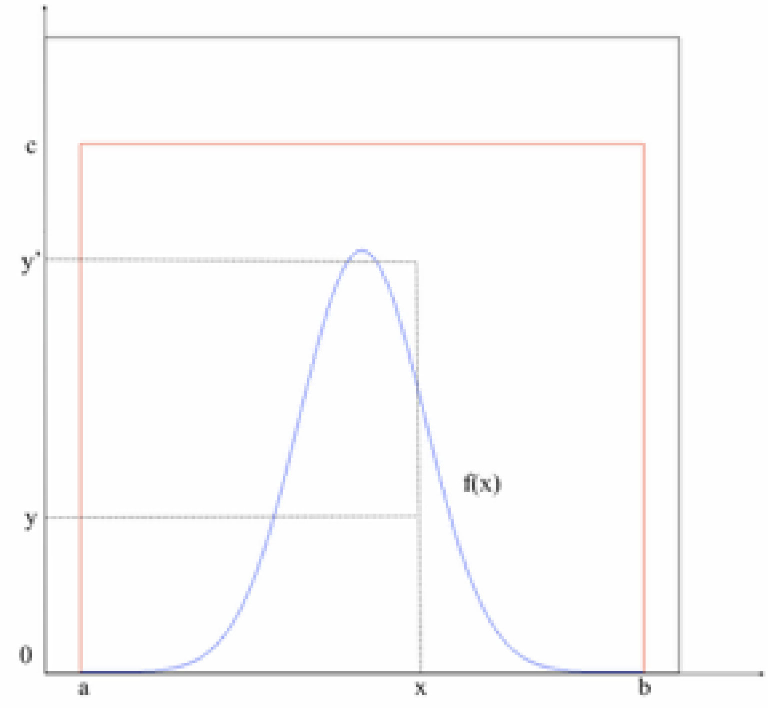
1. Generar dos números aleatorios u y v independientes y uniformemente distribuidos.
2. Hacer $x = a + (b - a)u$ para transformar u en una v.a. distribuida uniformemente en el soporte $[a, b]$.
3. Hacer $y = cv$, para transformar v en una v.a. distribuida uniformemente en el intervalo $[0, c]$.
4. Calcular $f(x)$. $Si\ y <= f(x)$, se acepta x como valor generado de la v.a. X. En caso contrario, se vuelve al paso 1.

Método de Aceptación-Rechazo

Los pasos 1, 2 y 3 generan puntos (X, Y) que están uniformemente distribuidos sobre un rectángulo de dimensiones $c \times (b - a)$.

- Si el punto cae bajo la curva $f(x)$, el paso 5 acepta el valor de x.
- En caso contrario, rechazamos el valor de x y repetimos.

La zona entre el rectángulo y la curva determina el rechazo de la observación. La probabilidad de rechazo será el cociente entre el área de dicha zona y el área del rectángulo. La región de aceptación tiene área 1 ya que $f(x)$ es una función de densidad.



Método de A-R (justificación)

El método de Aceptación-Rechazo genera valores según la distribución de la v.a. X.

Para poder demostrarlo, definamos en primer lugar la probabilidad condicionada.

Definimos la probabilidad de ocurrencia un suceso $x = A$ sabiendo que otro suceso $y = B$ ya ha ocurrido, como

$$P(x = A|y = B) = \frac{P(x = A, y = B)}{P(y = B)}$$

A esta probabilidad se le denomina probabilidad del suceso $x = A$ condicionada por el suceso $y = B$, o simplemente **probabilidad de $\boldsymbol{x}$ condicionada a $\boldsymbol{y}$**.

Método de A-R (justificación)

Consideramos $x_1, x_2 \in \mathbb{R}$ tales que $a < x_1 < x_2 < b$.

Denominamos por R a la región de $y = f(x)$ dentro del rectángulo $[x_1, x_2] \times [0, c]$.

La probabilidad de que el valor de la v.a. X se encuentre en dicha región es:

$$P\big(x_1 < X < x_2 | Y \leq f(x)\big) = \frac{P\big(x_1 < X < x_2, Y \leq f(x)\big)}{P\big(Y \leq f(x)\big)}$$

$$= \frac{P\big((x, y) \in R\big)}{P\big(Y \leq f(x)\big)}$$

$$= \frac{\frac{\int_{x_1}^{x_2} f(x)dx}{c(b-a)}}{\frac{1}{c(b-a)}} = \int_{x_1}^{x_2} f(x)dx$$

Así pues, la probabilidad condicionada inicial se corresponde con la probabilidad de que x caiga en $[x_1, x_2]$

Método de A-R (ejemplo)

Sea,

$$f(x) = \begin{cases} \frac{1}{2} + \frac{1}{2}x - \frac{3}{8}x^2, & 0 \leq x \leq 2 \\ 0, & \text{otro caso} \end{cases}$$

Para generar valores por A-R debemos calcular una función uniforme que supera a la densidad.

Métodos posibles:

- Buscamos el máximo de f por derivación.
- Aplicamos que la función es creciente/decreciente.
- Realizamos la representación gráfica de la función.

Método de A-R (ejemplo)

Sea,

$$f(x) = \begin{cases} \frac{1}{2} + \frac{1}{2}x - \frac{3}{8}x^2, & 0 \leq x \leq 2 \\ 0, & \text{otro caso} \end{cases}$$

Por derivación:

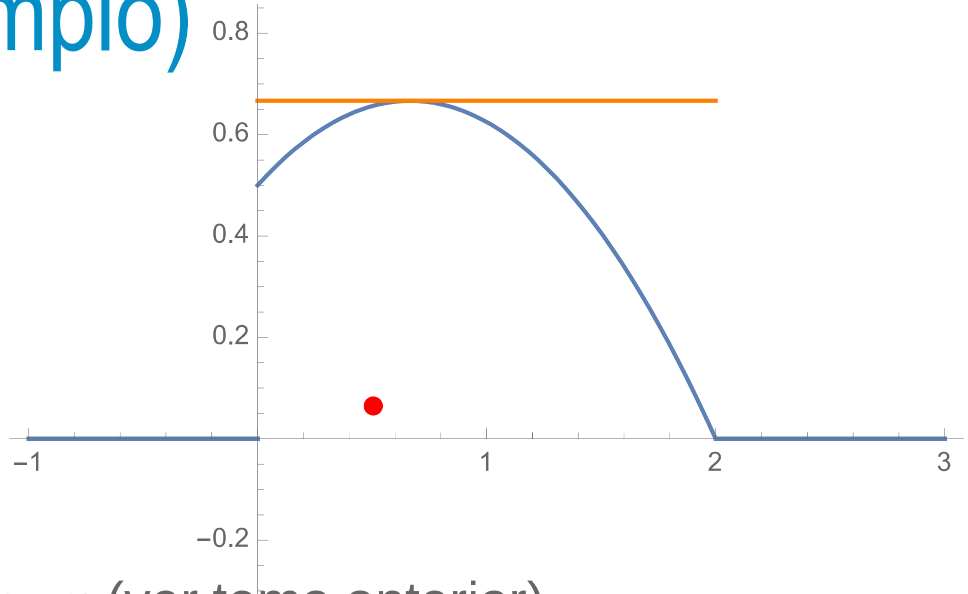
$$f'(x) = \frac{1}{2} - \frac{3}{4}x = 0 \to x = \frac{2}{3} \text{ (candidato a máximo)}$$
$$f''(x) = -3/4 < 0 \text{ (confirma candidato)}$$

Luego el máximo se toma $f\left(\frac{2}{3}\right) = \frac{2}{3}$.

Y comprobamos en los extremos: $f(0) = \frac{1}{2},\ f(2) = 0 < \frac{2}{3}$.

Por lo que $a = 0, b = 2$ y $c = 2/3$.

Método de A-R (ejemplo)



Sea,

$$f(x) = \begin{cases} \frac{1}{2} + \frac{1}{2}x - \frac{3}{8}x^2, & 0 \leq x \leq 2 \\ 0, & \text{otro caso} \end{cases}$$

Por lo que $a = 0, b = 2$ y $c = \frac{2}{3}$.

Generamos dos valores aleatorios u, v (ver tema anterior).

Por ejemplo, $u = 0.25, v = 0.1$.

Realizamos las 4 etapas:

1. $x_0 = a + u(b - a) = 0 + 0.25(2 - 0) = 0.5$
2. Generamos $y = cv = \frac{2}{3} \cdot 0.1 = 0.067$
3. Calculamos $f(x_0) = f(0.5) = 0.66$.
4. $y = 0.067 \leq 0.66 \Rightarrow$ **aceptamos 0.5 como valor de la v.a.**

Método de la convolución

Se base en la siguiente propiedad:

Si una v.a. a es resultado de sumar varias v.a. independientes

$$a = b_1 + b_2 + \ldots + b_n$$

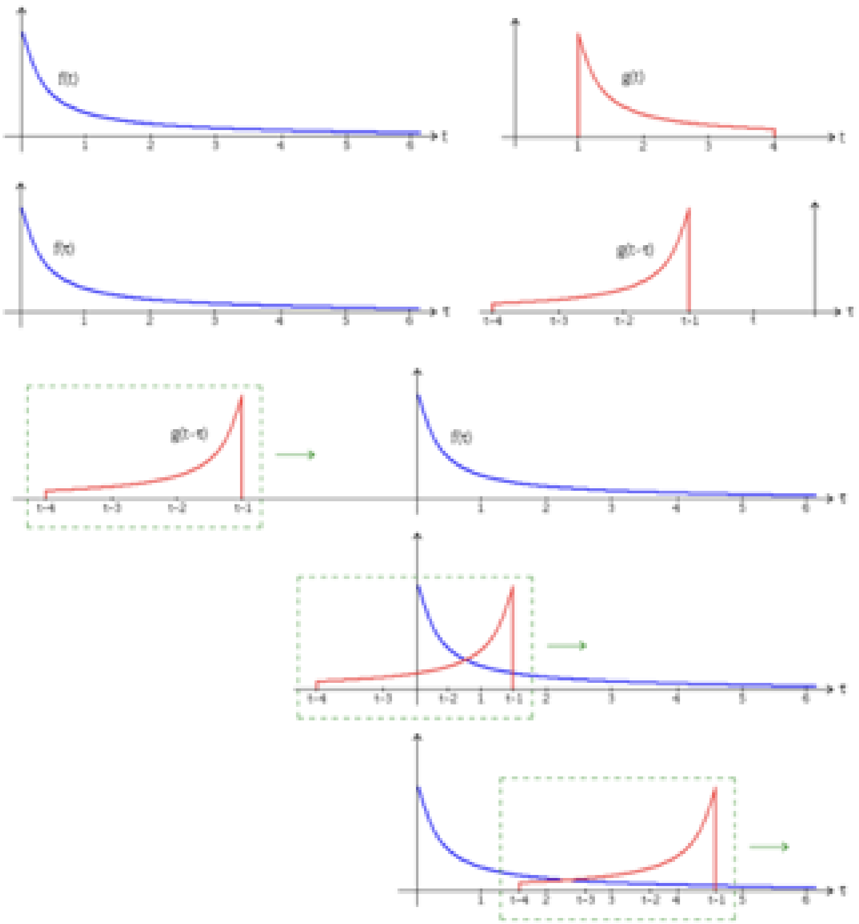
la función de densidad de la v.a. a se obtendrá como resultado de la convolución de las funciones de densidad de las v.a. $b_1, b_2, \ldots, b_n$.

Definición de convolución

La **convolución** de $f(x)$ y $g(x)$ se representa generalmente como $f(x) * g(x)$. Se define a través de la integral impropia del producto de la función $f(x)$ y $g(x)$ después de invertir y desplazar una de ellas una distancia t:

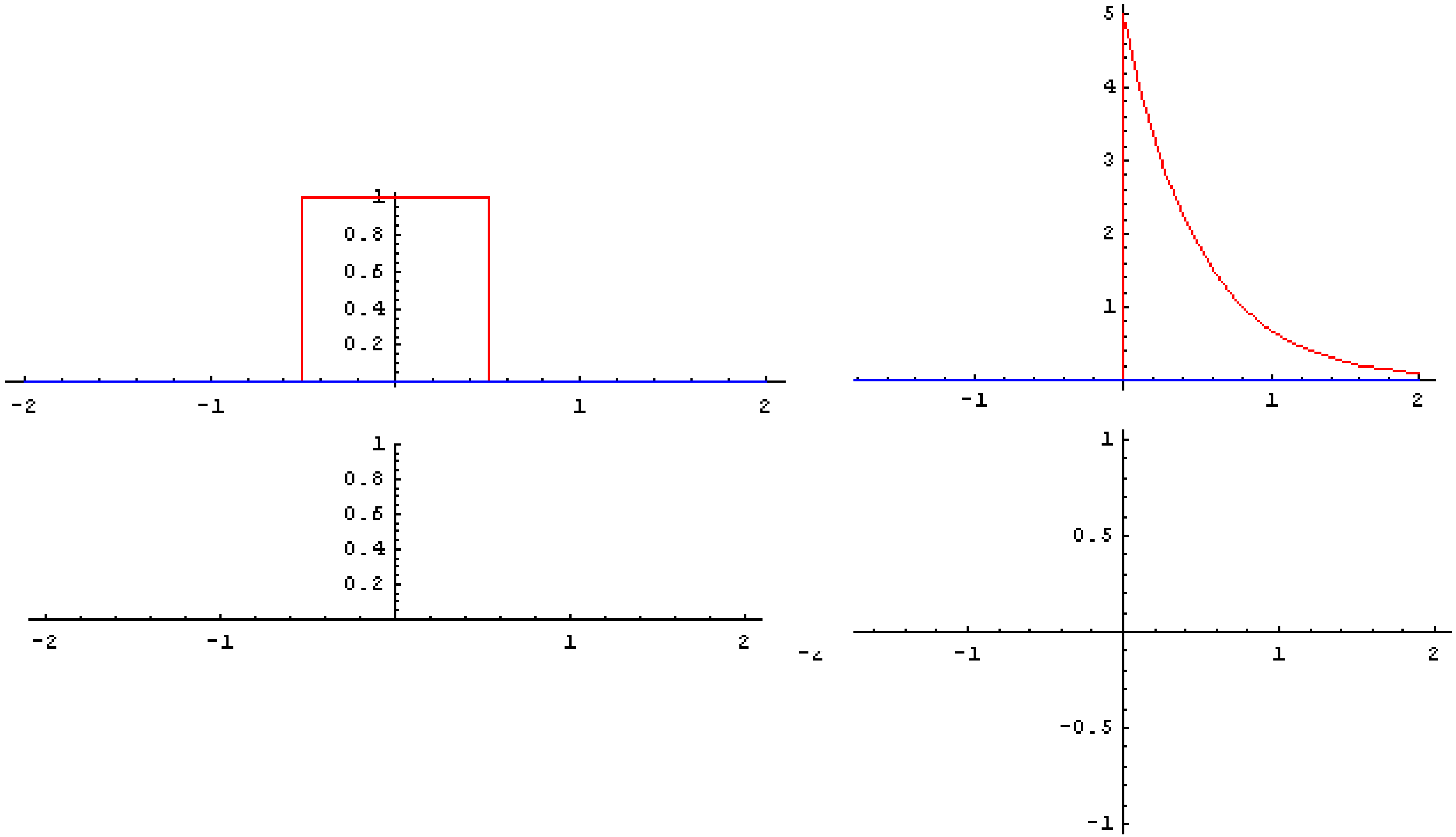
$$f(x) * g(x) = \int_{-\infty}^{\infty} f(\tau)\, g(x - \tau) d\tau$$

Método de la convolución



Fuente: wikipedia

Método de la convolución



Fuente: wikipedia

Ejemplo de cálculo de convolución

Sean Y y Z v.a. independientes exponencialmente distribuidas con función de densidad

$$f_Y(x) = f_Z(x) = f(x) = \lambda e^{-\lambda x}, \qquad x > 0$$

siendo $\lambda > 0$ el parámetro de la distribución exponencial. Entonces la función de densidad de la v.a. $X = Y + Z$ es

$$f_X(x) = f_{Y+Z}(x) = f_Y(x) * f_Z(x) = \int_{-\infty}^{\infty} f_Y(\tau) f_Z(x-\tau) d\tau =$$
$$= \int_0^x \lambda e^{-\lambda\tau} \lambda e^{-\lambda(x-\tau)} d\tau = \lambda^2 \int_0^x e^{-\lambda x} d\tau = \lambda^2 e^{-\lambda x} \int_0^x d\tau$$
$$= \lambda^2 e^{-\lambda x} x, \qquad x > 0$$

Es una función de densidad con nombre 2-Erlang.

Ejemplo de cálculo de convolución

Si queremos calcular una variable aleatoria 2-Erlang:

$$f(x) = \lambda^2 e^{-\lambda x} x, \qquad x > 0$$

Podemos calcular dos valores aleatorios de una exponencial y sumarlos.

$$x_1 = -\frac{\ln(1 - R_1)}{\lambda}, \qquad x_2 = -\frac{\ln(1 - R_2)}{\lambda}$$

$$E = x_1 + x_2$$

unir

LA UNIVERSIDAD
EN INTERNET

www.unir.net