Generación de variables aleatorias continuas
Método de la transformada inversa

Georgina Flesia

FaMAF

16 de abril, 2013

Generación de v.a. discretas

Existen diversos métodos para generar v.a. discretas:

- Transformada Inversa
- De aceptación-rechazo, o método de rechazo.
- De composición.
- Métodos mejorados según la distribución.

Método de la Transformada Inversa

- X: una variable aleatoria discreta con probabilidad de masa

$$P(X = x_j) = p_j, \qquad j = 0, 1, \ldots.$$

- $U \sim \mathcal{U}(0,1)$: simulación de una v.a. con distribución uniforme.
- Método de la transformada inversa:

$$X = \begin{cases} x_0 & \text{si } U < p_0 \\ x_1 & \text{si } p_0 \leq U < p_0 + p_1 \\ \vdots & \\ x_j & \text{si } p_0 + \cdots + p_{j-1} \leq U < p_0 + \cdots + p_{j-1} + p_j \\ \vdots & \end{cases}$$

$$\boxed{P(X = x_j) = P(p_0 + \cdots + p_{j-1} \leq U < p_0 + \cdots + p_{j-1} + p_j) = p_j.}$$

Método de la Transformada Inversa

$$F(x) = P(X \leq x): \text{ Función de distribución acumulada}$$

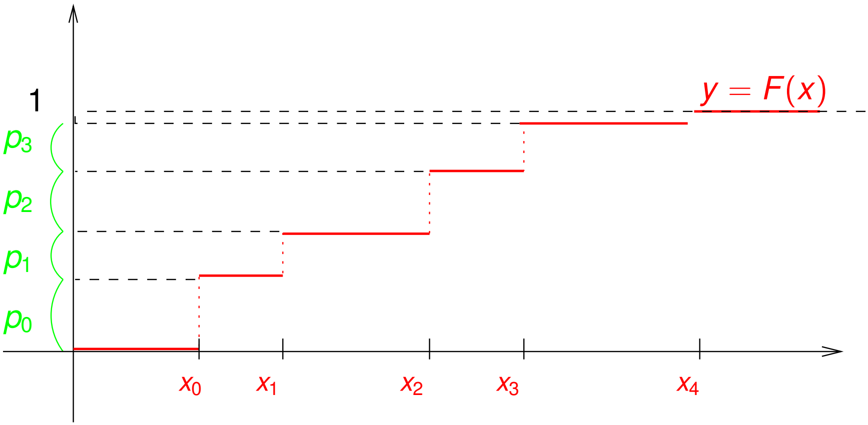
- F es una función creciente, escalonada, que toma valores entre 0 y 1.
- Si se ordenan los valores de la variable en forma creciente:

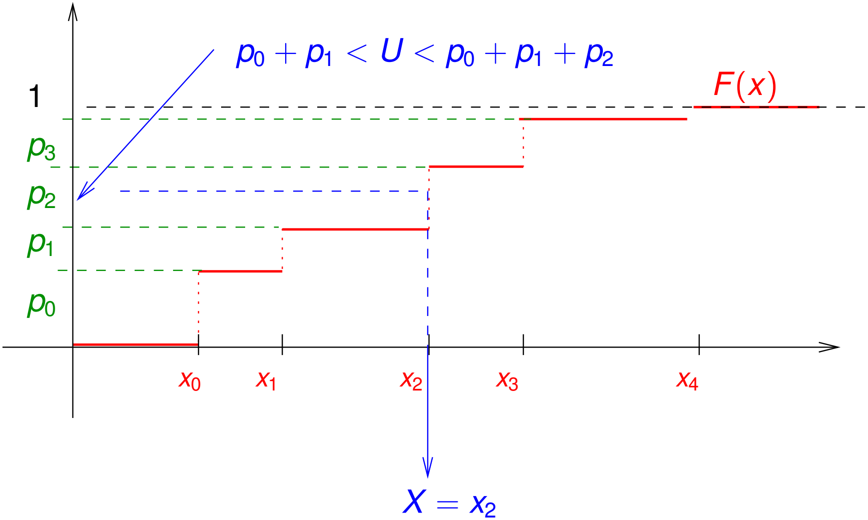
$$x_0 < x_1 < \cdots < x_n < \ldots,$$

entonces

$$F(x_j) = \sum_{k=0}^{j} p_k = p_0 + p_1 + \cdots + p_j.$$

Gráficamente




$p_0 + p_1 < U < p_0 + p_1 + p_2$
$F(x)$
1
p_3
p_2
p_1
p_0
x_0
x_1
x_2
x_3
x_4
$X = x_2$

Algoritmo

Si la v.a. toma un número finito de valores, el algoritmo es el siguiente:

Algorithm 1: Transformada Inversa

Generar $U \sim \mathcal{U}(0, 1)$;
if $U < p_0$ **then**
  $X \leftarrow x_0$ y terminar.
end
if $U < p_0 + p_1$ **then**
  $X \leftarrow x_1$ y terminar.
end
if $U < p_0 + p_1 + p_2$ **then**
  $X \leftarrow x_2$ y terminar.
end
$\vdots$

Algoritmo de la Transformada Inversa

Si $x_0 < x_1 < x_2 < \dots$, entonces $F(x_j) = \sum_{i=0}^{j} p_i$, y por lo tanto

$$\begin{array}{rcl} X \leftarrow x_0 & \text{si} & U < p_0 = F(x_0) \\ X \leftarrow x_j & \text{si} & F(x_{j-1}) \leq U < F(x_j) \end{array}$$

Se trata de hallar el intervalo $[F(x_{j-1}), F(x_j))$ donde se ubica U:

$$\boxed{U \in [F(x_{j-1}), F(x_j)) \qquad \Longleftarrow \qquad \text{Transformada Inversa}}$$

Ejemplo

$X : \{1, 2, 3, 4\}$.

$$p_0 = 0.20, \quad p_1 = 0.15, \quad p_2 = 0.25, \quad p_3 = 0.40$$

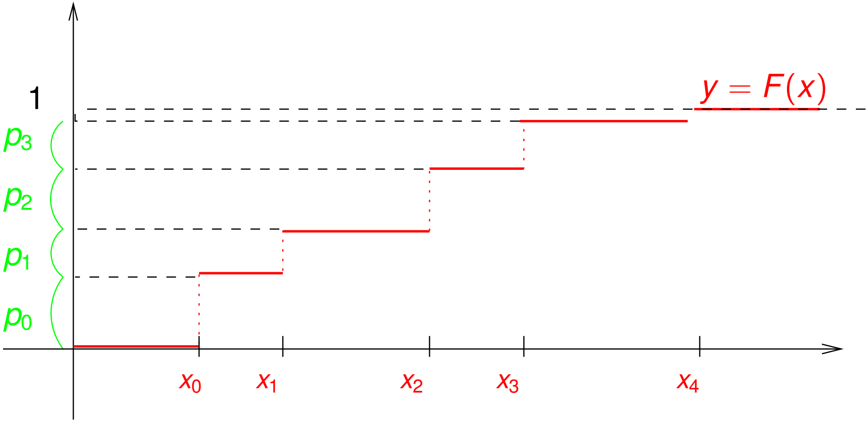
Algorithm 2:

Generar U;
if $U < 0.2$ **then**
| $X \leftarrow 1$ y terminar.
end
if $U < 0.35$ **then**
| $X \leftarrow 2$ y terminar.
end
if $U < 0.6$ **then**
| $X \leftarrow 3$
else
| $X \leftarrow 4$
end

Ejemplo

$X : \{1, 2, 3, 4\}$.

$$p_0 = 0.20, \quad p_1 = 0.15, \quad p_2 = 0.25, \quad p_3 = 0.40$$



Orden decreciente de p_i

Si se ordenan de manera decreciente las probabilidades de masa $p_0, p_1, \ldots,$ se puede obtener un algoritmo más eficiente:

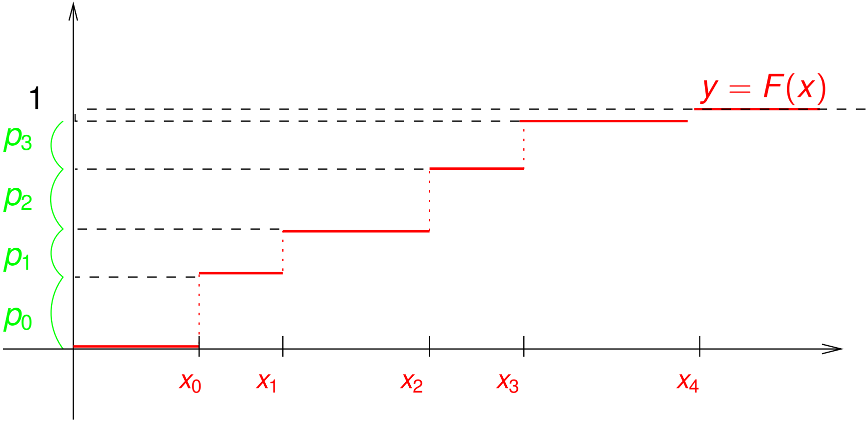
Algorithm 3: Ordenando p_i

Generar U;
if $U < 0.4$ **then**
| $X \leftarrow 4$ y terminar.
end
if $U < 0.65$ **then**
| $X \leftarrow 3$ y terminar.
end
if $U < 0.85$ **then**
| $X \leftarrow 1$
else
| $X \leftarrow 2$
end

Ejemplo

$X : \{1, 2, 3, 4\}$.

$$p_0 = 0.20, \quad p_1 = 0.15, \quad p_2 = 0.25, \quad p_3 = 0.40$$



Generación de variables aleatorias continuas con densidad

Decimos que X es una v.a. continua con densidad si existe $f : \mathbb{R} \mapsto \mathbb{R}$ positiva tal que

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

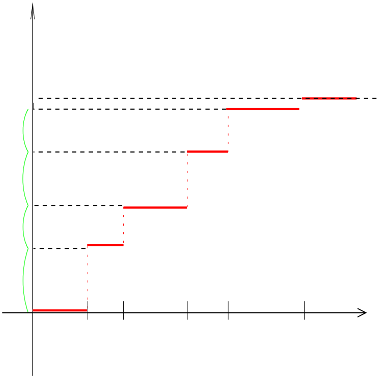
Estudiaremos los siguientes métodos de generación para una tal X:

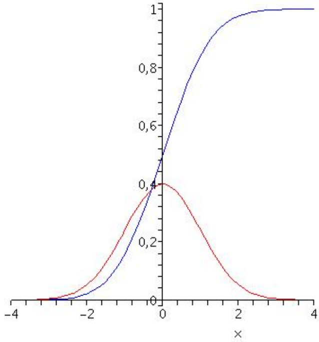
- Método de la transformada inversa.
- Método de aceptación y rechazo.
- Método de composición.

Método de la transformada inversa

Propiedades de $F(x) := \int_{-\infty}^{x} f(t)\,dt$.

- $F(x)$ es continua.
- $F(x)$ es no decreciente.
- $0 \leq F(x) \leq 1$, $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$





Método de la transformada inversa

Teorema

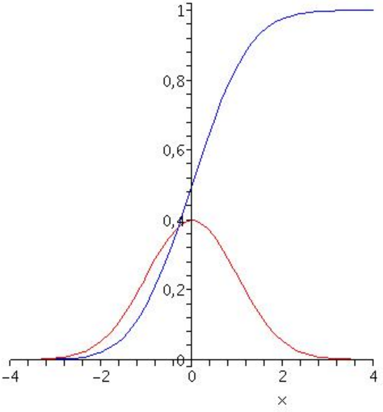
Si F es una función de distribución continua, inversible, y $U \sim \mathcal{U}(0,1)$, entonces

$$X = F^{-1}(U)$$

es una variable aleatoria con distribución F.

$$\begin{aligned} P(X \leq a) &= P(F^{-1}(U) \leq a) \\ &= P(F(F^{-1})(U) \leq F(a)) \qquad \text{por ser inversible} \\ &= P(U \leq F(a)) = F(a) \end{aligned}$$

Método de la transformada inversa



Método de la transformada inversa

- Es suficiente que F tenga inversa en $(0, 1)$.
- Problemas:
  - La inversa de F involucra funciones computacionalmente costosas. ($\sqrt[n]{f(x)}$, $\log(x)$, etc.)
  - La inversa de F no puede ser calculada explícitamente. (p. ej., distribución de la normal, de una gamma)
- Para ciertas distribuciones F pueden utilizarse otras estrategias, por ejemplo expresando a F como
  - distribución del mínimo y/o del máximo de v. a. independientes.
  - distribución de suma de variables aleatorias independientes
  - distribución de una v. a. condicional a otra,
- o existen métodos específicos (p. ej., para X con distribución normal.)

Veamos algunos ejemplos.

Aplicación del método de la transformada inversa

Ejemplo

Escribir un método para generar el valor de una v. a. X con función de densidad

$$f(x) = -\frac{x}{2} + 1, \qquad 0 \le x \le 2.$$

$$F(x) = -\frac{x^2}{4} + x, \qquad 0 \le x \le 2.$$

F no es inversible sobre $\mathbb{R}$, pero sólo nos interesa encontrar una inversa $F^{-1} : (0,1) \mapsto (0,2)$.

$$-\frac{x^2}{4} + x = u \quad \Rightarrow \quad x = \begin{cases} x = 2 + 2\sqrt{1-u} \\ \text{o} \\ x = 2 - 2\sqrt{1-u} \end{cases}$$

Algorithm 4:

Generar U;

X$\leftarrow 2 - 2\sqrt{U}$

Aplicación del método de la transformada inversa

Ejemplo

Escribir un método para generar el valor de una v. a. X con función de densidad

$$f(x) = \begin{cases} 1/4 & 0 \leq x \leq 2 \\ 1/2 & 2 < x < 3 \\ 0 & \text{c.c.} \end{cases}$$

$$F(x) = \begin{cases} 0 & x \leq 0 \\ x/4 & 0 \leq x \leq 2 \\ \dfrac{x-1}{2} & 2 < x < 3 \\ 1 & x \geq 3 \end{cases}$$

Generar U;
if $U < 1/2$ **then**
  X$\leftarrow 4U$
else
  X$\leftarrow 2U + 1$
end

Máximos y mínimos de v. a. independientes

Consideremos $X_1, X_2, \ldots, X_n$ v. a. independientes, con funciones de distribución $F_1, F_2, \ldots, F_n$, respectivamente.

$$F_i(a) = P(X_i \leq a).$$

- $X = \max\{X_1, X_2, \ldots, X_n\}$
- $F_X(a) = F_1(a) \cdot F_2(a) \ldots F_n(a)$

- $Y = \min\{X_1, X_2, \ldots, X_n\}$
- $1 - F_Y(a) = (1 - F_1(a)) \cdot (1 - F_2(a)) \ldots (1 - F_n(a))$

Máximos y mínimos de v. a. independientes

Si X es tal que

$$F_X(x) = x^n, \qquad 0 \leq x \leq 1, \quad 0 \text{ en otro caso.}$$

$$F_X(x) = \underbrace{x \cdot x \cdots x}_{n}$$

entonces $X = \max\{U_1, X_2, \ldots, U_n\}$ con $U_1, U_2, \ldots, U_n$ v. a. independientes, idénticamente distribuidas uniformes en [0,1]. Para generar X

Algorithm 5: $F_X(t) = t^n$

Generar $U_1, U_2, \ldots, U_n$;
X$\leftarrow \max\{U_1, U_2, \ldots, U_n\}$

- no requiere cálculo de una raíz n-ésima.
- se generan $n - 1$ uniformes adicionales.
- requiere de $n - 1$ comparaciones.

Generación de una v. a. exponencial

- Si $X \sim \mathcal{E}(\lambda)$, entonces $c \cdot X$ también es exponencial.
- $c \cdot X \sim \mathcal{E}(\frac{\lambda}{c})$.
- Calculamos la inversa de la función de distribución de $X \sim \mathcal{E}(1)$:

$$
\begin{aligned}
F_X(x) &= 1 - e^{-x} \\
u &= 1 - e^{-x} \\
1 - u &= e^{-x} \\
x &= -\log_e(1-u)
\end{aligned}
$$

Algorithm 6: $X \sim \mathcal{E}(1)$

Generar U;
$X \leftarrow -log(U)$

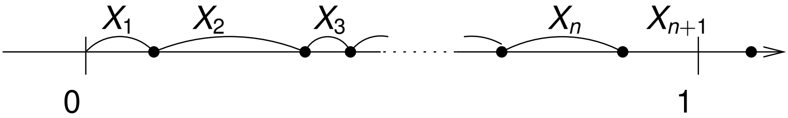
Algorithm 7: $X \sim \mathcal{E}(\lambda)$

Generar U;
$X \leftarrow -\frac{1}{\lambda} log(U)$

Generación de una v. a. Poisson $X \sim \mathcal{P}(\lambda)$

Para generar una variable Poisson, vamos a usar información sobre una estructura mas compleja, el proceso de Poisson. Recordemos que

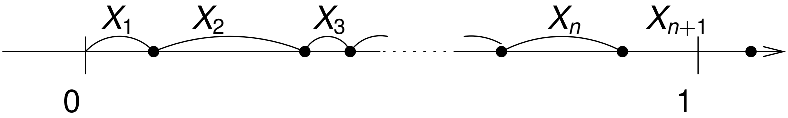
- En un proceso de Poisson homogéneo de parámetro λ,
  - los tiempos de llegada entre eventos son v. a. exponenciales de media $\frac{1}{\lambda}$.
  - el número de eventos en un intervalo de tiempo de longitud t es una v. a. Poisson de media $\lambda \cdot t$.
- $N(1)$ es una v. a. Poisson de media λ.



$$N(1) = \max\{n \mid X_1 + X_2 + \cdots + X_n \leq 1\}$$

Generación de una v. a. Poisson $X \sim \mathcal{P}(\lambda)$

Entonces, toda variable Poisson X de parámetro λ puede pensarse como la realización $N(1)$ de un proceso de Poisson de media λ.



$$N(1) = \max\{n \mid X_1 + X_2 + \cdots + X_n \leq 1\}$$

Generación de una v. a. Poisson $X \sim \mathcal{P}(\lambda)$

Empleando v.a. uniformes para generar las exponenciales, tenemos que

$$
\begin{aligned}
N(1) &= \max\{n \mid X_1 + X_2 + \cdots + X_n \leq 1\} \\
&= \max\{n \mid -\frac{1}{\lambda}(\log(U_1) + \log(U_2) + \cdots + \log(U_n)) \leq 1\} \\
&= \max\{n \mid -\frac{1}{\lambda}(\log(U_1 \cdot U_2 \cdots U_n)) \leq 1\} \\
&= \max\{n \mid \log(U_1 \cdot U_2 \cdots U_n) \geq -\lambda\} \\
&= \max\{n \mid U_1 \cdot U_2 \cdots U_n \geq e^{-\lambda}\}
\end{aligned}
$$

$$N(1) = \min\{n \mid U_1 \cdot U_2 \cdots U_n < e^{-\lambda}\} - 1$$

Generación de una v. a. Poisson $X \sim \mathcal{P}(\lambda)$

Algorithm 8: Transformada Inversa

Generar $U_1 \sim \mathcal{U}(0,1)$;
if $U_1 < e^{-\lambda}$ **then**
| $X \leftarrow 0$ y terminar.
end
Generar $U_2 \sim \mathcal{U}(0,1)$
if $U_1.U_2 < e^{-\lambda}$ **then**
| $X \leftarrow 1$ y terminar.
end
Generar $U_3 \sim \mathcal{U}(0,1)$
if $U_1.U_2.U_3 < e^{-\lambda}$ **then**
| $X \leftarrow 2$ y terminar.
end
$\vdots$

Generación de una v. a. con distribución Gamma

- Si $X_1, \ldots, X_n$ son v. a. exponenciales independientes, $X_i \sim \mathcal{E}(\lambda)$, entonces
$$X = X_1 + \cdots + X_n$$
es una v. a. gamma, de parámetros (n, λ).

$$\begin{aligned} F_X(x) &= P(X_1 + X_2 + \cdots + X_n \leq x) \\ &= P(-\frac{1}{\lambda}(\log(U_1 \cdot U_2 \cdots U_n)) \leq x) \\ &= P(-\frac{1}{\lambda}(\log(U_1 \cdot U_2 \cdots U_n)) \leq x\} \end{aligned}$$

Por lo cual

Algorithm 9: $X \sim \gamma(n, \lambda)$

Generar $U_1, \ldots, U_n \sim \mathcal{U}(0, 1)$;

$X \leftarrow -\frac{1}{\lambda}\log(U_1 \ldots U_n)$

Generación de una v. a. con distribución Gamma

Algorithm 10: Algoritmo: $X \sim \gamma(n, \lambda)$

Generar $U_1, \ldots, U_n \sim \mathcal{U}(0, 1)$;

$X \leftarrow -\frac{1}{\lambda} \log(U_1 \ldots U_n)$

- Emplea *n* uniformes.
- Calcula un único logaritmo.

- Para generar *n* exponenciales independientes, hacen falta generar *n* uniformes y calcular *n* logaritmos.

Generación de exponenciales a partir de una distribución gamma

Teorema

Si $X, Y \sim \mathcal{E}(\lambda)$, e independientes, entonces

$$f_{X|X+Y}(x \mid t) = \frac{1}{t}\mathbb{I}_{(0,t)}(x),$$

es decir, X condicional a $X + Y = t$ es uniforme en $(0, t)$.

Para generar X, Y exponenciales independientes, de parámetro λ podemos aplicar el siguiente algoritmo:

Algorithm 11: $X,\ Y \sim \mathcal{E}(\lambda)$

Generar $U_1,\ U_2 \sim \mathcal{U}(0,1)$;
$t \leftarrow -\frac{1}{\lambda}\log(U_1\ U_2)$;
Generar U_3;
$X \leftarrow t\,U_3$;
$Y \leftarrow t - X$

- Calcula un único logaritmo.
- Emplea 1 uniforme adicional.

Generación de exponenciales a partir de una distribución gamma

Para generar n exponenciales, se puede extender el método anterior generando

- una v. a. gamma, de parámetros (n, λ),
- $n - 1$ v. a. uniformes, adicionales.

Algorithm 12:

Generar $U_1, \ldots, U_n \sim \mathcal{U}(0,1)$;
$t \leftarrow -\dfrac{1}{\lambda}\log(U_1 \cdots U_n)$;
Generar $V_1, \ldots, V_{n-1}$ y ordenarlos de menor a mayor;
$X_1 \leftarrow t\, V_1$;
$X_2 \leftarrow t\,(V_2 - V_1)$;
$\vdots$;
$X_{n-1} \leftarrow t\,(V_{n-1} - V_{n-2})$;
$X_n \leftarrow t - t\, V_{n-1}$
