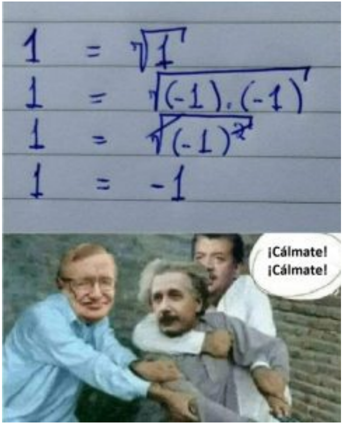


Figura: Homenaje a dos grandes físicos: ¿saben quienes son?

Probabilidad - Clase 23
Simulación de variables y vectores aleatorios

Ernesto Mordecki

Facultad de Ciencias, Universidad de la República. Montevideo, Uruguay

Curso de la Licenciatura en Matemática - 2020

Contenidos

Simulación de una variable uniforme

Un generador se define por la relación de recurrencia

$$X_{n+1} = (aX_n + c) \bmod m$$

donde X es una sucesión (pseudo) aleatoria[1], y

- m con $0 < m$ es el módulo (ej.: 2^{32})
- a tal que $0 < a < m$ es el multiplicador (ej: 1664525)
- c tal que $0 \leq c < m$ es le incremento, (ej: 1013904223)
- X_0, tal que $0 \leq X_0 < m$ es la semilla (la elige el usuario)
- En general m es primo, o c, m son co-primos, u otras variantes

[1] Tomado de Wikipedia `https://en.wikipedia.org/wiki/Linear_congruential_generator`

Simulación: método de la distribución inversa

Lema Sea X una v.a. con distribución F, que es continua biyectiva. Sea U una variable uniforme en $[0,1]$. Entonces, la v.a.

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

tiene distribución F (como X)

Demostración.

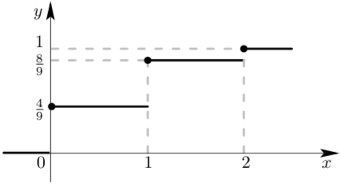
$$\mathbf{P}(Y \leq x) = \mathbf{P}(F^{-1}(U) \leq x) = \mathbf{P}(U \leq F(x)) = F(x).$$

Esto nos da un método para simular v.a., cuando se pueda obtener la distribución de X.

Si la F no es biyectiva, se define

$$F^{-1}(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\},$$

y el mismo lema funciona. Veamos una figura:



Aplicación: simulación de una v.a. exponencial

Sea $T : \Omega \to \mathbb{R}$ una variable aleatoria exponencial, tiene distribución

$$F_T(x) = P(T \leq x) = \begin{cases} 0, & 0 < x, \\ 1 - e^{-\alpha x}, & x \leq 0. \end{cases}$$

La distribución F_T es monótona creciente por lo que tiene inversa:

$$\begin{aligned} y &= 1 - e^{-\alpha x} \\ e^{-\alpha x} &= 1 - y \\ -\alpha x &= \log(1 - y) \\ x &= -\frac{1}{\alpha}\log(1 - y) \end{aligned}$$

Entonces

$$F^{-1}(p) = -\frac{1}{\alpha}\log(1-p)$$

Esto implica que

$$-\frac{1}{\alpha}\log(1-U) \sim \exp(\alpha)$$

Pero en realidad, el comando que se usa es

$$-\frac{1}{\alpha}\log U \sim \exp(\alpha)$$

Porque si U es uniforme en $[0,1]$, también lo es $1-U$:

$$P(1-U \leq x) = \mathbf{P}(U \geq 1-x) = 1-\mathbf{P}(U < 1-x) = 1-(1-x) = x.$$

Distribución triangular

Si U_1, U_2 son uniformes en $[0,1]$, entonces $X = U_1 + U_2$ tiene una distribución triangular. La densidad y la distribución de X son

$$f_X(x) = \begin{cases} x, & 0 \leq x \leq 1, \\ 2 - x, & 1 \leq x \leq 2, \end{cases}$$

$$F_X(x) = \begin{cases} x^2/2, & 0 \leq x \leq 1, \\ 1 - (2-x)^2/2, & 1 \leq x \leq 2. \end{cases}$$

Invertimos F. Primero

$$y = x^2/2, \quad 2y = x^2, \quad \sqrt{2y} = x.$$

Después

$$y = 1 - (2-x)^2/2, \quad 2 - 2y = (2-x)^2,$$

$$\sqrt{2-2y} = 2 - x, \quad 2 - \sqrt{2-2y} = x.$$

Entonces

$$F_X^{-1}(x) = \begin{cases} \sqrt{2x}, & 0 \leq x \leq 1/2, \\ 2 - \sqrt{2(1-x)}, & 1/2 \leq x \leq 1. \end{cases}$$

Aplicando el método de la transformación inversa, si U es uniforme, $F_X^{-1}(U)$ tiene distribución triangular en $[0, 2]$.

Distribución de Pareto

Una característica muy importante de las v.a. es el peso de sus "colas". Por eso entendemos cuan rápido va a cero

$$\mathbf{P}(X > x) \quad \text{cuando } x \to \infty$$

o (lo que es equivalente para estudiar si hay monotonía)

$$p(x) \quad \text{cuando } x \to \infty$$

El peso de la cola es una medida importante de riesgo, porque representa la probabilidad de obtener valores grandes, que en determinados contextos pueden representar eventos catastróficos, o grandes pérdidas económicas.

Ejemplos: colas livianas (light tails)

- Si la densidad tiene soporte compacto (uniforme, triangular) entonces los valores que toma están acotados: Una uniforme en $[a, b]$ nunca tomará un valor mayor que b.
- La distribución normal (estándar) tiene colas livianas:

$$\mathbf{P}(-3 \leq Z \leq 3) = 0.9973, \qquad \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

- La distribución exponencial tiene colas livianas (pero mas pesadas que una normal)

$$p(x) = \alpha e^{-\alpha x}.$$

Ejemplos: colas pesadas (heavy tails)

- Distribución de Cauchy (la vimos)

$$p(x) = \frac{1}{\pi(1 + x^2)}$$

- Distribución de Pareto (la definimos ahora): X tiene distribución de Pareto de parámetro α, si tiene densidad dada por

$$p(x) = \begin{cases} \frac{\alpha}{x^{\alpha+1}}, & \text{cuando } x \geq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Para simular variables Pareto, calculamos la distribución. Tomamos $x \geq 1$:

$$F(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^{x} t^{-\alpha-1} dt = \int_{1}^{x} t^{-\alpha-1} dt$$
$$= \left[t^{-\alpha}\right]_{-\infty}^{x} = 1 - \frac{1}{x^{\alpha}}.$$

Hallamos ahora la inversa de F

$$1 - \frac{1}{x^{\alpha}} = y, \qquad 1 - y = \frac{1}{x^{\alpha}}$$

$$\frac{1}{1-y} = x^{\alpha}, \qquad x = \frac{1}{(1-y)^{1/\alpha}}.$$

Entonces, si U es uniforme en $[0, 1]$, la v.a.

$$P = \frac{1}{U^{1/\alpha}}$$

tiene distribución de Pareto de parámetro α.

Simulación de un vectores gaussianos bidimensional (Box-Muller)

Como la función

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)dt,$$

no tiene fórmula elemental, no se puede aplicar el método de la transformación inversa.

Teorema (Método de Box-Muller) Sean U, V dos variables aleatorias uniformes independientes. Entonces

$$(X, Y) = \left(\sqrt{-2\log(U)}\sin(2\pi V), \sqrt{-2\log(U)}\cos(2\pi V)\right)$$

es un vector normal bivariado con[2]

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

[2] μ es el *a* de antes.

Demostración

Sea $A \subset \mathbb{R}^2$. Se tiene que

$$P((X,Y) \in A) = \iint_{(\sqrt{-2\log(u)}\sin(2\pi v), \sqrt{-2\log(u)}\cos(2\pi v)) \in A} dudv$$

Hacemos el cambio de variable

$$g(u,v) = (\sqrt{-2\log(u)}\sin(2\pi v), \sqrt{-2\log(u)}\cos(2\pi v)) = (x,y)$$

Despejando en términos de u y v tenemos que

$$g^{-1}(x,y) = \left(e^{\frac{-x^2-y^2}{2}}, \frac{1}{2\pi}\arctan(y/x) \right)$$

Calculemos el jacobiano: $\mathbb{J}_{g^{-1}}(x,y)$

$$\mathbb{J}_{g^{-1}}(x,y) = \begin{vmatrix} e^{\frac{-x^2-y^2}{2}}(-x) & e^{\frac{-x^2-y^2}{2}}(-y) \\ \frac{1}{2\pi(1+(y/x)^2)}(-y/x^2) & \frac{1}{2\pi(1+(y/x)^2)}(1/x^2) \end{vmatrix}$$
$$= \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}$$

Luego,

$$\begin{aligned} P((X,Y)\in A) &= \iint_{(\sqrt{-2\log(u)}\sin(2\pi v),\sqrt{-2\log(u)}\cos(2\pi v))\in A} dudv \\ &= \frac{1}{2\pi}\iint_{(x,y)\in A} e^{-\frac{1}{2}(x^2+y^2)} dxdy \end{aligned}$$

Figura: John Nash (1928 - 2015): matemático, premio Nobel de economía. Pueden ver la peli “A beatiful mind” (Una mente brillante)