(Parcial 1) Simulación

David Isaac López Romero
CU: 173993

4) Para la distribución triangular $Tri(a, c, b)$ con densidad $(a<c<b)$

$$f_x(x) = \frac{2(x-a)}{(b-a)(c-a)} \mathbb{1}_{\{a \le x \le c\}} + \frac{2(b-x)}{(b-a)(b-c)} \mathbb{1}_{\{c \le x \le b\}}$$

Hallar distribución inversa $F^{-1}(u)$ y obtener muestra de $n=3$ con la semilla.

Busco F.p.a. por casos: Si $x \in [a, c) \Rightarrow$

$$F_x(x) = \frac{2}{(b-a)(c-a)} \int_a^x (t-a)\,dt = \frac{2}{(b-a)(c-a)} \left\{\frac{t^2}{2} - at\right\}\Big]_a^x$$

$$= \frac{2}{(b-a)(c-a)} \left\{\frac{x^2}{2} - ax - \frac{a^2}{2} + a^2\right\} = \frac{2}{(b-a)(c-a)} \left\{\frac{x^2}{2} - ax + \frac{a^2}{2}\right\}$$

$$= \frac{x^2 - 2ax + a^2}{(b-a)(c-a)} = \frac{(x-a)^2}{(b-a)(c-a)} \quad \ldots (*)$$

Si $x \in [c, b) \Rightarrow$

$$F_x(x) = \frac{2}{(b-a)(c-a)} \int_a^c (t-a)\,dt + \frac{2}{(b-a)(b-c)} \int_c^x (b-t)\,dt$$

$$\overset{(*)}{=} \frac{(c-a)^2}{(b-a)(c-a)} + \frac{2}{(b-a)(b-c)} \left\{bt - \frac{t^2}{2}\right\}\Big]_c^x = \frac{(c-a)}{b-a} + \frac{2}{(b-a)(b-c)}\left\{bx - \frac{x^2}{2} - bc + \frac{c^2}{2}\right\}$$

$$= \frac{(c-a)(b-c) + 2bx - x^2 - 2bc + c^2}{(b-a)(b-c)} = \frac{bc - c^2 - ab + ac + 2bx - x^2 - 2bc + c^2}{(b-a)(b-c)}$$

$$= \frac{bc - ab + ac + 2bx - x^2}{(b-a)(b-c)} = \frac{bc - ab + ac + b^2 - b^2 + 2bx - x^2}{(b-a)(b-c)} \quad \ldots (*)$$

Notamos que $(b-a)(b-c) = b^2 - bc - ab + ac$

Por lo que, de $(*)$ resulta que

$$= \frac{(b-a)(b-c) - (b^2 - 2bx + x^2)}{(b-a)(b-c)} = 1 - \frac{(b-x)^2}{(b-a)(b-c)}$$

De esta forma,

$$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \dfrac{(x-a)^2}{(b-a)(c-a)} & \text{si } a \leq x < c \\ 1 - \dfrac{(b-x)^2}{(b-a)(b-c)} & \text{si } c \leq x < b \\ 1 & \text{si } b \leq x \end{cases}$$

Así pues, $F^{-1}(u)$ también es por casos

Si $x \in [a,c)$

$$y = \frac{(x-a)^2}{(b-a)(c-a)} \Leftrightarrow y(b-a)(c-a) = (x-a)^2$$

$$\underset{(x>a)}{\Leftrightarrow} x - a = +\sqrt{y(b-a)(c-a)} \Leftrightarrow x = a + \sqrt{y(b-a)(c-a)}$$

$$\therefore X = a + \sqrt{U(b-a)(c-a)} \text{ con } U \sim \text{Unif}(0,1)$$

Si $x \in [c,b)$

$$y = 1 - \frac{(b-x)^2}{(b-a)(b-c)} \Leftrightarrow (y-1)(b-a)(b-c) = -(b-x)^2$$

$$\underset{x<b}{\Leftrightarrow} b - x = \pm\sqrt{(1-y)(b-a)(b-c)} \Leftrightarrow x = b - \sqrt{(1-y)(b-a)(b-c)}$$

$$\therefore X = b - \sqrt{(1-U)(b-a)(b-c)} \text{ con } U \sim \text{Unif}(0,1)$$

Si $U \sim \text{Unif}(0,1) \rightarrow (1-U) \sim \text{Unif}(0,1)$ también,

$$\therefore X = b - \sqrt{U(b-a)(b-c)} \text{ con } U \sim \text{Unif}(0,1) /\!/$$

$$\therefore F^{-1}(u) = \begin{cases} a + \sqrt{u(b-a)(c-a)} & \text{si } 0 \leq u < \dfrac{c-a}{b-a} \quad (*) \\ b - \sqrt{u(b-a)(b-c)} & \text{si } \dfrac{c-a}{b-a} \leq u < 1 \end{cases}$$

$(u \in (0,1))$

Falta evaluar manualmente con $a=1,\ c=2,\ b=3$

set.seed(1968)

runif(3)

$u_1 = 0.1629056,\ u_2 = 0.4126848,\ u_3 = 0.2633616$

Primero, hallamos donde cambia la $F^{-1}(u)$ explícita.

$$u = \frac{2-1}{3-1} = \frac{1}{2}$$

$$\Rightarrow F^{-1}(u) = \begin{cases} 1+\sqrt{u(2)(1)} & \text{si } 0 \le u < \frac{1}{2} \\ 3-\sqrt{u(2)(1)} & \text{si } \frac{1}{2} \le u \le 1 \end{cases}$$

$$\therefore F^{-1}(u) = \begin{cases} 1+\sqrt{2u} & \text{si } 0 \le u < \frac{1}{2} \\ 3-\sqrt{2u} & \text{si } \frac{1}{2} \le u \le 1 \end{cases}$$

Para u_1 $\left(u_1 < \frac{1}{2}\right)$

$F^{-1}(u_1) = 1.570798739$

Para u_2 $\left(u_2 < \frac{1}{2}\right)$

$F^{-1}(u_2) = 1.908498542$

Para u_3 = $\left(u_3 < \frac{1}{2}\right)$

$F^{-1}(u_3) = 1.725756484$

4) Hay n cartas y n sobres. Cada sobre se diseñó para cada carta. Se meten las cartas al azar. Calcular probabilidad ninguna carta se ponga en su propio sobre considerando $n=10$. (En R).

5) Ver si el GLC alcanza un periodo máximo.

$a = 2,814,749,767,109$

$c = 59,482,661,568,307$

$m = 2^{48}$

Dem:

Debemos usar el teorema Hull & Dobell con el caso particular donde el módulo es de la forma $m = 2^k (\geq 4)$

Un GLC se define

$$Z_i \equiv (a Z_{i-1} + c) \text{ mód } m$$

Ahora, para el GLC que nos dieron,

un GLC tiene periodo completo $\Leftrightarrow$ c es impar y $a \equiv 1 \text{ mod } 4$
$(m = 2^k \geq 4)$

Vemos que c es impar, falta ver la segunda condición.

$a \equiv 1 \text{ mod } 4 \Rightarrow 2,814,749,767,108$ se divide por 4
($a-1$ se divide por 4)

$$\begin{array}{r} 703{,}687{,}441{,}777 \\ 4 \overline{)2{,}814{,}749{,}767{,}108} \\ 0\ 14 \\ 27 \\ 34 \\ 29 \\ 17 \\ 16 \\ 07 \\ 31 \\ 30 \\ 28 \\ 0 \end{array}$$

$\therefore a = 2{,}814{,}749{,}767{,}108$

$= 4(703{,}687{,}441{,}777) + 0$

El residuo de la división es 0

$\therefore$ El GLC tiene periodo máximo completo.

6) Probar si los números $2^\circ, 9^\circ, 16^\circ, \ldots,$ en la sucesión están autocorrelacionados con $\alpha = 0.05$

Autocorrelación de rezago (lag) j

$$\rho_j = \frac{Cov(X_i, X_{i+j})}{Desv(X_i)\,Desv(X_{i+j})} = \frac{Cov(X_i, X_{i+j})}{Var(X_i)}$$

$2^\circ, 9^\circ, 16^\circ, \ldots$ es una sucesión

$\left.\begin{array}{l} 9-2=7 \\ 16-9=7 \end{array}\right\}$ d- distancia entre cada número.

$X_{i+1} = X_i + 7$ es la Fórmula de la sucesión. $(j = 7)$

(Ver en R) Se aplicó la prueba Box-Pierce y, también, Se le aplicó la prueba de Ljung-Box para verificar simultáneamente varias autocorrelaciones, pues es una muestra pequeña.

Box-Pierce se obtuvo 0.8085

Ljung-Box se obtuvo 0.7544

$\therefore$ La sucesión está autocorrelacionada con $\alpha = 0.05$

8) Procedimiento aceptación-rechazo para muestras de

$$h_x(x) = \frac{e^{-x^2/2}}{1+|x|} \mathbb{1}_{\mathbb{R}}$$

Verificar eficiencia pues $\int_{-\infty}^{\infty} h_x(x)\,dx = 1{,}5413$

$h(x) = Ker(x)$

Opciones de v.a's a usar: Normal, Cauchy, Laplace

Consideremos $Y \sim N(0,1)$ como función mayorizante ($g_Y(y)$)

Buscamos c (mínima) tal que $\frac{h(x)}{g(x)} \leq c$

$$\frac{h(x)}{g(x)} = \frac{\frac{\cancel{e^{-x^2/2}}}{1+|x|}}{\frac{1}{\sqrt{2\pi}}\cancel{e^{-x^2/2}}} = \frac{\sqrt{2\pi}}{1+|x|}$$

$$\underset{x\in\mathbb{R}}{MAX}\left\{\frac{h(x)}{g(x)}\right\} = \underset{x\in\mathbb{R}}{MAX}\left\{\frac{\sqrt{2\pi}}{1+|x|}\right\}$$ (Notemos que $\frac{h(x)}{g(x)}$ es par basta optimizar un lado)

Si $x>0 \Rightarrow \frac{h(x)}{g(x)} = \frac{2\pi}{1+x} = 2\pi(1+x)^{-1}$

$$\Rightarrow \left(\frac{h(x)}{g(x)}\right)' = -2\pi(1+x)^{2} = -\frac{\sqrt{2\pi}}{(1+x)^2} < 0 \quad \forall x>0$$

Análogamente, $\left(\frac{h(x)}{g(x)}\right)' = \frac{\sqrt{2\pi}}{(1-x)^2} > 0 \quad \forall x<0$

Falta verificar el punto singular $x=0$

En $x=0$ $\frac{h(0)}{g(0)} = \sqrt{2\pi}$ que es el valor máximo pues la

Función decrece si $x \to -\infty$ y si $x \to +\infty$

Entonces, si $x=0$, resulta que la $c = \sqrt{2\pi}$

Concluimos que

$$\frac{h(x)}{g(x)} \leq \sqrt{2\pi} \Rightarrow h(x) \leq \sqrt{2\pi}\, g(x)$$

- Generamos $u \sim Unif(0,1)$
- Si $u \leq \frac{h(x)}{c\,g(x)} = \frac{\sqrt{2\pi}}{\sqrt{2\pi}\,(1+|x|)} = \frac{1}{1+|x|}$ $\therefore u \leq \frac{1}{1+|x|}$
- Finalmente, si $u \leq \frac{1}{1+|x|}$ se acepta, sino, se rechaza

Sabemos que el número promedio de iteraciones para aceptar $\frac{1}{c}$. La c que tenemos es muy grande pues envuelve a $h(x)$ la cual $\int_{-\infty}^{\infty} h(x)dx = 1.5413$. Le faltan las constantes para que integre 1. La c es para el Kernel de esa v.a., no para cada función de densidad de probabilidad.

Como $c = \sqrt{2\pi} \Rightarrow \frac{1}{\sqrt{c}} = \frac{1}{\sqrt{2\pi}} \approx 0.3989$

Entonces aproximadamente 4 de cada 10 ensayos se acepta la u, lo cual en cuestión de eficiencia de tiempo es muy bueno.