

# SIMULACIÓN, ALGUNOS MÉTODOS Y APLICACIONES

Hugo Carrasco \*

## Resumen

La palabra **simulación**, actualmente, es de uso corriente y cotidiano; su significado es, según la Real Academia, *representar una cosa, fingiendo o imitando lo que no es*.

En estadística, simular es esencialmente una técnica de muestreo controlada, que es usada en conjunto con un modelo, para obtener respuestas aproximadas a preguntas respecto de complicados problemas probabilísticos. Esto es muy usado cuando las técnicas analíticas y numéricas no son capaces de dar respuestas. Cuando los modelos tienen una formulación matemática, pero las predicciones analíticas no son posibles, entonces es frecuente que simular se transforme en una poderosa herramienta, no solo para describir el modelo en si mismo, sino para investigar como puede cambiar el comportamiento ante cambios de los parámetros del modelo.

Analizaremos el método Monte Carlo, desarrollado por John von Neumann y S. M. Ulam, durante el proyecto Manhattan

A partir del método Monte Carlo desarrollaremos algunos algoritmos particulares para variables aleatorias no uniformes, primeramente, y dada la importancia, trabajaremos con variables aleatorias con distribución Normal, luego, nos concentraremos en algunas otras variables absolutamente continuas (por ejemplo Exponencial, Gamma, Chi-cuadrado), y por último, veremos algunas técnicas para algunas variables aleatorias discretas (por ejemplo Bernoulli, Binomial, Poisson).

Para casos más generales desarrollaremos algunos de los métodos de transformación de variables uniformes.

Por último analizaremos algunas implementaciones informáticas en  $\mathbf{R}$  y las potencialidades de estos métodos.

**Palabras clave:** *simulación, método Monte Carlo, variables aleatorias*

---

\*Departamento de Métodos Matemáticos-Cuantitativos, email: carrasco@ccee.edu.uy

# 1. Método de Monte Carlo

El método de Monte Carlo o simulación cruda, es la generación de números aleatorios u observaciones de una variable aleatoria uniforme, en general lo haremos en el intervalo  $[0,1]$ . Estos números son seleccionados de manera que cada uno de ellos tenga igual probabilidad de salir, sin tener en cuenta cuantas veces ya haya salido. Se pueden generar de diversas maneras, utilizando tablas, bolilleros, computadoras, etc..

En libros, relativamente antiguos, se hace referencia a tablas de números aleatorios, que eran usadas en selección de muestras o para establecer diseños experimentales; actualmente es muy raro ver a un Estadístico usando una tabla impresa de números aleatorios, ocasionalmente tendrá acceso a alguna de esas tablas por medio de su computador, aunque en general su computador es usado para *generar números aleatorios*.

Las computadoras no pueden generar números aleatorios, y esto no es conveniente para conectar la computadora con eventos aleatorios externos. Para la mayor parte de las aplicaciones estadísticas, esto no es un problema ya que existen muchos métodos de generación de números pseudo-aleatorios, cuyas muestras se comportan como si fueran aleatorios. Existen infinidad de estos métodos, en general aplican fórmulas recurrentes basándose en un valor inicial, estos métodos no serán tratados en el presente trabajo, solo supondremos que tenemos la posibilidad de obtener los números aleatorios.

En general, no se hacen distinciones precisas respecto a los términos *simulación*, *Monte Carlo* y *remuestreo*.

## 2. Métodos particulares para variables aleatorias no uniformes

En este capítulo trataremos algunos métodos de simulación para algunas distribuciones particulares, para nada es exhaustivo, pues solo veremos los que nos resultaron de mayor interés y simpatía.

### 2.1. Variables aleatorias con distribución Normal

Dada la importancia para la estadística de las variables aleatorias con distribución normal es que haremos un desarrollo de varias técnicas para la simulación de este tipo de variables aleatorias.

Sabemos que si  $X \sim N(\mu, \sigma^2)$  podemos decir que  $X = \sigma Z + \mu$ , donde  $Z \sim N(0, 1)$ . Por esta simple relación es suficiente ver métodos para generar variables aleatorias normales estándar,  $N(0, 1)$

### 2.1.1. Usando teorema central del límite

Usando teorema central del límite frecuentemente encontramos la distribución normal, y esto es la base de mucha de la teoría estadística; la idea será usar estos resultados para simular variables aleatorias normales [3] [4].

Por ejemplo, podemos simular  $n$  variables aleatorias independientes, con distribución  $U[0, 1]$ ,  $U_1, U_2, \dots, U_n$ , y sea  $S = \sum_{i=1}^n U_i$ , por el teorema central del límite, cuando  $n \rightarrow \infty$ , la distribución de  $S$  tiende a una normal.

En la práctica podemos establecer que para valores finitos de  $n$ , se tiene que  $S$  tiene una distribución aproximadamente normal. El problema que surge es ¿cuál sería un  $n$  mínimo desde el cual podamos decir que la aproximación es aceptable?

Si  $n = 2$  es inaceptable, pues  $S$  tiene distribución triangular. Desde  $n = 3$  en adelante  $S$  tiene una agradable forma de campana, el resultado para el valor  $n = 12$  es de una aproximación aceptable, y como  $E(U_i) = 1/2$  y  $Var(U_i) = 1/12$ , tenemos que  $S$  tiene distribución aproximadamente  $N(6, 1)$ , de donde  $Z = S - 6$  tendrá distribución  $N(0, 1)$ .

Una implementación en **R** de este método para generar  $m$  observaciones de  $Z$  es:

```
norm_tcl ← function(m)
z ← rep(0,m)
for(i in 1:m) {
  u ← runif(12);
  z[i] ← sum(u)-6;
}
return(z)
```

Claramente este método, además de ser aproximado, es un tanto lento, ya que para cada simulación de  $N$  necesitaremos simular doce uniformes.

### 2.1.2. Métodos de transformaciones polares

En estos métodos utilizaremos notación en coordenadas polares para llegar a los resultados requeridos. Cualquier punto del plano, de  $\mathbb{R}^2$ , se puede escribir en sus coordenadas cartesianas  $(x, y)$  o en sus coordenadas polares  $(r, \theta)$ , donde

$$r = \sqrt{x^2 + y^2} \quad \theta = \arctan\left(\frac{y}{x}\right)$$

En forma análoga se puede transformar un punto en coordenadas polares  $(r, \theta)$ , en sus coordenadas cartesianas  $(x, y)$ , como  $x = r \cos \theta$ ,  $y = r \sin \theta$  (ver la figura 1)

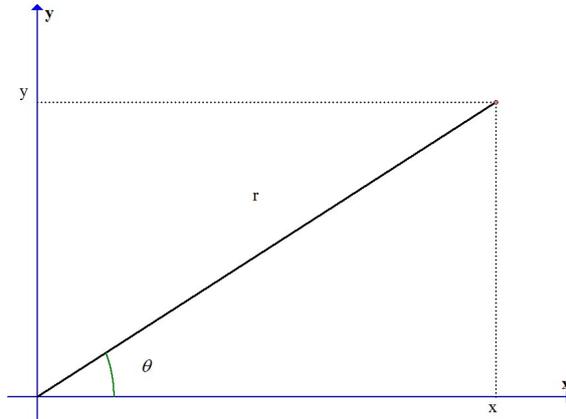


Figura 1 : Transformación a coordenadas polares

■ **Método directo**

Sean  $X$  y  $Y$  variables aleatorias independientes con distribución  $N(0, 1)$ . El par  $(X, Y)$  define un punto en el plano. La densidad bivariada de normales independientes se obtiene como el producto de densidades, así, tenemos

$$f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$$

ahora escrito en sus coordenadas polares

$$X = R \cos \Theta \quad Y = R \sin \Theta$$

podemos hacer el cambio de variable, a través del cálculo del jacobiano

$$J^{-1} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r$$

$$\Rightarrow f_{R,\Theta}(r, \theta) = \underbrace{\frac{1}{2\pi}}_{f_\Theta} \underbrace{r e^{-\frac{r^2}{2}}}_{f_R} = f_\Theta(\theta) f_R(r)$$

Así tenemos que  $R$  y  $\Theta$  son variables aleatorias independientes,  $\Theta$  es  $U[0, 2\pi]$  y  $R$  tiene una distribución exponencial de parámetro  $1/2$  (esto último se puede obtener fácilmente haciendo un cambio de variable  $T = R^2$ , obteniendo que  $f_T(t) = \frac{1}{2} e^{(-T/2)}$ )

■ **Método de Box-Muller**

Sean  $U_1$  y  $U_2$  variables aleatorias independientes con distribución uniforme en  $[0, 1]$ , entonces Box y Muller (1958) [3] [4] mostraron que

$$X = \sqrt{-2L(U_1)} \cos(2\pi U_2)$$

$$Y = \sqrt{-2L(U_1)} \sin(2\pi U_2)$$

son variables aleatorias independientes con distribución  $N(0, 1)$ . Para demostrar esto debemos partir de que

$$f_{U_1, U_2}(u_1, u_2) = \begin{cases} 1 & \text{si } 0 \leq u_1 \leq 1, 0 \leq u_2 \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

por otra parte:  $u_1 = e^{-\frac{1}{2}(x^2+y^2)}$ ,  $u_2 = \frac{1}{2\pi} \arctan\left(\frac{y}{x}\right)$

$$\begin{aligned} J^{-1} &= \begin{vmatrix} \frac{\partial u_1}{\partial x} & \frac{\partial u_1}{\partial y} \\ \frac{\partial u_2}{\partial x} & \frac{\partial u_2}{\partial y} \end{vmatrix} = \begin{vmatrix} -xu_1 & -yu_1 \\ \frac{y}{2\pi(x^2+y^2)} & \frac{-x}{2\pi(x^2+y^2)} \end{vmatrix} = \\ &= \frac{x^2 u_1}{2\pi(x^2+y^2)} + \frac{y^2 u_1}{2\pi(x^2+y^2)} = \frac{u_1}{2\pi} = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} \end{aligned}$$

con lo que queda probado.

Entonces para simular dos variables aleatorias independientes  $N(0, 1)$ , alcanza con generar dos variables aleatorias independientes  $U[0, 1]$  y hacer la transformación correspondiente.

Una implementación en **R** de este método para generar  $m$  observaciones de  $X$  e  $Y$  es:

```
norm_BM ← function(m) {
z ← matrix(data=0,nr=m,nc=2)
for(i in 1:m) {
u ← runif(2);
z[i,1] ← sqrt(-2*log(u[1]))*cos(2*pi*u[2])
z[i,2] ← sqrt(-2*log(u[1]))*sin(2*pi*u[2])
}
return(z)
}
```

#### ■ Método de Marsaglia

La idea de este método, propuesto por Marsaglia, Bray y MacLauren (1964) [3] [4], es nuevamente a partir de dos variables uniformes, ahora en  $[-1, 1]$ , según los valores simulados de la uniforme se seguirá adelante o se rechazará.

Sea  $U \sim U[0, 1]$ , entonces  $V = 2U - 1$  será uniforme en  $[-1, 1]$ .

Consideremos  $V_1$  y  $V_2$  dos variables aleatorias independientes uniformes en  $[-1, 1]$ , entonces queda definido un punto en el plano de coordenadas  $(V_1, V_2)$ , con coordenadas polares  $(\tilde{R}, \Theta)$ , que se obtienen como:

$$\tilde{R}^2 = V_1^2 + V_2^2, \quad \text{tg } \Theta = \frac{V_2}{V_1}$$

Cada punto generado aleatoriamente estará dentro del cuadrado de lado 2, el punto será rechazado si cae fuera del círculo de radio 1, inscrito dentro del cuadrado (ver la figura 2).

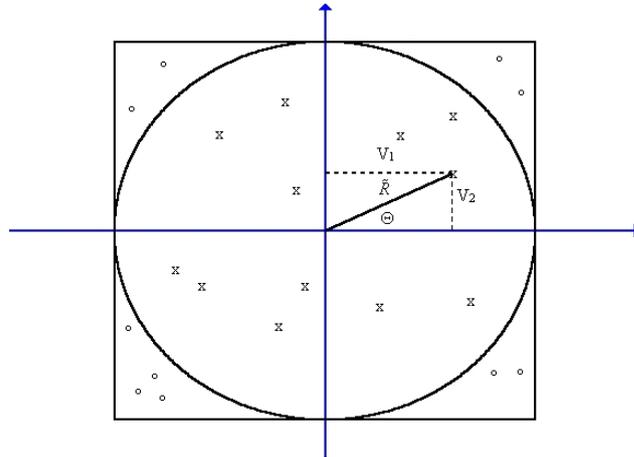


Figura 2 : Aceptación y rechazo en el método de Marsaglia

Es claro, después de todo lo que hemos visto, que  $\tilde{R}$  y  $\Theta$  son independientes, y además que  $\Theta \sim U[0, 2\pi]$  y que  $\tilde{R} \sim U[0, 1]$ . También es claro que la pareja  $(\tilde{R}, \Theta)$  son los requeridos por el método de Box-Muller, y podemos simplemente escribir que:

$$\sin \theta = \frac{V_2}{\tilde{R}} = V_2 (V_1^2 + V_2^2)^{-\frac{1}{2}} \quad \cos \theta = \frac{V_1}{\tilde{R}} = V_1 (V_1^2 + V_2^2)^{-\frac{1}{2}}$$

Así que  $X$  e  $Y$  son variables aleatorias  $N(0, 1)$  independientes

$$X = (-2L(V_1^2 + V_2^2))^{\frac{1}{2}} V_2 (V_1^2 + V_2^2)^{-\frac{1}{2}}$$

$$Y = (-2L(V_1^2 + V_2^2))^{\frac{1}{2}} V_1 (V_1^2 + V_2^2)^{-\frac{1}{2}}$$

Ahora si llamamos  $W = V_1^2 + V_2^2$ , podemos resumir en

$$X = V_2 \left( \frac{-2L(W)}{W} \right)^{\frac{1}{2}}$$

$$Y = V_1 \left( \frac{-2L(W)}{W} \right)^{\frac{1}{2}}$$

Claramente la proporción de rechazos es igual a la porción de área del cuadrado que no está dentro del círculo respecto al área total del cuadrado, y esto es  $1 - \pi/4$ .

Una implementación en **R** de este método para generar  $m$  observaciones de  $X$  e  $Y$  es:

```

norm_Mars ← function(m) {
z ← matrix(data=0,nr=m,nc=2)
for(i in 1:m) {
u ← runif(2);
v ← u
v[1] ← 2*u[1] -1
v[2] ← 2*u[2] -1
w ← v[1]2+v[2]2
if (w <=1) {
z[i,1] ← v[2]*sqrt(-2*log(w)/w)
z[i,2] ← v[1]*sqrt(-2*log(w)/w)
}
}
return(z)
}

```

## 2.2. Variables aleatorias Exponenciales, Gammas y Chi-cuadrados

Las variables aleatorias con distribución exponencial y gamma son de uso frecuente en modelos de tiempo de espera, y es una consecuencia natural de los procesos de Poisson [4].

El camino simple para obtener variables aleatorias con función de densidad exponencial, del tipo  $e^{-x}$ , para  $x \geq 0$ , es hacer:

$$X = -L(U) \quad \text{con} \quad U \sim U[0, 1]$$

Es fácil ver que si  $Y = \frac{X}{\lambda}$ , tenemos que la densidad de  $Y$  es:

$$f_Y(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Una implementación en **R** de este método para generar  $m$  observaciones de  $Y$  exponencial de parámetro  $\lambda$  es:

```

exp_dis ← function(m, lambda) {
y ← rep(0,m)
for(i in 1:m) {
u ← runif(1);
y[i] ← -log(u)/lambda
}
}
return(y)
}

```

Por otra parte, si tenemos  $Y_1, Y_2, \dots, Y_n$  variables aleatorias independientes con función de densidad  $f_Y$  tenemos que si

$$G = \sum_{i=1}^n Y_i \quad \Rightarrow \quad G \sim \Gamma(n, \lambda)$$

Este resultado se prueba hallando la densidad conjunta.

Así para generar una variable aleatoria con distribución  $\Gamma(n, \lambda)$ , con  $n$  un número natural, nosotros solo necesitamos hacer

$$G = -\frac{1}{\lambda} \sum_{i=1}^n L(U_i)$$

donde  $U_1, U_2, \dots, U_n$  son variables aleatorias  $U[0, 1]$  independientes.

Una implementación en **R** de este método para generar  $m$  observaciones de  $G$  Gamma de parámetros  $n$  y  $\lambda$  es:

```
gamma_dis <- function(m, n, lambda) {
  g <- rep(0,m)
  for(i in 1:m) {
    u <- runif(n);
    g <- sum(log(u))/lambda
  }
  return(g)
}
```

Una variable aleatoria  $X \sim \chi_n^2$  se puede pensar de dos maneras:

- Como una suma de  $n$  variables aleatorias con distribución  $N(0, 1)$  al cuadrado

$$X = \sum_{i=1}^n Z_i^2$$

Este resultado se obtiene del cálculo de la densidad conjunta.

Por lo cuál podríamos simular una variable aleatoria con distribución  $\chi_n^2$  a partir de simulaciones de normales independientes.

- Como una variable aleatoria con distribución  $\Gamma(\frac{n}{2}, \frac{1}{2})$ :
  - si  $n$  es par, utilizando el método recién desarrollado para las Gammas se obtendría sin mayor inconveniente.
  - Si  $n$  es impar, por el procedimiento para las Gammas, podemos obtener una variable aleatoria con distribución  $\chi_{n-1}^2$ , y utilizando la propiedad de las  $\chi^2$  expuesta en el punto anterior, podemos llegar a la  $\chi_n^2$ , adicionando una  $Z^2$ , donde  $Z \sim N(0, 1)$  es independiente de las variables anteriores.

## 2.3. Variables Variables aleatorias Bernoulli, Binomiales y Poisson

- **Variables aleatorias con distribución Bernoulli**

Una variable aleatoria  $B \sim Ber(p)$  toma solo dos valores,  $B = 1$  con probabilidad  $p$  y  $B = 0$  con probabilidad  $1 - p$ . Así para poder simular  $B$ , alcanza con una  $U \sim U[0, 1]$ , asignado a  $B$  el valor 1, si  $U \leq p$ , y asignando el valor 0 si  $U > p$  [4].

- **Variables aleatorias con distribución Binomial**

Una variable aleatoria  $X \sim Bin(n, p)$ , puede escribirse como

$$X = \sum_{i=1}^n B_i$$

donde  $B_1, B_2, \dots, B_n$  son variables aleatorias  $Ber(p)$  independientes. De esta forma y utilizando el punto anterior se puede simular  $X$ , simulando  $n$  variables aleatorias  $U[0, 1]$  independientes [4].

- **Variables aleatorias con distribución Poisson**

Las variables aleatorias con distribución Poisson de parámetro  $\lambda$  pueden ser generadas como consecuencia del siguiente resultado [4]:

Supongamos que  $E_i$ , con  $i \geq 1$  es una sucesión de variables aleatorias independientes, con distribución exponencial y función de densidad

$$f_{E_i}(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Sea  $S_0 = 0$  y  $S_k = \sum_{i=1}^k E_i$  para  $k \geq 1$ , como ya vimos  $S_k \sim \Gamma(k, \lambda)$ . Entonces la variable aleatoria  $K$ , definida implícitamente por las inecuaciones  $S_K \leq 1 < S_{K+1}$  tiene distribución Poisson con parámetro  $\lambda$ .

Entonces para simular  $K$ , debemos simular  $S_k$ , hasta alcanzar o rebasar el valor 1, recordamos que

$$S_k = -\frac{1}{\lambda} \sum_{i=1}^k L(U_i) = -\frac{1}{\lambda} L\left(\prod_{i=1}^k U_i\right)$$

donde las  $U_1, U_2, \dots, U_k$  son variables aleatorias uniformes en  $[0, 1]$  e independientes. La comparación  $S_k > 1$ , la podemos escribir como

$$-\frac{1}{\lambda} L\left(\prod_{i=1}^k U_i\right) > 1 \Leftrightarrow L\left(\prod_{i=1}^k U_i\right) < -\lambda \Leftrightarrow \prod_{i=1}^k U_i < e^{-\lambda}$$

### 3. Métodos generales de transformación de variables uniformes

Existen una gran variedad de métodos para obtener observaciones simuladas de distintas distribuciones a partir de transformaciones de variables uniformes, describiremos a continuación un par de ellas, que tienen en común una descripción simple y aplicabilidad directa.

#### 3.1. Transformación en probabilidad

La Transformación en probabilidad [1] es la utilización de la preimagen de la función de distribución de una variables aleatorias cualquiera. Luego a partir de observaciones de una variable con distribución uniforme en  $[0, 1]$ , obtendremos, mediante las preimagenes, observaciones de la variable que deseamos.

##### 3.1.1. Caso Absolutamente Continuo

Dada una variable aleatoria  $X$  con función de distribución  $F_X$ , si  $F_X$  es continua, y definimos la variable aleatoria  $U$  como resultado de la transformación  $U = F_X(X)$ , tenemos que la variable  $U$  tiene distribución uniforme en el intervalo  $[0, 1]$  [2].

Claramente  $U$  es una variable aleatoria, por ser la composición de una función con una variable aleatoria, también por esto el recorrido de  $U$  es el intervalo  $[0, 1]$ , por otra parte la función de distribución de  $U$  la obtenemos de

$$\begin{aligned} F_U(u) &= P(U \leq u) = P(F_X(X) \leq u) = P(F_X^{-1}(F_X(X)) \leq F_X^{-1}(u)) = \\ &= P(X \leq F_X^{-1}(u)) = F_X(F_X^{-1}(u)) = u \end{aligned}$$

De esta forma partiendo de valores de una uniforme en  $[0, 1]$ , obtenemos valores de una variable absolutamente continua cualquiera, solo conociendo su función de distribución. (ver figura 3)

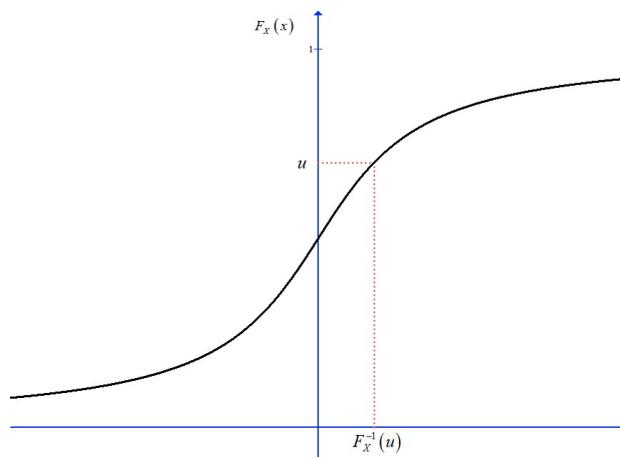


Figura 3 : Transformación en probabilidad, caso absolutamente continuo

### 3.1.2. Caso Discreto (recorrido finito)

Si  $X$  es una variable aleatoria discreta (de recorrido finito), también podemos usar el método [3], claro está que no pasará por la obtención de la inversa de la función de distribución, ya que no existe. Supongamos que la variable tiene por recorrido

$$Rec(X) = \{x_1, x_2, x_3, \dots, x_n\}, \text{ con } x_1 < x_2 < x_3 < \dots < x_n$$

y que su función de cuantía es

$$p_X(x) = \begin{cases} P(X = x) & \text{si } x \in Rec(X) \\ 0 & \text{en otro caso} \end{cases}$$

Con función de distribución

$$F_X(x) = \sum_{\substack{x_i \in Rec(X) \\ x_i \leq x}} p_X(x)$$

Para la aplicación del método, primero obtenemos  $u$  la observación de la variable aleatoria  $U$ , con distribución uniforme en  $[0, 1]$ , entonces, según la distribución, diremos que la preimagen es  $x$ , si  $x$  satisface

$$F_X(x^-) < u \leq F_X(x)$$

Esta última relación hace que el valor encontrado pertenezca al  $Rec(X)$  (ver figura 4).

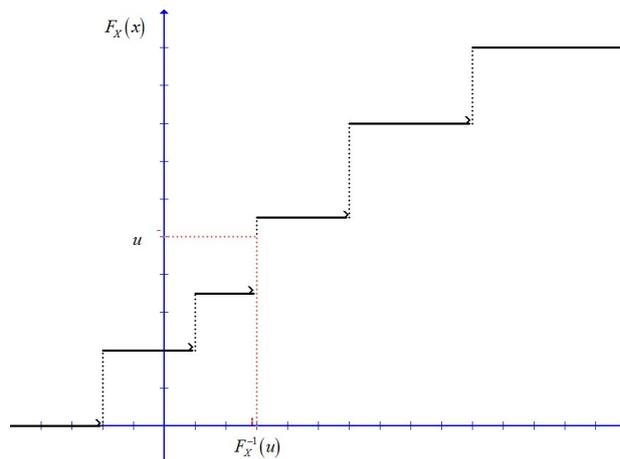


Figura 4 : Transformación en probabilidad, caso discreto

### 3.2. Método de Aceptación y Rechazo

Supongamos que tenemos una variable aleatoria  $X$  con función de densidad  $f_X(x)$ , y supongamos que la función es distinta de cero solo en un intervalo finito [3] [4], supongamos

$[\alpha, \beta]$ , entonces es fácil construir un rectángulo con base en dicho intervalo y altura igual al máximo de la función de densidad, supongamos  $\delta$ .

Ahora usando dos variables uniformes en  $[0, 1]$  e independientes,  $U_1$  y  $U_2$ , podemos obtener puntos al azar en el rectángulo, a partir de sus coordenadas cartesianas  $(\alpha + (\beta - \alpha)U_1, \delta U_2)$ . Los puntos que caen por encima de  $f_X(x)$  son rechazados, mientras que para los que caen por debajo, sus abscisas  $\alpha + (\beta - \alpha)U_1$ , se toman como una realización de  $X$  (ver figura 5).

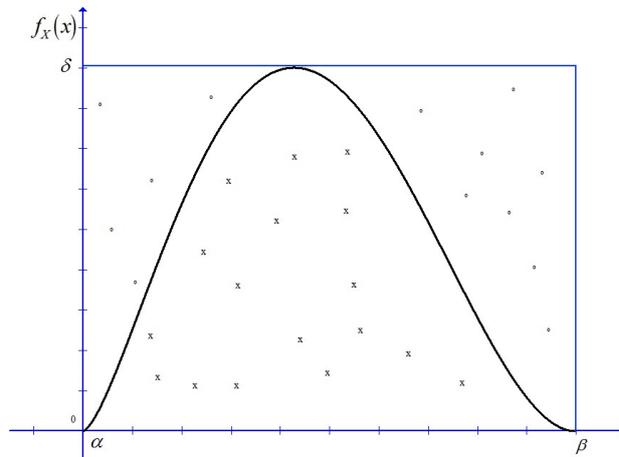


Figura 5 : Método de aceptación y rechazo

Como el área del rectángulo es  $\delta(\beta - \alpha)$ , y el área debajo de la curva es 1, la probabilidad de aceptación de un punto es  $\frac{1}{\delta(\beta - \alpha)}$ , y esta es grande, cuando  $\delta$  es muy pequeño, de igual manera si  $\delta$  es grande el método es poco eficiente y no se sugiere usar.

Una forma alternativa del método que soluciona este problema, y también la limitación de que el rango de la variable aleatoria sea finito se describe a continuación.

Consideremos una función de densidad  $h_W(x)$ , con el mismo rango de  $f_X(x)$ , pero para la cual sea relativamente más fácil simular, sea  $W$  una variable aleatoria con densidad  $h_W(x)$ , y sea  $Y$  una variable aleatoria con densidad condicionada a  $W = x$ , uniforme en  $[0, g(x)]$ , donde  $g(x) = k \cdot h_W(x)$ , con  $k > 1$ , llamado factor de separación. Ahora simulamos la variable  $W$ , obtenemos la observación  $x$ , ahora simulamos  $Y$ , que será  $U \cdot g(x)$ , donde  $U$  es una variable aleatoria independiente uniforme en  $[0, 1]$ , aceptaremos la observación  $X = x$  como una realización de la variable aleatoria con densidad  $f_X(x)$  si y solo si  $Y < f_X(x)$  (ver figura 6).

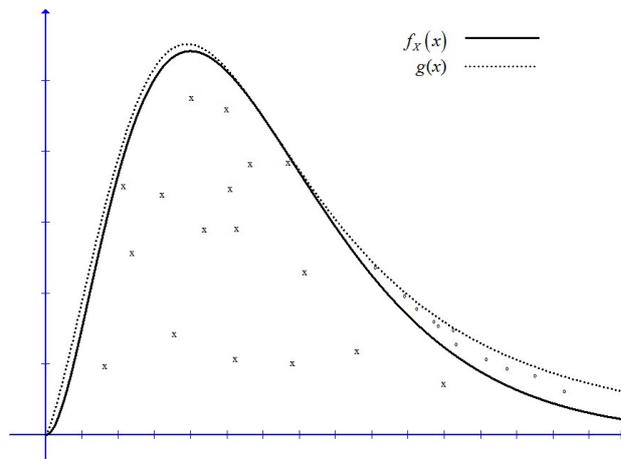


Figura 6 : Método de aceptación y rechazo mejorado

En el siguiente ejemplo vemos como se puede elegir  $k$ , el factor de separación. Sea  $X$  una variable aleatoria con función de densidad

$$f_X(x) = \begin{cases} \sqrt{\frac{2}{\pi}} e^{-\frac{x^2}{2}} & \text{si } x \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

y sea

$$g(x) = k \cdot e^{-x} \quad \text{para } x \geq 0$$

Un camino para elegir  $k$  es considerar la condición de que las funciones tengan algún punto de contacto

$$\begin{aligned} k \cdot e^{-x} &= \sqrt{\frac{2}{\pi}} e^{-\frac{x^2}{2}} \quad \Rightarrow \quad k \cdot \sqrt{\frac{\pi}{2}} = e^{x - \frac{x^2}{2}} \\ \Rightarrow x^2 - 2x + 2L\left(k \cdot \sqrt{\frac{\pi}{2}}\right) &= 0 \end{aligned}$$

Para que esta ecuación tenga una sola raíz debe ser

$$2L\left(k \cdot \sqrt{\frac{\pi}{2}}\right) = 1 \quad \Rightarrow \quad k^2 \cdot \frac{\pi}{2} = e$$

de donde  $k = \sqrt{\frac{2e}{\pi}}$  y el punto de contacto se da en  $x = 1$ , que es el punto de inflexión de la densidad (ver figura 7).

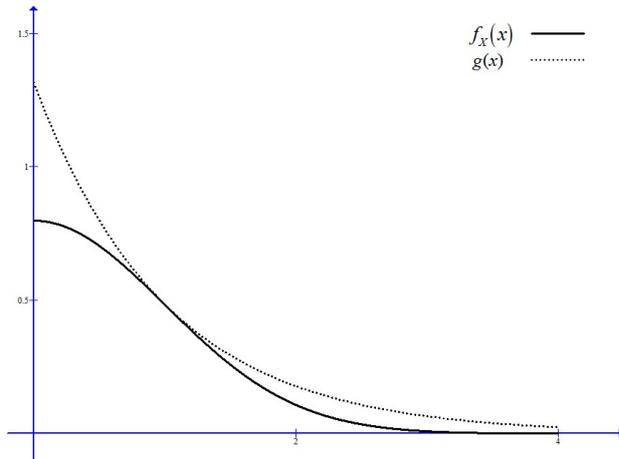


Figura 7 : Ejemplo del método de aceptación y rechazo mejorado

### 3.2.1. Método de Aceptación y Rechazo para variables aleatorias discretas

Hay varias formas para que el método de aceptación y rechazo pueda ser usado para distribuciones discretas [3]. Una de las ventajas de estos métodos es que pueden ser fácilmente adaptables a cambios en la distribución.

Consideremos la variable aleatoria  $X_s$  con recorrido

$$Rec(X_s) = \{x_1, x_2, x_3, \dots, x_n\}, \text{ con } x_1 < x_2 < x_3 < \dots < x_n$$

tal que su función de cuantía es

$$p_{si} = P(X_s = x_i) = \frac{a_{si}}{a_{s1} + a_{s2} + \dots + a_{sn}} \quad \text{para } i = 1, \dots, n$$

(Si  $\sum_{i=1}^n a_{si} = 1$ , el numerador  $a_{si}$  es la probabilidad  $p_{si}$ , la cuantía de  $x_i$ ).

Supongamos que existe un  $a_i^*$  tal que  $a_i^* \leq a_{si}$ , y  $b > 0$  tal que  $\sum_{i=1}^n a_{si} \geq b$  con  $s = 1, 2, \dots$

Sea  $a^* = \max a_i^*$ , y sea  $P_{si} = \frac{a_{si}}{a^*}$ , para  $i = 1, \dots, n$ .

Generamos  $u$  desde una distribución uniforme en  $[0, 1]$ , y sea  $i = [n \cdot u]$  (es la parte entera de  $k \cdot u$ ) y  $r = i - n \cdot u$ ; si  $r \leq P_{si}$  entonces  $i$  es aceptada como una realización de  $X_s$ , de otra forma es rechazada.

## 4. Algunas conclusiones

En las secciones 2 y 3 hicimos una aproximación respecto de algunos de los métodos de simulación, y analizamos algunas implementaciones en  $\mathbf{R}$ . Vimos que en general llevar a la programación las diferentes rutinas no es un proceso de gran complejidad. Las implementaciones siempre fueron para la obtención de  $m$  observaciones independientes de la variable aleatoria en cuestión, pensando en la construcción de la función de distribución empírica de la variable aleatoria. Esto toma mayor importancia cuando no conocemos totalmente la distribución de la o las variables aleatorias implicadas en un problema de estudio, allí, la posibilidad de éxito de la implementación estará en tener una buena aproximación al problema a través del modelo matemático que explica la dinámica entre sus distintos componentes.

## Referencias

- [1] Catedra de Estadística I, *Probabilidad, fundamentos de teoría*, Facultad de Ciencias Económicas y de Administración, 1992.
- [2] Gibbons Jean Dickinson, *Nonparametric statistical inference*, pp. 22–23, Marcel Dekker, Inc., New York, 1985.
- [3] Gentle James E., *Random number generation and monte carlo methods*, Springer, New York, 1998.
- [4] Morgan Byron J. T., *Elements of simulation*, Chapman & Hall, London, 1984.