

3

SIMULACIÓN CON VARIABLES ALEATORIAS PROVENIENTES DE DISTRIBUCIONES DISCRETAS Y CONTINUAS

En este capítulo se describen algunos modelos de probabilidad, discretos y continuos, para la simulación de valores de variables aleatorias que se ajustan a ellos. De manera integrada, se diseñan los algoritmos correspondientes y se generan valores aleatorios en cada caso (algunas veces de forma manual, y otras con la ayuda de programas de computador); estos se presentan en tablas construidas para tal fin. En el proceso de simulación se utilizan los métodos descritos en el capítulo anterior.

3.1 Uso de modelos aleatorios discretos en simulación

3.1.1 Modelo de Bernoulli

Una variable aleatoria X tiene la distribución de Bernoulli si su función de probabilidad es

$$f(x, p) = p^x (1 - p)^{1-x}$$

Con:

$$x = 0, 1$$

p : probabilidad de éxito

$1 - p$: probabilidad de fracaso

Su valor esperado y su varianza están dados por:

$$E(X) = p$$
$$Var(X) = p(1 - p)$$

Para generar los valores x de una variable aleatoria Bernoulli X con probabilidad de éxito p , se puede proceder a través de un algoritmo conformado por los siguientes pasos (Barceló, 1996):

1. Generar r_i
2. Asignar, $x_i = 0$ si $r_i > p$
 $x_i = 1$ si $r_i \leq p$

Donde los r_i son números aleatorios con distribución uniforme en el intervalo (0,1). Como ejemplo, en la Tabla 3.1 se presentan los resultados obtenidos al desarrollar un proceso para simular 10 valores x_i de una variable X con distribución de Bernoulli con parámetro $p = 0,7$. La función de probabilidad correspondiente está dada por:

$$f(x, 0.7) = (0.7)^x (0.3)^{1-x}$$

El valor esperado y la varianza son, respectivamente,

$$E(X) = 0.7$$

$$Var(X) = 0.7(0.3) = 0.21$$

Tabla 3.1. Valores simulados para una variable Bernoulli con $p=0.7$

Números aleatorios	P	x_i
0,914	$> 0,7$	0
0,775	$> 0,7$	0
0,495	$< 0,7$	1
0,534	$< 0,7$	1
0,585	$< 0,7$	1
0,185	$< 0,7$	1
0,521	$< 0,7$	1
0,079	$< 0,7$	1
0,968	$> 0,7$	0
0,346	$< 0,7$	1

(Los valores de esta tabla fueron generados por los autores con el software libre R).

Siguiendo un procedimiento similar, en la Tabla 3.2 se indican los resultados de simular 10 valores x_i de una variable X con distribución de Bernoulli de parámetro $p = 0.3$.

La función de probabilidad respectiva corresponde a la siguiente expresión:

$$f(x, 0.3) = (0.3)^x (0.7)^{1-x}$$

El valor esperado y la varianza para este caso son:

$$E(X) = 0.3$$

$$Var(X) = 0.3(0.7) = 0.21$$

Tabla 3.2. Valores simulados para una variable Bernoulli con $p=0.3$

Números aleatorios	p	x_i
0,534	$> 0,3$	0
0,391	$> 0,3$	0
0,695	$> 0,3$	0
0,996	$> 0,3$	0
0,925	$> 0,3$	0
0,278	$< 0,3$	1
0,407	$> 0,3$	0
0,363	$> 0,3$	0
0,37	$> 0,3$	0
0,52	$> 0,3$	0

(Los valores de esta tabla fueron generados por los autores con el software libre R).

3.1.2 Modelo binomial

Una variable aleatoria X tiene la distribución binomial de parámetros n y p , donde p es la probabilidad de éxito si su función de probabilidad es:

$$f(x) = \binom{n}{x} p^x q^{n-x}$$

Con $x = 0, 2, 3, \dots, n$

$$q = 1 - p$$

$$E(X) = np$$

$$Var(X) = npq$$

Para realizar el proceso de simulación se aplica el siguiente algoritmo:

$$x_i = (x_{i-1}) + 1 \quad \text{si } x_i \leq p$$

$$x_i = (x_{i-1}) \quad \text{si } x_i > p$$

$$i = 1, 2, 3, \dots, n; \quad x_0 = 0$$

A continuación se presenta un ejemplo para simular 10 valores x de una variable X con distribución binomial de parámetro $n = 5$, $p = 0,6$. En los datos que se presentan a continuación, el símbolo $>$ indica que el número aleatorio es mayor que 0,6, y, por lo tanto, se asigna cero (0), y el símbolo \leq indica que el número aleatorio es menor o igual que 0,6, y, por consiguiente, se asigna uno (1).

0,602 > 0	0,026 = 1	0,373 = 1	0,841 > 0
0,714 > 0	0,463 = 1	0,367 = 1	0,753 > 0
0,092 = 1	0,927 > 0	0,658 > 0	0,863 > 0
0,268 = 1	0,02 = 1	0,99 > 0	0,132 = 1
0,091 = 1	0,231 = 1	0,232 = 1	0,478 = 1
<hr style="width: 50%; margin: 0 auto;"/> x=3	<hr style="width: 50%; margin: 0 auto;"/> x=4	<hr style="width: 50%; margin: 0 auto;"/> x=3	<hr style="width: 50%; margin: 0 auto;"/> x=2
0,941 > 0	0,047 = 1	0,156 = 1	0,38 = 1
0,871 > 0	0,96 > 0	0,618 > 0	0,138 = 1
0,982 > 0	0,485 = 1	0,958 > 0	0,273 = 1
0,134 = 1	0,42 = 1	0,01 = 1	0,025 = 1
0,984 > 0	0,934 > 0	0,972 > 0	0,381 = 1
<hr style="width: 50%; margin: 0 auto;"/> x=1	<hr style="width: 50%; margin: 0 auto;"/> x=3	<hr style="width: 50%; margin: 0 auto;"/> x=2	<hr style="width: 50%; margin: 0 auto;"/> x=5
0,769 > 0	0,556 = 1		
0,186 = 1	0,31 = 1		
0,367 = 1	0,705 > 0		
0,488 = 1	0,966 > 0		
0,383 = 1	0,409 = 1		
<hr style="width: 50%; margin: 0 auto;"/> x=4	<hr style="width: 50%; margin: 0 auto;"/> x=3		

Como resultado del proceso de simulación se tienen los valores que se indican a continuación:

$$x = 3,4,3,2,1,3,2,5,4,3$$

Para el caso considerado, la función de probabilidad está dada por la expresión

$$f(x, 0.6) = \binom{5}{x} (0.6)^x (0.4)^{5-x}$$

El valor esperado y la varianza son los siguientes:

$$E(X) = 5(0.6) = 3$$

$$Var(X) = 5(0.6)(0.4) = 1.2$$

En Blanco (2004) se describe una manera alternativa para simular los valores x de una variable X con distribución binomial mediante la aplicación del siguiente algoritmo:

1. Entrada:
Parámetros n, p
Número *máximo* de simulaciones.
2. Inicialización:
 $q := 1 - p$
 P es un vector de q^{n+1} con componentes $P(i)$ definidas e inicializadas de la siguiente manera:
 $P(0) := q^n$
3. Para $j = 1$ (incrementando en 1) hasta n hacer:
$$P(j) := \frac{n-j+1}{j} * \frac{p}{q} * P(j-1)$$

 $suma$ es un vector de R^{n+1} con componentes $suma(i)$ definidas e inicializadas de la siguiente manera:
 $suma(0) := P(0)$
4. Para $j = 1$ (incremento en 1) hasta n hacer:
 $suma(j) := suma(j-1) + P(j)$
5. Para $j = 1$ (incrementando en 1) hasta *máximo* hacer:
 $i := 0$
6. Mientras $RND \geq suma(i)$ hacer:
 $i := i + 1$

Escribir i
7. Finalizar.

3.1.3 Modelo de Poisson

Una variable aleatoria X tiene distribución de Poisson de parámetro $\lambda > 0$ si su función de probabilidad es:

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Con $x = 0, 2, 3, \dots, n, \dots$

El valor esperado y la varianza para la variable en consideración son:

$$E(X) = \lambda$$

$$Var(X) = \lambda$$

La función de distribución correspondiente a la variable aleatoria X se puede escribir de la siguiente forma:

$$F(x) = \sum_{i=0}^x \frac{\lambda^i e^{-\lambda}}{i!}$$

$$F(x) = e^{-\lambda} \sum_{i=0}^x \frac{\lambda^i}{i!}$$

Una forma de realizar el proceso de simulación consiste en formar clases o intervalos para cada valor de i ; luego, se generan números aleatorios y se observa en cuál clase quedan ubicados.

En el siguiente ejemplo, con base en 60 datos correspondientes al número de llegadas a la cola por minuto en un cajero automático, se determinó que dichos datos se ajustaban a una distribución de Poisson con parámetro $\lambda = 0,5$. Se desea simular 40 valores x de esa variable aleatoria X que sigue una distribución de Poisson. La función de probabilidad corresponde a la expresión

$$f(x) = \frac{(0,5)^x e^{-0,5}}{x!}$$

El valor esperado y la varianza para la variable en consideración son:

$$E(X) = 0,5$$

$$Var(X) = 0,5$$

Su función de distribución es:

$$F(x) = e^{-0,5} \sum_{i=0}^x \frac{(0,5)^i}{i!}$$

$$x = 0 \rightarrow F(0) = e^{-0,5} \sum_{i=0}^0 \frac{(0,5)^i}{i!} = 0,6065$$

$$x = 1 \rightarrow F(1) = e^{-0,5} \sum_{i=0}^1 \frac{(0,5)^i}{i!} = 0,6065(1 + 0,5) = 0,9097$$

$$x = 2 \rightarrow F(2) = e^{-0,5} \sum_{i=0}^2 \frac{(0,5)^i}{i!} = 0,6065(1 + 0,5 + 0,125) = 0,9855$$

$$x = 3 \rightarrow F(3) = e^{-0,5} \sum_{i=0}^3 \frac{(0,5)^i}{i!} = 0,6065(1 + 0,5 + 0,125 + 0,208) = 0,9981$$

$$x = 4 \rightarrow F(4) = e^{-0,5} \sum_{i=0}^4 \frac{(0,5)^i}{i!} = 0,9997$$

Con los valores obtenidos se forman los intervalos aleatorios [0.000, 0.6065]; [0.6066, 0.9097]; [0.9098, 0.9855]; [0.9856, 0.9981]; [0.9982, 0.9997] y [0.9998, 1.000]. Se generan 40 números aleatorios con distribución uniforme en (0,1) con buenas propiedades estadísticas y se ubican en el intervalo correspondiente; de acuerdo con el mencionado intervalo se asumen los valores para x , como se observa en la Tabla 3.3.

Los valores simulados resultantes son:

$$x = 0,0,0,1,1,0,0,0,2,0,0,0,0,0,0,0,2,2,0,0,0,0,0,0,1,0,0,0,0, \\ 1,0,1,0,0,1,2,0,1,3$$

En Blanco (2004) se menciona otra forma de simular los valores x de una variable X con distribución Poisson, la cual consiste en utilizar el siguiente algoritmo:

1. Entrada:

Parámetros λ

Número *máximo* de simulaciones

2. Para $j = 1$ (incrementando en 1) hasta *máximo* hacer:

Inicialización:

$i := 0$

$prob := e^{-\lambda}$

$suma := prob$

3. Verificación

Mientras $RND \geq suma$ hacer:

$i := i + 1$

$prob := \lambda * \frac{prob}{i}$

$suma := suma + prob$

Escribir i

4. Finalizar.

Tabla 3.3. Valores simulados para una variable Poisson de parámetro 0.5

X	Intervalo aleatorio	Números aleatorios generados	Valores variable X
0	0.000-0.6065	0.1160, 0.1000, 0.2830, 0.5238, 0.4132, 0.3206 0.2583, 0.6011, 0.0245, 0.2240, 0.5980, 0.2994 0.4980, 0.0040, 0.5012, 0.3370, 0.2856, 0.4197 0.1230, 0.5480, 0.4902, 0.3869, 0.5674, 0.4246 0.0354, 0.1982, 0.1965, 0.3461	0,0,0,0,0,0 0,0,0,0,0,0 0,0,0,0,0,0 0,0,0,0,0,0 0,0,0,0
1	0.6066-0.9097	0.8525, 0.6714, 0.6219, 0.7637 0.7222, 0.6122, 0.6211	1, 1, 1, 1 1, 1, 1
2	0.9098-0.9855	0.9158, 0.9316, 0.9273, 0.9249	2, 2, 2, 2
3	0.9856-0.9981	0.9962	3
4	0.9982-0.9997		
5	0.9990-1.000		

(Los valores de esta tabla fueron generados por los autores con el software libre R).

3.1.4 Aplicación del método de la transformada inversa para modelos discretos

En este caso se quiere simular valores x de una variable aleatoria X con una distribución dada cuya función de probabilidad se puede expresar de la siguiente forma:

$$P(X = x_i) = p_i \text{ con } i = 0, 1, 2, 3, \dots$$

Además,

$$\sum_i p_i = 1$$

Se supone que el rango R_x de la variable X asume los valores $x_0, x_1, x_2, x_3, \dots$. Para realizar el proceso de simulación se generan valores u correspondientes a una variable aleatoria U con distribución uniforme en el intervalo $(0,1)$; estos se involucran en el proceso como se indica a continuación:

$$X = \begin{cases} x_0 & \text{si } u < p_0 \\ x_1 & \text{si } p_0 \leq u < p_0 + p_1 \\ \bullet \\ \bullet \\ x_i & \text{si } \sum_{j=1}^{i-1} p_j \leq u < \sum_{j=1}^i p_j \\ \bullet \\ \bullet \end{cases}$$

Debido a que la variable aleatoria U tiene distribución uniforme en $(0,1)$, se cumple que

$$P(a \leq U < b) = b - a$$

Con $0 < a < b < 1$. Por la definición de la función de probabilidad resulta que

$$P(X = x_i) = P\left(\sum_{j=1}^{i-1} p_j \leq u < \sum_{j=1}^i p_j\right) = p_i$$

Lo anterior implica que si los x_i están ordenados de manera ascendente, entonces se cumple que

$$X = x_i \text{ siempre y cuando } F(x_{i-1}) \leq u < F(x_i)$$

El contexto anterior permite plantear el siguiente algoritmo para simular los valores x de una variable aleatoria discreta X :

Generar un número aleatorio u

Si $u < p_0$, hacer $x = x_0$ y terminar

Si $u < p_0 + p_1$, hacer $x = x_1$ y terminar; se continúa el proceso

Si $u < p_0 + p_1 + \dots + p_p$, hacer $x = x_i$ y terminar, y así sucesivamente.

A continuación se presenta un ejemplo en el cual se trabaja con el anterior algoritmo. Sea X una variable aleatoria con una distribución cuya función de probabilidad está dada por:

$$f(x) = P(X = x) = \begin{cases} 0.12 & \text{si } x = 1 \\ 0.18 & \text{si } x = 2 \\ 0.30 & \text{si } x = 3 \\ 0.25 & \text{si } x = 4 \\ 0.15 & \text{si } x = 5 \end{cases}$$

Su función de distribución es:

$$F_x(x) = \begin{cases} 0 & \text{si } x < 1 \\ 0.12 & \text{si } 1 \leq x < 2 \\ 0.30 & \text{si } 2 \leq x < 3 \\ 0.60 & \text{si } 3 \leq x < 4 \\ 0.85 & \text{si } 4 \leq x < 5 \\ 1 & \text{si } x \geq 5 \end{cases}$$

Los valores x de una variable aleatoria discreta X que se pueden obtener mediante los números aleatorios u de una variable aleatoria discreta U con distribución uniforme en el intervalo $(0,1)$ se pueden obtener mediante la siguiente expresión:

$$x = \begin{cases} 1 & \text{si } u < 0.12 \\ 2 & \text{si } 0.12 \leq u < 0.30 \\ 3 & \text{si } 0.30 \leq u < 0.60 \\ 4 & \text{si } 0.60 \leq u < 0.85 \\ 5 & \text{si } 0.85 \leq u < 1 \end{cases}$$

Con los 10 números aleatorios de la Tabla 3.4 se simularon los 10 valores x para la variable aleatoria en consideración.

Tabla 3.4 Valores simulados de la variable X

Números aleatorios	X
0,564	3
0,106	1
0,685	4
0,096	1
0,988	5
0,45	3
0,757	4
0,213	2
0,729	4
0,367	3

(Los valores de esta tabla fueron generados por los autores con el software libre R).

Adicionalmente, es pertinente aclarar que este método es aplicable a los modelos de Bernoulli, Binomial, Poisson o cualquier modelo discreto (Barceló, 1996).

3.2 Simulación de variables aleatorias continuas

A continuación se indican algunas distribuciones usuales y el método de simulación que frecuentemente resulta más apropiado para simular valores de variables aleatorias que se modelen a través de dichas distribuciones; comúnmente, estas se describen en la mayoría de textos dedicados al tema de la simulación. Luego, a manera de ejemplos se presentan algunas de ellas y los procedimientos correspondientes para realizar la simulación de una cantidad determinada de valores aleatorios.

Para simular valores de variables con distribución uniforme, exponencial, weibull, triangular, empírica, discreta arbitraria o geométrica se utiliza el método de la transformada inversa; para variables con distribución *erlang* o binomial se usa el método de convolución; para variables con distribución gamma, normal, Poisson se puede trabajar con el método del rechazo; la distribución normal también se puede simular mediante transformación directa, las 12 uniformes y el método de Box-Muller.

3.2.1 Modelo uniforme

Una variable aleatoria X sigue la distribución uniforme en el intervalo (a, b) si su función de densidad de probabilidad está dada por:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{Si } x \in (a, b) \\ 0 & \text{en otro caso} \end{cases}$$

El valor esperado y la varianza están dados por:

$$E(X) = \frac{a+b}{2}$$

$$Var(X) = \frac{(b-a)^2}{12}$$

En particular, una variable aleatoria X sigue la distribución uniforme en el intervalo $(0,1)$ que se denota con $U(0,1)$, si su función de densidad de probabilidad está dada por:

$$f(x) = \begin{cases} 1 & \text{Si } x \in (0,1) \\ 0 & \text{en otro caso} \end{cases}$$

En este caso, el valor esperado y la varianza toman los valores siguientes:

$$E(X) = \frac{1}{2}$$

$$Var(X) = \frac{1}{12}$$

La función de distribución para la variable aleatoria X que sigue la distribución uniforme en el intervalo (a, b) es:

$$F_x(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x < b \\ 1 & \text{si } x \geq b \end{cases}$$

Para simular los valores de una variable aleatoria discreta con distribución uniforme en el intervalo se generan los números aleatorios de una variable aleatoria discreta con distribución uniforme en el intervalo $(0,1)$, y mediante el método de la transformada se aplica la siguiente expresión:

$$u = F(x)$$

O equivalentemente,

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

Luego para $a \leq x < b$ se tiene que

$$u = \frac{x-a}{b-a}$$

De aquí resulta que

$$x = a + u(b-a) = F^{-1}(u)$$

Un algoritmo para simular valores aleatorios de una variable con distribución uniforme en el intervalo (a, b) es el siguiente:

1. Entrada:
 Parámetros a, b
 Número *máximo* de simulaciones
2. Para $j = 1$ (incrementando en 1) hasta *máximo* hacer:
 Generación número aleatorio u
 $x = a + u(b-a)$
3. Terminar.

La función de distribución para la variable aleatoria X que sigue la distribución uniforme en el intervalo $(-2, 2) = (a, b)$ es:

$$F_X(x) = \begin{cases} 0 & \text{si } x < -2 \\ \frac{x+2}{4} & \text{si } -2 \leq x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

A manera de ejemplo, se desea simular 10 valores x de una variable aleatoria discreta X con distribución uniforme en el intervalo $(-2, 2)$. Para ello, se generan 10 números aleatorios u provenientes de una distribución uniforme en el intervalo $(0,1)$ y se reemplazan en la expresión siguiente:

$$x = a + u(b - a) = -2 + u(2 - (-2))$$

De manera simplificada, resulta la expresión

$$x = -2 + 4u$$

Al sustituir los números aleatorios u , resultan los valores simulados que se presentan en la Tabla 3.5.

Tabla 3.5. Valores simulados de una uniforme en $(-2,2)$

Números aleatorios	$x = -2 + 4u$
0,13964	-1,44144
0,4313	-0,2748
0,61218	0,44872
0,29075	-0,837
0,15573	-1,37708
0,6995	0,798
0,3463	-0,6148
0,44564	-0,21744
0,05242	-1,79032
0,1032	-1,5872

(Los valores de esta tabla fueron generados por los autores con el software libre R)

3.2.2 Modelo Exponencial

Una variable aleatoria X tiene distribución exponencial de parámetro $\lambda > 0$ si su función de densidad es:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

El valor esperado y la varianza de la variable X son:

$$E(X) = \frac{1}{\lambda}$$

$$Var(X) = \frac{1}{\lambda^2}$$

La función de distribución para la variable aleatoria X que sigue la distribución exponencial de parámetro $\lambda > 0$ es:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

Para simular los valores x de una variable aleatoria discreta X con distribución exponencial se generan los números aleatorios u de una variable aleatoria discreta U con distribución uniforme en el intervalo $(0,1)$, y mediante el método de la transformada inversa se aplica la siguiente igualdad:

$$u = F(x)$$

La cual implica la expresión

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

Luego para $\lambda > 0$, $\lambda \geq 0$ se tiene que

$$u = 1 - e^{-\lambda x}$$

Despejando x resulta

$$x = \frac{-1}{\lambda} \text{Ln}(1 - u)$$

Un algoritmo para simular valores aleatorios de una variable con distribución exponencial de parámetro es el siguiente:

1. Entrada:
Parámetro $\lambda > 0$
Número *máximo* de simulaciones
2. Para $j = 1$ (incrementando en 1) hasta *máximo* hacer:

Generación número aleatorio u

$$x := \frac{-1}{\lambda} \text{Ln}(1 - u)$$

3. Terminar.

Se requiere simular 10 valores x de una variable aleatoria discreta X con distribución exponencial de parámetro $\lambda = 2$.

En este caso, la función de densidad de probabilidad es:

$$f(x) = \begin{cases} 2e^{-2x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

El valor esperado y la varianza de la variable X son:

$$E(X) = \frac{1}{2}$$

$$Var(X) = \frac{1}{2^2} = \frac{1}{4}$$

La función de distribución para la variable aleatoria X es:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-2x} & \text{si } x \geq 0 \end{cases}$$

Para simular los 10 valores x de una variable aleatoria discreta X en cuestión, se generan 10 números aleatorios u provenientes de una distribución uniforme en el intervalo $(0,1)$ y se reemplazan en la expresión siguiente:

$$x = \frac{-1}{\lambda} \text{Ln}(1-u) = \frac{-1}{2} \text{Ln}(1-u)$$

De forma más simple, se tiene la expresión

$$x = -0.5 \text{Ln}(1-u)$$

Al reemplazar los números aleatorios u resultan los valores simulados que se presentan en la Tabla 3.6.

Observación. La expresión

$$x = \frac{-1}{\lambda} \text{Ln}(1-u)$$

es equivalente a la expresión

$$x = \frac{-1}{\lambda} \text{Ln}(u)$$

debido a que si u toma valores en el intervalo $(0,1)$, entonces también $1-u$ toma valores en el intervalo $(0,1)$. Esta última expresión es también utilizada para simular valores de una variable aleatoria con distribución exponencial de parámetro λ .

Tabla 3.6. Valores simulados de una exponencial de parámetro 2

Números aleatorios	$x = -0.5\text{Ln}(1-u)$
0,486	0,3327
0,187	0,1035
0,317	0,1906
0,295	0,1747
0,788	0,7755
0,129	0,069
0,982	2,0086
0,597	0,4544
0,742	0,6773
0,068	0,0352

(Los valores de esta tabla fueron generados por los autores con el software libre R).

3.2.3 Modelo continuo no usual

Los valores de una variable aleatoria X se ajustan a una función definida de la siguiente manera:

$$f(x) = \begin{cases} 2x & \text{si } 0 \leq x \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

La representación de la anterior función se presenta en la Figura 3.1, en la cual se observa una función continua a trozos; con base en esta función, se requiere simular 10 valores de la variable aleatoria X . Para esto, inicialmente se debe mostrar que $f(x)$ es una función de densidad de probabilidad para X .

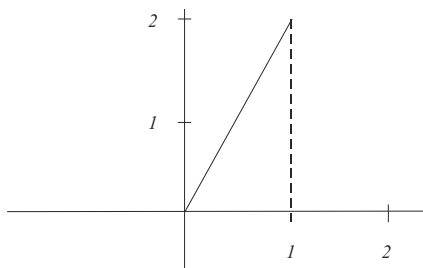


Figura 3.1. Representación gráfica de $f(x)$

Fuente: Los autores, con la ayuda del software libre R.

1. De la Figura 3.1 se observa que $f(x) \geq 0$

$$2. \quad \int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^0 f(x)dx + \int_0^1 f(x)dx + \int_1^{\infty} f(x)dx$$

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^0 (0)dx + \int_0^1 2x dx + \int_1^{\infty} (0)dx = x^2 \Big|_0^1 = 1^2 - 0^2 = 1$$

De las partes 1. y 2. se tiene que $f(x)$ sí corresponde a una función de densidad de probabilidad para la variable aleatoria X .

Utilizando la expresión

$$u = F(x)$$

resulta

$$u = F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^0 f(t)dt + \int_0^x f(t)dt$$

$$u = F(x) = \int_0^x 2t dt = t^2 \Big|_0^x = x^2$$

$$u = x^2$$

Para $0 \leq x \leq 1$ se tiene que

$$x = \sqrt{u}$$

Ahora se generan 10 números aleatorios u provenientes de una distribución uniforme en el intervalo $(0,1)$ y se reemplazan en la expresión anterior; los resultados se presentan en la Tabla 3.7.

Tabla 3.7. Valores simulados de un modelo no usual

Números aleatorios	$x = \sqrt{u}$
0,141156	0,37571
0,042938	0,20721
0,621706	0,78848
0,153592	0,39191
0,715214	0,8457
0,928301	0,96348
0,578133	0,76035
0,261857	0,51172
0,72483	0,85137
0,037092	0,19259

(Valores generados por los autores con el software libre R).

3.2.4 Aplicación del método de composición

Se necesita generar valores de una variable aleatoria X cuya función de densidad se define de la siguiente forma:

$$f(x) = \begin{cases} x & \text{si } 0 < x < 1 \\ 2-x & \text{si } 1 \leq x < 2 \\ 0 & \text{en otro caso} \end{cases}$$

Por la forma como está definida la función de densidad, hay dos subáreas correspondientes a la $f(x)$ original; cada una equivale a 0.5. Se puede tomar $p_1 = 0.5$ y $p_2 = 0.5$

Las funciones de densidad $f_1(x)$ para cada subárea deben satisfacer:

$$f(x) = 0.5f_1(x) + 0.5f_2(x)$$

Luego se puede tomar

$$f_1(x) = \begin{cases} 2x & \text{si } 0 < x < 1 \\ 0 & \text{en otro caso} \end{cases}$$

$$f_2(x) = \begin{cases} 2(2-x) & \text{si } 1 \leq x < 2 \\ 0 & \text{en otro caso} \end{cases}$$

Las correspondientes distribuciones de probabilidad son:

$$F_1(x) = x^2$$

$$F_2(x) = 4x - x^2 - 3$$

Se generan dos números aleatorios U_1, U_2 con distribución uniforme en $(0,1)$, luego, se aplica el siguiente algoritmo:

Si $0 < U_1 \leq 0.5$, entonces $x^2 = U_1 \Rightarrow x = \sqrt{U_1}$

En caso contrario:

$$4x - x^2 - 3 = U_2 \Rightarrow x = 2 - \sqrt{1 - U_2}$$

A continuación se ilustra el proceso para los tres primeros números aleatorios U_1 , y el uso del número aleatorio U_2 , que se indican en la Tabla 3.8.

$$U_1 = 0.559355 \Rightarrow x = 2 - \sqrt{1 - 0.333891} = 1.183845$$

$$U_1 = 0.982616 \Rightarrow x = 2 - \sqrt{1 - 0.620771} = 1.384184$$

$$U_1 = 0.325262 \Rightarrow x = \sqrt{0.641131} = 0.800707$$

Lo anterior es concordante con los valores x en el dominio de definición de la función de densidad $f(x)$ en los intervalos $(0,1)$ y $(1,2)$.

Tabla 3.8. Valores simulados por el método de composición

u_1	u_2	X
0,559355	0,333891	1,183845
0,982616	0,620771	1,384184
0,325262	0,641131	0,800707
0,535503	0,752135	1,50214
0,507678	0,204952	1,108345
0,281995	0,143259	0,378496
0,347206	0,445881	0,667743
0,921085	0,369064	1,205685
0,408148	0,48528	0,69662
0,095739	0,099268	0,315068

(Valores generados por los autores con el software libre R).

3.2.5 Simulación con una distribución Beta

Inicialmente se define la distribución Beta de manera general, para luego realizar un ejemplo de simulación de valores de una variable aleatoria con distribución Beta para un caso específico de sus parámetros.

Una variable aleatoria X tiene distribución Beta de parámetros $a > 0$ y $b > 0$ si su función de densidad de probabilidad es:

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} & \text{si } 0 < x < 1 \\ 0 & \text{si } x \notin (0,1) \end{cases}$$

Donde $B(a, b)$ es la función Beta definida a través de la siguiente expresión:

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$$

El valor esperado y la varianza de la variable aleatoria X son:

$$E(X) = \frac{a}{a+b}$$

$$Var(X) = \frac{ab}{(a+b+1)(a+b)^2}$$

Como caso particular, se requiere simular valores x de la variable aleatoria X con distribución Beta de parámetros $a > 0$ y $b = 1$.

En este caso, la función de densidad de probabilidad toma la forma:

$$f(x) = \begin{cases} \frac{1}{B(a,1)} x^{a-1} (1-x)^{1-1} & \text{si } 0 < x < 1 \\ 0 & \text{si } x \notin (0,1) \end{cases}$$

Es decir,

$$f(x) = \begin{cases} \frac{1}{B(a,1)} x^{a-1} & \text{si } 0 < x < 1 \\ 0 & \text{si } x \notin (0,1) \end{cases}$$

Donde,

$$B(a,1) = \int_0^1 x^{a-1} dx = \frac{x^a}{a} \Big|_0^1 = \frac{1}{a}$$

Luego la función de densidad se puede escribir así:

$$f(x) = \begin{cases} ax^{a-1} & \text{si } 0 < x < 1 \\ 0 & \text{si } x \notin (0,1) \end{cases}$$

Su función de distribución es:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^0 0 dt + \int_0^x f(t) dt = \int_0^x at^{a-1} dt = x^a \quad \text{si } 0 < x < 1$$

Esto es,

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x^a & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Luego para simular los valores x de la variable aleatoria X con distribución Beta de parámetros $a > 0$ y $b = 1$ resulta pertinente utilizar el método de la trasformada inversa. Para esto, se generan valores u de una variable aleatoria U con distribución uniforme en el intervalo $(0,1)$ y se utiliza la siguiente igualdad:

$$u = F(x) = x^a \quad \text{si } 0 \leq x < 1$$

Lo anterior implica que

$$u = x^a \quad \Rightarrow \quad x = F^{-1}(u) \quad \Rightarrow \quad x = u^{\frac{1}{a}}$$

A manera de ilustración, en la Tabla 3.9 se presentan 20 valores x de la variable aleatoria X , considerando el caso en que $a = 2$, $a = 3$, $a = 4$ y $a = 5$, usando 20 números aleatorios provenientes de una distribución uniforme en $(0,1)$.

Tabla 3.9. Valores simulados de una distribución Beta

U	$a=2$	$a=3$	$a=4$	$a=5$
0,559355	0,7479	0,82394	0,864812	0,890302
0,982616	0,99127	0,994171	0,995625	0,996499
0,325262	0,570317	0,687719	0,755194	0,798816
0,535503	0,731781	0,812058	0,855442	0,882577
0,507678	0,712515	0,797743	0,844106	0,873208
0,281995	0,531032	0,655763	0,728719	0,776333
0,347206	0,589242	0,70285	0,767621	0,809315
0,921085	0,959732	0,972971	0,979659	0,983694
0,408148	0,638865	0,741776	0,79929	0,835918
0,095739	0,309417	0,45747	0,556253	0,625486
0,333891	0,577833	0,693748	0,760153	0,80301
0,620771	0,78789	0,853055	0,887632	0,909047
0,641131	0,800707	0,862281	0,894822	0,914933
0,752135	0,867257	0,909422	0,931266	0,944624
0,204952	0,452716	0,589591	0,672842	0,728334
0,143259	0,378496	0,523248	0,61522	0,677992
0,445881	0,667743	0,763964	0,817156	0,850832
0,369064	0,607506	0,7173	0,779427	0,819257
0,48528	0,69662	0,785834	0,834638	0,865363
0,099268	0,315068	0,463024	0,561309	0,630031

(Valores generados por los autores con el software libre R)

3.2.6 Aplicación del método del rechazo

Se requiere simular los valores x de una variable aleatoria X con función de densidad continua $f(x)$ definida de la siguiente manera:

$$f(x) = \begin{cases} 12x(1-x)^2 & \text{si } 0 < x < 1 \\ 0 & \text{si } x \notin (0,1) \end{cases}$$

De la expresión anterior se observa que la función de densidad de la variable aleatoria X se anula fuera del intervalo $(0,1)$; luego para aplicar el método del rechazo es conveniente considerar una función $g(x)$ tal que $f(x) \leq g(x)$ para todo $0 < x < 1$. Para cumplir con el anterior requerimiento se ha de considerar una función de densidad $g(x)$, definida de la siguiente forma:

$$g(x) = \begin{cases} 1 & \text{si } 0 < x < 1 \\ 0 & \text{si } x \notin (0,1) \end{cases}$$

La constante que verifique la siguiente desigualdad $\frac{f(x)}{g(x)} \leq c$ se puede determinar buscando el máximo de

$$\frac{f(x)}{g(x)} = 12x(1-x)^2 = 12x - 24x^2 + 12x^3$$

Derivando e igualando a cero se obtiene que

$$12 - 48x + 36x^2 = 0$$

De donde se tiene que $x = 1$ o $x = \frac{1}{3}$; usando el criterio de la segunda derivada se observa que en $x = \frac{1}{3}$ la segunda derivada es menor que cero, garantizando la existencia de un máximo en $x = \frac{1}{3}$, cuyo valor es:

$$12 \left(\frac{1}{3} \right) \left(1 - \frac{1}{3} \right)^2 = \frac{48}{27} = \frac{16}{9} = c$$

Ahora,

$$\frac{f(x)}{g(x)} \leq c = \frac{16}{9}$$

Por lo tanto,

$$\frac{f(x)}{cg(x)} = \frac{12x(1-x)^2}{\left(\frac{16}{9}\right)} = \frac{108x(1-x)^2}{16} = \frac{27x(1-x)^2}{4}$$

En concordancia con lo expuesto en la sección 2.7.2, el algoritmo que se ha de seguir es:

1. Se genera un valor y de la variable aleatoria Y con densidad $g(x)$ que corresponde a una densidad uniforme en el intervalo $(0,1)$.
2. Se genera un valor u de una variable aleatoria U con distribución uniforme en el intervalo $(0,1)$.
3. Si se cumple la desigualdad:

$$u \leq \frac{27y(1-y)^2}{4} \text{ entonces hacer } x = y.$$

De lo contrario, regresar al paso 1.

En la Tabla 3.10, la primera columna contiene 20 valores generados para una variable aleatoria con función de densidad $g(x)$; la segunda columna presenta 20 números aleatorios con distribución uniforme en $(0,1)$ que se utilizan para comparar con los resultados de la expresión indicada en la tercera columna y permiten generar los valores x de una variable aleatoria X con función de densidad continua $f(x)$ que aparecen en la cuarta columna. La letra N que aparece en la cuarta columna indica que no se generó el valor x porque la desigualdad no se cumplió. En resumen, solo se pudieron generar 13 valores x en este caso.

Tabla 3.10. Valores simulados en la aplicación del método del rechazo

y	u	$27y(1-y)^2/4$	$x = y$
0,559355	0,761206	0,73311	0,559355
0,982616	0,803134	0,002004	N
0,325262	0,702858	0,999557	0,325262
0,535503	0,047913	0,779887	0,535503
0,507678	0,394261	0,830597	0,507678
0,281995	0,22862	0,981296	0,281995
0,347206	0,726215	0,998719	0,347206
0,921085	0,882465	0,038719	N
0,408148	0,182222	0,965045	0,408148
0,095739	0,27969	0,528421	N
0,333891	0,695545	0,999998	0,095739
0,620771	0,902572	0,602613	N
0,641131	0,122813	0,557343	0,641131
0,752135	0,507697	0,31191	N
0,204952	0,117283	0,874465	0,204952
0,143259	0,924725	0,709782	N
0,445881	0,517283	0,924121	0,445881
0,369064	0,271406	0,99169	0,369064
0,48528	0,644636	0,867837	0,48528
0,099268	0,643786	0,543631	N

(Valores generados por los autores con el software libre R)

Ejercicios capítulo 3

- 3.1 Simular de forma manual 10 valores de una variable aleatoria con distribución de Bernoulli de parámetros $p=0.15$. Usando el computador, simular 200 valores de la mencionada variable.
- 3.2 Simular 50 valores de una variable aleatoria con distribución binomial de parámetros $n = 18$ y $p = 0.72$.
- 3.3 Simular 30 valores de una variable aleatoria con distribución de Poisson de parámetro 0.8. En el computador, simular 500 valores de dicha variable.
- 3.4 Simular 40 valores de una variable aleatoria discreta cuya función de probabilidad se indica en el ejercicio 2.5.
- 3.5 Simular 35 valores de una variable aleatoria continua cuya función de densidad de probabilidad se presenta en el ejercicio 2.6.
- 3.6 Simular 100 valores de una variable aleatoria con distribución uniforme en el intervalo $(-1, 3)$. Mediante el computador, simular 5000 valores de dicha variable.
- 3.7 Simular 25 valores de una variable aleatoria con distribución exponencial de parámetro 10. Usando el computador, simular 1000 valores de la mencionada variable.
- 3.8. Simular 20 valores de una variable aleatoria con distribución Beta de parámetros $a=9$ y $b=1$. A través del computador, simular 500 valores de la mencionada variable aleatoria.
- 3.9 Usando el método del rechazo, obtener 15 valores de una variable aleatoria con función de densidad dada por

$$f(x) = \begin{cases} 30x^2(1-x)^2 & \text{Si } x \in (0,1) \\ 0 & \text{en otro caso} \end{cases}$$