

METODOS DE GENERACIÓN DE VARIABLES ALEATORIAS

INTRODUCCION

Dependiendo de la situación del mundo real, hemos visto cómo las probabilidades pueden ser basadas tanto en frecuencias relativas como en juicios subjetivos. A pesar del origen de estas probabilidades, en el caso de los modelos probabilísticos computarizados una mejor aproximación, para describir las probabilidades, es usar distribuciones de probabilidad teóricas tales como la normal o exponencial negativa.

Hasta ahora los ejemplos de simulación que se han presentado sólo usan distribuciones de probabilidades discretas para las variables aleatorias. Sin embargo, en muchas simulaciones es más realista y práctico usar variables aleatorias continuas. En esta sección presentaremos y describiremos algunos procedimientos para generar cantidades aleatorias a partir de distribuciones continuas. El principio básico es muy semejante al caso discreto. Como en el método discreto, generamos primero un número aleatorio $U(0,1)$ y luego lo transformamos en una cantidad aleatorio de acuerdo con la distribución especificada. Sin embargo, el proceso para llevar a cabo la transformación es bastante distinto del caso discreto.

Hay muchos métodos diferentes para generar cantidades aleatorias continuas. La selección de un algoritmo determinado dependerá de la distribución de la cual deseamos generar, teniendo en cuenta factores tales como la exactitud de las variables aleatorias, las eficacias de cómputo y de almacenamiento, y la complejidad del algoritmo. Los dos algoritmos que más se usan son el método de transformación inversa y el de aceptación o rechazo. Entre esos dos, es posible generar variables aleatorias a partir de casi cualquiera de las distribuciones que más se usan. Presentaremos una descripción detallada de ambos algoritmos, junto con ejemplos de cada método. Además, daremos dos métodos para generar variables aleatorias a partir de una distribución normal.

MÉTODO DE TRANSFORMACIÓN INVERSA PARA DISTRIBUCIONES CONTINUAS

Este método se usa, por lo general, para distribuciones cuya función de distribución acumulada se pueda obtener *en forma* cerrada. Las distribuciones de este tipo son por ejemplo la exponencial, uniforme, triangular y de Weibull. El método de transformación inversa es relativamente fácil de describir y de ejecutar. Consiste en los tres pasos siguientes:

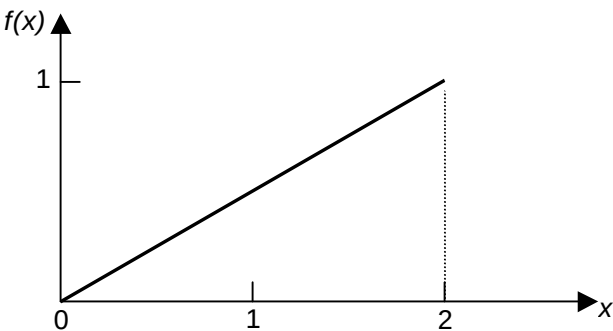
- Paso 1.** Dada una función de densidad de probabilidad $f(x)$ para una variable aleatoria X , obtener la función de probabilidad acumulada $F(x)$ como
- $$F(x) = \int_{-\infty}^x f(t) dt .$$
- Paso 2.** Generar un número aleatorio r .
- Paso 3.** Hacer $F(x)=r$ y despejar x . La variable x es entonces un número aleatorio procedente de la distribución cuya función de probabilidad es $f(x)$.

Ejemplo 1: Función rampa

A continuación veamos la mecánica del algoritmo mediante un ejemplo. Para ello se tiene la distribución presentada por:

$$f(x) = \left\{ \begin{array}{ll} \frac{x}{2} & 0 \leq x \leq 2 \\ 0 & \text{en cualquier otro caso} \end{array} \right.$$

Una función de este tipo se llama función rampa. Se puede presentar en forma gráfica como se ve en la siguiente figura:



El área bajo la curva, $f(x) = x/2$, representa la probabilidad de ocurrencia de la variable aleatoria X . Supongamos que en este caso X representa los tiempos de servicio de un cajero automático. Para obtener valores aleatorios a partir de esta distribución mediante el método de transformación inversa, calculamos primero la función de densidad acumulada como

$$F(x) = \int_0^x \frac{t}{2} dt = \frac{x^2}{4}$$

Esta función de distribución acumulada se representa mediante la función

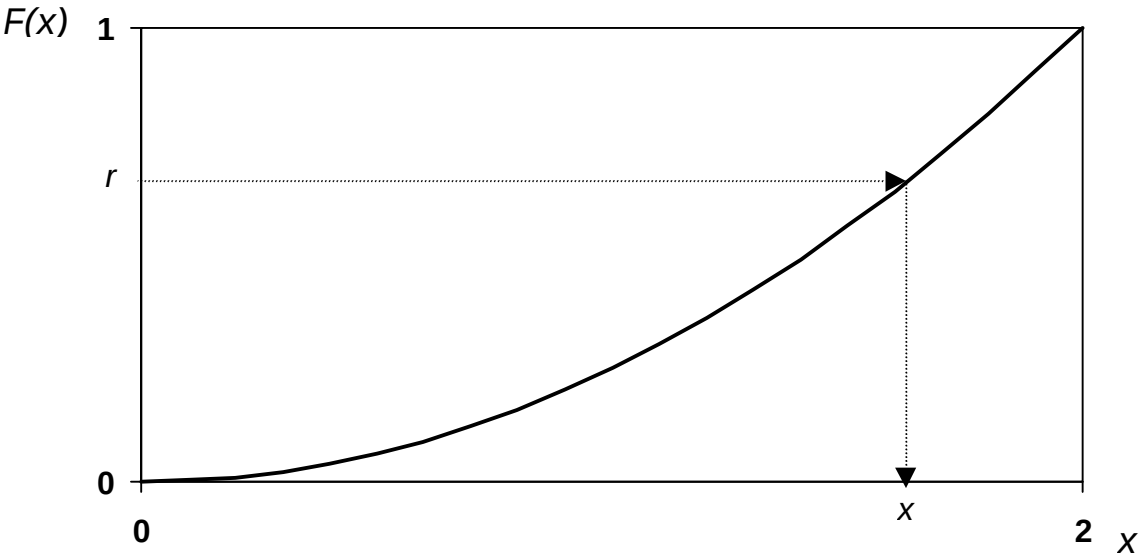
$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{4} & 0 \leq x \leq 2 \\ 1 & x > 0 \end{cases}$$

A continuación, en el paso 2, generamos un número aleatorio r . Por último, en el paso 3, hacemos que $F(x) = r$ y calculamos x .

$$\frac{x^2}{4} = r \quad \text{de aquí se obtiene que} \quad x = \pm 2\sqrt{r}$$

Como los tiempos de servicio sólo se definen para valores positivos de x , no es posible un valor $x = -2\sqrt{r}$. Esto nos deja con $x = 2\sqrt{r}$ como solución de x . A esta ecuación se le llama **generador de valor aleatorio** o **generador de proceso**. Así, para obtener un tiempo de servicio, generamos primero un número aleatorio y luego lo transformamos por medio de la ecuación anterior. Cada ejecución de la ecuación nos dará un tiempo de servicio de la distribución dada. Por ejemplo, si se obtiene un número aleatorio $r = 0.64$, se genera un tiempo de servicio $x = 2\sqrt{0.64} = 1.6$

En forma gráfica, el método de la transformación inversa se puede presentar como se ve en la siguiente figura.



Vemos en esta gráfica que el intervalo de valores para la variable aleatoria es $0 \leq x \leq 2$, el cual coincide con las probabilidades acumuladas $0 \leq F(x) \leq 1$. En otras palabras, para cualquier valor de $F(x)$ en el intervalo $[0,1]$ existe un valor correspondiente de la variable aleatorio, representado por x . Como un número aleatorio también se define en el intervalo entre 0 y 1, esto indica que un número aleatorio se puede traducir en forma directa a un valor correspondiente de x mediante la relación $r = F(x)$. La solución para x en términos de r se conoce como el cálculo de la inversa de $F(x)$ y se representa por $x = F^{-1}(r)$. De aquí el nombre de transformación inversa.

Ejemplo 2: Distribución exponencial

La distribución exponencial tiene importantes aplicaciones en la representación matemática de los sistemas de cola de espera. La función de distribución de probabilidad para esta distribución está representada por:

$$f(x) = \begin{cases} l e^{-l x} & x \geq 0. \quad \lambda > 2 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Generamos la función de probabilidad acumulada y la igualamos al número aleatorio generado

$$F(x) = \int_0^x l e^{-l t} dt = 1 - e^{-l x} = r$$

Despejando x se obtiene la función buscada

$$x = -\frac{1}{l} \ln(1-r)$$

Para simplificar nuestros cálculos, podemos reemplazar $(1-r)$ por r . Como r es número aleatorio, $(1-r)$ también será aleatorio. Esto quiere decir que no hemos cambiado nada, a excepción de la manera de escribir el número aleatorio $U(0,1)$. Así, nuestro generador de proceso para la distribución exponencial será ahora:

$$x = -\frac{1}{l} \ln(r)$$

Por ejemplo, si $\lambda = 0.5$ y $r = 0.456$, resulta $x = 1.571$

Ejemplo 3: Distribución uniforme general (a,b)

Si se desea simular números aleatorios con distribución uniforme entre a y b ; la función de densidad es

$$f(x) = \frac{1}{b-a} \quad a \leq x \leq b$$

La distribución acumulada de esta función de 0 a un valor x es:

$$F(x) = \frac{x}{b-a} \quad a \leq x \leq b$$

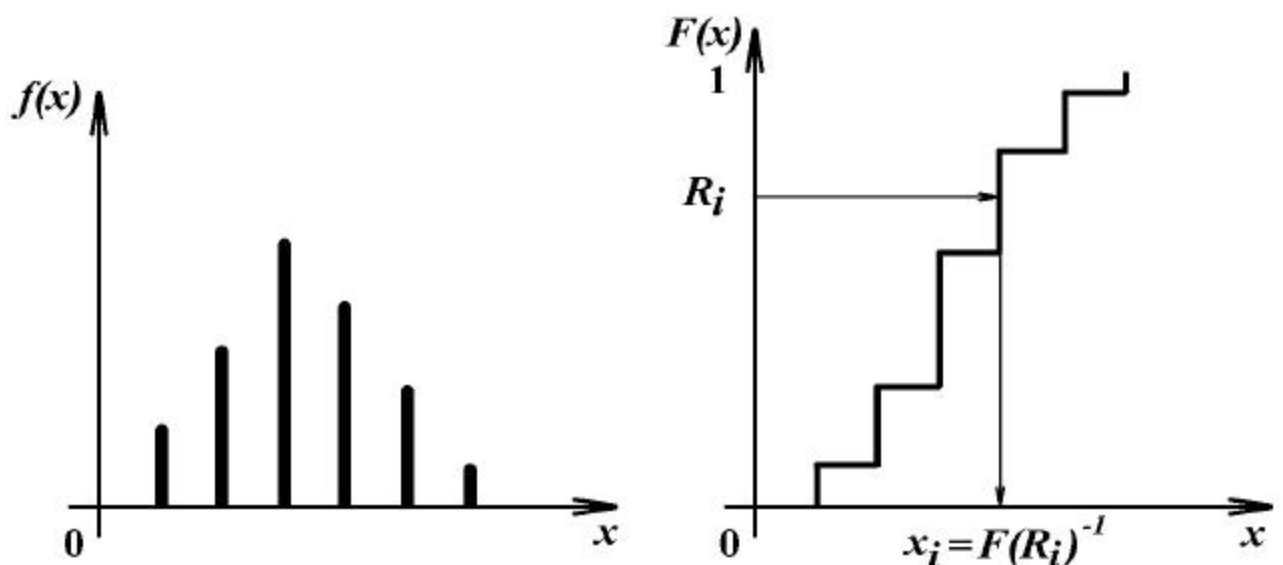
Igualando la acumulada de la función $F(x)$ al número aleatorio r y encontrando la transformada inversa (despejando x) se obtiene:

$$x = a + (b - a)r$$

MÉTODO DE TRANSFORMACIÓN INVERSA PARA DISTRIBUCIONES DISCRETAS

Se utiliza cuando se desea simular variables aleatorias de tipo discreto, como la distribución Bernoulli, binomial, Poisson, discreta general, etc. El procedimiento es similar al continuo pero el valor de $F(x)$ se encuentra acumulando las probabilidades de los eventos individuales $p(x)$. También en este caso, $F(x)$ está definida en el intervalo 0 a 1; se genera un número aleatorio r y se determina el valor de la variable aleatoria cuya distribución acumulada es igual a r .

La siguiente figura muestra en forma gráfica el procedimiento anterior para una función cualquiera $p(x)$ discreta.



La dificultad de este método radica en que no existe una expresión final sencilla, como en el caso de la continua.

La metodología es la siguiente:

- Paso 1.** Calcular todos los valores de $p(x)$ para la distribución propuesta.
- Paso 2.** Calcular la acumulada $F(x)$ para cada valor de x .
- Paso 3.** Generar un valor r . verificar en $F(x)$ a qué intervalo de x pertenece y ese será el número aleatorio generado por la distribución propuesta.

Ejemplo 1: Distribución Bernoulli

La distribución Bernoulli es la más sencilla y está definida como:

$$p(x) = p^x (1 - p)^{1-x} \quad \text{para } x = 0, 1$$

donde

$$\begin{aligned} x = 0 & \text{ (no ocurre el evento)} \\ x = 1 & \text{ (ocurre el evento)} \end{aligned}$$

Calculando las probabilidades de los eventos, se tiene:

$$p(0) = 1 - p \quad \text{y} \quad p(1) = p$$

la función acumulada es

$$F(0) = 1 - p \quad \text{y} \quad F(1) = 1$$

se genera un número aleatorio r entre 0 y 1, y

- si el número r se encuentra entre 0 y $1 - p$ el número aleatorio generado con la distribución de Bernoulli es 0
- si el número r se encuentra entre $1 - p$ y 1 el número aleatorio generado con la distribución de Bernoulli es 1.

La misma metodología puede emplearse para generar valores de una variable con distribución Poisson o binomial, con la diferencia de que en el cálculo de las probabilidades $p(x)$ del paso 1 hay que utilizar la distribución de probabilidades respectiva.

OTROS MÉTODOS ESTADÍSTICOS

En ocasiones, no es posible aplicar el método de la transformada inversa a distribuciones de probabilidad, debido principalmente a que algunas de ellas no tienen forma de integrarse, como el caso de la distribución normal, gamma, etc. En tal caso se requiere la aplicación de algunas propiedades estadísticas como el teorema del límite central o la propiedad de convolución, que permiten generar una distribución mediante la suma de distribuciones más sencillas, mediante la transformada z, etc. Estos son procedimientos especiales para lograr expresiones matemáticas que generen las variables aleatorias deseadas. A continuación se presentan algunas de las expresiones más utilizadas en la simulación de sistemas para generar variables aleatorias.

RESUMEN PARA DISTRIBUCIONES CONTINUAS

Se presentan ahora las expresiones finales para generar variables aleatorias con las distribuciones de probabilidad más usuales.

i. Distribución uniforme general

Obtenida a partir de la transformada inversa

$$U_i = a + (b - a) r_i$$

donde

$$\begin{aligned} a &= \text{límite inferior de la distribución uniforme.} \\ b &= \text{límite superior de la distribución uniforme.} \end{aligned}$$

U_i = número aleatorio de la distribución uniforme.
 r_i = número aleatorio con distribución uniforme entre 0 y 1.

ii. **Distribución exponencial**

Obtenida a partir del método de la transformada inversa.

$$E_i = -\frac{1}{l} \ln(r_i)$$

donde:
 $1/\lambda$ = media de la distribución exponencial.
 E_i = número aleatorio con distribución exponencial

iii. **Distribución normal**

Obtenida a partir del método de convolución

$$N_i = \left(\sum_{i=1}^{12} r_i - 6 \right) s + m$$

o bien mediante el método directo:

$$N_i = \left[\left(\sqrt{-2 \ln(1 - r_i)} \right) \cos(2\pi r_{i+1}) \right] s + m$$

$$N_i = \left[\left(\sqrt{-2 \ln(1 - r_i)} \right) \sin(2\pi r_{i+1}) \right] s + m$$

donde

 μ = media de la distribución normal.
 σ = desviación estándar de la distribución normal.
 N_i = número aleatorio de la distribución normal.

iv. **Distribución de Erlang**

Obtenida a partir del método de convolución

$$ER_i = -\frac{1}{kl} \ln \left(\prod_{i=1}^k r_i \right)$$

donde:
 $1/\lambda$ = valor esperado.
 k = parámetro de forma.
 ER_i = número aleatorio de la distribución de Erlang

RESUMEN PARA DISTRIBUCIONES DISCRETAS

i. **Distribución de Bernoulli**

Obtenida a partir del método de transformada inversa.

Si $0 \leq r_i < 1-p$	$BE_i = 0$
Si $1-p \leq r_i < 1$	$BE_i = 1$

donde

 p = probabilidad de ocurrencia del evento $x = 1$
 $1-p$ = probabilidad de ocurrencia del evento $x = 0$
 BE_i = número aleatorio con distribución Bernoulli.

ii. **Distribución binomial**

Obtenida a partir del método de convolución.

$$BI_i = \sum_{j=1}^n BE_j$$

donde

BE_i = número aleatorio con distribución Bernoulli.

N = número del evento máximo de la distribución binomial.

p = probabilidad de éxito de la distribución binomial que se involucra al generar los Bernoulli.

BI_i = número aleatorio con distribución binomial.

iii. Distribución Poisson

Obtenida a partir de la propiedad que mantiene con la distribución exponencial.

Paso 1 : Hacer $N = 0$, $T = 1$ y obtener r_i

Paso 2 : Calcular $T = T * r_i$

Paso 3 : Si la T calculada es mayor a $e^{-\alpha}$, calcular otro r_i y regresar al paso 2, incrementando N en 1
Si la T calculada es menor a $e^{-\alpha}$, entonces $P_i = N$.

donde:

α = media de la distribución de Poisson

P_i = número aleatorio de distribución Poisson.

N = Contador.

T = Contador.

Esta distribución es muy utilizada en las simulaciones de cola, ya que posee las siguientes propiedades:

- La probabilidad de que la entrada de un elemento se produzca en un intervalo de tiempo, es proporcional a la magnitud del intervalo.
- Admitimos que dos entradas no se producen al mismo tiempo. La probabilidad de que tal hecho suceda es muy pequeña.

Respetando lo anterior, se puede demostrar que la ley de entrada es Poisson.

Ejemplos:

A partir de un generador de números aleatorios uniformes entre 0 y 1 se obtuvieron los valores 0.7814 y 0.5643. A partir de ellos simular:

- a) Una variable aleatoria con distribución uniforme entre 15 y 19

$$U_i = (19 - 15) \cdot r_i + 15 = (4 \cdot 0.7814) + 15 = 18.1256$$

- b) Una variable aleatoria con distribución exponencial con media $1/\lambda = 5$.

$$E = -5 \ln(r_i) = -5 \ln(0.7814) = 1.233$$

- c) Una variable aleatoria con distribución Bernoulli con $p = 0.25$

Como $r_i = 0.7814$ se encuentra entre 0.25 y 1, entonces : $Be_i = 0$.

- d) Una variable aleatoria con distribución normal con media 10 y desviación estándar 2.

$$N_i = \left[\left(\sqrt{-2 \ln(1 - r_i)} \right) \sin(2\pi r_{i+1}) \right] 2 + 10$$

$$N_i = \left[\left(\sqrt{-2 \ln(1 - 0.7814)} \right) \sin(2\pi 0.5643) \right] 2 + 10 = 9.4477$$

- e) Una variable aleatoria con distribución Poisson con media $\alpha = 0.5$.

1. $N = 0$, $T = 1$

2. $T = T * r_i$, $T = 0.7814$

3. Es $T > e^{-0.5} = 0.606$, Afirmativo

4. $N = N + 1 = 1$

5. $T = T * r_{i+1}$, $T = 0.7814 * 0.5643 = 0.44$

6. Es $T > e^{-0.5} = 0.606$, Negativo

$$P_i = 1$$

GENERACION DE VARIABLES ALEATORIAS USANDO PLANILLA DE CALCULO

Las hojas de cálculo permiten poner en práctica modelos de simulación de tamaño reducido y moderado, con relativa facilidad y rapidez. Por tal motivo es conveniente conocer las herramientas que posee para generar los números aleatorios. Las funciones que permiten generar dichos números, en las planillas Excel son las siguientes:

ALEATORIO()
Devuelve un número aleatorio mayor o igual que 0 y menor que 1, distribuido uniformemente.

DISTR.NORM.INV(prob;media;desv_estándar)
Prob es una probabilidad asociada a la distribución normal.
Media es la media aritmética de la distribución.
Desv_estándar es la desviación estándar de la distribución.

Devuelve, para una probabilidad dada, el valor de la variable aleatoria siguiendo una distribución acumulativa normal para la media y desviación estándar especificadas.

DISTR.GAMMA.INV(prob;alfa;beta)
Prob es la probabilidad asociada con la distribución gamma.
Alfa es un parámetro de la distribución.
Beta es un parámetro de la distribución. Si beta = 1, DISTR.GAMMA.INV devuelve el valor de la variable aleatoria siguiendo una distribución gamma estándar.

Devuelve, para una probabilidad dada, el valor de la variable aleatoria siguiendo una distribución gamma acumulativa. Si $p = \text{DISTR.GAMMA}(x, \dots)$, entonces $\text{DISTR.GAMMA.INV}(p, \dots) = x$. Utilice esta función para estudiar variables cuya distribución podría ser asimétrica.

DISTR.BETA.INV(probabilidad;alfa;beta;A;B)
Probabilidad es una probabilidad asociada con la distribución beta.
Alfa es un parámetro de la distribución.
Beta es un parámetro de la distribución.
A es un límite inferior opcional del intervalo de x.
B es un límite superior opcional del intervalo de x.

Devuelve, para una probabilidad dada, el valor de la variable aleatoria siguiendo una distribución beta. Es decir, si el argumento probabilidad = $\text{DISTR.BETA}(x; \dots)$, entonces $\text{DISTR.BETA.INV}(\text{probabilidad}; \dots) = x$. La distribución beta acumulada puede emplearse en la organización de proyectos para crear modelos con fechas de finalización probables, de acuerdo con un plazo de finalización y variabilidad esperados.

DISTR.T.INV(probabilidad;grados_de_libertad)
Probabilidad es la probabilidad asociada con la distribución t de Student dos colas.
Grados_de_libertad es el número de grados de libertad para diferenciar la distribución.

Devuelve, para una probabilidad dada, el valor de la variable aleatoria siguiendo una distribución t de Student para los grados de libertad especificados.

Otras funciones estadísticas de interés son:

DISTR.EXP(x;lambda;acum)
POISSON(x;media;acumulado)
DIST.WEIBULL(x;alfa;beta;acumulado)
DISTR.LOG.INV(probabilidad;media;desv_estándar)