Método de Transformada Inversa

Temas Selectos I

Dr. Rubén Blancas Rivera

Universidad de las Américas Puebla

Otoño 2025

Content

Introducción

- Mostraremos que a partir de variables aleatorias uniformes en el intervalo (0, 1), podemos generar variables aleatorias con distribución arbitraria dada.
- Recordemos la definición de función de distribución:

$$F(x) := \mathbb{P}(X \leq x)$$

Transformada Inversa

Suponga que queremos generar el valor de una variable aleatoria discreta X con función de probabilidad de masa:

$$Pr[X = x_j] = p_j, \qquad j = 0, 1, ..., \tag{1}$$

donde $\sum_j p_j = 1$.

Para esto, generamos un número aleatorio U; es decir, U está distribuido uniformemente en $(0,1)$, y sea

$$X = \begin{cases} x_0 \quad si & U < p_0 \\ x_1 \quad si & p_0 \leq U < p_0 + p_1 \\ \vdots & \\ x_j \quad si & \sum_{i=0}^{j-1} p_i \leq U < \sum_{i=0}^{j} p_j \\ \vdots & \end{cases}$$

Como $P[a \leq U \leq b] = b - a$ para $0 < a < b < 1$, tenemos que

$$P[X = x_j] = P\left[\sum_{i=1}^{j-1} p_i \leq U \leq \sum_{i=1}^{j} p_i\right] = p_j$$

y entonces X tiene la distribución deseada.

1. 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.
   - Si $U < p_0 + p_1 + p_2$ hacer $X = x_2$ y terminar.
   - $\vdots$
2. Si los x_i para $i \geq 0$, están ordenados de modo que

   $x_0 < x_1 < x_2 < \vdots$ y si F denota la función de distribución de X, entonces $F(x_k) = \sum_{i=0}^{k} p_i$ y entonces

$$X \text{ será igual a } x_j \qquad \text{si } F(x_{j-1}) \leq U \leq F(x_j).$$

Algoritmo

Algorithm 1 Estimación de π

Entrada: Función de distribución acumulada discreta $F(x)$, con soporte $x_0, x_1, \ldots, x_n, \cdots$
Salida: Una variable aleatoria X con f.d. F
procedure TransformadaInversaDiscreta
  Generar un número aleatorio $U \sim \mathcal{U}(0,1)$
  Inicializar $x \leftarrow x_0$
  while $F(x) < U$ **do**
    $x \leftarrow$ siguiente valor en el soporte
  end while
  return x
end procedure

Ejemplo

Si queremos simular una variable aleatoria X tal que

$p_1 = 0.20 \quad p_2 = 0.15 \quad p_3 = 0.25 \quad p_4 = 0.40$ donde $p_j = P[X = j]$

entonces podríamos generar U y hacer lo siguiente:

1. Si $U < 0.20$ hacemos $X = 1$ y terminamos
2. Si $U < 0.35$ hacemos $X = 2$ y terminamos
3. Si $U < 0.60$ hacemos $X = 3$ y terminamos En caso contrario, hacemos $X = 4$

Ejemplo

Sin embargo, un procedimiento más eficaz es el siguiente:

- Si $U < 0.40$ hacemos $X = 4$ y terminamos
- Si $U < 0.65$ hacemos $X = 3$ y terminamos
- Si $U < 0.85$ hacemos $X = 1$ y terminamos
- En caso contrario, hacemos $X = 2$

Variable Aleatoria Discreta Uniforme

X tiene una distribución discreta uniforme sí

$$P[X=j]=\frac{1}{n}, \quad j=1,...,n.$$

Aplicando el método de la transformada inversa podemos simular X al generar U y luego hacer que

$$X=j \quad \text{si} \quad \frac{j-1}{n} \leq U \leq \frac{j}{n}.$$

Variable Aleatoria Discreta Uniforme

Por lo tanto, X será igual a j si $j - 1 \leq nU < j$; o bien, en otras palabras,

$$X = [nU] + 1$$

donde $[x]$ es la parte entera de x (es decir, el mayor entero menor o igual a x).

Simulación Variable Aleatoria Geométrica

X es una variable aleatoria geométrica con parámetro p si

$$\mathbb{P}\{X = i\} = p\,q^{i-1}, \quad i \geq 1,$$

donde $q = 1 - p$.
La variable X puede interpretarse como el tiempo de la primera ocurrencia de un éxito cuando se realizan ensayos independientes, cada uno de los cuales resulta en éxito con probabilidad p.

Simulación Variable Aleatoria Geométrica

En efecto, se tiene que

$$\begin{aligned}\sum_{i=1}^{j-1} \mathbb{P}\{X = i\} &= 1 - \mathbb{P}\{X > j - 1\} \\ &= 1 - \mathbb{P}\{\text{los primeros } j - 1 \text{ ensayos son todos fracasos}\} \\ &= 1 - q^{j-1}, \quad j \geq 1.\end{aligned}$$

Simulación Variable Aleatoria Geométrica

Podemos generar el valor de X generando un número aleatorio U uniformemente distribuido en $(0, 1)$ y fijando X igual a aquel valor j para el cual

$$1 - q^{j-1} \leq U < 1 - q^j,$$

o, equivalentemente, para el cual

$$q^j < 1 - U \leq q^{j-1}.$$

Es decir, podemos definir

$$X = \min\{ j : q^j < 1 - U \}.$$

Simulación Variable Aleatoria Geométrica

Por lo tanto, utilizando el hecho de que el logaritmo es una función monótona —de modo que

$$a < b \iff \log(a) < \log(b),$$

se obtiene que X puede expresarse como

$$X = \min\{\, j : j\log(q) < \log(1-U)\,\}\,.$$

Equivalentemente,

$$X = \min\left\{ j : \; j > \frac{\log(1-U)}{\log(q)} \right\}.$$

Simulación Variable Aleatoria Geométrica

Obsérvese que en la última desigualdad el signo cambia porque

$$\log(q) < 0 \quad \text{para } 0 < q < 1.$$

Así, utilizando la notación $\mathrm{Int}(\cdot)$ (parte entera), podemos expresar

$$X \;=\; \mathrm{Int}\left(\frac{\log(1-U)}{\log(q)}\right) + 1.$$

Simulación Variable Aleatoria Geométrica

Finalmente, al notar que $1 - U$ también está uniformemente distribuido en $(0, 1)$, se sigue que

$$X \equiv \mathrm{Int}\left(\frac{\log(U)}{\log(q)}\right) + 1$$

también es geométrica con parámetro p.

Simulación Variable Aleatoria Geométrica

Algoritmo para generar una variable aleatoria geométrica:

1. Sea p el parámetro de la distribución geométrica, $0 < p < 1$.
2. Generar un número aleatorio U uniformemente distribuido en $(0, 1)$.
3. Calcular
$$X = \left\lfloor \frac{\log(U)}{\log(1-p)} \right\rfloor + 1.$$
4. Devolver X.

Simulación Variables Aleatoria Poisson

La variable aleatoria X es *Poisson* con media λ si

$$p_i \;=\; \mathbb{P}\{X = i\} \;=\; e^{-\lambda}\,\frac{\lambda^i}{i!}, \qquad i = 0, 1, \ldots$$

La clave para usar el método de la transformada inversa para generar una variable aleatoria de este tipo es la siguiente identidad

$$p_{i+1} \;=\; \frac{\lambda}{i+1}\,p_i, \qquad i \geq 0.$$

Simulación Variables Aleatoria Poisson

- Al usar la recurrencia anterior para calcular las probabilidades de Poisson a medida que se necesiten, el algoritmo de la transformada inversa para generar una variable aleatoria Poisson con media λ como sigue.
- Note que la cantidad i se refiere al valor que se está considerando; $p = p_i$ es la probabilidad de que X sea igual a i, y $F = F(i)$ es la probabilidad de que $X \leq i$.

Simulación Variables Aleatoria Poisson

Algorithm 2 Algoritmo de la transformada inversa para la distribución Poisson(λ)

Generar un número aleatorio $U \sim U(0,1)$.
Inicializar $i = 0, \quad p = e^{-\lambda}, \quad F = p$.
if $U < F$ **then**
    Asignar $X = i$ y detener.
end if
loop
    Actualizar $p = \dfrac{\lambda}{i+1} p, \quad F = F + p, \quad i = i + 1$.
    if $U < F$ **then**
        Asignar $X = i$ y detener.
    end if
end loop

Simulación Variables Aleatorias Poisson

El número de búsquedas necesarias por este algoritmo es aproximadamente 1 más que la diferencia absoluta entre la variable aleatoria X y su media λ. Dado que, para λ grande, una distribución de Poisson es (por el teorema central del límite) aproximadamente normal con media y varianza iguales a λ, se sigue que

$$\text{Número promedio de búsquedas} \approx 1 + \mathbb{E}[\,|X - \lambda|\,], \qquad X \sim N(\lambda, \lambda).$$

Número esperado de comparaciones

- Se verifican los valores $0, 1, \ldots, X$, es decir $X + 1$ comparaciones.
- Valor esperado para Poisson de media λ:

$$\mathbb{E}[\text{comparaciones}] \approx 1 + \lambda$$

- Eficiente para λ pequeño, lento para λ grande.

Optimización para λ grande

1. Sea $I = \text{Int}(\lambda)$.
2. Calcular $F(I)$ usando la recurrencia.
3. Generar U.
4. Si $U \leq F(I)$, buscar hacia abajo desde I.
5. Si $U > F(I)$, buscar hacia arriba desde $I + 1$.

- Reduce comparaciones promedio al centrarse cerca de la moda de Poisson $Int(\lambda) - 1$.

Busquedas necesarias

El número de búsquedas necesarias por este algoritmo es aproximadamente 1 más que la diferencia absoluta entre la variable aleatoria X y su media λ. Dado que, para λ grande, una distribución de Poisson es (por el teorema central del límite) aproximadamente normal con media y varianza iguales a λ, se sigue que

$$\text{Número promedio de búsquedas} \approx 1+\mathbb{E}[\,|X-\lambda|\,], \qquad X \sim N(\lambda,\lambda).$$

Busquedas necesarias

Entonces,

$$= 1 + \sqrt{\lambda}\,\mathbb{E}\left[\frac{|X-\lambda|}{\sqrt{\lambda}}\right] = 1 + \sqrt{\lambda}\,\mathbb{E}[\,|Z|\,], \qquad Z \sim N(0,1).$$

Puesto que $\mathbb{E}[\,|Z|\,] = \sqrt{\frac{2}{\pi}} \approx 0.798$ se obtiene

$$\text{Número promedio de búsquedas} \;\approx\; 1 + 0.798\,\sqrt{\lambda}.$$

Generación de variables aleatorias binomiales

Supongamos que queremos generar el valor de una variable aleatoria binomial (n, p), es decir, X tal que

$$\mathbb{P}\{X = i\} = \binom{n}{i} p^i (1-p)^{n-i}, \qquad i = 0, 1, \dots, n.$$

Para ello, empleamos el método de la transformada inversa utilizando la identidad recursiva

$$\mathbb{P}\{X = i+1\} = \frac{n-i}{i+1} \frac{p}{1-p} \mathbb{P}\{X = i\}.$$

Con i denotando el valor que se está considerando, pr $= \mathbb{P}\{X = i\}$ la probabilidad de que $X = i$, y $F = F(i)$ la probabilidad de que $X \leq i$, el algoritmo puede expresarse de la siguiente manera:

Algorithm 3 Transformada inversa para generar una variable binomial (n, p)

```
Generar un número aleatorio U ~ U(0,1).
Inicializar c = p/(1-p), i = 0, pr = (1-p)^n, F = pr.
if U < F then
    Asignar X = i y detener.
end if
loop
    Actualizar pr = c(n-i)/(i+1) pr,   F = F + pr,   i = i + 1.
    if U < F then
        Asignar X = i y detener.
    end if
end loop
```

- El algoritmo anterior primero verifica si $X = 0$, luego si $X = 1$, y así sucesivamente.
- Por lo tanto, el número de búsquedas que realiza es $1 + X$. En promedio, tomará $1 + np$ búsquedas generar X.
- Dado que una variable binomial (n, p) representa el número de éxitos en n ensayos independientes, donde cada ensayo es un éxito con probabilidad p, se sigue que dicha variable también puede generarse restando de n el valor de una variable binomial $(n, 1 - p)$.
- Por ello, cuando $p > 1/2$, podemos generar una variable binomial $(n, 1 - p)$ en lugar de (n, p) para mayor eficiencia.

Transformada Inversa para Variables Aleatorias Continuas

Consideremos una variable aleatoria continua con función de distribución F. Un método general para generar dicha variable aleatoria —llamado *método de transformación inversa*— se basa en la siguiente proposición.

Proposición.

Sea U una variable aleatoria uniforme en $(0,1)$. Para cualquier función de distribución continua F, la variable aleatoria X definida por

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

tiene distribución F. Aquí, $F^{-1}(u)$ se define como aquel valor x tal que $F(x) = u$.

Transformada Inversa para Variables Aleatorias Continuas

Demostración.

Sea F_X la función de distribución de $X = F^{-1}(U)$. Entonces,

$$F_X(x) = \mathbb{P}\{X \leq x\} = \mathbb{P}\{F^{-1}(U) \leq x\}. \qquad (*)$$

Dado que F es una función de distribución, $F(x)$ es monótona creciente en x, por lo que la desigualdad "$a \leq b$" es equivalente a "$F(a) \leq F(b)$". Por lo tanto, a partir de la ecuación (*), tenemos

$$F_X(x) = \mathbb{P}\big(F(F^{-1}(U)) \leq F(x)\big) = \mathbb{P}\{U \leq F(x)\} = F(x).$$

pues U es uniforme en $(0, 1)$

Demostración.

Demostración.

Así, la proposición anterior muestra que podemos generar una variable aleatoria X con función de distribución continua F generando un número aleatorio U y luego tomando

$$X = F^{-1}(U).$$

Ejemplo

Ejemplo

Supongamos que queremos generar una variable aleatoria X con función de distribución

$$F(x) = x^n, \quad 0 < x < 1.$$

Si dejamos $x = F^{-1}(u)$, entonces

$$u = F(x) = x^n \quad \text{o, de manera equivalente,} \quad x = u^{1/n}.$$

Ejemplo

Ejemplo

Por lo tanto, podemos generar dicha variable aleatoria X generando un número aleatorio $U \sim U(0,1)$ y luego tomando

$$X = U^{1/n}.$$

Simulación Variable Aleatoria Exponencial

Si X es una variable aleatoria exponencial con tasa $\lambda = 1$, su función de distribución viene dada por

$$F(x) = 1 - e^{-x}, \quad x \geq 0.$$

Si dejamos $x = F^{-1}(u)$, entonces

$$u = F(x) = 1 - e^{-x} \quad \Rightarrow \quad 1 - u = e^{-x} \quad \Rightarrow \quad x = -\log(1-u).$$

Simulación Variable Aleatoria Exponencial

Por lo tanto, podemos generar una variable aleatoria exponencial con parámetro 1 generando un número aleatorio $U \sim U(0,1)$ y tomando

$$X = F^{-1}(U) = -\log(1-U).$$

Un pequeño ahorro de tiempo se obtiene al notar que $1-U$ también es uniforme en $(0,1)$, y por lo tanto $-\log(1-U)$ tiene la misma distribución que $-\log U$. Es decir, el logaritmo negativo de un número aleatorio uniforme es exponencial con tasa 1.

Simulación Variable Aleatoria Exponencial

Además, si X es exponencial con media 1, entonces para cualquier constante positiva c, la variable cX es exponencial con media c. Por lo tanto, una variable aleatoria exponencial X con tasa λ (media $1/\lambda$) puede generarse generando un número aleatorio U y tomando

$$X = -\frac{1}{\lambda} \log U.$$

Simulación Variable Aleatoria Exponencial

- El método anterior también nos proporciona otro algoritmo para generar una variable aleatoria Poisson.
- Para comenzar, recordemos que un *proceso de Poisson* con tasa λ se obtiene cuando los tiempos entre eventos sucesivos son variables aleatorias exponenciales independientes con tasa λ.

Simulación Variable Aleatoria Exponencial

Para tal proceso, $N(1)$, el número de eventos hasta el tiempo 1, tiene distribución Poisson con media λ. Sin embargo, si denotamos por X_i, $i = 1, 2, \ldots,$ los tiempos entre llegadas sucesivas, entonces el n-ésimo evento ocurre en el tiempo

$$\sum_{i=1}^{n} X_i,$$

y así, el número de eventos hasta el tiempo 1 puede expresarse como

$$N(1) = \max\left\{ n : \sum_{i=1}^{n} X_i \leq 1 \right\}.$$

Simulación Variable Aleatoria Exponencial

Por lo tanto, utilizando los resultados anteriores, podemos generar $N = N(1)$, una variable aleatoria Poisson con media λ, generamos números aleatorios $U_1, U_2, \ldots$ y tomando

$$
\begin{aligned}
N &= \max\left\{ n : \sum_{i=1}^{n} -\frac{1}{\lambda} \log U_i \leq 1 \right\} \\
&= \max\left\{ n : \sum_{i=1}^{n} \log U_i \leq -\lambda \right\} \\
&= \max\left\{ n : \log(U_1 \cdots U_n) \leq -\lambda \right\} \\
&= \max\left\{ n : U_1 \cdots U_n \geq e^{-\lambda} \right\}.
\end{aligned}
$$

Simulación Variable Aleatoria Exponencial

Por lo tanto, una variable aleatoria Poisson N con media λ puede generarse generando sucesivamente números aleatorios hasta que su producto sea menor que $e^{-\lambda}$, y luego tomando N igual a uno menos el número de números aleatorios generados. Es decir,

$$N = \min\left\{n : U_1 \cdots U_n < e^{-\lambda}\right\} - 1.$$

Simulación Variable Aleatoria Gamma

Supongamos que queremos generar el valor de una variable aleatoria *gamma* (n, λ). Dado que la función de distribución F de tal variable aleatoria está dada por

$$F(x) = \int_0^x \lambda e^{-\lambda y} \frac{(\lambda y)^{n-1}}{(n-1)!} \, dy.$$

Simulación Variable Aleatoria Gamma

- No es posible dar una expresión cerrada para su inversa.
- Sin embargo, utilizando el resultado de que una variable aleatoria *gamma* (n, λ),
- X, puede considerarse como la suma de n variables exponenciales independientes, cada una con tasa λ, se obtiene un método práctico para su simulación.

Simulación Variable Aleatoria Gamma

Específicamente, podemos generar una variable aleatoria *gamma* (n, λ) generando n números aleatorios $U_1, \ldots, U_n$ y luego tomando

$$X = -\frac{1}{\lambda}\log U_1 - \cdots - \frac{1}{\lambda}\log U_n = -\frac{1}{\lambda}\log(U_1 \cdots U_n),$$

donde el uso de la identidad

$$\sum_{i=1}^{n} \log x_i = \log(x_1 \cdots x_n)$$

ahorra tiempo de cómputo, ya que requiere solo un cálculo de logaritmo en lugar de n.

Máximo de Variables Aleatorias i.i.d.

Supongamos que $X_1, \ldots, X_n$ son variables aleatorias continuas independientes e idénticamente distribuidas, cuya función de distribución F es invertible, y que queremos generar

$$M = \max_{i=1,\ldots,n} X_i.$$

Para ello, notemos que la función de distribución de M, llamémosla G, es

$$G(x) = \mathbb{P}(M \leq x) = \mathbb{P}(X_1 \leq x, \ldots, X_n \leq x) = F^n(x).$$

En consecuencia, si $M = G^{-1}(U)$, donde $U \sim U(0,1)$, entonces

$$M = F^{-1}(U^{1/n}).$$

Supongamos que la variable aleatoria continua X puede generarse mediante el método de la transformada inversa, y que queremos generar X condicional a que su valor se encuentre en el intervalo (a, b), para valores dados $a < b$.
Dado que $F^{-1}(U)$ tiene la distribución de X, se sigue que

$$F^{-1}(U) \mid a < F^{-1}(U) < b$$

tiene la distribución condicional que deseamos.

Generación de variables aleatorias continuas condicionales

Como F es una función creciente, $a < F^{-1}(U) < b$ es equivalente a

$$F(a) < U < F(b),$$

y por lo tanto queremos generar una variable aleatoria distribuida como $F^{-1}(U)$ condicional a $F(a) < U < F(b)$.

Generación de variables aleatorias continuas condicionales

Pero U condicional a $F(a) < U < F(b)$ se distribuye como una variable uniforme en el intervalo $(F(a), F(b))$. Por lo tanto, el resultado deseado se obtiene generando un número aleatorio $U \sim U(0,1)$, definiendo

$$U_{a,b} = F(a) + (F(b) - F(a))\,U$$

y tomando

$$X = F^{-1}(U_{a,b}).$$

Generación condicional de una variable uniforme

Supongamos que $X \sim U(0,1)$ y queremos generar X condicional a que $X \in (0.2, 0.5)$.
La función de distribución de X es $F(x) = x$. Según el método de transformada inversa condicional, generamos un número aleatorio $U \sim U(0,1)$ y definimos

$$U_{0.2,0.5} = F(0.2) + (F(0.5) - F(0.2))\,U = 0.2 + (0.5 - 0.2)U = 0.2 + 0.3U.$$

Luego, tomamos

$$X = F^{-1}(U_{0.2,0.5}) = U_{0.2,0.5} = 0.2 + 0.3U.$$

De esta manera, X queda distribuida uniformemente en el intervalo $(0.2, 0.5)$.

Geenración condicional de una variable exponencial

Supongamos que $X \sim \text{Exp}(\lambda)$, es decir, con función de distribución

$$F(x) \;=\; 1 - e^{-\lambda x}, \quad x > 0.$$

Queremos generar X condicional a que $X \in (a, b)$, donde $0 < a < b$.
De acuerdo con el método, generamos un número $U \sim U(0,1)$ y definimos

$$U_{a,b} = F(a) + (F(b) - F(a))\,U = \left(1 - e^{-\lambda a}\right) + \left(e^{-\lambda a} - e^{-\lambda b}\right)U.$$

Luego, tomamos

$$X = F^{-1}(U_{a,b}) = -\frac{1}{\lambda}\ln\left(1 - U_{a,b}\right).$$

Por ejemplo, si $\lambda = 2$, $a = 1$ y $b = 3$, entonces

$$U_{1,3} = \left(1 - e^{-2}\right) + \left(e^{-2} - e^{-6}\right)U,$$

y

$$X = -\tfrac{1}{2}\ln\left(1 - U_{1,3}\right).$$

De esta forma obtenemos una variable exponencial con tasa λ condicionada a tomar valores en el intervalo (a, b).