
ESTADÍSTICA 2007/2008

TEMA 10: SIMULACIÓN DE VARIABLES ALEATORIAS

DESCRIPCIÓN DEL TEMA:

- 10.1. Introducción.
- 10.2. Método de las transformaciones.
- 10.3. Método de inversión.
- 10.4. Método de aceptación-rechazo.
- 10.5. Simulaciones Montecarlo.

OBJETIVOS PRINCIPALES:

- Comprender los métodos de simulación de variables aleatorias básicos y sus aplicaciones.
 - Conocer las ideas básicas que subyacen a las técnicas de tipo Montecarlo.
-

10.1. Introducción.

- Los ordenadores pueden simular probabilidades y variables aleatorias *numéricamente*.
 - **Aplicaciones:**
 - Aproximar cantidades que resultan difíciles de calcular matemáticamente.
 - Simular sistemas físicos o biológicos complejos.
 - Implementar algoritmos complejos para procesar imágenes, reconocer sonidos, etc.
 - Encriptar datos o generar claves.
 - Generar opciones aleatorias en juegos de ordenador, etc.
-

Punto de partida

- Prácticamente todos los lenguajes de programación disponen de un *generador de números aleatorios* que permiten generar una secuencia U_1, U_2, \dots de números aleatorios aproximadamente independientes y con una distribución aproximadamente $U(0, 1)$.
- Un ejemplo del algoritmo de generación:

$$x_n = ax_{n-1} \mod m$$

(x_0 es un parámetro del algoritmo),

$$m = 2^{31} - 1, \quad a = 7^5 = 16\,807.$$

Requerimientos sobre un generador de números aleatorios

- Buena distribución;
 - Período grande;
 - Posibilidad de repetir (por ejemplo, para hacer un test del programa). *Posibilidad de almacenar el estado de generador.*
-

-
- Posibilidad de transferencia de un lenguaje de programación a otro y de un ordenador al otro, con exactamente los mismos resultados.
 - Eficiencia en el sentido del consumo de tiempo.
-
- Podemos suponer que disponemos de una sucesión de variables aleatorias $U(0, 1)$ e independientes: $U_1, U_2, U_3, U_4, U_5, \dots$
 - A partir de la sucesión anterior vamos a estudiar cómo generar valores aleatorios de otras distribuciones de probabilidad.
 - Veremos diferentes métodos para lograr este objetivo:
 1. Método de las transformaciones
 2. Método de inversión
 3. Método de aceptación-rechazo
-

1. Método de las transformaciones

- Para simular una v.a. X bastará encontrar una v.a. que se pueda expresar como una función de U_1, U_2, \dots, U_n y que tenga la misma distribución que X . Es decir, se trata de encontrar una transformación

$$Y = f(U_1, \dots, U_n) \sim X.$$

Ejemplos:

- **Uniforme sobre (a, b)**

$$X = a + (b - a)U_1 \sim \text{U}(a, b), \quad U_1 \sim \text{U}(0, 1).$$

$$\text{U}(10, 12)$$

-
- **Uniforme sobre $\{1, \dots, n\}$**

$$X = \lceil nU_1 \rceil \sim \mathbf{U}(\{1, \dots, n\}).$$

$$\mathbf{U}(\{1, 2, 3, \dots, 10\})$$

- **Bernoulli de parámetro p**

$$\mathbf{B}(1; p) \sim X_1 = \begin{cases} 1 & \text{si } U_1 \leq p, \\ 0 & \text{si } U_1 > p. \end{cases}$$

$$\mathbf{B}(1; 0,5)$$

- **Binomial de parámetros n, p**

$$\mathbf{B}(1; p) \sim X_i = \begin{cases} 1 & \text{si } U_i \leq p, \\ 0 & \text{si } U_i > p. \end{cases} \quad X = X_1 + \dots + X_n \sim \mathbf{B}(n; p).$$

- **Geométrica de parámetro p**

$$B(1; p) \sim X_i = \begin{cases} 1 & \text{si } U_i \leq p, \\ 0 & \text{si } U_i > p. \end{cases} \quad X = \min\{i : X_i = 1\} \sim G(p).$$

- **Normal tipificada (La transformada de Box-Muller)**

Sean $U_1 \sim U(0, 1)$ y $U_2 \sim U(0, 1)$ independientes. Tomemos,

$$Z_1 := \sqrt{2 \log(1/U_1)} \cos(2\pi U_2), \quad Z_2 := \sqrt{2 \log(1/U_1)} \sin(2\pi U_2).$$

Se tiene

$$Z_1 \sim N(0; 1), \quad Z_2 \sim N(0; 1) \quad \text{y} \quad Z_1, Z_2 \text{ independientes.}$$

Desventaja: Necesidad de calcular \sin , \cos , \log
(consume mucho tiempo).

- **Normal**

$$\text{Si } Z \sim N(0; 1), \text{ entonces } X = \mu + \sigma Z \sim N(\mu; \sigma).$$

- **Simulación de la Normal bivalente**

Sean $Z_1, Z_2 \sim N(0; 1)$ independientes, definimos

$$X = \mu_X + \sigma_X Z_1, \quad Y = \mu_Y + \sigma_Y(\rho Z_1 + \sqrt{1 - \rho^2} \cdot Z_2),$$

$$\rho = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y} = \frac{\text{cov}(X, Y)}{\sqrt{V(X)} \sqrt{V(Y)}}$$

$\implies (X, Y)$ normal bivalente con correlación ρ :

$$(X, Y) \sim N(\boldsymbol{\mu}; \boldsymbol{\Sigma}) \text{ con } \boldsymbol{\mu} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_X^2 & \sigma_{X,Y} \\ \sigma_{X,Y} & \sigma_Y^2 \end{bmatrix}$$

Observamos que:

– Si $\rho = 0$, entonces

$$X = \mu_X + \sigma_X Z_1, \quad Y = \mu_Y + \sigma_Y Z_2 \implies X \text{ e } Y \text{ independientes.}$$

Bajo normalidad: Incorrelación \iff Independencia.

-
- Las medias condicionadas

x fijo \implies

$$y := E[Y|X = x] = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X} (x - \mu_X) \quad (\text{rectas verdes})$$

y fijo \implies

$$x := E[X|Y = y] = \mu_X + \rho \frac{\sigma_X}{\sigma_Y} (y - \mu_Y). \quad (\text{rectas violetas})$$

- Si $\rho = 1$, entonces

$$Y = \mu_y + \sigma_y \left(\frac{X - \mu_x}{\sigma_x} \right) \implies \text{Dependencia lineal positiva perfecta}$$

- Si $\rho = -1$, entonces

$$Y = \mu_y - \sigma_y \left(\frac{X - \mu_x}{\sigma_x} \right) \implies \text{Dependencia lineal negativa perfecta}$$

- **Ejemplo (Normal bivalente):**

Supongamos que queremos simular valores de:

$$N(\boldsymbol{\mu}; \boldsymbol{\Sigma}) \quad \text{con} \quad \boldsymbol{\mu} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad (\rho = 1/\sqrt{2} \approx 0,7071).$$

Sean $Z_1, Z_2 \sim N(0; 1)$ independientes, definimos:

$$X = 1 + \sqrt{2}Z_1, \quad Y = 2 + \frac{1}{\sqrt{2}}(Z_1 + Z_2).$$

$$(X, Y) \sim N\left(\boldsymbol{\mu} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \boldsymbol{\Sigma} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\right).$$

2. Método de inversión

- Dada una v.a. X con f.d. F , definimos **la inversa generalizada de F o función cuantil de X** , F^{-1} , mediante:

$$F^{-1}(t) := \inf\{x : F(x) \geq t\}, \quad 0 < t < 1.$$

● **Observación:**

- La función cuantil es no decreciente y continual por la izquierda.
- Si F es biyectiva, entonces F^{-1} es la función inversa de F .
- Si X es una v.a. discreta que toma los valores $x_1 < x_2 < x_3 < \dots$, entonces

$$F^{-1}(t) = \min \left\{ x_j : \sum_{k=1}^j P(x_k) \geq t \right\}.$$

-
- **Método de inversión:** El método de inversión se basa en el siguiente resultado:

Dada una v.a. X con función de distribución F y $U \sim U(0, 1)$, entonces

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

● **Variables discretas**

Sea X una v.a. discreta que toma los valores $x_1 < x_2 < x_3 < \dots$, entonces

$$X \sim F^{-1}(U_1) = \min \left\{ x_j : \sum_{k=1}^j P(x_k) \geq U_1 \right\}.$$

Ejemplo 1 Sea X una v.a. discreta con función de probabilidad:

$$P(2) = 0,2, \quad P(3) = 0,2, \quad P(5) = 0,4, \quad P(7) = 0,1, \quad P(11) = 0,1.$$

Dividimos $[0, 1]$ en subintervalos:

$$0 \quad 0,2 \quad 0,4 \quad 0,8 \quad 0,9 \quad 1.$$

Les asignamos números:

$$2 \quad 3 \quad 5 \quad 7 \quad 11.$$

- Si $U_1 = 0,133$, entonces tomamos 2 como un valor aleatorio de X .
 - Si $U_1 = 0,254$, entonces tomamos 3 como un valor aleatorio de X .
 - Si $U_1 = 0,495$, entonces tomamos 5 como un valor aleatorio de X .
 - Si $U_1 = 0,891$, entonces tomamos 7 como un valor aleatorio de X .
 - Si $U_1 = 0,902$, entonces tomamos 11 como un valor aleatorio de X .
-

Ejemplo 2: Binomial de parámetros n, p

$$X = \min \left\{ j : \sum_{k=0}^j \binom{n}{k} p^k (1-p)^{n-k} \geq U_1 \right\} \sim \mathbf{B}(n; p).$$

- **Geométrica de parámetro p**

$$\begin{aligned} \mathbf{G}(p) \sim X &= \min \left\{ j : \sum_{k=0}^j p(1-p)^k \geq U_1 \right\} \\ &= \min \left\{ j : 1 - (1-p)^{j+1} \geq U_1 \right\} \\ &\sim \min \left\{ j : (1-p)^{j+1} \leq U_2 \right\} \\ &= \left\lceil \frac{\log U_2}{\log(1-p)} \right\rceil. \end{aligned}$$

- **Exponencial de parámetro 1**

Si $X \sim \text{Exp}(1)$, su función de distribución es: $F(x) = 1 - e^{-x}$, luego

$$F^{-1}(t) = \log \left(\frac{1}{1-t} \right) \quad \text{y} \quad \log \left(\frac{1}{1-U_1} \right) \sim \text{Exp}(1).$$

$$\log \left(\frac{1}{U_1} \right) \sim \log \left(\frac{1}{1-U_1} \right) \sim \text{Exp}(1).$$

Ejemplo: Consideremos una variable aleatoria X con densidad de probabilidad:

$$f(x) := \begin{cases} x^{-2} & \text{si } x > 1, \\ 0 & \text{en otro caso.} \end{cases}$$

- (a) Dada una sucesión $\mathcal{U} := \{U_1, U_2, U_3, \dots\}$ de variables aleatorias (mutuamente) independientes y uniformes en $(0, 1)$, descríbase el método de inversión para obtener valores aleatorios de la variable X a partir de la sucesión \mathcal{U} .
 - (b) Los dos primeros valores de la sucesión de variables uniformes han sido 0.515968 y 0.298085. Calcúlense dos valores aleatorios de X a partir de los números uniformes anteriores.
-

3. Método de aceptación-rechazo (Acceptance-rejection sampling) (*John von Neumann (1951)*)

- **Rejection sampling.**

Supongamos que queremos simular un valor de la variable X con densidad f tal que:

- f se concentra en el intervalo $[a, b]$.
- $f(x) \leq k$, para alguna constante k .

- **Idea:** Lanzamos dardos aleatoriamente en el rectángulo $[a, b] \times [0, k]$. Si el dardo cae dentro de la gráfica de f , tomamos el valor de la abscisa como un valor de la v.a. X .
-

- **Algoritmo:**

- (1) Generar un valor $U_1 \sim U[a, b]$.
- (2) Generar un valor $U_2 \sim U[0, k]$ independiente de U_1 .
- (3) Entonces:
 - (3.1) Si $U_2 \leq f(U_1)$, entonces $X = U_1$.
 - (3.2) Si $U_2 > f(U_1)$, entonces volver al paso (1).

- **Fundamento:** $X \sim U_1 | \{U_2 \leq f(U_1)\}$.

Ejemplo: Método de aceptación-rechazo

- **Normal tipificada:** $Z \sim N(0; 1)$ con densidad f .
 - $f(x) \approx 0$, para $|x| \geq 10$. Luego, podemos tomar $a = -10$, $b = 10$.
 - $|f(x)| \leq f(0) = 1/\sqrt{2\pi} = k$, $x \in \mathbb{R}$.
 - **Algoritmo:**
 - (1) Generar un valor $U_1 \sim U[-10, 10]$.
 - (2) Generar un valor $U_2 \sim U[0, 1/\sqrt{2\pi}]$ independiente de U_1 .
 - (3) Entonces:
 - (3.1) Si $U_2 \leq f(U_1)$, entonces $Z = U_1$.
 - (3.2) Si $U_2 > f(U_1)$, entonces volver al paso (1).
-

-
- En esta situación, la probabilidad de aceptar un valor de U_1 será $\sqrt{2\pi}/20 \approx 0,125331$. Es deseable encontrar métodos que permitan generar algoritmos más eficientes, es decir, que aumenten la probabilidad de aceptación.
-

Ejemplo: Consideremos una variable aleatoria X con distribución beta de parámetros 2 y 2, es decir, X es una variable continua con densidad de probabilidad:

$$f(x) := \begin{cases} 6x(1-x) & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

- (a) Supongamos que disponemos de una sucesión $(U_1, V_1), (U_2, V_2), \dots$ de pares de variables (mutuamente) independientes y con la misma distribución uniforme en el intervalo $(0, 1)$. Escribese el algoritmo de aceptación-rechazo (caso 1) (Montecarlo) para generar valores de la variable X a partir de los valores de la sucesión de uniformes.
 - (b) Los dos primeros valores de la sucesión de variables uniformes han sido $(0,479663; 0,142918)$ y $(0,0398101; 0,298392)$. ¿Cuáles son las salidas del algoritmo para este par?
 - (c) ¿Cuál es la probabilidad de ejecutar el algoritmo y no generar un valor aleatorio de X en dos ocasiones antes de generar 4 valores aleatorios de X ?
-

Simulaciones Montecarlo

- La técnicas de tipo Montecarlo son muy útiles para:
 - Aproximar el valor de una integral o una constante complicada de forma numérica, *sobre todo para integrar funciones de muchas variables.*
 - Calcular la distribución (aproximada) de una variable aleatoria para la que el cálculo analítico es muy complejo.
 - Resolver problemas estadísticos complejos.

Aplicaciones:

- Física estadística;
 - Física de grandes energías.
-

-
- **Problema:** Sea $\mathbf{X} = (X_1, \dots, X_n)$ un vector aleatorio y supongamos que queremos calcular (de forma aproximada) el valor de $E\phi(\mathbf{X})$, donde $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una función conocida.

► **Solución: Simulación Montecarlo**

- (1) Generar una sucesión de v.a.s $\mathbf{X}_1, \mathbf{X}_2, \dots$ independientes entre sí y con la misma distribución que \mathbf{X} .
- (2) Calcular $\mathbf{Y}_i := \phi(\mathbf{X}_i)$, $i \geq 1$.
- (3) La ley (fuerte) de grandes números asegura que:

$$\lim_{n \rightarrow \infty} \frac{\mathbf{Y}_1 + \dots + \mathbf{Y}_n}{n} = E\mathbf{Y}_1 = E\phi(\mathbf{X}) \quad \text{casi seguramente.}$$

Ejemplo

Calcular un valor aproximado para el número π mediante el método Montecarlo.

- Tomemos $X, Y \sim U(-1, 1)$. Definimos:

$$Z := \begin{cases} 1 & \text{si } X^2 + Y^2 \leq 1, \\ 0 & \text{en otro caso.} \end{cases}$$

- Tenemos $Z \sim B(1; \pi/4)$ y $EZ = \pi/4$.
- Consideremos Z_1, Z_2, \dots variables aleatorias independientes como Z . La ley de grandes números asegura que:

$$\frac{4}{n} (Z_1 + \dots + Z_n) \xrightarrow{c.s.} \pi, \quad n \rightarrow \infty.$$

Propuesta de problemas a entregar

Problema:

Supongamos que un ordenador con doble núcleo trabaja para ejecutar dos rutinas de un determinado programa. La distribución conjunta del tiempo de ejecución de las dos rutinas sigue una distribución de probabilidad normal de media $(5, 5)$ (en minutos), desviaciones típicas de 1 minuto y correlación 0.7. El programa no se puede ejecutar hasta que ambas rutinas se han ejecutado.

Calcúlese una estimación (mediante el método Montecarlo) del tiempo esperado que se tarda en poder ejecutar el programa.
