Breve Introdução à Simulação Estocástica

Frederico Caeiro

DM, Faculdade de Ciências e Tecnologia - Universidade Nova de Lisboa

8 de Fevereiro de 2010

1 Introdução

A simulação (estocástica) é uma técnica que permite estudar o comportamento de modelos. Esta técnica é usada frequentemente em situações em que é muito difícil, ou até mesmo impossível, obter uma solução exacta do modelo. Iremos apenas abordar a simulação estocástica pelo método de Monte Carlo[1] (MMC). No MMC a simulação é feita a partir de sequências de números aleatórios. O MMC é usado em diversas áreas, sendo por exemplo aplicado em:

- Resolução de integrais e equações diferenciais;
- Controlo de tráfego;
- Previsão de Índices financeiros;
- Modelação de materiais;
- Aerodinâmica;

É possível simular uma qualquer distribuição, a partir da simulação de números aleatórios uniformes no intervalo $]0,1[$. Um processo de gerar os valores aleatórios do intervalo $]0,1[$, consiste em colocar numa urna 10 bolas numeradas de 0 a 9 e proceder à extracção, com reposição, de n bolas. Os n algarismos extraídos constituem as primeiras casas decimais do número aleatório. Por exemplo, se extrairmos os algarismos 4, 8, 7, 0, 7, o número aleatório é 0.48707. Este procedimento não é adequado em situações que necessitam muitos valores aleatórios.

Actualmente, a simulação é feita com computadores. Como os computadores não conseguem gerar rapidamente grandes quantidades de números aleatórios, a simulação é feita

[1] A designação de método de "Monte Carlo" foi adoptada for cientistas do projecto Manhattan. O nome é uma referência a um casino do Mónaco.

utilizando números gerados por um processo determinístico, mas com um comportamento semelhante aos números efectivamente aleatórios. Estes números designam-se por números pseudo-aleatórios (NPA's). As sequências de NPA's são gerada por um método gerador (algoritmo) e necessitam de uma "semente", isto é, de um número que inicia o processo determinístico. Uma determinada semente, associada a um determinado algoritmo, produz sempre a mesma sequência de NPA's.

2 Geração de números pseudo-aleatórios uniformes

Qualquer programa computacional que sirva para gerar NPA's deve satisfazer os seguintes requisitos:

- Deve ser rápido;
- Deve ser portátil;
- Deve ter ciclos longos;
- Os números gerados devem poder ser reproduzidos (com a mesma semente);
- Os números gerados devem apresentar um comportamento independente e uniforme;

Observação: E costume designar por comprimento do ciclo, ou período, o número de valores consecutivos de cada ciclo.

Existem diversos algoritmos para gerar números pseudo-aleatórios. Como não é objectivo desta disciplina estudá-los em detalhe, iremos apenas referir um dos métodos mais importante: o método congruencial (misto).

Definição 1 (Método Congruencial). *O método Congruencial é um dos métodos mais usados para a gerar números pseudo-aleatórios. Este método permite-nos gerar uma sequência de números inteiros no intervalo* $[0, m-1]$, *através da seguinte fórmula recursiva:*

$$x_0 = \text{“semente”}$$
$$x_i = (a\,x_{i-1} + c)(\operatorname{mod} m), \qquad i = 1, 2, \ldots$$

onde a, c *e* m *são números inteiros* ($a < m$, $c < m$) *e* $x(\operatorname{mod} m)$ *representa o resto da divisão inteira de* x *por* m. *O comprimento do ciclo é sempre menor ou igual a* m.

Exemplo 2 (Método Congruencial). *Considere o método Congruencial com $a = 5$, $c = 7$ e $m = 8$. Se escolhermos a semente $x_0 = 5$, os primeiros 20 números gerados são:*

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
x_i	0	7	2	1	4	3	6	5	0	7	2	1	4	3	6	5	0	7	2	1

Neste exemplo, o comprimento do ciclo é 8.

Observações:

- A conversão de números inteiros no intervalo $[0, m-1]$, para valores no intervalo contínuo $]0, 1[$ é feita através da transformação:

$$u_i = \frac{x_i + 1/2}{m}, \qquad i = 1, 2, \ldots$$

- Para gerar NPA's com qualidade aceitável, é necessário escolher convenientemente as constantes a, c e m do método Congruencial. Para obtermos um ciclo de comprimento m é necessário impor algumas condições aos valores das constantes a, c e m. Por exemplo, o comprimento do ciclo é igual a m (máximo valor possível), e independente da semente x_0, se $a = 4c + 1$, com $c \in \mathbb{N}$, c é ímpar e $m = 2^k$, com $k \in \mathbb{N}$.

Tabela 1: Valor das constantes a, c e m de algumas implementações do método Congruencial

	m	a	c
Turbo Pascal 3	2^{32}	129	907633385
Borland Delphi, Virtual Pascal	2^{32}	134775813	1
Numerical Recipes in C	2^{32}	1664525	1013904223
Borland C/C++	2^{32}	22695477	1
ANSI C: Watcom, Digital Mars, CodeWarrior	2^{32}	1103515245	12345
Microsoft Visual/Quick C/C++	2^{32}	214013	2531011

3 Geração de números pseudo-aleatórios de uma qualquer distribuição

3.1 Método da Transformação Inversa

O Teorema da transformação uniformizante assegura que é possível gerar números de uma qualquer distribuição absolutamente contínua, a partir da números aleatórios da distribuição uniforme no intervalo $]0,1[$. Isto é, se X é uma v.a. contínua com distribuição F, o teorema assegura-nos que $F(X) \sim U(0,1)$. Logo, para obter uma observação de X, basta calcular $F^{-1}(u)$ onde u é um número do intervalo $]0,1[$. A resolução da equação $x = F^{-1}(u)$ pode ser feita analiticamente, ou através de métodos numéricos (quando não existe expressão analítica para F^{-1}).

Exemplo 3. *Suponha que estamos interessados em gerar NPA's da distribuição Exponencial, com função de distribuição* $F(x) = 1 - e^{-\lambda x}, \;\; x > 0$. *É fácil de verificar que*

$$F(x) = u \quad \Leftrightarrow \quad x = -\frac{1}{\lambda}\ln(1-u),$$

ou seja, $F^{-1}(u) = -\frac{1}{\lambda}\ln(1-u)$. *Logo para gerar uma sequência de* n *NPA's da distribuição Exponencial, é necessário calcular*

$$x_i = -\frac{1}{\lambda}\ln(1-u_i), \qquad i = 1,2,\ldots,n,$$

onde u_1, u_2,...,u_n *representa uma sequência de NPA's uniformes no intervalo* $]0,1[$.

Observação: Se $X \sim U(0,1)$, então também $1 - X \sim U(0,1)$. Assim, os n NPA's da distribuição Exponencial, também podem ser obtidos a partir da equação:

$$x_i = -\frac{1}{\lambda}\ln(u_i), \qquad i = 1,2,\ldots,n.$$

Exemplo 4. *Suponha que estamos interessados em gerar NPA's da distribuição Triangular, com função densidade:*

$$f(x) = \begin{cases} x, & 0 < x \leq 1 \\ 2 - x, & 1 < x < 2 \\ 0, & \textit{caso contrário} \end{cases}$$

Para utilizar o método da transformação inversa, necessitamos de obter a função de distribuição desta variável aleatória. Assim é fácil de verificar que,

$$F(x) = \begin{cases} 0, & x \le 0 \\ \frac{x^2}{2}, & 0 < x \le 1 \\ 1 - \frac{(2-x)^2}{2}, & 1 < x < 2 \\ 1, & x \ge 2 \end{cases}$$

É fácil de verificar que

$$F(x) = u \quad \Leftrightarrow \quad x = \begin{cases} \sqrt{2u}, & 0 < u \le \frac{1}{2} \\ 2 - \sqrt{2(1-u)}, & \frac{1}{2} < u < 1 \end{cases}$$

Logo para gerar uma sequência de n NPA's da distribuição Triangular, é necessário calcular

$$x_i = \begin{cases} \sqrt{2u_i}, & 0 < u_i \le \frac{1}{2} \\ 2 - \sqrt{2(1-u_i)}, & \frac{1}{2} < u_i < 1 \end{cases}, \qquad i = 1, 2, \ldots, n,$$

onde u_1, u_2,. . . ,u_n representa uma sequência de NPA's uniformes no intervalo $]0,1[$.

3.2 Método da Transformação Inversa para distribuições discretas

Seja X uma variável aleatória discreta com função de probabilidade,

$$p_j = P(X = x_j), \quad j = 1, 2, \ldots, k \qquad \text{com} \qquad \sum p_j = 1.$$

Então a sua função de distribuição é dada por:

$$F(x) = \begin{cases} 0, & x < x_1 \\ p_1, & x_1 \le x < x_2 \\ p_1 + p_2, & x_2 \le x < x_3 \\ \vdots & \\ 1, & x \ge x_k \end{cases}$$

Para gerar um NPA w desta distribuição, a partir de um valor u uniforme, basta escolher:

$$w = \begin{cases} x_1, & 0 < u \le p_1 \\ x_2, & p_1 < u \le p_1 + p_2 \\ x_3, & p_1 + p_2 < u \le p_1 + p_2 + p_3 \\ \vdots & \\ x_k, & \sum_{i=1}^{k-1} p_i < u < 1 \end{cases}$$

Exemplo 5. *Considere a variável aleatória X com distribuição $Bin(2, 0.5)$. A função de distribuição desta v.a. é*

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{4}, & 0 \leq x < 1 \\ \frac{3}{4}, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

Para gerar um NPA w desta distribuição, a partir de um valor u uniforme, basta fazer:

$$w = \begin{cases} 0, & 0 < u \leq \frac{1}{4} \\ 1, & \frac{1}{4} < u \leq \frac{3}{4} \\ 2, & \frac{3}{4} < u < 1 \end{cases}$$

Por exemplo, se $u = 0.67411$, então o valor gerado para a distribuição $Bin(2, 0.5)$ é $w = 1$.

4 Exercícios

1. Uma das distribuições, usada no estudo do tempo de vida de componente mecânicos, é a distribuição $Weibull(\alpha, \beta)$, com função de distribuição:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\alpha x^{\beta}}, & x > 0 \end{cases}$$

onde $\alpha > 0$ e $\beta > 0$.

(a) Indique como gerar observações da distribuição $Weibull(3, 5)$, a partir de NPA's uniformes no intervalo]0,1[.

(b) Suponha que foram gerados os valores $u_1 = 0.2517$, $u_2 = 0.7437$ e $u_3 = 0.1911$ da $U(0, 1)$. Calcule os valores gerados da $Weibull(3, 5)$.

2. Considere a distribuição Cauchy (trata-se da distribuição t de Student com 1 grau de liberdade) com função densidade

$$f(x) = \frac{1}{\pi(1 + x^2)}, \quad x \in \mathbb{R}.$$

(a) Verifique que a sua função de distribuição é $F(x) = \frac{1}{\pi}\arctan(x) + \frac{1}{2}$, $x \in \mathbb{R}$.

(b) Indique como gerar observações da distribuição Cauchy, a partir de valores gerados para a distribuição $U(0, 1)$.

(c) Imagine que gerámos os NPA's $u_1 = 0.6235$ e $u_2 = 0.4515$ da $U(0, 1)$. Quais os valores gerados para a distribuição Cauchy?

3. (Teste de P.E. 2007/08) Considere a v.a. X com função densidade, $f(x) = 2x$, $0 < x < 1$. Sejam $u_1 = 0.2517$ e $u_2 = 0.7437$ dois números pseudo-aleatórios da distribuição $U(0,1)$. Usando o método da Transformação Inversa, calcule dois números pseudo-aleatórios da v.a. X. Observação: $P(X \leq x) = x^2, \quad 0 < x < 1$.

4. Considere a variável aleatória discreta com função de probabilidade dada por:

$$X \left\{ \begin{array}{cccc} 0 & 1 & 2 & 3 \\ 0.4 & 0.2 & 0.1 & 0.3 \end{array} \right.$$

   (a) Indique como gerar valores desta v.a., a partir de valores gerados da distribuição $U(0,1)$.

   (b) Considere a sequência de valores gerados da $U(0,1)$:

$$0.9982, \quad 0.9373, \quad 0.1916, \quad 0.5737, \quad 0.9726, \quad 0.2889, \quad 0.4082.$$

   Quais os valores gerados para X?

Solução dos exercícios

1. (a) Seja u um NPA da distribuição $U(0,1)$. Então, $x = \sqrt[5]{-\frac{\ln(1-u)}{3}}$ é um NPA da distribuição Weibull(3,5).

   (b) Os NPA gerados são: $x_1 = 0.6267$, $x_2 = 0.8538$ e $x_3 = 0.5887$.

2. (a)

   (b) Seja u um NPA da distribuição $U(0,1)$. Então, $x = \tan((u - 1/2)\pi)$ é um NPA da distribuição Cauchy.

   (c) Os NPA gerados são: $u_1 = 0.4087$ e $u_2 = -0.1536$.

3. $x_1 = \sqrt{0.2517} = 0.5017 \qquad x_2 = \sqrt{0.7437} = 0.8624$

4. $x_1 = 3, \quad x_2 = 3, \quad x_3 = 0, \quad x_4 = 1, \quad x_5 = 3, \quad x_6 = 0, \quad x_7 = 1$