PGMAT0061/MATG17 - Técnicas Computacionais em Estatística "Simulação Estocástica"

Professor responsável: Paulo Henrique Ferreira da Silva

Universidade Federal da Bahia
Instituto de Matemática e Estatística
Departamento de Estatística

12 de Novembro de 2024

Introdução

- "Simular": significa imitar, reproduzir a realidade;
- Justificativa: muitas vezes, o fenômeno de interesse é muito caro ou demanda um tempo muito grande.
- Aplicação (Estatística): estudar as propriedades de novos modelos/estimadores.
- Instrumento poderoso!!! $\Longrightarrow$ porém, por definição, impreciso

Introdução

- O computador digital não pode gerar números "aleatórios". Os geradores usados são determinísticos.
- von Neumann (1951):

> *"Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin."*

John von Neumann nos anos 40.

Introdução

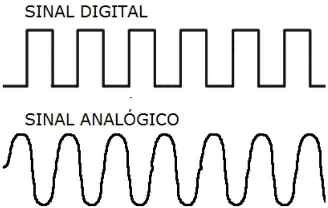
- Geração de números "verdadeiramente aleatórios": é feita a partir de processos físicos, isto é, tem por base fenômenos como:
  - ✓ ruído térmico;
  - ✓ efeito fotoelétrico;
  - ✓ decaimento radioativo (via contador Geiger);

  - ✓ outro fenômeno quântico.

- Precisa-se de um amplificador para trazer o resultado do processo físico para o nível macroscópico, e de um transdutor para converter este resultado em sinais digitais.

Geradores para Distribuição $U(0, 1)$

- Números "pseudo-aleatórios"
- Sinal digital $\times$ analógico:



Geradores para Distribuição $U(0, 1)$

- Alternativa para gerar números aleatórios: pegar um número (ou sequência) em um espaço de tempo.
- Antigamente, usava-se tabelas de números aleatórios, comumente apresentadas em livros de Estatística.

TABLE 2.1: Part of a Table of Random Numbers

Line/Col.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	10480	15011	01536	02011	81647	91646	69179	14194
2	22368	46573	25595	85393	30995	89198	27982	53402
3	24130	48360	22527	97265	76393	64809	15179	24830
4	42167	93093	06243	61680	07856	16376	39440	53537
5	37570	39975	81837	16656	06121	91782	60468	81305
6	77921	06907	11008	42751	27756	53498	18602	70659

Source: Abridged from William H. Beyer, ed., *Handbook of Tables for Probability and Statistics*, 2nd ed., © The Chemical Rubber Co., 1968. Used by permission of the Chemical Rubber Co.

Geradores para Distribuição $U(0, 1)$

Dois métodos <u>básicos</u> de geração de $U(0, 1)$:

- ✓ Congruencial;
- ✓ *Feedback shift register*.

Método Congruencial

- Lehmer (1951)
- São os mais conhecidos e utilizados.
- Método congruencial linear misto: é definido pela recursão:

$$x_i = (a\,x_{i-1} + c) \bmod m, \qquad i \geq 1$$

em que:

✓ a, c e m : são parâmetros (fixos);

✓ x_0 : semente ou raiz.

- Note que: $0 \leq x_i \leq m - 1$, ou ainda, $0 \leq x_i < m$.

Método Congruencial

- **OBS.:** Não existem valores contínuos em uma máquina digital. Tudo que deriva de sinal digital é discreto.
- Neste caso, além do problema da geração discreta, pode existir periodicidade.
- **Exemplo:**

$$x_i\text{'s}: \ 157, 632, 1.050, 29.009, \ldots, \underline{157}$$

- Período: número de iterações que uma sequência leva para se repetir.

Método Congruencial

- Para obter números entre 0 e 1, basta fazer: $u_i = x_i / m$.
- São desejados períodos longos para minimizar "buracos".
- Escolha dos parâmetros:
  - ✓ m : próximo ao máximo da máquina (inteiro);
  - ✓ c : não é relevante para periodicidade da sequência;
  - ✓ a : a seguir.
- **Exemplos:** $m = 2^{31} - 1$ e $a = 7^5 = 16.807$ (máquina de 32 bits)
  $m = 2^{35} - 31$ e $a = 5^5 = 3.125$ (máquina de 36 bits)

$\Longrightarrow$ propriedades desejáveis

Método Congruencial

- **Exemplo:** Se $x_0 = 5$ e $x_i = 3\,x_{i-1} \mod 150$, encontre $x_1, x_2, \ldots, x_{10}$.

Método Congruencial

- **Exemplo:** Se $x_0 = 5$ e $x_i = 3\,x_{i-1}$ mod 150, encontre $x_1, x_2, \ldots, x_{10}$.

i	x_i	$u_i = x_i/150$
1	15	0,1
2	45	0,3
3	135	0,9
4	$\underline{105}$	0,7
5	15	0,1
6	45	0,3
7	135	0,9
8	$\underline{105}$	0,7
9	15	0,1
10	45	0,3

$\Longrightarrow$ período $= 4$

Método Congruencial

- **Exemplo:** Se $x_0 = 5$ e $x_i = 3\,x_{i-1} \bmod 150$, encontre $x_1, x_2, \ldots, x_{10}$.

No *software* R:

```
> a = 3
> m = 150
> n = 11        # tamanho da amostra + 1
> x = numeric(n)
> x[1] = 5      # semente ou raiz
> for(i in 1:(n-1)){
>       x[i+1] = (a*x[i])%%m
> }
> x2 = x[-1]
> u = x2/m
> cbind(x2,u)
```

Método Congruencial

- **Exercício:** Se $x_0 = 3$ e $x_i = (5\,x_{i-1} + 7) \mod 200$, encontre $x_1, x_2, \ldots, x_{10}$.

Método Congruencial

- **OBS.:** a é dita "raiz primitiva" (período observado é o maior possível) se:

$$a^i \bmod m \neq 1, \qquad i = 1, 2, \ldots, m-2$$

- **Exemplo:** Considere $x_0 = 19$ e $x_i = 3\,x_{i-1} \bmod 31$.

i	x_i
1	26
2	16
3	17
$\vdots$	$\vdots$
29	27
30	<u>19</u>
31	26
$\vdots$	$\vdots$

$\Longrightarrow$ o período é 31-1 $= 30$ (maior possível, já que 3 é uma raiz primitiva de 31).