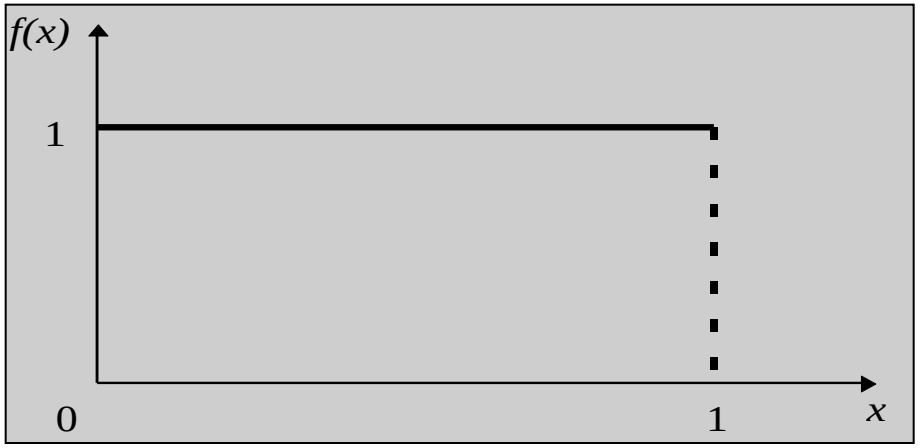
GERAÇÃO DE NÚMEROS ALEATÓRIOS

Propriedades dos Números Aleatórios

- Uma seqüência de números aleatórios, x_1, x_2,..., deve possuir duas importantes propriedades: uniformidade e independência.
- Todo número aleatório x_i é uma amostra independente de uma distribuição uniforme e contínua no intervalo de zero a 1.
- Sua função densidade de probabilidade de x é dada por:

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{outro valor} \end{cases}$$



Métodos Geradores de Números Aleatórios

- A técnica empregada mais comum faz uso de uma relação recursiva na qual, o próximo número na seqüência é uma função do último ou dois últimos números gerados, isto é:

$$x_n = f(x_{n-1}, x_{n-2}, \ldots)$$

Exemplo

$$x_n = (5x_{n-1} + 1) \bmod 16$$

- Para dar início ao processo de geração é preciso definir um valor inicial para $x_{o.}$
- Iniciando a série com $x_o = 5$ obtém-se x_1 da forma que segue:

$$x_1 = (5 \text{ x } 5 + 1) \bmod 16 = 26 \bmod 16 = 10$$

Os primeiros 32 números obtidos por meio deste procedimento são:

10, 3, 0, 1, 6, 15, 12, 13, 2, 11, 8, 9, 14, 7, 4, 5, 10, 3, 0, 1, 6, 15, 12, 13, 2, 11, 8, 9, 14, 7, 4, 5.

Observa-se que a série se repete após os primeiros 16 números.

Exemplo...

- Observa-se que os valores de *x* são inteiros entre 0 e 15. Dividindo-os por 16, teremos uma seqüência de números aleatórios com valores entre 0 e 1.

0,6250	0,1875	0,0000	0,0625	0,3750	0,9375	0,7500	0,8125
0,1250	0,6875	0,5000	0,5625	0,8750	0,4375	0,2500	0,3125
0,6250	0,1875	0,0000	0,0625	0,3750	0,9375	0,7500	0,8125
0,1250	0,6875	0,5000	0,5625	0,8750	0,4375	0,2500	0,3125

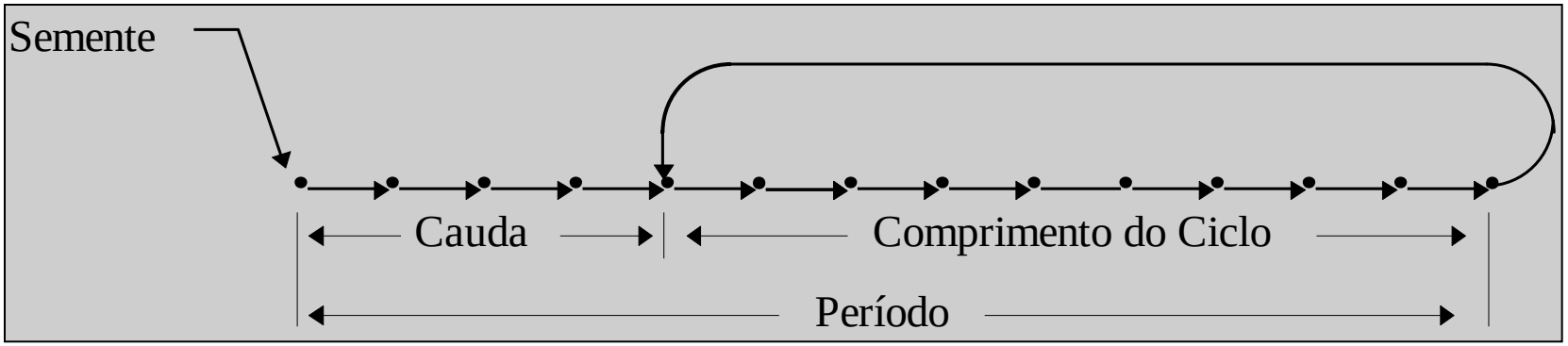
- Fica claro que, conhecida a função *f*, podemos gerar novamente a seqüência sempre que fornecermos o valor inicial de x_0. Este valor, usado para iniciar a seqüência é conhecido por ***semente***.

Observações

- A função *f* é **determinística**.
- Dada a semente, podemos afirmar, com 100% de certeza, qual serão os números na seqüência.
- Embora estes números sejam considerados aleatórios, no sentido de serem aprovados quando submetidos a testes estatísticos de aleatoriedade, são, de fato, *pseudo-aleatórios*.
- O objetivo em qualquer método de geração é produzir uma seqüência de números aleatórios entre zero e 1, os quais simulem ou imitem, as propriedades dos verdadeiros números aleatórios.

Observações

- Outra importante observação sobre o exemplo apresentado, é que somente os 16 primeiros valores são únicos.
- O 17º é igual ao primeiro e o restante da seqüência é apenas uma repetição cíclica dos primeiros 16 números.
- Dito de outra forma, o gerador utilizado possui um *comprimento de ciclo* igual a 16 valores.
- Alguns geradores não repetem uma parte inicial do ciclo, chamada de *cauda*. Neste caso, o comprimento de seu *período* é dado pela soma do comprimento *L* da cauda mais o comprimento *C* do ciclo.



Propriedades Desejadas

- As **propriedades desejadas** em um gerador de números aleatórios são as seguintes:
  - ***Deve ser computacionalmente eficiente**:* Uma vez que as simulações necessitam da geração de, até mesmo, milhares de números aleatórios em cada rodada, o tempo para processar cada geração deve ser mínimo;
  - ***O período deve ser muito longo**:* Um período curto pode fazer com que haja a reciclagem da seqüência de números aleatórios, resultando em uma repetição da seqüência de eventos. Isto pode, conseqüentemente, limitar o período utilizável de uma rodada de simulação.
  - ***Os sucessivos valores devem ser independentes e uniformemente distribuídos**:* A correlação entre os diversos valores gerados deve ser pequena. A correlação, se significativa, indica dependência.

Método Congruente

- Desenvolvido pelo Prof. D. H. Lehmer, em 1951, quando dos experimentos executados pelo computador ENIAC no MIT.
- Segundo ele, os restos de sucessivas potências de um número possuíam boas características de aleatoriedade.
- Obtinha o *n-ésimo* número de uma seqüência, tomando o resto da divisão da *n-ésima* potência de um inteiro ***a*** por um outro inteiro ***m***.

$$x_n = a^n \bmod m$$

Método Congruente Multiplicativo

- Uma expressão equivalente usada para o cálculo de x_n após calcular x_{n+1} é dada por:

$$x_n = a x_{n-1} \bmod m$$

- Os parâmetros a e m são chamados de ***multiplicador*** e ***módulo*** respectivamente.
- As escolhas de Lehmer para estes parâmetros foram $a = 23$ e $m = 10^8 + 1$.
- Segundo Jain (1991), tais valores foram baseados na facilidade de implementação no ENIAC, que era uma máquina de oito dígitos decimais.

Método Congruente Linear (MCL)

- Muitas das propostas atuais são generalizações da proposta de Lehmer e seguem a seguinte fórmula:

$$x_n = (ax_{n-1} + b) \bmod m$$

  - Os valores de x_n são inteiros entre 0 e m-1. As constantes a e b são positivas.
  - De maneira geral, a escolha dos valores de a, b, e m afetam o período e a autocorrelação na seqüência.

Exercício

$$x_n = (ax_{n-1} + b) \bmod m$$

- Use o MCL para gerar uma seqüência de números aleatórios entre zero e 1, com os seguintes parâmetros:
  - $x_0 = 27$, $a = 17$, $b = 43$ e $m = 100$.
- Qual o intervalo em que os valores são gerados? Porque?
- Observe também, que estarão sendo gerados inteiros aleatórios e não números aleatórios.
- Para transformá-los em valores entre 0 e 1, emprega-se:

$$R_i = X_i/m, \text{ para } i = 1, 2, ...$$

Resposta Exercício 1 $x_n = (ax_{n-1} + b) \bmod m$

- A seqüência de valores para x_i e subseqüentes R_i, é apresentada abaixo:

$x_0 = 27$

$x_1 = (17 \,.\, 27 + 43) \bmod 100 = 502 \bmod 100 = 2$

$R_1 = 2 / 100 = 0{,}02$

$x_2 = (17 \,.\, 2 + 43) \bmod 100 = 77 \bmod 100 = 77$

$R_2 = 77 / 100 = 0{,}77$

$x_3 = (17 \,.\, 77 + 43) \bmod 100 = 1352 \bmod 100 = 52$

$R_3 = 52 / 100 = 0{,}52$

Quais serão os próximos três valores da seqüência?

Resposta Exercício 1 $x_n = (ax_{n-1} + b) \bmod m$

- A seqüência de valores para x_i e subsequentes R_i, é apresentada abaixo:

$x_0 = 27$

$x_1 = (17 \,.\, 27 + 43) \bmod 100 = 502 \bmod 100 = 2$

$R_1 = 2 / 100 = 0{,}02$

$x_2 = (17 \,.\, 2 + 43) \bmod 100 = 77 \bmod 100 = 77$

$R_2 = 77 / 100 = 0{,}77$

$x_3 = (17 \,.\, 77 + 43) \bmod 100 = 1352 \bmod 100 = 52$

$R_3 = 52 / 100 = 0{,}52$

Quais serão os próximos três valores da seqüência?

0,27; 0,02; 0,77

Exercício 2: E se x_0 for igual a 13?

Resposta Exercício 2

Para x_0 igual a 13:

Xn-1	a	b	m	Xn
13	17	43	100	64
64				31
31				70
70				33
33				4
4				11
11				30
30				53
53				44
44				91
91				90
90				73
73				84
84				71
71				50
50				93
93				24
24				51
51				10
10				13
13				64
64				31
31				70
70				33
33				4
.				.

Relações entre *a*, *b*, *m* e X_0

- O módulo de *m* deve ser grande. Uma vez que os valores de *x* estarão entre 0 e *m* -1, o período nunca será maior do que *m*; Numa máquina de 32 bits, o valor máximo do período será: $2^{bit-1}-1 = 2^{31}-1 = 2.147.483.647$
- Para que a computação de mod *m* seja eficiente, *m* deve ser uma potência de 2, isto é, 2^k. Neste caso, o mod *m* poderá ser obtido truncando-se o resultado à direita por *k* bits.
- Se *b* for diferente de zero, o máximo período possível *m* é obtido se e somente se:
  - os inteiros *m* e *b* sejam primos, um em relação ao outro, isto é, não possuam nenhum outro fator além de 1;
  - todo número primo que é um fator de *m*, é também um fator de *a*-1;
  - *a*-1 é um múltiplo de 4, se o inteiro *m* é múltiplo de 4.
- Se $b = 0$, e *m* potência de 2, o maior período possível será $P = m / 4$, considerando que: *x*0 (semente) seja um número impar e o multiplicador (*a*) seja dado por $a = 8k + 3$ ou $a = 8k + 5$, para algum $k = 0, 1, 2, ...$

Práticas nas relações entre *a*, *b*, *m* e X_0

- Embora sem garantia completa, considerando apenas os MCL, pode-se dizer que é sempre possível obter um gerador de ciclo completo para qualquer ***m*** (melhor se m for um número primo) e qualquer semente x_0 se:
  - ***b*** terminar com os dígitos 1, 3, 7 ou 9
  - ***a*** terminar em 01, 21, 41 64 ou 81

- Realize um exercício no Excel considerando o MCL e diferentes valores ***m*** e x_0. Escolhendo ***a*** e ***b*** de acordo com as regras acima.