

Octave para Cálculo Numérico - 2020

© Gustavo C. Buscaglia

ICMC - Sala 4-219, Ramal 738176, gustavo.buscaglia@gmail.com

Simulações e variáveis aleatórias

- O Cálculo Numérico resolve **modelos da realidade** que são **complexos** demais para serem resolvidos **analiticamente**.
- A complexidade pode vir de **incertezas nos dados, alta dimensão das incógnitas, não-linearidades**, etc.
- Para isto são utilizadas **aproximações do modelo** e **algoritmos eficientes**. O sistema real é **simulado**, no sentido de determinar a solução do modelo aproximado.
- Nessas aulas veremos como utilizar Octave para lidar com o primeiro tipo de complexidade, na qual a resposta surge de uma modelo probabilístico da realidade.
- Revisaremos os conceitos de probabilidade necessários, discutiremos as funções de Octave mais relevantes e estudaremos exemplos básicos. Para mais informação, o livro Simulation, de S. Ross, é uma boa referência.

Atenção: Como já foi discutido, que o modelo numérico seja **consistente** com o modelo real (i.e., que a diferença entre eles seja pequena) não é suficiente, a aproximação deve ser estável (i.e., a diferença entre soluções de modelos parecidos deve ser pequena).

1. Muitas situações da vida profissional envolvem decisões a partir de eventos que tem variabilidade, tais como
 - a chegada de clientes a uma loja,
 - o conteúdo de açúcar da colheita de cana,
 - o tempo de ruptura de uma material sob uma certa tensão, etc.
2. Para tomar essas decisões racionalmente, é construído um modelo probabilístico. Isto é,
 - identifica-se a variável de interesse (que podem ser várias) e os possíveis valores que ela pode tomar, e
 - estima-se uma probabilidade para cada um daqueles valores.

3. Por exemplo, consideremos o comando `rand()`, que gera um número aleatório no intervalo $[0, 1]$. *Na verdade... não é tão assim. É uma função determinística que gera números representados em ponto flutuante. Mas não aprofundaremos no que a função `rand()` tem por trás.*
4. Nesse exemplo simples, uma decisão possível seria “se quando executar `rand()` o resultado for $\leq a$, molho as plantas, se não, não molho”.

Tem algum sentido esse exemplo? Imagine que um AirBnB cujas plantas devem ser em média molhadas a cada 4 dias. Como conseguir isto sem obrigar os inquilinos a levar um registro de quando foi a ultima irrigação?

Assim, o dono do apartamento pode pedir “execute `rand()` uma vez por dia e, por favor, se o resultado for $\leq a$ molhe as plantas”. Resulta claro que ele deve decidir o valor de a , e para fazer isto racionalmente deve conhecer, ou ao menos estimar, a probabilidade com o qual a função `rand()` gerará os diversos valores possíveis, em particular aquela com a qual aqueles valores serão $\leq a$.

5. A função `rand()` é um exemplo de **variável aleatória** (uma quantidade (número) de interesse cujo valor só é conhecido uma vez realizado o experimento).
6. A **função** `rand()` não é um **número**, mas depois de executarmos `a=rand()` a variável `a` passa a conter um número. Esse número é uma **realização** da variável aleatória `rand`.
7. Uma variável aleatória é uma função. Pode depender de alguns parâmetros (variáveis computacionais), por exemplo `function z=func(x,y)`. Mas mesmo fixando `x` e `y`, o resultado de fazer `a=func(x,y)` não é sempre o mesmo e não dá para prever com certeza.
8. No caso da função `rand()`, ela não depende de **nada**. O resultado de executar `a=rand(n,m)` é apenas gerar uma matriz $n \times m$ com realizações.
9. Seja $X = \text{rand}()$. X é uma variável **contínua**, porque toma qualquer valor no intervalo real $[0, 1]$. Atenção: No computador nada é contínuo, mas *quase*.

10. Sendo X uma variável aleatória, para todo $x \in \mathbb{R}$, existe uma função não decrescente

$$F_X(x) = \text{Prob}\{X \leq x\}$$

que é a distribuição cumulativa de X . É claro que $F(x) \rightarrow 0$ para $x \rightarrow -\infty$ e $F_X(x) \rightarrow 1$ para $x \rightarrow +\infty$.

Para variáveis aleatórias contínuas a função F_X é diferenciável, isto é, existe $f_X : \mathbb{R} \rightarrow [0, +\infty)$ tal que

$$F_X(x) = \int_{-\infty}^x f_X(y) dy ,$$

é dizer, $f_X = F'_X$ e $\int_{-\infty}^{+\infty} f_X(x) dx = 1$.

11. O **valor esperado** ou **média** de X é a média ponderada com f_X ,

$$\mathbb{E}(X) = \mu_X = \int_{-\infty}^{+\infty} x f_X(x) dx .$$

12. Na simulação numérica a estimação das probabilidades dos diversos valores da variável aleatória X é feita a partir da **Lei fraca dos grandes números**.

Teorema (lei fraca dos grandes números): Sejam X_1, X_2, \dots variáveis aleatórias independentes e identicamente distribuídas ($X_1 \sim X_2 \sim \dots \sim X$), cuja média é μ_X . Então, para qualquer $\epsilon > 0$,

$$\text{Prob} \left\{ \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mu_X \right| > \epsilon \right\} \rightarrow 0 \quad \text{quando } n \rightarrow \infty .$$

13. Em termos de computação, $X_1 \sim X_2 \sim \dots \sim X$ quer dizer que X_1, X_2, \dots , são calculadas **chamando a mesma função**. Por exemplo,

```
x=rand(1,n);   ou      for i=1:n x(i)=rand(); end
```

14. O teorema indica que a probabilidade de que uma variável M definida como média aritmética de n variáveis independentes $\sim X$ tome um valor diferente de μ_X tende a zero quando n tende a infinito. **Todas as realizações de M tem essencialmente o mesmo resultado, μ_X , se n é suficientemente grande.**

15. Assim, **com uma realização é suficiente!** (em princípio).

16. Em Octave podemos conferir que a lei se cumpre com a variável $X = \text{rand}()$ (cada coluna é o resultado de uma execução do comando)

```
mean(rand(10,1))
ans = 0.54559 0.47543 0.44478 0.41047
mean(rand(1000,1))
ans = 0.49399 0.49837 0.50859 0.50050
mean(rand(100000,1))
ans = 0.49896 0.49856 0.50039 0.49955
```

Quando n cresce, obtemos sempre essencialmente 0.5. Concluimos que a variável aleatória $X = \text{rand}()$ cumpre a lei dos grandes números, o que sugere que sucessivas chamadas a ela são independentes e que X tem média 0.5.

17. A partir de uma variável aleatória podemos definir outras por transformações (algébricas, etc.). Por exemplo, se $X = \text{rand}()$ então $Y = X^2$ também é uma variável aleatória. A continuação mostramos como utilizando transformações podemos estimar com simulações a cumulativa F_X .

Estimação de F_X e f_X

18. Em geral, a maneira de calcular a probabilidade de um evento é definir uma variável que vale 1 quando o evento acontece e 0 quando não. Vejamos a maneira de calcular $\text{Prob}\{X \leq a\}$.

Sendo X uma variável aleatória, definimos outra variável Y como

$$Y = \begin{cases} 1 & \text{se } X \leq a \\ 0 & \text{se não} \end{cases}$$

Y é uma variável **discreta**, que só pode tomar os valores 0 e 1.

19. Uma implementação possível é

```
function y=Y(a)
y=(rand()<=a);
end
```

Nesse caso a função depende de a , mas para um valor fixo de a o resultado não é determinístico, pode ser 0 ou 1.

20. Seja Y uma variável aleatória que só pode tomar k valores, a_1, a_2, \dots, a_k . Então a média ou valor esperado de Y é definido como

$$\mathbb{E}(Y) = \mu_Y = \sum_{i=1}^k a_i \text{Prob}\{Y = a_i\} .$$

21. O valor esperado de Y é

$$\mathbb{E}(Y) = 0 \cdot \text{Prob}\{Y = 0\} + 1 \cdot \text{Prob}\{Y = 1\} = \text{Prob}\{Y = 1\} = \text{Prob}\{X \leq a\} = \mathbf{F_X(a) !!}$$

22. Assim, podemos aproximar $F_X(a)$ como a média de n variáveis $\sim Y(a)$, para n suficientemente grande. Sendo $M = (Y_1 + Y_2 + \dots + Y_n)/n$ como antes, para n grande cada realização será uma boa aproximação de $F_X(a)$.

```
mean(rand(10,1)<0.3)
ans = 0.40000 0.30000 0.40000 0.40000
mean(rand(1000,1)<0.3)
ans = 0.31700 0.30900 0.28500 0.33000
mean(rand(100000,1)<0.3)
ans = 0.29973 0.30007 0.30154 0.29769
```

Concluimos que se $X = \text{rand}()$, então $F_X(0.3) \simeq 0.3$.

23. O procedimento utilizado acima é um **estudo de simulação**. Para determinar o valor de uma quantidade θ de um modelo probabilístico,

- se escolhe **uma variável aleatória X cujo valor esperado seja θ** ;
- se **simula o modelo** n vezes, obtendo valores x_1, x_2, \dots, x_n ;
- a **média amostral** $m = \bar{x} = \sum_{i=1}^n x_i/n$ é utilizada como estimação, ou aproximação, de θ ;
- a **variança amostral**

$$\sigma_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

orienta a escolha de n , tipicamente parando a geração de amostras quando σ_x/\sqrt{n} é menor que uma tolerância adequada (**teorema do limite central**).

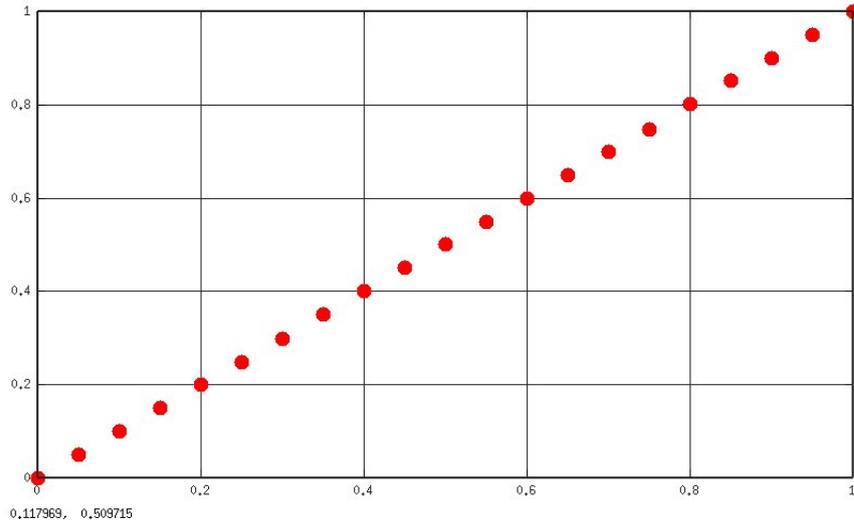
A parte essencial é a capacidade de gerar uma variável aleatória X com a distribuição indicada pelo modelo.

24. **Exercício:** Provar numericamente que, para a função `rand()`,

$$F_X(a) = \text{Prob}\{X \leq a\} = a .$$

Simplesmente repetimos o procedimento anterior para diversos a ,

```
n=1000; a=0.:0.05:1; m=length(a); x=rand(n,1);
for i=1:m
    y=(x<=a(i));
    F(i)=mean(y);erro(i)=var(y)/sqrt(n);
end
max(F-a),
plot(a,F,"ro-","markersize",8,"markerfacecolor","r")
hold on
plot(a,F+2*erro,"b-")
plot(a,F-2*erro,"b-")
plot(a,a,"k-","linewidth",2)
hold off
```



A máxima diferença entre $F_X(a)$ e a é $6.2e-04$, concluímos que a variável $X = \text{rand}()$ tem cumulativa $F_X(x) \simeq x$.

25. **Exercício:** Calcular uma aproximação de $F_Y(x)$ nos pontos $z_i = 0, 0.1, \dots, 3$, sendo a variável $Y = X_1 + X_2 + X_3$ e $X_1 \sim X_2 \sim X_3 \sim X = \text{rand}()$.

Ajuda: $x = \text{rand}(3, 1)$; $y = \text{sum}(x)$;

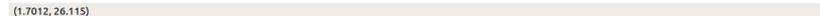
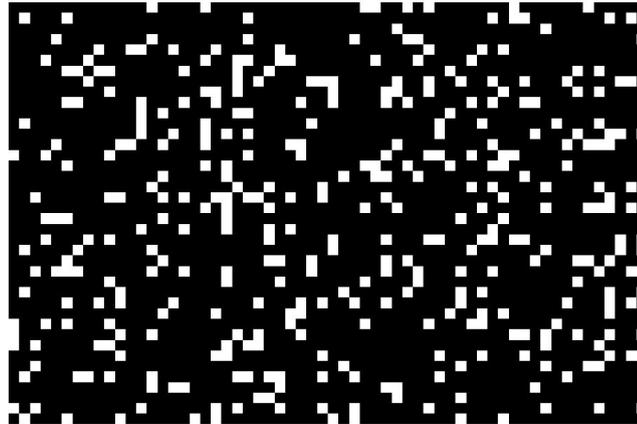
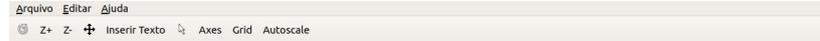
Gerar uma variável com distribuição de Bernoulli

26. Resulta claro que se desejamos simular uma variável estocástica Z que toma valor 1 com probabilidade p e 0 com probabilidade $1 - p$, podemos gerar uma amostra z fazendo

```
z=(rand()<p);
```

ou m amostras de n componentes

```
m=40; n=60; p=0.15; z=(rand(m,n)<p);  
imshow(z)
```



Amostra de 600×400 realizações da distribuição de Bernoulli com parâmetro $p = 0.15$.

Gerar uma variável com distribuição normal

27. Existem funções para gerar amostras de variáveis aleatórias com várias distribuições além da uniforme.
28. Uma de especial importância é a distribuição normal $N(\mu, \sigma)$ de média μ e desvio padrão σ . A função `randn()` gera uma instância de $N(0, 1)$.
29. Podemos obter uma variável com distribuição $N(\mu, \sigma)$ fazendo

```
x=mu+sigma*randn();
```

30. **Exercício:** Desenhar uma aproximação da função de densidade de probabilidade f_X da função amostrada com o comando do item anterior.

Exemplo: Uma fila de tempo discreto

31. Consideremos uma loja. Vamos simular ela em tempo discreto. Suponhamos que na a probabilidade de uma pessoa entrar na loja, cada minuto, é $p = 0.05$, e que a probabilidade de duas entrarem no mesmo minuto é zero. Construir uma função que gere a variável aleatória NC (número de clientes que entraram na loja em 300 minutos).

```
t=1; NC(1)=0; p=0.05;
for t=2:301
    NC(t)=NC(t-1);
    if (rand())<p)
        NC(t)=NC(t)+1;
    end
end
```

32. **Exercício:** Modifique a função anterior para uma loja tal que p vale 0.1 nos primeiros e últimos 60 minutos, e o resto do tempo é 0.05.

```
t=1; NC(1)=0; p0=0.05; p1=0.1;
for t=2:301
if ((t<=61) || (t>=242))
p=p1;
else
p=p0;
end
NC(t)=NC(t-1);
if (rand()<p)
NC(t)=NC(t)+1;
end
end
```

33. **Exercício:** Sabendo que cada cliente demora 20 minutos para ser atendido, calcule o número de clientes satisfeitos depois de 300 minutos.

```
t=1; NC(1)=0; p=0.05; NS(1)=0; tF=0;
for t=2:301
    NC(t)=NC(t-1);
    if (rand()<p)
        NC(t)=NC(t)+1;
    end
    NS(t)=NS(t-1);
    if (tF>0)
        tF=tF-1;
        if (tF==0)
            NS(t)=NS(t)+1;
        end
    end
    n(t)=NC(t)-NS(t);
    if ((n(t)>0)&&(tF==0))
        tF=20;
    end
end
```

34. Consideremos um ponto de serviço que modelaremos em tempo discreto, minuto a minuto. Cada minuto, a chegada de um cliente é aleatória com probabilidade de p de sucesso, e não chega mais de um cliente cada minuto. Existe um só atendente, se ele esteve ocupado os clientes esperam ser atendidos em fila. O tempo que leva satisfazer cada pedido é também aleatório. Os pedidos podem demorar 10, 20 ou 30 minutos, com igual probabilidade. A fila inicia vazia cada dia às 8 horas, e deixa de aceitar novos clientes às 16 horas. Estimar o número médio de clientes atendidos por dia, o tamanho médio da fila ao meio-dia, o tempo médio que o atendente está desocupado, o tempo médio que cada cliente passa no ponto. Estude a dependência dessas quantidades com p . Interpretar os resultados.

35. Como organizar a simulação?

- Variáveis: Tempo t , Número N_C de clientes chegados (até t), Número N_S de clientes satisfeitos (até t), n número de clientes na fila (a tempo t), t_D número de minutos que o atendente está desocupado (até t), t_F tempo até completar o pedido atual.

- Inicializar:

```
t=1; NC(1)=0; NS(1)=0; tD(1)=0; n(1)=0; p=0.2; tF=0;
```

- Avanço no tempo:

```
for t=2:600
    NC(t)=NC(t-1);
    if((rand()<p)&&(t<481)); ##novo cliente
        NC(t)=NC(t)+1;
    endif
    if (tF>0)
        tF=tF-1;
        if (tF == 0) ##finalizado
            NS(t)=NS(t-1)+1;
        end
    end
    n(t)=NC(t)-NS(t); #pedidos na fila
    if ((n(t)>0) & (tF==0)) #tempo necessario para novo pedido
        aux=rand();
        if (aux < 1/3) tF=10;
        elseif (aux < 2/3) tF=20;
        else tF=30;
    end
end
end
```

36. **Exercício:** Completar o código anterior e utilizar ele para calcular todas as quantidades de interesse do modelo, em particular o tempo médio desde a chegada do cliente até sair satisfeito. Estudar a dependência dessas quantidades com p , especialmente para valores próximos de 0.05.

37. **Exercício:** Um sistema requiere n máquinas para funcionar. Elas quebram, a cada unidade de tempo, com probabilidade p . Por isto, s máquinas adicionais são mantidas para reposição. Deseja-se estimar o tempo médio de “colapso” do sistema, isto é, o valor esperado do tempo T que o sistema funciona até dever ser parado porque não há n máquinas disponíveis. O tempo de reparo de cada máquina é t_R (fixo). Estudar a dependência de $\mathbb{E}(T)$ com s . Considere $n = 10$, $p = 0.01$, $t_R = 40$. Quantas máquinas adicionais são necessárias para que $\mathbb{E}(T)$ seja maior que 1000?

Nota: Uma variante usual adiciona que a oficina de reparo só pode reparar uma máquina por vez, aí o problema se complica um pouco porque tem uma fila de reparo parecida às de exercícios anteriores.

38. **Exercício:** Modificar o código anterior para calcular, para cada s , a probabilidade de $T < 1000$.

39. **Exercício:** Modificar o código anterior para o caso em que o tempo de reparo não é constante, mas pode tomar os valores 20, 40 e 60 com igual probabilidade?

40. **Exercício:** Para o problema do dono do AirBnB, simule o comportamento do sistema e escolha um valor de a tal que em média as plantas sejam irrigadas a cada 4 dias. Calcule também o valor de a para que a probabilidade de que as plantas passem mais de 10 dias sem irrigação seja menor que 5%.

Discussão sobre covariância, correlação, independência

41. Os testes anteriores não são suficientes para concluir que a função `rand` gera números adequadamente. É necessário examinar outras propriedades. Em particular, aquilo de que o resultado não dependa de **nada**.
42. Pelo método que está por trás de `rand()`, a principal suspeita é que possa existir dependência entre sucessivas realizações de X .

Se $x = \{x_i\}$ ($i = 1, \dots, N_x$) é uma amostra de X , e seja y uma subamostra de x definida como $y = \{x_i \in x, x_{i-1} < 0.3\}$ esperamos que:

- As médias amostrais de x e y sejam aproximadamente iguais, com valor 0.5, quando N_x é suficientemente grande

$$\bar{x} = \frac{1}{N_x} \sum_{i=1}^{N_x} x_i \simeq \frac{1}{2}, \bar{y} = \frac{1}{N_y} \sum_{i=1}^{N_y} y_i \simeq \frac{1}{2},$$

Octave: Funções `find`, `mean`, `var` e operações lógicas.

```
x=rand(1,10001);  
aux=(x < 0.3);  
ind=find(aux);  
y=x(ind+1);  
res=[mean(x) mean(y)]
```

0.49430 0.49526

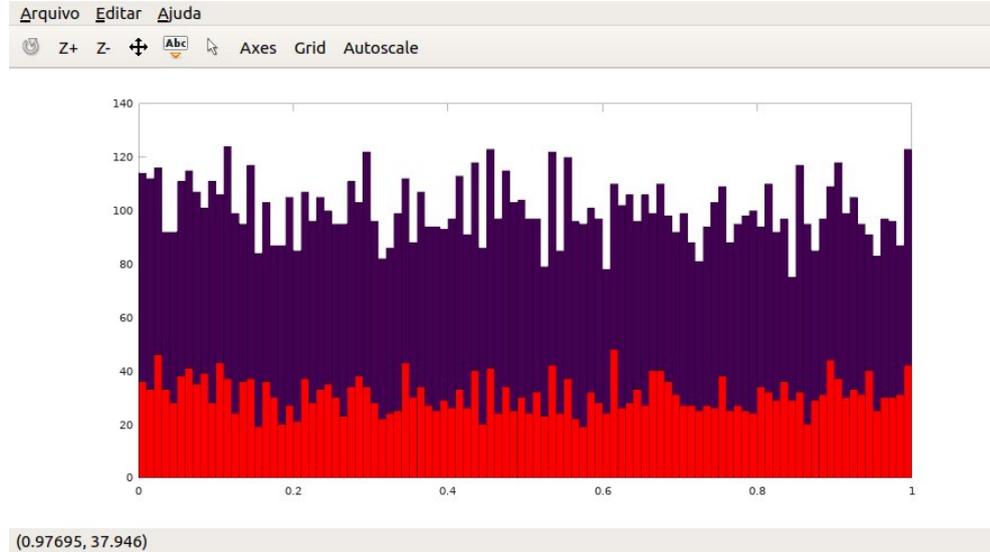
- As variâncias ($\text{Var}(x) = \sum_i (x_i - \bar{x})^2 / (N_x - 1)$) devem ambas valer aproximadamente $\text{Var}(X) = \int_0^1 (x - 1/2)^2 dx = 1/12$,

```
res=[var(x) var(y)]
```

0.083595 0.083150

- De fato, ambos histogramas devem ser parecidos (uniformes)

```
hist(x,100)  
hold on  
hist(y,100,"r")
```



- A **covariança** de duas variáveis estocásticas X e Y , denotada $\text{Cov}(X, Y)$, se define como

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] .$$

Se as variáveis são independentes $\text{Cov}(X, Y) = 0$.

O estimador de $\text{Cov}(X, Y)$, dadas amostras x e y das variáveis (com $N_x = N_y = N$) é implementado pela função `cov(x,y)`, sendo ele

$$\text{cov}(x, y) = \frac{1}{N-1} \sum_{i=1}^N (x_n - \bar{x})(y_n - \bar{y}) .$$

- A covariança entre valores sucessivos gerados por `rand()` deve portanto ser próxima de zero. Funções **cov** e **corr**.

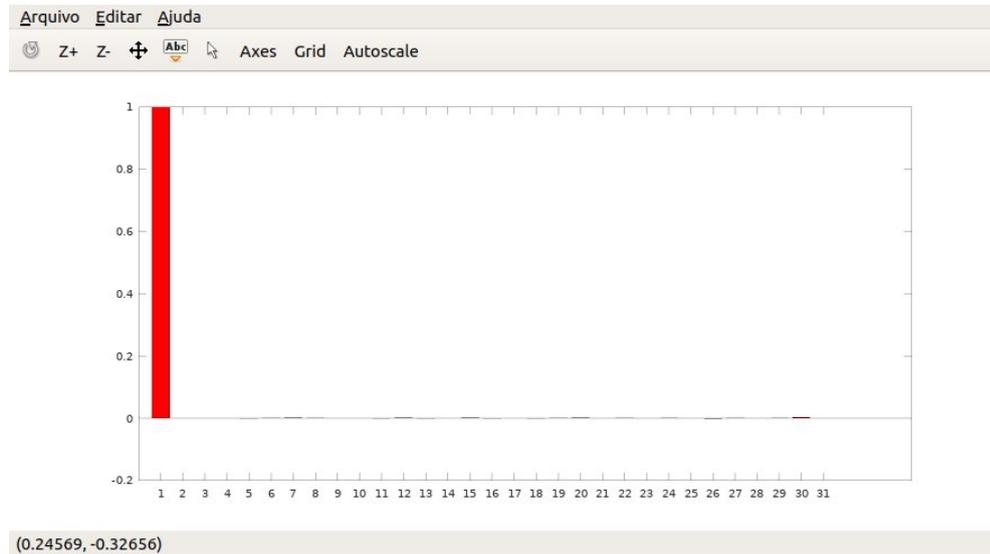
```
n=1000000; x=rand(1,n);
x0=x(1:n-2); x1=x(2:n-1); x2=x(3:n);
res = [cov(x0,x0) cov(x0,x1) cov(x0,x2)]
      8.3401e-02   5.2876e-05  -7.1730e-05
```

- O coeficiente de autocorrelação de x ,

$$R(k) = \frac{1}{(N_x - k - 1)\text{Var}(x)} \sum_{i=1}^{N_x-k} (x_i - \bar{x})(x_{i+k} - \bar{x})$$

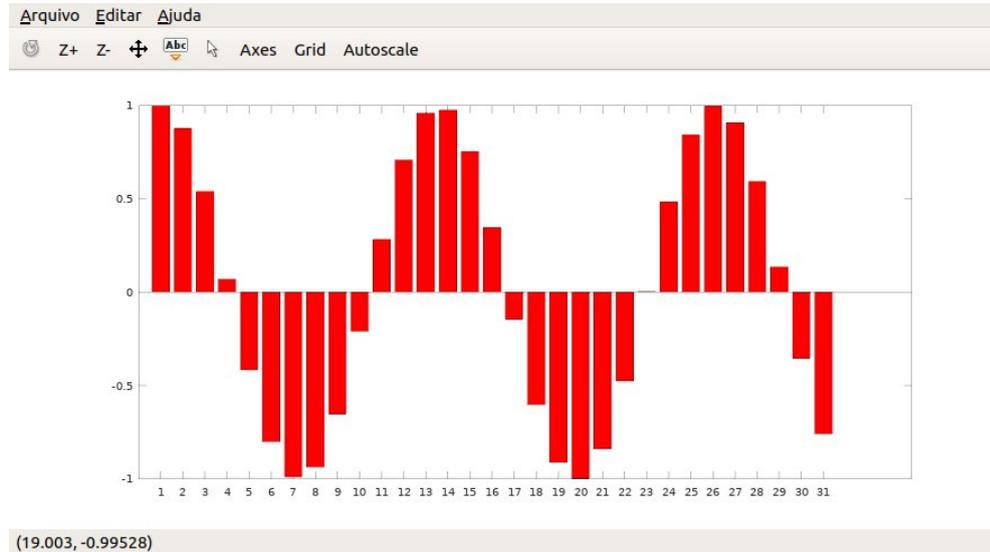
deveria ser aproximadamente $R(k) \simeq \delta_{1k}$.

```
n=1000000; x=rand(1,n); x=x-mean(x); vx=var(x);
for i=0:30 r(i+1)=x(1:n-i)*x(i+1:n)'/vx/(n-i-1); end
r(1:8)
    1.0    6.5e-04   -1.5e-03    1.0e-03    1.1e-03    1.3e-03   -2.0e-04    2.0e-03
bar([1:31],r,"r")
```



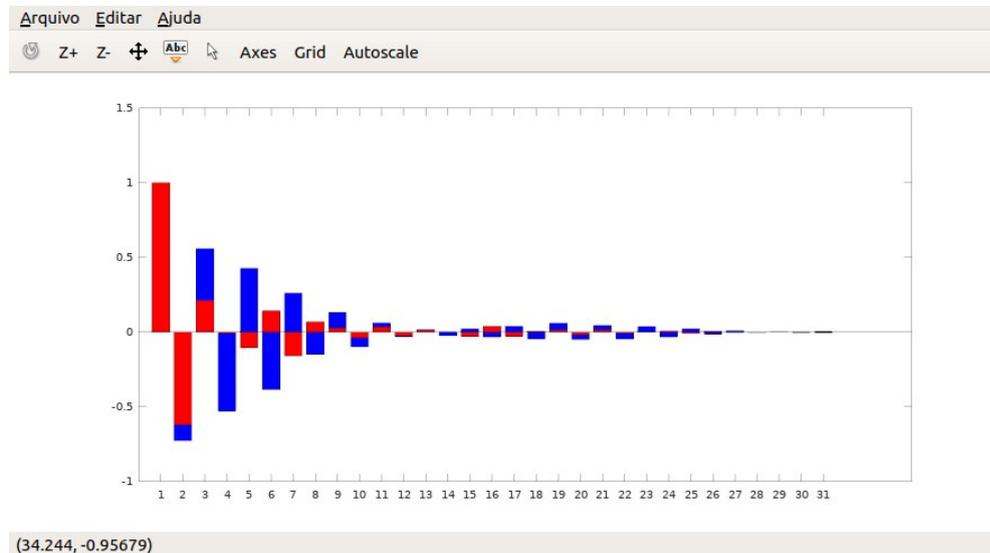
Podemos comparar com uma variável determinística

```
n=1000000; x=sin([1:n]*0.5); x=x-mean(x); vx=var(x);  
for i=0:30 r(i+1)=x(1:n-i)*x(i+1:n)'/vx/(n-i-1); end  
bar([1:31],r,"r")
```



Podemos comparar com uma variável caótica

```
n=1000000; a=3.8; x(1)=0.1;
for i=2:n x(i)=a*x(i-1)*(1-x(i-1)); end
x=x-mean(x); vx=var(x);
for i=0:30 r(i+1)=x(1:n-i)*x(i+1:n)'/vx/(n-i-1); end
bar([1:31],r,"r")
```



Coeficiente de autocorrelação R . Em azul $a = 3.7$, em vermelho $a = 3.8$.

43. Concluimos do anterior que `rand()` gera valores com distribuição uniforme, com independência entre sucessivos chamados.