Modelos para Séries Temporais

Ricardo Ehlers
ehlers@icmc.usp.br

Departamento de Matemática Aplicada e Estatística
Universidade de São Paulo

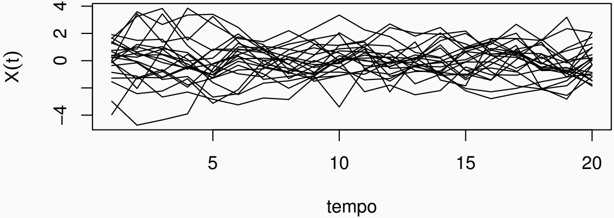
- Os modelos usados para descrever séries temporais são processos estocásticos.
- Neste caso assume-se que a série temporal (observada) é uma realização de um processo estocástico subjacente.

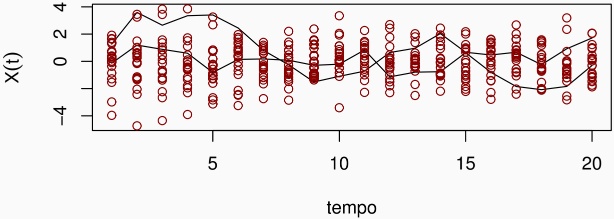
Um processo estocástico $\{X(t), t \in T\}$ é uma coleção de variáveis aleatórias.

- Para cada t, $X(t)$ é uma variável aleatória.
- Se $T \subseteq \mathbb{R}$ temos um *processo estocástico em tempo contínuo*, por exemplo, $\{X(t), t \geq 0\}$.
- Caso contrário, temos um *processo estocástico em tempo discreto*, por exemplo, $\{X_t, t = 1, 2, \ldots\}$.

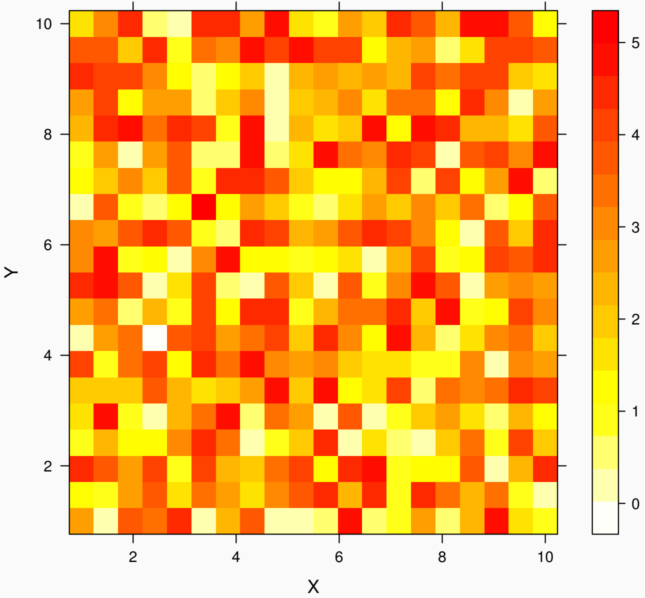
- O *espaço de estados* de um processo estocástico é o conjunto de todos os possíveis valores que as variáveis aleatórias $X(t)$ podem assumir.
- O espaço de estados que pode ser discreto (e.g. o número de chamadas que chegam a uma central telefônica a cada 2 horas) ou contínuo (e.g. a temperatura do ar em uma localidade observada em intervalos de 1 hora).

Um Processo Estocástico Simulado.





Processo estocástico simulado em 2 dimensões



Descrição de um processo estocástico

- Uma maneira de descrever um processo estocástico é através da distribuição de probabilidade conjunta de $X(t_1), \ldots, X(t_k)$ para qualquer conjunto de tempos $t_1, \ldots, t_k$ e qualquer valor de k.

- Esta é uma tarefa extremamente complicada e na prática costuma-se descrever um processo estocástico através das funções média, variância e autocovariância.

Estas funções são definidas a seguir para o caso contínuo sendo que definições similares se aplicam ao caso discreto.

- Função média:

$$\mu(t) = E[X(t)]$$

- Função variância:

$$\sigma^2(t) = Var[X(t)]$$

- Função de autocovariância:

$$\gamma(t_1, t_2) = E[X(t_1) - \mu(t_1)][X(t_2) - \mu(t_2)]$$

- Note que a função de variância é um caso especial da função de autocovariância quando $t_1 = t_2$.
- Momentos de ordem mais alta do processo também ser definidos mas são raramente utilizados na prática e as funções $\mu(t)$ e $\gamma(t_1, t_2)$ são em geral suficientes.

Definição

Um processo estocástico é dito ser estritamente estacionário se a distribuição de probabilidade conjunta de $X(t_1), \ldots, X(t_k)$ é a mesma de $X(t_1 + \tau), \ldots, X(t_k + \tau)$.

Ou seja, o deslocamento da origem dos tempos por uma quantidade τ não tem efeito na distribuição conjunta que portanto depende apenas dos intervalos entre $t_1, \ldots, t_k$.

- Para $k = 1$ a estacionariedade estrita implica que a distribuição de $X(t)$ é a mesma para todo t de modo que, se os dois primeiros momentos forem finitos,

$$\mu(t) = \mu \quad \text{e} \quad \sigma^2(t) = \sigma^2, \ \forall t.$$

- Para $k = 2$ a distribuição conjunta de $X(t_1)$ e $X(t_2)$ depende apenas da distância $t_2 - t_1$. A função de autocovariância também depende apenas de $t_2 - t_1$ e pode ser escrita como $\gamma(\tau)$ sendo,

$$\gamma(\tau) = E[X(t) - \mu][X(t+\tau) - \mu] = Cov[X(t), X(t+\tau)]$$

- $\gamma(\tau)$ é chamado de coeficiente de autocovariância na defasagem τ.

- Note que o tamanho de $\gamma(\tau)$ depende da escala de $X(t)$.
- Para efeito de interpretação, é mais útil padronizar a função de autocovariância dando origem a uma função de autocorrelação

$$\rho(\tau) = \gamma(\tau)/\gamma(0)$$

que mede a correlação entre $X(t)$ e $X(t+\tau)$.
- O argumento τ será discreto se a série temporal for discreta e contínuo se a série temporal for contínua.
- Na prática é muito difícil usar a definição de estacionariedade estrita e costuma-se definir estacionariedade de uma forma menos restrita.

Definição

Um processo estocástico $\{X(t), t \in T\}$ é dito ser estacionário de segunda ordem ou fracamente estacionário se a sua função média é constante e sua função de autocovariância depende apenas da defasagem,

$$\begin{aligned} E[X(t)] &= \mu, \ \forall t, \\ Cov[X(t), X(t+\tau)] &= \gamma(\tau), \forall t. \end{aligned}$$

- Nenhuma outra suposição é feita a respeito dos momentos de ordem mais alta.
- Fazendo $\tau = 0$ segue que $Var[X(t)] = \gamma(0)$, $\forall t$.
- Tanto a média quanto a variância precisam ser finitos.
- Esta definição mais fraca de estacionariedade será utilizada daqui em diante.

Definição

Um processo estocástico é dito ser um processo Gaussiano se, para qualquer conjunto $t_1, t_2, \ldots, t_n$ as variáveis aleatórias $X(t_1), X(t_2), \ldots, X(t_n)$ tem distribuição normal multivariada.

- A distribuição normal multivariada fica completamente caracterizada pelo primeiro e segundo momentos, $\mu(t)$ e $\gamma(t_1, t_2)$.
- Estacionariedade fraca implica em estacionariedade estrita para processos normais.
- Por outro lado, μ e $\gamma(\tau)$ podem não descrever adequadamente processos que sejam muito "não-normais".

Se um processo estocástico estacionário $X(t)$ tem média μ e variância σ^2 então

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)} = \frac{\gamma(\tau)}{\sigma^2}$$

e portanto $\rho(0) = 1$.

As seguintes propriedades são facilmente verificáveis.

- $\rho(\tau) = \rho(-\tau)$.
- $-1 < \rho(\tau) < 1$.

Embora um processo estocástico tenha uma estrutura de autocovariância única o contrário não é verdadeiro em geral. É possível encontrar vários processos com a mesma função de autocorrelação.

Alguns Processos Estocásticos

Alguns processos estocásticos são utilizados com frequência na especificação de modelos para séries temporais.

Definição

Um processo em tempo discreto é chamado *puramente aleatório* se consiste de uma sequência de variáveis aleatórias $\{\epsilon_t\}$ independentes e identicamente distribuidas.

Isto implica nas seguintes propriedades,

- $E(\epsilon_t) = E(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots) = \mu$
- $Var(\epsilon_t) = Var(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots) = \sigma^2_\epsilon$
- $\gamma(k) = Cov(\epsilon_t, \epsilon_{t+k}) = 0, \quad k = \pm 1, \pm 2, \dots.$

- Como a média e a função de autocovariância não dependem do tempo o processo é estacionário em segunda ordem.
- A função de autocorrelação é simplesmente

$$\rho(k) = \begin{cases} 1, & k = 0 \\ 0, & k = \pm 1, \pm 2, \ldots. \end{cases}$$

- Um processo puramente aleatório é as vezes chamado de *ruído branco* e pode ser útil por exemplo na construção de processos mais complicados.
- As propriedades acima podem ser entendidas como ausência de *correlação serial* e *homocedasticidade condicional* (variância condicional constante).

Passeio Aleatório

Definição

Seja $\{\epsilon_t\}$ um processo discreto puramente aleatório com média μ e variância σ_ϵ^2. Um processo $\{X_t\}$ é chamada de *passeio aleatório* se,

$$X_t = X_{t-1} + \epsilon_t.$$

Fazendo-se substituições sucessivas obtém-se que

$$\begin{aligned} X_t &= X_{t-2} + \epsilon_{t-1} + \epsilon_t \\ &= X_{t-3} + \epsilon_{t-2} + \epsilon_{t-1} + \epsilon_t \\ &\vdots \\ &= X_0 + \sum_{j=1}^{t} \epsilon_j \end{aligned}$$

Iniciando o processo em $X_0 = 0$ não é difícil verificar que,

$$\begin{aligned} E(X_t) &= \sum_{j=1}^{t} E(\epsilon_j) = t\mu \\ Var(X_t) &= \sum_{j=1}^{t} Var(\epsilon_j) = t\sigma_\epsilon^2. \end{aligned}$$

A função de autocovariância é dada por,

$$\begin{aligned} &Cov(X_t, X_{t-k}) = \\ &Cov(\epsilon_1 + \cdots + \epsilon_{t-k} + \epsilon_{t-k+1} + \cdots + \epsilon_t, \epsilon_1 + \cdots + \epsilon_{t-k}) = \\ &(t-k)\sigma_\epsilon^2 \end{aligned}$$

e a função de autocorrelação fica,

$$\rho_t(k) = \sqrt{\frac{t-k}{t}}.$$

- Como a média, a variância e as autocovariâncias dependem de t este processo é não estacionário.

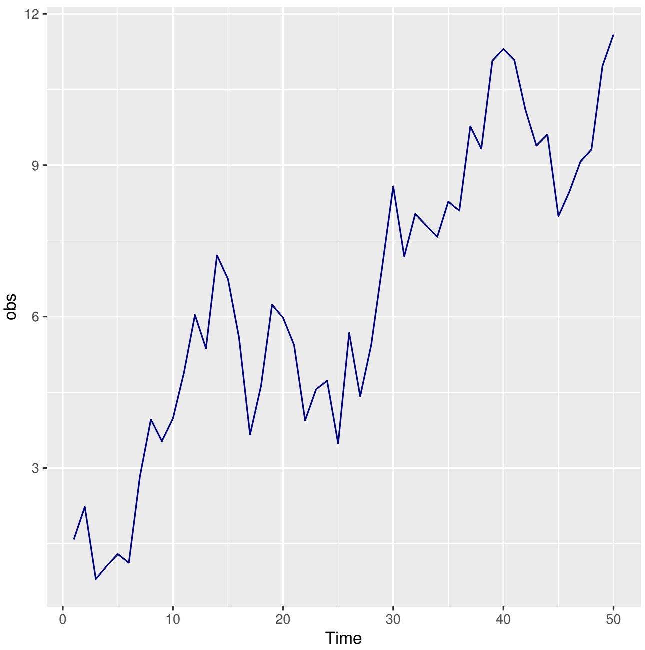
- A primeira diferença de um passeio aleatório é estacionária já que,

$$\nabla X_t = X_t - X_{t-1} = \epsilon_t.$$

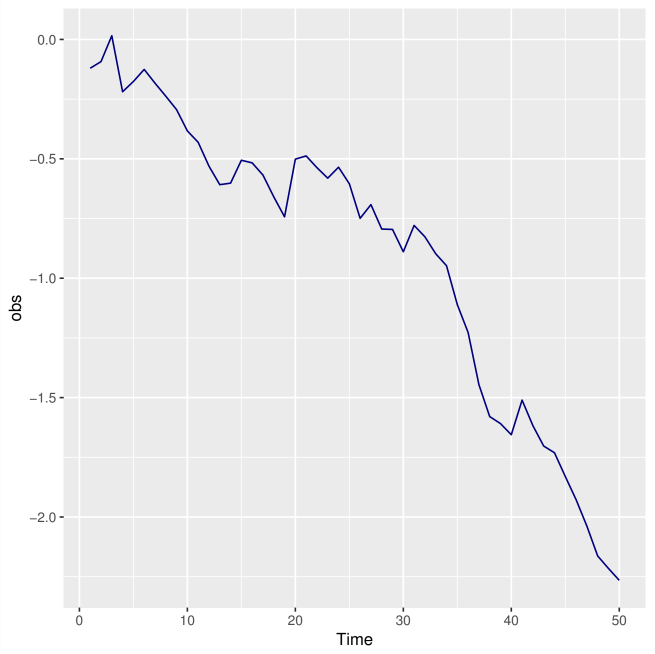
- O passeio aleatório também é conhecido como *tendência estocástica* pois uma realização do processo pode parecer uma tendência quando na verdade é uma soma de pequenas flutuações estocásticas.

- Os exemplos mais conhecidos de séries temporais que se comportam como um passeio aleatório são os preços de ações em dias sucessivos.

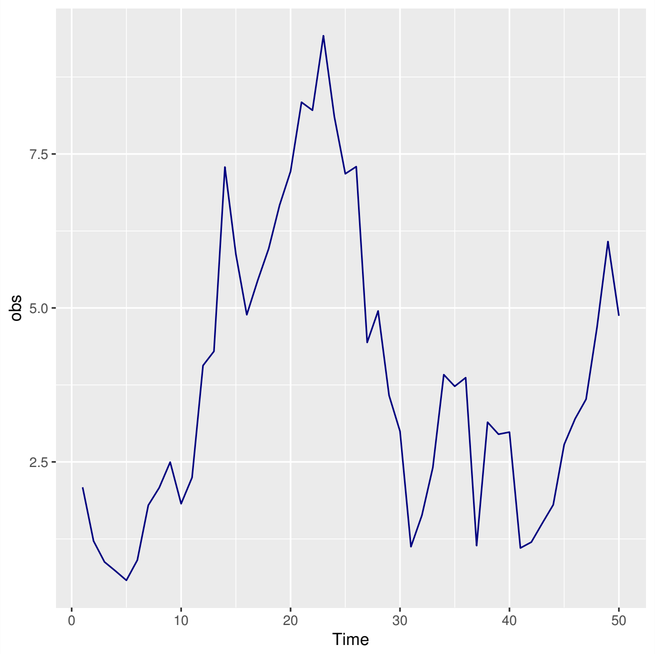
50 valores simulados de um passeio aleatório com $\sigma_\epsilon^2 = 1$



50 valores simulados de um passeio aleatório com $\sigma_\epsilon^2 = 0.1$



50 valores simulados de um passeio aleatório com $\sigma_\epsilon^2 = 1.5$



Passeio aleatório e retornos

- Suposição usual sobre preços de ativos financeiros,

$$P_t = \mu + P_{t-1} + \sigma\epsilon_t,\ \epsilon_t \sim N(0,1),$$

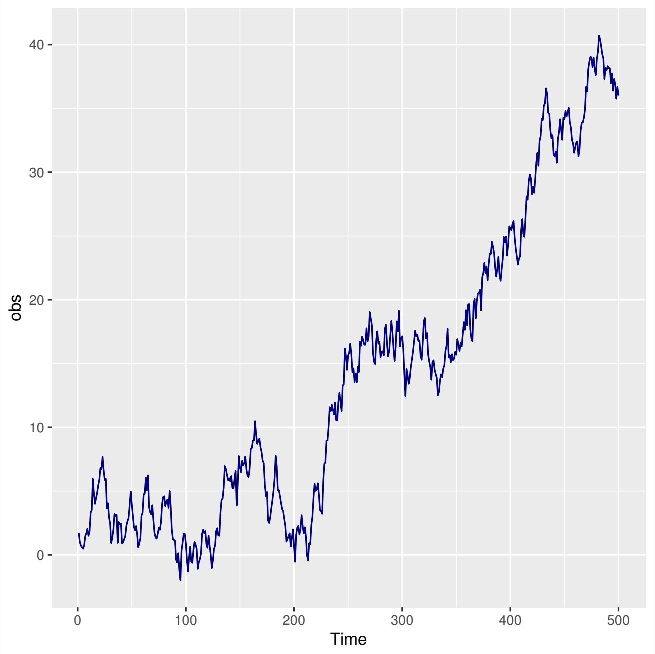
  sendo os ϵ_t independentes.
- Neste caso, $P_t|P_{t-1} \sim N(\mu + P_{t-1}, \sigma^2)$.
- Redefinindo em termos de log-retornos,

$$\log\left(\frac{P_t}{P_{t-1}}\right) = r_t = \mu + \sigma\epsilon_t,$$

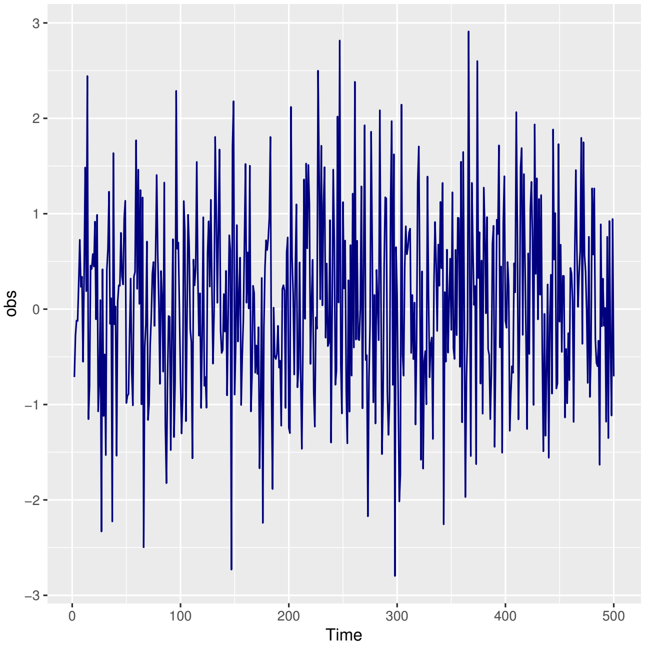
  i.e. $\log(P_t)$ segue um passeio aleatório.
- Uma suposição mais realista é,

$$r_t = \sigma_t\epsilon_t,\ \epsilon_t \sim N(0,1).$$

500 valores simulados de um passeio aleatório com $\sigma_\epsilon^2 = 1$ e $\mu = 0.005$



Primeira diferença dos valores simulados de um passeio aleatório com $\sigma_\epsilon^2 = 1$ e $\mu = 0.005$



Tipos de Modelos

Alguns tipos de modelos de séries temporais.

- Modelos paramétricos (número finito de parâmetros). A análise é feita no dominio do tempo.
  - modelos de erro (ou de regressão),
  - modelos ARMA, ARIMA, ARFIMA (memória longa),
  - modelos estruturais e não lineares.
- Modelos não paramétricos (número infinito de parâmetros). A análise é feita no dominio das frequências.
- Forma geral,

$$X_t = f(t) + \epsilon_t, \ t = 1, \ldots, n$$

sendo $f(t)$ o sinal e ϵ_t um ruido.

Modelos de regressão ou de erro

Suponha que o sinal $f(t)$ é uma função deterministica do tempo e ϵ_t é uma sequencia aleatória independente de $f(t)$. Além disso,

$$\begin{aligned} E(\epsilon_t) &= 0, \ \forall t \\ E(\epsilon_t^2) &= \sigma_\epsilon^2, \ \forall t \\ E(\epsilon_t \epsilon_s) &= 0, \ \forall t \neq s. \end{aligned}$$

A forma mais simples é,

$$X_t = \alpha + \beta t + \epsilon_t \tag{1}$$

onde α e β são constantes a serem estimadas.

O nível médio da série é $m_t = \alpha + \beta t$ que é algumas vezes chamado de *termo de tendência* ou *tendência global* (i.e. vale para toda a série), em oposição a *tendência local*.

De um modo geral, uma forma de se lidar com dados não sazonais que contenham uma tendência consiste em ajustar uma função polinomial,

$$X_t = \beta_0 + \beta_1 t + \cdots + \beta_p \; t^p + \epsilon_t.$$

- Uma função linear ou quadrática seria apropriada no caso de uma tendência monotonicamente crescente ou decrescente.
- Caso contrário polinômios de ordem mais alta devem ser ajustados.

Outras possíveis formas de tendência são crescimentos descritos por

- uma curva Gompertz,

$$\log x_t = a + br^t$$

  onde a, b e r são parâmetros com $0 < r < 1$, ou

- uma curva Logística,

$$x_t = a/(1 + be^{-ct})$$

  onde a, b e c são parâmetros.

Estas duas últimas são chamadas curvas S e se aproximam de uma assíntota quando $t \to \infty$. Neste caso o ajuste pode levar a equações não lineares.

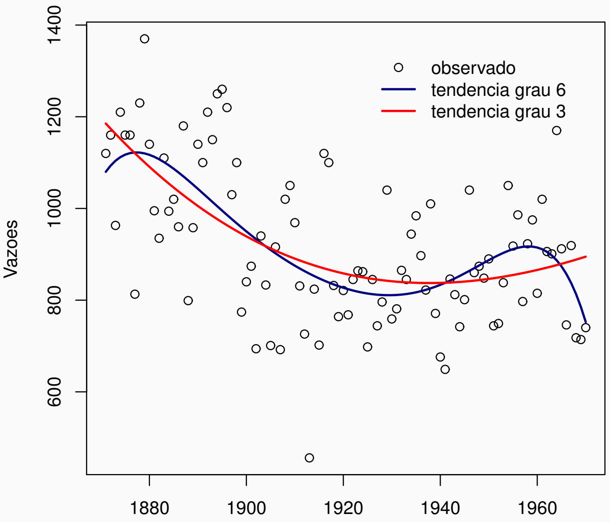
Seja qual for a curva utilizada, a função ajustada fornece uma medida da tendência da série, enquanto os resíduos (valores observados – valores ajustados) fornecem uma estimativa de flutuações locais.

Exemplo. Medições anuais de vazões do Rio Nilo em Ashwan entre 1871 e 1970.

```
> Nile

Time Series:
Start = 1871
End = 1970
Frequency = 1
  [1] 1120 1160  963 1210 1160 1160  813 1230 1370 1140  995
 [12]  935 1110  994 1020  960 1180  799  958 1140 1100 1210
 [23] 1150 1250 1260 1220 1030 1100  774  840  874  694  940
 [34]  833  701  916  692 1020 1050  969  831  726  456  824
 [45]  702 1120 1100  832  764  821  768  845  864  862  698
 [56]  845  744  796 1040  759  781  865  845  944  984  897
 [67]  822 1010  771  676  649  846  812  742  801 1040  860
 [78]  874  848  890  744  749  838 1050  918  986  797  923
 [89]  975  815 1020  906  901 1170  912  746  919  718  714
[100]  740
```

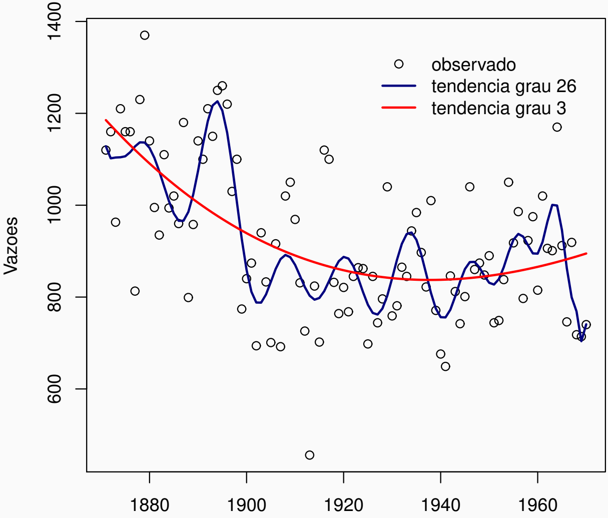
Polinômios de graus 3 e 6 foram ajustados via método de mínimos quadrados.

```
Call:
lm(formula = y ~ X)

Coefficients:
(Intercept)           X1           X2           X3
  8.571e+02   -2.369e+00    7.465e-02   -2.304e-04

Call:
lm(formula = y ~ X)

Coefficients:
(Intercept)           X1           X2           X3
  8.254e+02   -3.178e+00    1.601e-01    1.282e-03
         X4           X5           X6
  1.856e-05   -5.447e-07   -2.807e-08
```



observado
tendencia grau 6
tendencia grau 3
Vazoes
1400
1200
1000
800
600
1880
1900
1920
1940
1960

Não adianta aumentar muito o grau do polinômio.



Para lidar com flutuações cíclicas ou sazonais pode-se adotar um *polinômio harmônico*,

$$f(t) = \sum_{j=1}^{h} (\alpha_j \cos \lambda_j t + \beta_j \text{sen} \lambda_j t).$$

sendo $\lambda_j = 2\pi j/p$, p é o periodo de $f(t)$ e o valor máximo de h é $p/2$.

Um processo com tendência e flutuações cíclicas ou sazonais pode ser representado como,

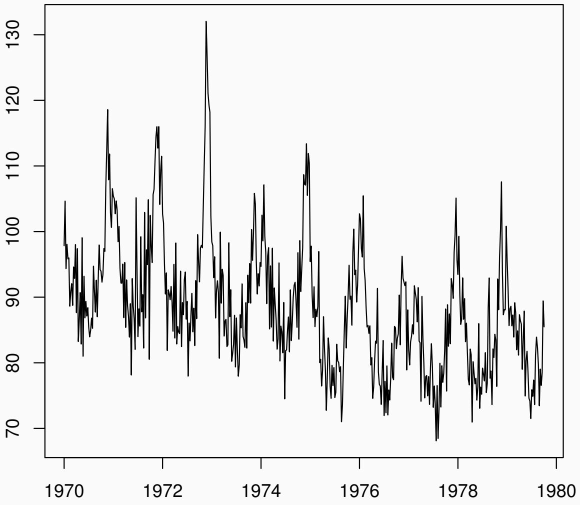
$$X_t = f(t) + \epsilon_t = T_t + S_t + \epsilon_t.$$

Exemplo. Dados semanais de mortalidade cardiovascular em Los Angeles (estudos de poluição). Dados obtidos em http://www.stat.pitt.edu/stoffer/tsa4 e também disponiveis no pacote `astsa` (*Applied Statistical Time Series Analysis*).

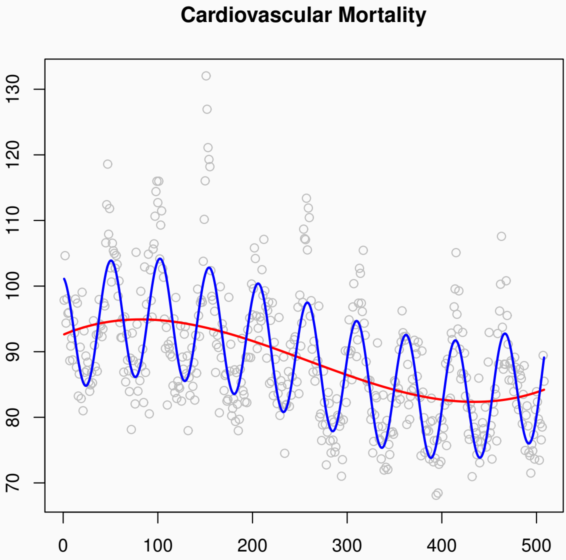
```
> library(astsa)
```

```
Start = c(1970, 1)
End = c(1979, 40)
Frequency = 52
  [1]  97.85 104.64  94.36  98.05  95.85  95.98  88.63
  [8]  90.85  92.06  88.75  94.60  92.86  98.02  87.64
 [15]  97.40  83.24  86.60  90.69  82.86  99.06  81.00
 [22]  93.18  86.86  89.35  87.13  88.39  85.38  83.96
 [29]  84.95  86.81  85.25  94.72  90.96  87.73  92.58
 [36]  87.02  92.68  97.95  94.19  93.95  92.29  93.48
 [43]  97.41  96.98 106.60 112.41 118.59 107.90 111.82
 [50] 102.78 100.63 106.54 105.39 104.99 102.72 104.67
 ...
```

Cardiovascular Mortality



Polinomio de grau 3 com 1 harmonico



```
> tempo = 1:length(cmort)
> tempo2 = tempo^2
> tempo3 = tempo^3
> c = cos(2*pi*tempo/52)
> s = sin(2*pi*tempo/52)
> lm(cmort~ tempo + tempo2 + tempo3)

Call:
lm(formula = cmort ~ tempo + tempo2 + tempo3)

Coefficients:
(Intercept)        tempo       tempo2       tempo3
  9.258e+01    6.095e-02   -4.428e-04    5.716e-07


> lm(cmort~ tempo + tempo2 + tempo3 + c + s)

Call:
lm(formula = cmort ~ tempo + tempo2 + tempo3 + c + s)

Coefficients:
(Intercept)        tempo       tempo2       tempo3
  9.242e+01    7.204e-02   -5.184e-04    6.943e-07
          c            s
  8.901e+00   -1.803e+00
```

Variáveis indicadoras (dummy)

Uma variável indicadora (ou *dummy*) é uma variável categórica que assume dois possíveis valores. Assumiremos que estes valores são 1, se um determinado evento ocorre, e zero caso contrário.

No contexto de séries temporais o uso de variáveis *dummy* pode ser útil se por exemplo,

- Deseja-se fazer previsões de vendas diárias levando em conta se o dia é um feriado.
- Levar em conta se uma observação é um *outlier* ao invés de simplesmente remover a observação.

Se houver mais de duas categorias pode-se criar várias variáveis *dummy*, sendo uma a menos do que o número de categorias.

Variáveis dummy sazonais

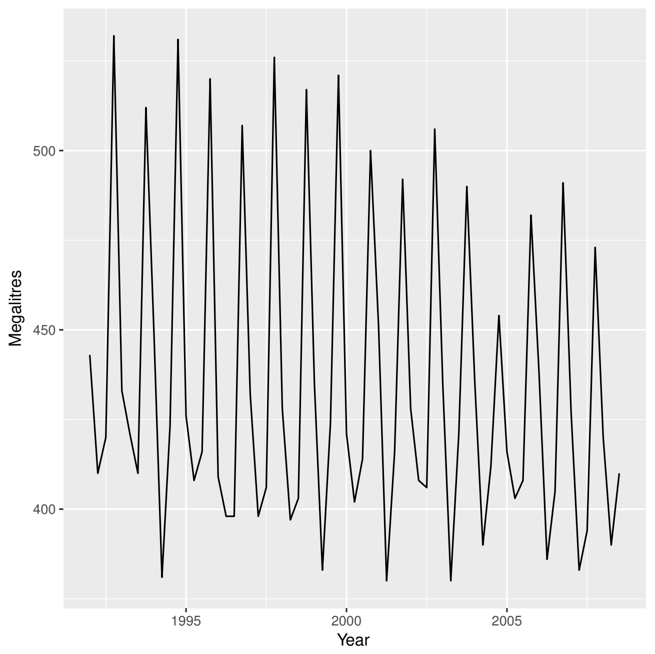
Suponha que uma série temporal tem periodicidade trimestral. As seguintes variáveis dummy podem ser criadas,

Ano	Trimestre	d1	d2	d3
1	1	1	0	0
	2	0	1	0
	3	0	0	1
	4	0	0	0
2	1	1	0	0
	2	0	1	0
	3	0	0	1
	4	0	0	0
⋮	⋮	⋮	⋮	⋮

- Note que foram criadas 3 variáveis indicadoras para representar 4 trimestres.
- O efeito do quarto trimestre nas previsões é capturado pelo intercepto da regressão.
- O coeficiente de cada variável dummy é interpretado como o efeito daquele trimestre em relação ao trimestre omitido.

Exemplo. Produção trimestral de cerveja (em megalitros) na Austrália. Fonte: *Australian Bureau of Statistics*.

	Qtr1	Qtr2	Qtr3	Qtr4
1992	443	410	420	532
1993	433	421	410	512
1994	449	381	423	531
1995	426	408	416	520
1996	409	398	398	507
1997	432	398	406	526
1998	428	397	403	517
1999	435	383	424	521
2000	421	402	414	500
2001	451	380	416	492
2002	428	408	406	506
2003	435	380	421	490
2004	435	390	412	454
2005	416	403	408	482
2006	438	386	405	491
2007	427	383	394	473
2008	420	390	410	



Megalitres
500
450
400
1995
2000
2005
Year

Os dados y_t serão modelados com uma tendência linear e variáveis dummy trimestrais,

$$y_t = \beta_0 + \beta_1 t + \beta_2 d_{2,t} + \beta_3 d_{3,t} + \beta_4 d_{4,t} + \epsilon_t$$

sendo

$$d_{i,t} = \begin{cases} 1, \text{se } t \text{ for o trimestre } i \text{ e} \\ 0, \text{caso contrário.} \end{cases}$$

A variável relativa ao primeiro trimestre foi omitida, portanto o coeficiente de $d_{i,t}$ mede a diferença entre o trimestre i e o primeiro trimestre.

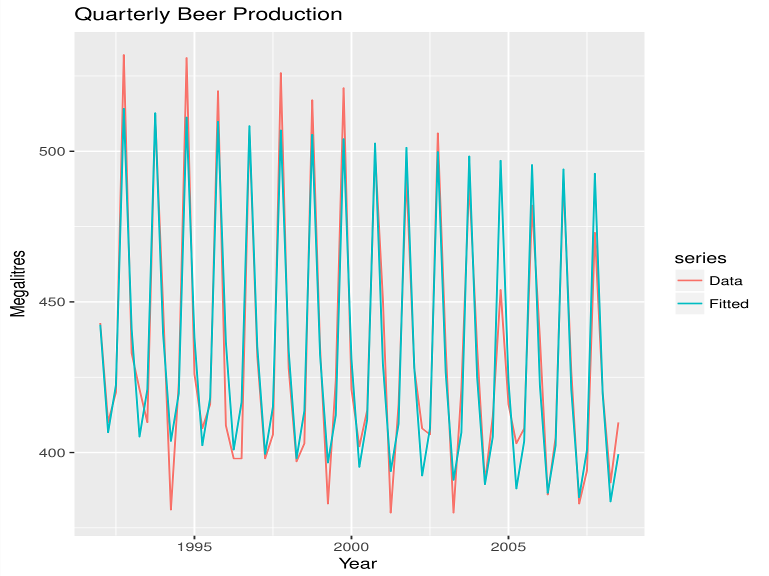
Regressão com tendência linear e dummies trimestrais.

```
Call:
tslm(formula = beer2 ~ trend + season)

Residuals:
    Min      1Q  Median      3Q     Max
-42.916  -7.877  -0.070   7.594  21.494

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 442.78341    3.98067 111.233  < 2e-16 ***
trend        -0.35886    0.07866  -4.562 2.45e-05 ***
season2     -35.40585    4.26869  -8.294 1.22e-11 ***
season3     -19.28229    4.27086  -4.515 2.89e-05 ***
season4      72.79268    4.33485  16.792  < 2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 12.44 on 62 degrees of freedom
Multiple R-squared:  0.923,       Adjusted R-squared:  0.918
F-statistic: 185.8 on 4 and 62 DF,  p-value: < 2.2e-16
```

- Foram definidas variáveis dummy para o segundo, terceiro e quarto trimestres.

- Os coeficientes estimados destas variáveis medem a diferença em relação ao primeiro trimestre.

- Por exemplo, o segundo trimestre tem uma produção em média 35.4 megalitros menor do que o primeiro trimestre.


Quarterly Beer Production
Megalitres
500
450
400
1995
2000
2005
Year
series
Data
Fitted

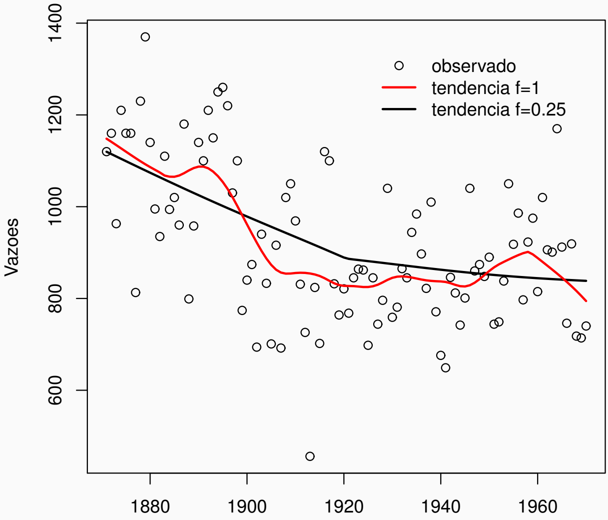
Regressão Local

- estimar para cada t uma equação de regressão polinomial diferente, por exemplo

$$\hat{x}_t = \hat{\alpha}(t) + \hat{\beta}(t)t.$$

- As estimativas de α e β dependem do tempo o que dá o caráter local das retas de regressão.
- `loess` (*Locally weighted regression scatter plot smoothing*): a cada passo aplica a regressão local anterior, calcula $x_t - \hat{x}_t$ e aplica novamente a regressão local dando peso menor às observações com resíduos maiores. Repete-se até atingir convergência.

Medições anuais de vazões do Rio Nilo em Ashwan entre 1871 e 1970 (pontos), tendência estimada via função `lowess`



Em certas aplicações, particularmente em séries de vendas de produtos, outras variáveis preditoras podem ser incluidas no modelo.

- Efeito do número de dias úteis (em cada mês ou trimestre).
- Variáveis com valores defasados, e.g. gasto com propaganda m meses atrás.
- Efeito de feriados móveis, que pode durar vários dias à frente.
- etc.

Diferenciação

Um tipo especial de filtro, muito útil para remover uma componente de tendência polinomial, consiste em *diferenciar* a série até que ela se torne *estacionária*.

- Para dados não sazonais, a primeira diferença é em geral suficiente para induzir estacionariedade aproximada. A nova série $y_2, \ldots, y_n$ é formada a partir da série original $x_1, \ldots, x_n$ como
$$y_t = x_t - x_{t-1} = \nabla x_t.$$
- Note que isto nada mais é do que um filtro (assimétrico) com coeficientes -1, 1 e 0.

Modelo com tendência linear,

$$\begin{aligned} X_t &= \alpha + \beta t + \epsilon_t \\ \nabla X_t &= \beta + \epsilon_t - \epsilon_{t-1} \end{aligned}$$

O processo ∇X_t é estacionário mas não é *inversivel*.

Modelo com tendência quadrática,

$$
\begin{aligned}
X_t &= \beta_0 + \beta_1 t + \beta_2 t^2 + \epsilon_t \\
\nabla X_t &= \beta_1 - \beta_2 + 2\beta_2 t + \epsilon_t - \epsilon_{t-1}
\end{aligned}
$$

O processo ∇X_t tem uma tendência linear e não é estacionário.

Modelos ARIMA

Definição

Processos lineares estacionários,

$$X_t - \mu = \psi_0 \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots$$

sendo $\psi_0 = 1$, $\mu = E(X_t)$ e $\psi_1, \psi_2, \dots$ tais que,

$$\sum_{k=0}^{\infty} \psi_k^2 < \infty.$$

Casos particulares,

- Processo autoregressivo de ordem p, AR(p).
- Processo de médias móveis de ordem q, MA(q).
- Processo autoregressivo e de médias móveis de ordem p e q, ARMA(p, q).

- Processos lineares não estacionários homogêneos. Supõe que a série é não estacionária em nível e/ou inclinação e pode se tornar estacionária por diferenciação.
- Processos de memória longa. São processos estacionários cuja função de autocorrelação decai muito lentamente (decaimento hiperbólico). É necessário tomar diferenças fracionárias.

Estes processos são denotados como ARIMA(p, d, q) e podem ser generalizados incluindo-se um operador sazonal.

Modelos Estruturais (espaço de estados)

Admitem variabilidade nas componentes de tendência e sazonalidade da decomposição clássica. Em sua forma mais simples,

$$\begin{aligned} X_t &= Q_t + \epsilon_t \\ Q_{t+1} &= Q_t + \nu_t \end{aligned}$$

sendo ϵ_t e ν_t ruidos brancos não correlacionados com $E(\epsilon_t) = E(\nu_t) = 0$, $Var(\epsilon_t) = \sigma^2_\epsilon$, $Var(\nu_t) = \sigma^2_\nu$, $\forall t$ e $Q_1 = q$.

Adicionando uma inclinação estocástica, temos o modelo de tendência linear local,

$$\begin{aligned} X_t &= Q_t + \epsilon_t \\ Q_{t+1} &= Q_t + \beta_t + \nu_t \\ \beta_{t+1} &= \beta_t + \eta_t \end{aligned}$$

sendo ϵ_t, ν_t e η_t ruidos brancos não correlacionados com $E(\epsilon_t) = E(\nu_t) = E(\eta_t) = 0$, $Var(\epsilon_t) = \sigma^2_\epsilon$, $Var(\nu_t) = \sigma^2_\nu$ e $Var(\eta_t) = \sigma^2_\eta$, $\forall t$.

Pode-se adicionar componentes sazonais deterministicas ou estocásticas.

Modelos não lineares

Alguns tipos de não linearidade,

- Modelos com mudanças estruturais. Modelos bilineares, modelos com limiar (Threshold, Markov switching, etc).
- Modelo com variância condicional evoluindo no tempo. Familia ARCH-GARCH e modelos de voltilidade estocástica.

Suavização

Quando $f(t)$ não pode ser aproximada por alguma função simples do tempo utiliza-se procedimentos não paramétricos de suavização (ou alisamento).

Definição

Um filtro linear converte uma série $\{x_t\}$ em outra $\{y_t\}$ através da seguinte operação linear

$$y_t = \sum_{j=-q}^{q} a_j x_{t+j}$$

sendo $\{a_j\}$ um conjunto de pesos.

- Como queremos estimar a média local os pesos devem ser tais que $\sum_{j=-q}^{q} a_j = 1$, garantindo assim que

$$\min\{x_t\} < y_t < \max\{x_t\}.$$

- Esta operação é chamada *média móvel de ordem q*.

- Por exemplo, se $q = 2$ temos que

$$y_t = a_2 x_{t-2} + a_1 x_{t-1} + a_0 x_t + a_1 x_{t+1} + a_2 x_{t+2}.$$

- O caso mais simples é quando todos os pesos tem o mesmo valor,

$$a_j = 1/(2q+1),\ j = -q, \ldots, q.$$

Neste caso, o valor suavizado de x_t é dado por

$$y_t = \frac{1}{2q+1} \sum_{j=-q}^{q} x_{t+j}.$$

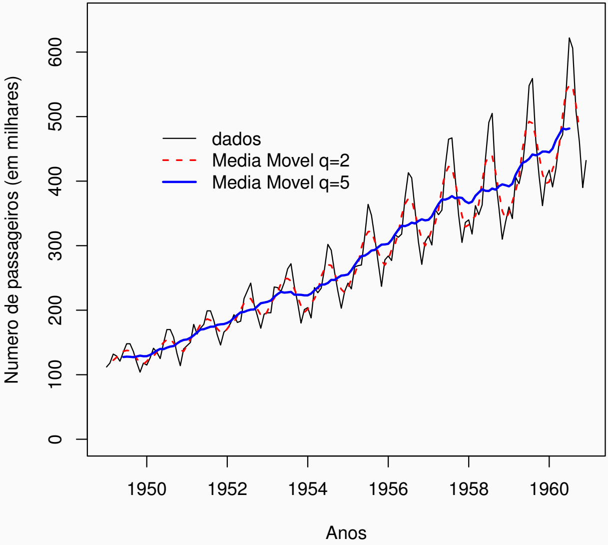
Qualquer que seja o filtro utilizado, y_t é uma estimativa da tendência no tempo t e $x_t - y_t$ é uma série livre de tendência.

Exemplo. Totais mensais de passageiros de linhas aéreas internacionais nos EUA e suas médias móveis de ordem 2 e 5.

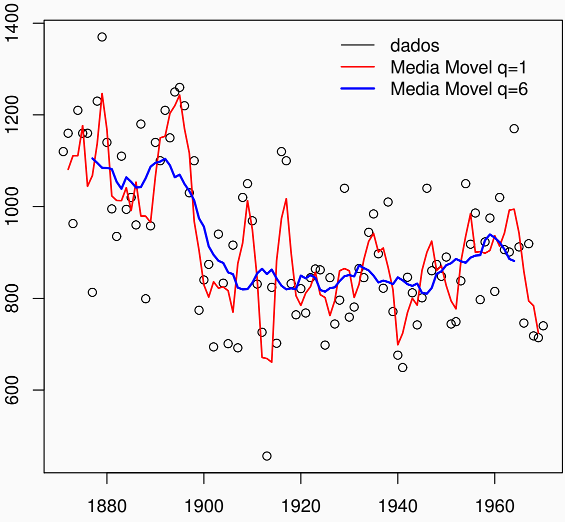
	Obs	q=2	q=5
Jan 1949	112	NA	NA
Feb 1949	118	NA	NA
Mar 1949	132	122.4	NA
Apr 1949	129	127.0	NA
May 1949	121	133.0	NA
Jun 1949	135	136.2	127.4545
Jul 1949	148	137.6	128.0000
Aug 1949	148	137.2	127.7273
Sep 1949	136	131.0	127.1818
Oct 1949	119	125.0	128.2727
Nov 1949	104	118.4	129.5455
Dec 1949	118	116.4	128.6364
Jan 1950	115	120.8	128.7273
Feb 1950	126	127.0	130.7273
Mar 1950	141	128.4	133.8182

```
         Obs   q=2      q=5
Oct 1959 407 439.2 442.0909
Nov 1959 362 410.8 445.8182
Dec 1959 405 396.4 445.8182
Jan 1960 417 398.8 444.6364
Feb 1960 391 418.6 450.3636
Mar 1960 419 432.0 463.3636
Apr 1960 461 455.6 472.5455
May 1960 472 501.8 481.5455
Jun 1960 535 539.2 480.1818
Jul 1960 622 548.6 481.5455
Aug 1960 606 546.4       NA
Sep 1960 508 517.4       NA
Oct 1960 461 479.4       NA
Nov 1960 390    NA       NA
Dec 1960 432    NA       NA
```

Observações e médias móveis.



Exemplo. Medições anuais de vazões do Rio Nilo em Ashwan entre 1871 e 1970 e suas média móveis de ordem 1 e 6.



Distribuições de Retornos

Seja P_t o preço de um ativo no tempo t.

- Os retornos $R_t = (P_t - P_{t-1})/P_{t-1}$ em geral não são estacionários.
- Os log-retornos $r_t = \log(1 + R_t)$ em geral são estacionários.
- Os log-retornos em geral são tem caudas pesadas e apresentam assimetrias.

Assimetria e Curtose

Para uma variável aleatória X tal que $E(X) = \mu$ e $Var(X) = \sigma^2$ define-se os coeficientes de assimetria e curtose como,

$$\begin{aligned} A(X) &= E\left(\frac{(X-\mu)^3}{\sigma^3}\right) \\ K(X) &= E\left(\frac{(X-\mu)^4}{\sigma^4}\right) \end{aligned}$$

- A distribuição normal tem assimetria 0 e curtose igual a 3.
- Distribuições com caudas pesadas $K(X) > 3$ e a quantidade $K(X) - 3$ é chamada de excesso de curtose.

Substituindo os momentos teóricos de X pelos seus equivalente amostrais

$$m_j = \frac{1}{n}\sum_{t=1}^{n}(X_t - \overline{X})^j$$

os estimadores da assimetria e curtose são dados por,

$$\hat{A} = \frac{m_3}{\sqrt{m_2^3}} \quad \text{e} \quad \hat{K} = \frac{m_4}{\sqrt{m_2^2}}$$

respectivamente.