SÉRIES TEMPORAIS

Professores:
Paulo Maciel
Ricardo Massa

Aluno:
Anderson Elias - *aen@cin.ufpe.br*
Ayhalla Riceli – *carlp@cin.ufpe.br*

ROTEIRO

- Introdução
- Técnicas Descritivas
  - Recursos Gráficos
  - Decomposição
  - Filtros Lineares
- Modelos Probabilísticos
  - Processo Estacionário
- Previsão
  - Alisamento Exponencial Simples
- Exercício

SÉRIES TEMPORAIS

INTRODUÇÃO

- Uma série temporal é uma sequencia de observações de uma variável ao longo do tempo.
- Uma característica importante deste tipo de dados é que as observações vizinhas são dependentes e o interesse é analisar e modelar esta dependência.

INTRODUÇÃO

- Para este estudo, a ordem dos dados é muito importante.
- O tempo pode ser substituído por outra variável como profundidade, espaço, etc.
- Requer uso de técnicas específicas.

SÉRIES TEMPORAIS

INTRODUÇÃO

- Surge em vários campos de conhecimento como:
  - **Economia** (preços diários de ações, taxa mensal de desemprego, produção industrial);
  - **Medicina** (eletrocardiograma, eletroencefalograma);
  - **Epidemiologia** (número mensal de novos casos de meningite);
  - **Metereologia** (precipitação pluviométrica, temperatura diária, velocidade do vento),
  - Entre outros.

SÉRIES TEMPORAIS

TERMINOLOGIA

- **Discreta**

Uma série temporal é dita ser *discreta* quando as observações são feitas em tempos específicos.

Definindo o conjunto $T = \{t_1, \ldots, t_n\}$ a série temporal será denotada por $\{X_t : t \in T\}$.
Por simplicidade podemos fazer $T = \{1, 2, \ldots, n\}$.

Ex: Exportações mensais de 1970 a 1980
{01/1970, 02/1970, . . . , 11/1980, 12/1980}.

SÉRIES TEMPORAIS

TERMINOLOGIA

- **Contínua**

Uma série temporal é dita ser *contínua* quando as observações são feitas continuamente no tempo.

Definindo o conjunto $T = \{t : t_1 < t < t_2\}$ a série temporal será denotada por $\{X(t) : t \in T\}$.

Ex: Registro da maré no Recife durante 1 ano T = [0, 24] se unidadede tempo é a hora.
Notação: X(t)

SÉRIES TEMPORAIS

TERMINOLOGIA

- **Multivariada**

Uma série temporal também pode ser ***multivariada***.

Se **k** variáveis são observadas a cada tempo (por exemplo discreto) denota-se por $\{X_{1t}, \ldots, X_{kt}, t \in T\}$.

Ex: Vendas semanais $X_1(t)$ e gastos com propaganda $X_2(t)$.

Neste caso várias séries correlacionadas devem ser analisadas conjuntamente, ou seja em cada tempo tem-se um vetor de observações.

SÉRIES TEMPORAIS

OBJETIVOS

SÉRIES TEMPORAIS

Em algumas situações o objetivo é prever valores futuros enquanto em outras, a relação de uma série com outras séries pode ser o interesse principal.

Alguns dos principais objetivos em séries temporais:

- **Descrição -** Descrever propriedades da série, Ex: padrão de tendência, existência de variação sazonal ou cíclica, observações, etc.

- **Explicação -** Usar a variação em uma série para explicar a variação em outra série.

Centro de Informática

OBJETIVOS

SÉRIES TEMPORAIS

Continuando...

- **Predição -** Predizer valores futuros com base em valores passados. Aqui assume-se que o futuro envolve incerteza, ou seja as previsões não são perfeitas. Porém devemos tentar reduzir os erros de previsão.

- **Controle -** Os valores da série temporal medem a “qualidade” de um processo de manufatura e o objetivo é o controle do processo. Um exemplo é o controle estatístico de qualidade onde as observações são representadas por Gráficos de Controle.

Centro de Informática

ABORDAGENS

- **Técnicas Descritivas -** Técnicas gráficos, identificação de padrões, etc.
- **Modelos Probabilísticos -** Seleção, comparação e adequação de modelos, estimação, predição. Ferramenta básica é a função de autocorrelação.
- **Métodos não paramétricos -** (alisamento ou suavização).
- **Outras Abordagens -** Modelos de espaço de estados, modelos não lineares, séries multivariadas, estudos longitudinais, processos de longa dependência, modelos para volatilidade, etc.

SAZONALIDADE

Muitas séries temporais exibem um comportamento que tende a se repetir a cada **s** períodos de tempo.

Ex: É natural esperar que as vendas mensais de brinquedos terão um pico no mês de dezembro e talvez um pico secundário em outubro. Este padrão possivelmente se repetirá ao longo de vários anos.

SÉRIES TEMPORAIS

SAZONALIDADE

Tipos de Sazonalidade

- **Aditiva -** A série apresenta flutuações sazonais mais ou menos constantes não importando o nível global da série.

  EX: No exemplo dos brinquedos, suponha que o aumento esperado nas vendas nos meses de dezembro é de 1 milhão de reais em relação à média anual. Então as previsões para os meses de dezembro dos próximos anos deve somar a quantia de 1 milhão de reais à uma média anual para levar em conta esta flutuação sazonal.

SÉRIES TEMPORAIS

Centro de Informática

SAZONALIDADE

Tipos de Sazonalidade

- **Multiplicativa -** O tamanho das flutuações sazonais varia dependendo do nível global da série.

  EX: Suponha agora que o aumento esperado nos meses de dezembro seja de 30%. Então o aumento esperado (em valor absoluto) de vendas em dezembro será pequeno ou grande dependendo da média anual de vendas ser baixa ou alta. Nas previsões para os próximos meses de dezembro deve-se multiplicar a média anual pelo fator 1,3.

SÉRIES TEMPORAIS

Centro de Informática

TENDENCIAS

Globalmente, uma série pode exibir tendência de crescimento (ou decrescimento) com vários possíveis padrões.

Podermos pensar em tendência como uma mudança de longo prazo no nível médio da série.

$$X_t = \alpha + \beta_t + \epsilon_t$$

Onde α e β são constantes serem estimadas e ϵ_t denota um erro aleatório com média zero.

A forma de lidar com dados não sazonais que contenham tendência consistem em ajustar a função polinominal.

$$X_t = \beta_0 + \beta_1 + \dots \beta_p t^p + \epsilon_t$$

SÉRIES TEMPORAIS

TENDENCIAS

- **Crescimento linear -** Por exemplo, a cada ano o aumento esperado nas vendas de um certo brinquedo é de 1 milhão de reais.
- **Crescimento exponencial -** Por exemplo, a cada ano as vendas de um certo brinquedo aumentam de um fator 1,3.
- **Crescimento amortecido -** Por exemplo, as vendas de um certo brinquedo tem uma aumento esperado de 70% sobre o ano anterior. Se o aumento esperado for de 1 milhão de reais no primeiro ano, no segundo ano será de 700 mil reais, no terceiro ano será de 490 mil reais e assim por diante.

SÉRIES TEMPORAIS

TÉCNICAS DESCRITIVAS

Em uma análise de uma série temporal, a representação gráfica dos dados sequenciais ao longo de um período de tempo pode revelar padrões de comportamento importantes.

Tendências de crescimento (ou decrescimento), alterações estruturais, entre outros são muitas vezes facilmente identificados.

O gráfico temporal deve ser sempre o primeiro passo e antecede qualquer análise.

SÉRIES TEMPORAIS

TÉCNICAS DESCRITIVAS

Recursos Gráficos

Usaremos o banco de dados do R para gerar gráficos das séries: AirPassengers.

Importar o pacote: tseries *(Time series analysis and computational finance)*

```
data()
dados <- data(package='tseries')
```

SÉRIES TEMPORAIS

TÉCNICAS DESCRITIVAS

Recursos Gráficos

Verificando alguns dados do pacote:

help(AirPassengers)

The classic Box & Jenkins airline data. Monthly totals of international airline passengers, 1949 to 1960.

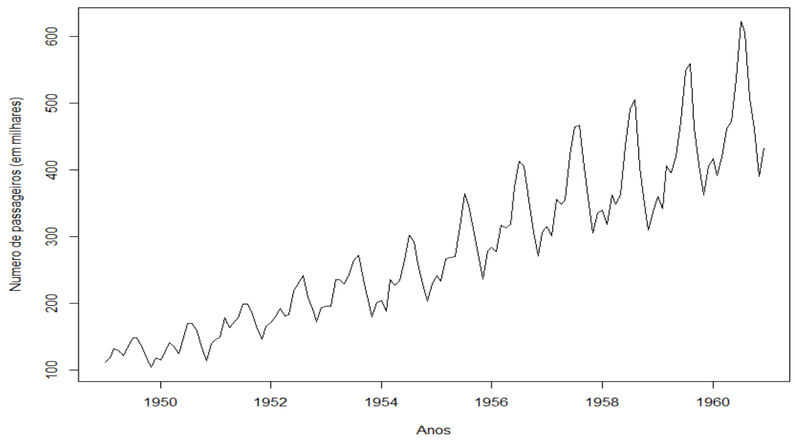
Fazendo um gráfico da série

```
x = AirPassengers
par(mfrow=c(1,1))
is.ts(x)
plot(x,xlab='Anos',ylab='Numero de passageiros (em milhares)')
```

SÉRIES TEMPORAIS

TÉCNICAS DESCRITIVAS

Recursos Gráficos



SÉRIES TEMPORAIS

TÉCNICAS DESCRITIVAS

SÉRIES TEMPORAIS

Decomposição

Muitas das propriedades observadas em uma série temporal X_t podem ser captadas assumindo-se a seguinte forma.

$$X_t = T_t + C_t + R_t$$

Onde T_t é um componente de tendência, C_t é um componente cíclica ou sazonal e R_t é um componente aleatório ou ruído (Seria a parte não explicada, que espera-se ser puramente aleatória).

$$\ldots = C_{t-2s} = C_{t-s} = C_t = C_{t+s} = C_{t+2s} = \ldots$$

Assim, variações periódicas podem ser captadas por este componente.

TÉCNICAS DESCRITIVAS

SÉRIES TEMPORAIS

Decomposição

A função ``decompose'' estima as componentes de tendência e sazonalidade (aditiva ou multiplicativa) via medias moveis. Com os comandos abaixo obtém-se um gráfico da serie original com a tendência estimada superposta bem como o gráfico da componente aleatória para checar a adequação do modelo.

```
x = AirPassengers
par(mfrow=c(1,1))
is.ts(x)
plot(x,xlab='Anos',ylab='Numero de passageiros (em milhares)')
```

Aditiva

```
m = decompose(x)

par(mfrow=c(1,1))
plot(x)
lines(m$trend,col='blue')
plot(m$random)

plot(m)
```

Mutiplicativa

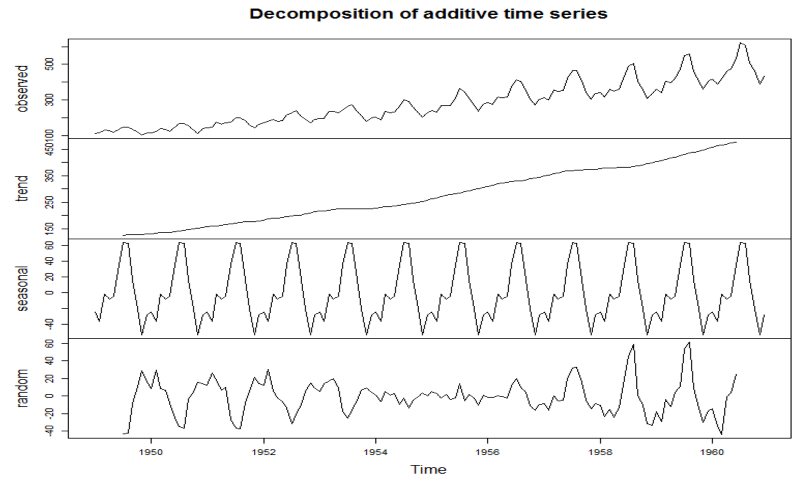
Para sazonalidade multiplicativa usar a opção type='m'

```
m = decompose(x,type='m')
par(mfrow=c(1,1))
plot(x)
lines(m$trend,col='blue')
plot(m$random)

plot(m)
```

TÉCNICAS DESCRITIVAS

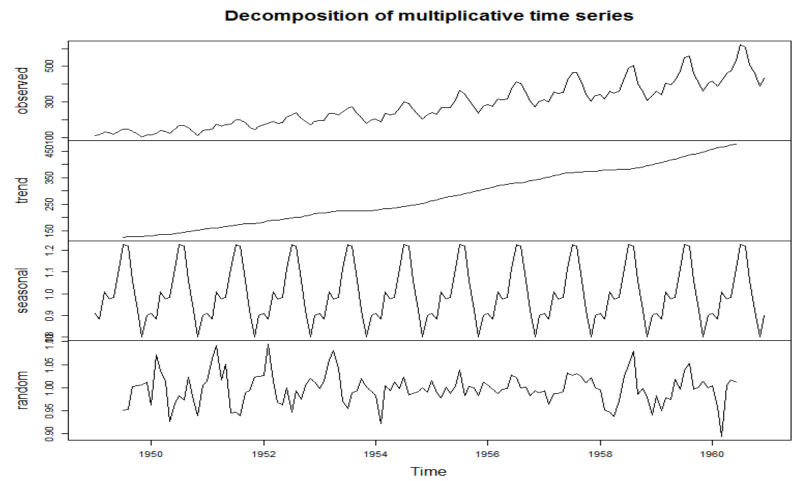
Decomposição



SÉRIES TEMPORAIS

TÉCNICAS DESCRITIVAS

Decomposição

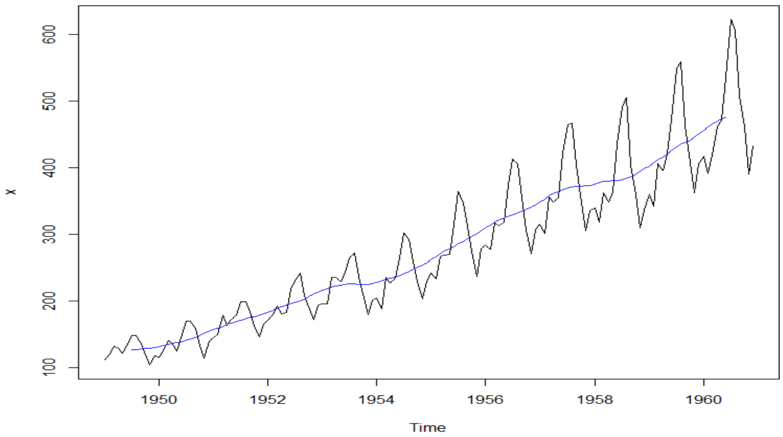


SÉRIES TEMPORAIS

TÉCNICAS DESCRITIVAS

SÉRIES TEMPORAIS

Decomposição (Tendência – Crescimento Linear)



TÉCNICAS DESCRITIVAS

SÉRIES TEMPORAIS

Filtros Lineares

São usados para analisar tendências nas séries convertendo uma série $\{x_t\}$ em outra $\{y_t\}$

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

$\{a_j\}$ é um conjunto de pesos e desejamos estimar a média local, os pesos devem ser tais que $\sum_{j=-q}^{s} a_j = 1,$ garanta que $\min\{x_t\} < y_t < \max\{x_t\}$, sendo assim considerada **média móvel**.

y_t é uma estimativa da tendência no tempo t e $x_t - y_t$ é uma série livre de tendência.

TÉCNICAS DESCRITIVAS

SÉRIES TEMPORAIS

Filtros Lineares

Utilizando a ferramenta R para aplicar filtros lineares em séries temporais.

A série AirPassengers contem os totais mensais de passageiros de linhas aereas internacionais nos EUA, entre 1949 e 1960 (Box, Jenkins and Reinsel, 1976).

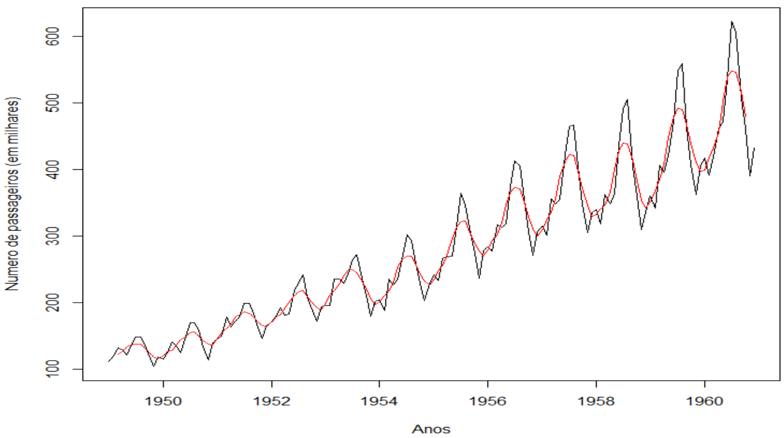
Aplicando um filtro linear media móvel para ``estimar'' a tendência.

```
x = AirPassengers
q = 2
coefs = rep(1/(2*q+1),2*q+1) # vetor de coeficientes do filtro
y = filter(x=x,filter=coefs,sides=2,method="convolution")
plot(x,xlab='Anos',ylab='Numero de passageiros (em milhares)')
lines(y,col='red')
```

TÉCNICAS DESCRITIVAS

SÉRIES TEMPORAIS

Filtros Lineares



MODELOS PROBABILÍSTICOS

- Modelos Probabilísticos
  - Processos Estacionários
  - Função de Autocorrelação
  - Processos Estocásticos
    - Sequência Aleatória
    - Processos Aleatórios
    - Processos Móveis
    - Processos Autoregressivos
    - Modelos Misto ARMA
    - Modelos Mistos ARIMA

SÉRIES TEMPORAIS

MODELOS PROBABILÍSTICOS

São modelos adequados para dados de séries temporais. Também conhecidos como processos estocásticos.

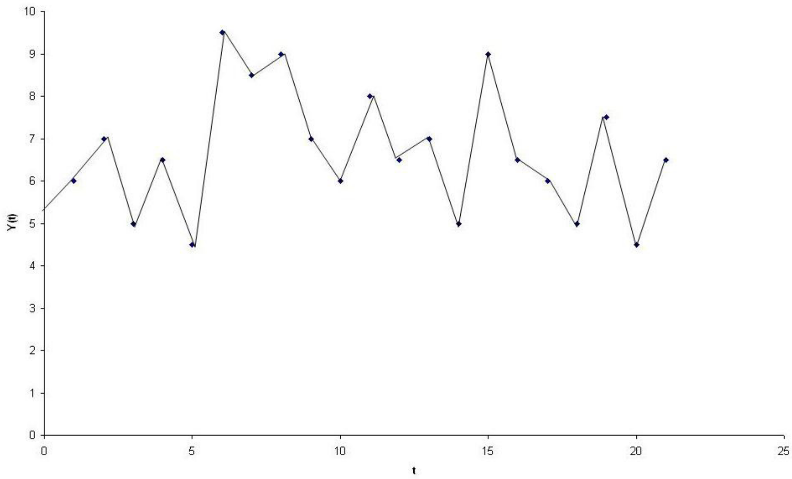
PROCESSOS ESTOCÁSTICOS: Coleção de variáveis aleatórias ordenadas no tempo e definidas em um conjunto de pontos (T), podem ser contínuos ou discretos.

A variável aleatória no tempo *t* é denotada por $X(t)$, em caso contínuo; e, por X_t, em caso discreto.

SÉRIES TEMPORAIS

PROCESSOS ESTACIONÁRIOS

A mais comum em séries temporais é a de estacionariedade. Basicamente isso significa que o comportamento da série não se altera com o passar do tempo



SÉRIES TEMPORAIS

PROCESSOS ESTACIONÁRIOS

- Uma série temporal é dita estacionária se a distribuição de probabilidade conjunta de:

$$\boxed{X(t_1), \ldots, X(t_k)}$$

é a mesma de

$$\boxed{X(t_1 + \tau), \ldots, X(t_k + \tau)}$$

- O deslocamento da origem dos tempos por uma quantidade τ não tem efeito na distribuição conjunta, depende apenas dos intervalos $t_1, \ldots, t_k$

SÉRIES TEMPORAIS

PROCESSOS ESTACIONÁRIOS

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

$$\boxed{\mu(t) = \mu \quad \text{e} \quad \sigma^2(t) = \sigma^2}$$

São constantes que não dependem de t.

SÉRIES TEMPORAIS

PROCESSOS ESTACIONÁRIOS

Para $k = 2$ a distribuição conjunta de $X(t_1) \text{ e } X(t_2)$ depende apenas da distância $t_2 - t_1$.

A função de autocovariância $\gamma(t_1, t_2)$ também depende apenas de $t_2 - t_1$ pode ser escrita como $\gamma(\tau)$ onde:

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

SÉRIES TEMPORAIS

PROCESSOS ESTACIONÁRIOS

Definição 3.1. *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, i.e.*

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

SÉRIES TEMPORAIS

PREVISÃO

Forma pelo qual se faz previsões de valores futuros através de modelos ajustados.

- Devido à incerteza presente;
- Deve também ser econômico;
- Descrição deve ser relativamente simples e flexível para poder se adaptar ao futuro (incerto) e facilitar aprendizado.
- Aprendizado é processamento de informação através do modelo.
- Previsão é hipótese, conjectiva ou especulação sobre o futuro.

SÉRIES TEMPORAIS

PREVISÃO

SÉRIES TEMPORAIS

Se t é o período atual e estamos interessados em prever valores de X_{t+1}, X_{t+2}, ...X_{t+n} . A previsão de X_{t+k}, para k=1,2 ... Será denotada $\hat{x}_t(k)$ e é definida como a esperança condicional de X_{t+k} dados todos os valores passados.

Temos um número finito de observações onde obtemos:
$\hat{x}_t(k) = E(X_{t+k}|x_t, x_{t-1}, \ldots)$ equação denominada *função de previsão* e o inteiro *k* é chamado de *horizonte de previsão.*

Uma prática comum para chegar a performance preditiva do modelo é observar x_1 , ... ,x_n onde as previsões podem ser realizadas dentro do período amostral e comparadas com os valores observados.

PREVISÃO

SÉRIES TEMPORAIS

Alisamento Exponencial Simples

Dada uma série temporal x_1 ,...x_n , é razoável tomar estimativa de x_{n+1} como uma soma ponderada das observações passadas.

$$\hat{x}_n(1) = a_0 x_n + a_1 x_{n-1} + \ldots$$

Onde $\{a_j\}$ são os pesos.
Os pesos decaem geometricamente a uma taxa constante dadas por:

$$a_j = \alpha(1-\alpha)^j, \quad j = 0, 1, \ldots$$

Onde $0 < \alpha < 1$ é chamada de *constante de alisamento* e a previsão 1 passo a frente em t=n fica:

$$\hat{x}_n(1) = \alpha x_n + \alpha(1-\alpha)x_{n-1} + \alpha(1-\alpha)^2 x_{n-2} + \ldots.$$

Haverá um número finito de observações passadas e a soma acima será também finita.

O parâmetro α está controlando o grau de envelhecimento uma vez que o conteúdo informativo de uma observação decai com sua idade.

PREVISÃO

SÉRIES TEMPORAIS

Alisamento Exponencial Simples

A equação pode ser reescrita de forma recursiva, colocando (1- α) em evidência, obtendo assim:

$$\begin{aligned}\hat{x}_n(1) &= \alpha x_n + (1-\alpha)[\alpha x_{n-1} + \alpha(1-\alpha)x_{n-2} + \alpha(1-\alpha)^2 x_{n-3} + \ldots] \\ &= \alpha x_n + (1-\alpha)\hat{x}_{n-1}(1)\end{aligned}$$

Podendo ainda ser reescrita na forma de erro, $e_n = x_n - \hat{x}_{n-1}(1)$ definindo o erro de previsão 1 passo à frente no tempo n temos:

$$\hat{x}_n(1) = \hat{x}_{n-1}(1) + \alpha e_n$$

Ou seja, a previsão para *t=n+1* é igual a previsão para *t=n* que foi feita em *t=n-1* mas uma proporção de erro cometido.

A previsão k=passos a frente é a mesma:

$$\hat{x}_n(k) = \hat{x}_n(1),\ k = 2, 3, \ldots.$$

PREVISÃO

SÉRIES TEMPORAIS

Alisamento Exponencial Simples

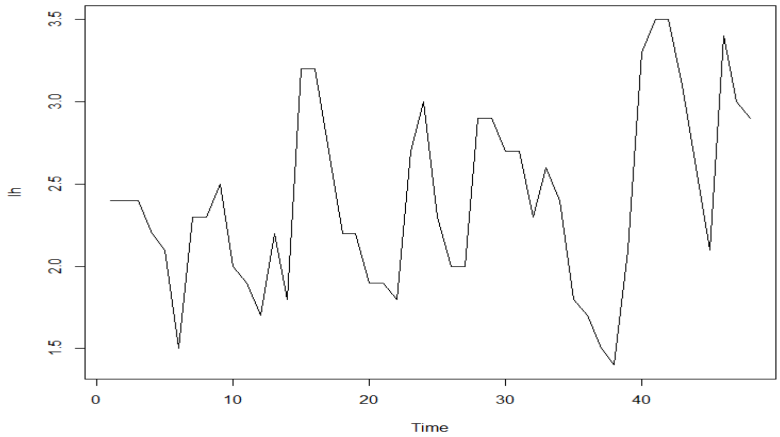
Aplicando o método de alisamento exponencial simples a serie ***lh****(luteinizing hormone in blood samples at 10 mins intervals from a human female 48 samples)* do banco de dados do R.

```
> lh
Time Series:
Start = 1
End = 48
Frequency = 1
 [1] 2.4 2.4 2.4 2.2 2.1 1.5 2.3 2.3 2.5 2.0 1.9 1.7 2.2 1.8 3.2 3.2
[17] 2.7 2.2 2.2 1.9 1.9 1.8 2.7 3.0 2.3 2.0 2.0 2.9 2.9 2.7 2.7 2.3
[33] 2.6 2.4 1.8 1.7 1.5 1.4 2.1 3.3 3.5 3.5 3.1 2.6 2.1 3.4 3.0 2.9
```

PREVISÃO

SÉRIES TEMPORAIS

Alisamento Exponencial Simples



PREVISÃO

SÉRIES TEMPORAIS

Alisamento Exponencial Simples

Especificação de α

O valor de α reflete a influência das observações passadas nas previsões. O critério utilizado é a minimização da soma de quadrado dos erros de previsão.

$$
\begin{aligned}
\hat{x}_0(1) &= x_1, \\
\hat{x}_1(1) &= \alpha x_1 + (1-\alpha)\hat{x}_0(1), \quad e_2 = x_2 - \hat{x}_1(1) \\
\hat{x}_2(1) &= \alpha x_2 + (1-\alpha)\hat{x}_1(1), \quad e_3 = x_3 - \hat{x}_2(1) \\
&\vdots \\
\hat{x}_{n-1}(1) &= \alpha x_{n-1} + (1-\alpha)\hat{x}_{n-2}(1), \quad e_n = x_n - \hat{x}_{n-1}(1)
\end{aligned}
$$

O procedimento se repete para valores de α variando entre 0 e 1.

PREVISÃO

SÉRIES TEMPORAIS

Alisamento Exponencial Simples

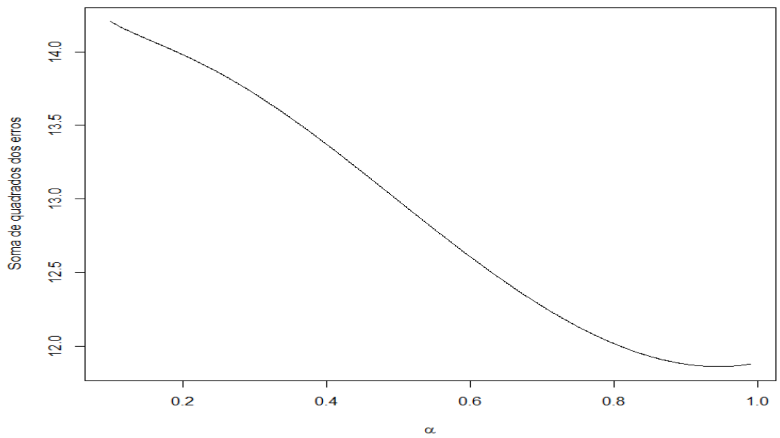
Podemos selecionar o valor da constante de alisamento que minimiza a soma dos quadrados dos erros de previsão.

```
data(lh)
x=lh
int = seq(0.1,0.99,0.001)
e=NULL
for (alfa in int){
  e2=0
  prev.ant = x[1]
  for (i in 2:length(x)){
    prev = alfa*x[i-1] + (1-alfa)*prev.ant
    prev.ant = prev
    e2 = e2 + (x[i]-prev)**2
  }
  e=c(e,e2)
}
plot(int,e,type='l',xlab=expression(alpha),
     ylab='Soma de quadrados dos erros')
```

PREVISÃO

SÉRIES TEMPORAIS

Alisamento Exponencial Simples



PREVISÃO

SÉRIES TEMPORAIS

Alisamento Exponencial Simples (Utilizando uma Função)

Obtendo a soma dos quadrados dos erros de previsão 1 passo a frente em função de α. Valor observados (pontos) e previsões 1 passo a frente (linhas) usando o valor ótimo de α.

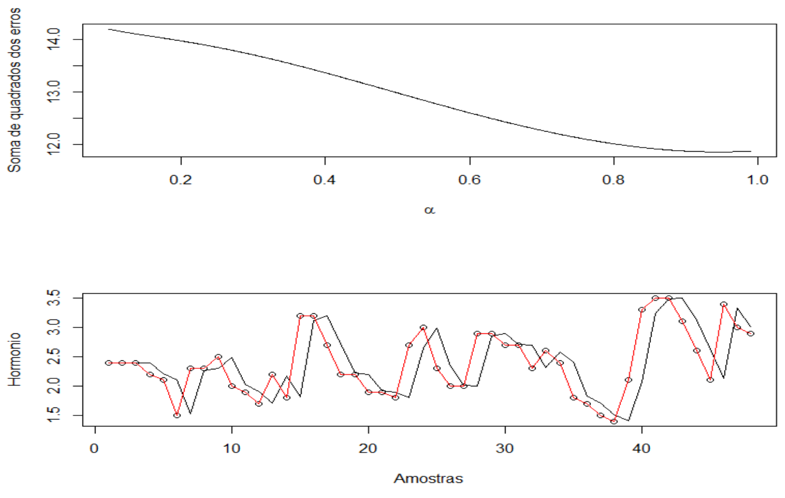
```
AES = function(x,interval){
  e=NULL
  for (alpha in interval){
    e2=0
    prev = x[1]
    for (i in 2:length(x)){
      prev = c(prev,alpha*x[i-1] + (1-alpha)*prev[i-1])
      e2 = e2 + (x[i]-prev[i])**2
    }
    e=c(e,e2)
  }
  plot(interval,e,type='l', xlab=expression(alpha),
       ylab='Soma de quadrados dos erros')
  e.min=min(e)
  alpha=interval[e==e.min]
  prev = x[1]
  for (i in 2:length(x)) prev = c(prev,alpha*x[i-1] + (1-alpha)*prev[i-1])
  return(list(alpha=alpha,sq2=e.min,prev=prev))
}

par(mfrow=c(2,1))
m = AES(lh, seq(0.1, 0.99, 0.001))
plot(1:48, m$prev, ylab="Hormonio", xlab="Amostras", type="l")
points(lh)
lines(lh,col='red')
```

PREVISÃO

SÉRIES TEMPORAIS

Alisamento Exponencial Simples (Utilizando uma Função)



OBRIGADO !

SÉRIES TEMPORAIS