

# Séries temporais

Pedro Monteiro de Almeida Junior

## Conteúdo programático

### 1. Conceitos iniciais em séries temporais

- Objetivos de séries temporais
- Análise descritiva com séries temporais
- Tendência
- Sazonalidade
- Estacionariedade
- exemplos práticos de aplicações em séries temporais

### 2. Modelos de Suavização exponencial

- Médias móveis simples
- Suavização exponencial simples
- Suavização exponencial de Holt
- Suavização exponencial sazonal de Holt-Winters

### 3. Modelos AR, MA, ARMA e ARIMA

- Modelos Autoregressivos (AR)
- Modelos de médias móveis (MA)
- Modelos Autoregressivos e de médias móveis (ARMA)
- Modelos Autoregressivos e de médias móveis integrados (ARIMA)
- Identificação das ordens dos modelos
- critério para seleção dos modelos

- Previsão e medidas de avaliação
- diagnóstico dos modelos
- Exemplos práticos no R

#### 4. Projetos Práticos

- Modelos de séries temporais não gaussianas (GARMA)
- Modelos ARCH e GARCH
- Modelos de séries temporais não paramétricos
- Modelo ETS
- Exemplos práticos no R
- Introdução ao Python e aplicações

## Referências bibliográficas

- [1] SHUMWAY R. H. STOFFER, D. S. Time series analysis and its applications with R examples. 3ªed. New York: Springer, 2010.
- [2] MORETTIN, P.A. Análise de Séries Temporais. S. Paulo, Edgard Blucher, 2006.
- [3] BROCKWELL, P. & DAVIS, R. Introduction to Time Series and Forecasting. Springer, 1996.
- [4] HAMILTON, J.D. Time Series Analysis.
- [5] Forecasting: Principles and Practice (3rd ed). *Rob J Hyndman and George Athanasopoulos*. Monash University, Australia. [Versão online](#)

## Sites com exemplos em séries temporais

<https://www.kaggle.com/>

<https://data.world/datasets/time-series>

<https://medium.com/analytics-vidhya/10-time-series-datasets-for-practice-d14fec9f21bc>

<https://dados.gov.br/dataset>

<http://ipeadata.gov.br/Default.aspx>

<https://campinagrande.pb.gov.br/portal-da-transparencia/>

<http://dados.recife.pe.gov.br/>

<https://transparencia.joaopessoa.pb.gov.br/#/>

## Conceitos iniciais em séries temporais

Uma série temporal é qualquer conjunto de observações ordenadas ao longo do tempo. A seguir, temos alguns exemplos de séries temporais:

- valores diários de precipitação de chuva em campina grande;
- quantidade de assaltos semanais em uma determinada região;
- índices diários da Bolsa de Valores de São Paulo;
- valor gasto com gasolina anualmente por órgãos públicos;
- registros do fluxo de carros em um estacionamento por hora;

Estes são alguns, dos mais diversos exemplos nos quais podemos aplicar as técnicas de séries temporais. Note que, essa técnica pode ser aplicada para diferentes medidas de tempo: Segundos, minutos, hora, dia, semana, mês e ano.

### Objetivos principais

As análises usando séries temporais têm como principais objetivos:

- (i) Fazer previsões de valores futuros da série;
- (ii) Descrever o comportamento da série;
- (iii) Identificar periodicidades relevantes nos dados;

### Domínio temporal e espectral

Existem, basicamente, dois enfoques usados na análise de séries temporais. No primeiro, a análise é feita sobre o domínio temporal (será o enfoque do curso) e no domínio de frequência, que temos a análise espectral com aplicações principalmente na área da engenharia e física.

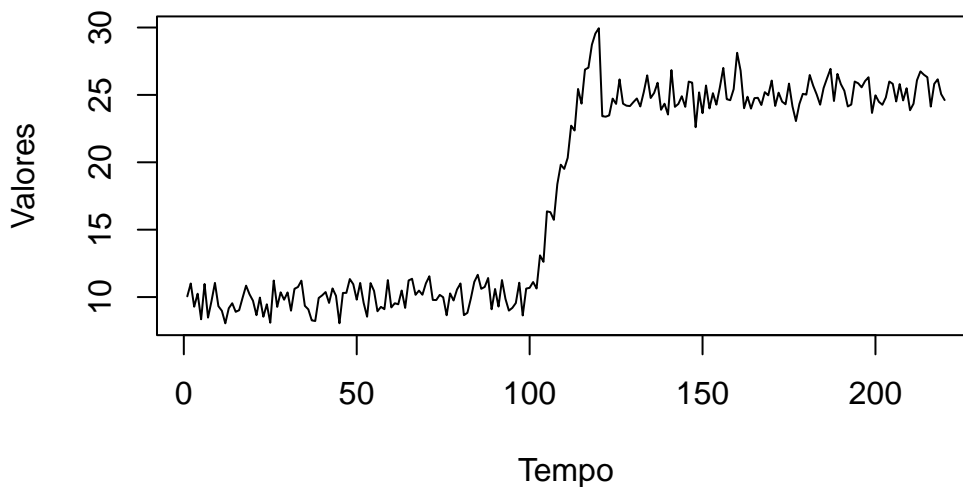
### Notação

Seja  $Z_t$  uma variável aleatória no instante  $t$  ( Por exemplo, a temperatura no tempo  $t$  ). De modo geral, uma série temporal poderá ser representado por um vetor  $Z_{t_1}, \dots, Z_{t_n}$ .

## Tendência

Uma das principais características para se observar em uma série temporal é a tendência. Ou seja, identificar se a série apresenta um comportamento crescente ou decrescente ao longo do tempo. Por exemplo, no gráfico abaixo podemos identificar que em determinado tempo a série apresentou um comportamento de tendência crescente até obter uma certa estabilidade novamente.

### Exemplo de série com tendência



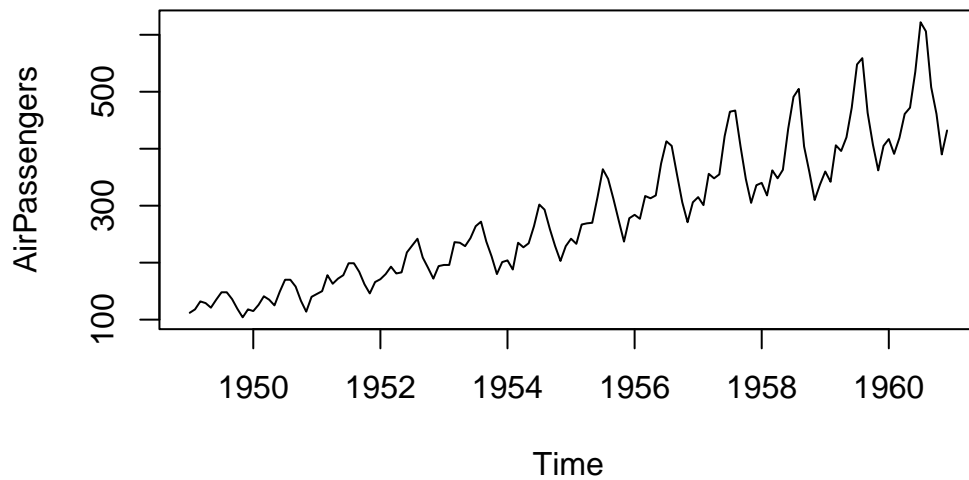
## Sazonalidade

O componente de sazonalidade também é outra característica importante para observar-se em uma série temporal. Esse componente determina se existe um comportamento sazonal ou cíclico, ou seja, se uma série observada apresenta um comportamento que se repete de tempos em tempos. A seguir, é fornecido um exemplo de uma série temporal que apresenta esse tipo de comportamento.

```
## Carregando os pacotes
library(TSA)
library(forecast)

data("AirPassengers")
plot.ts(AirPassengers, main = 'Quantidade de passageiros de avião entre 1949-1960')
```

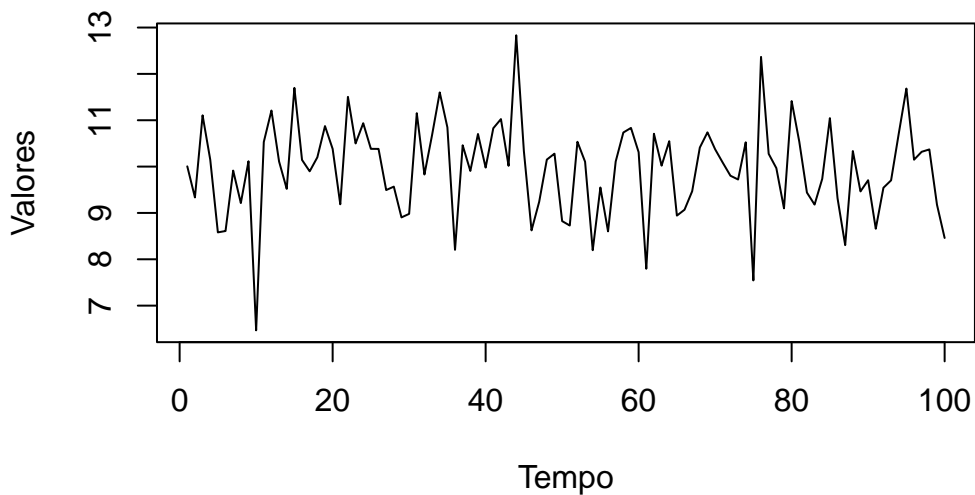
## Quantidade de passageiros de avião entre 1949–1960



### Estacionariedade

Uma das suposições que podemos fazer em uma série temporal é a de que ela é **estacionária**. Isto é, ao longo do tempo seus valores estão aleatoriamente próximos de uma média constante, refletindo de alguma forma um equilíbrio estável. No gráfico a seguir, temos um exemplo de uma série temporal estacionária.

## Exemplo de uma série estacionária



Contudo, em várias situações práticas, vamos ter que lidar com **séries não estacionárias**. É bastante comum, por exemplo, em séries econômicas e financeiras a presença de tendências de longos períodos ou curtos períodos (o que geralmente caracteriza uma mudança de nível). Dessa forma, em alguns modelos estatísticos de séries temporais, vamos ter que “tratar” a não estacionariedade das séries através de transformações. A transformação mais comum seria tomar diferenças sucessivas da série original, até obter uma série estacionária.

### Operador Diferença

o operador diferença (  $\Delta$  ) de primeira ordem pode ser definido como

$$\Delta Z_t = Z_t - Z_{t-1}$$

a segunda diferença é

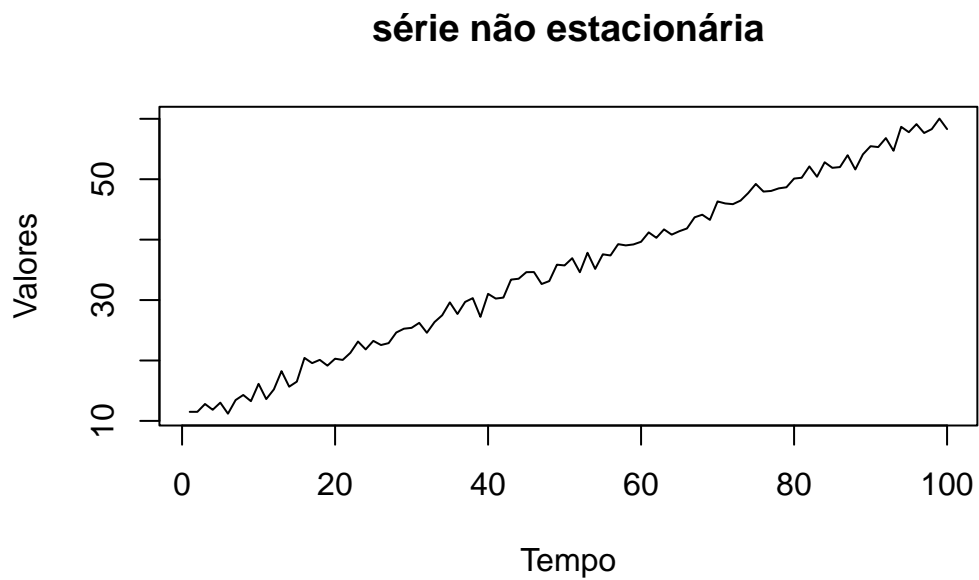
$$\Delta^2 Z_t = \Delta[\Delta Z_t] = \Delta[Z_t - Z_{t-1}]$$

De modo geral, a n-ésima diferença de  $Z_t$  é

$$\Delta^n Z_t = \Delta[\Delta^{n-1} Z_t]$$

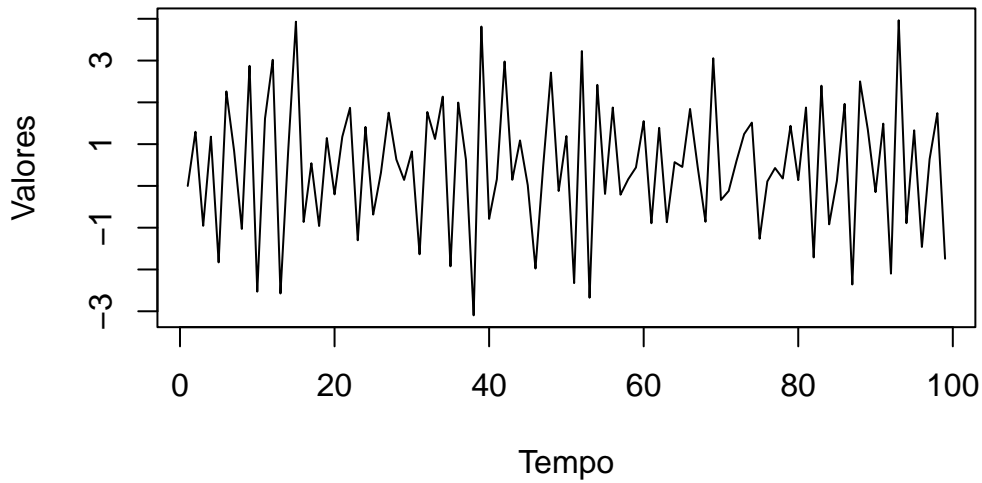
Na prática, é comum tomar uma ou duas diferenças para tornar a série temporal estacionária. O exemplo a seguir, ilustra como o operador diferença pode tornar uma série claramente não estacionária em uma série estacionária.

```
### Exemplo simulado no R:  
x1 = c()  
for (i in 1:100) x1[i] = rnorm(1, mean = 10+(i/2), sd = 1)  
x2 = diff(x1, 1)  
  
plot.ts(x1, xlab = "Tempo", ylab = "Valores", main = "série não estacionária" )
```



```
plot.ts(x2, xlab = "Tempo", ylab = "Valores", main = "série aplicando o operador \ndiferen
```

### série aplicando o operador diferença de primeira ordem



#### Modelos para séries temporais

Os modelos em séries temporais controlados por leis probabilísticas, são baseados na teoria de um processo estocástico. A construção desses modelos depende de vários fatores, tais como a natureza dos dados, objetivo da análise e comportamento da série. Dessa forma, o leque de opções de modelos que podemos aplicar são muitos, inclusive modelos não probabilísticos também podem ser usados (Redes neurais, deep learning, etc ...).

#### Processos estocásticos

**Definição:** Seja  $\tau$  um conjunto arbitrário. Um processo estocástico é uma família  $Z = \{Z_t, t \in \tau\}$ , tal que, para cada  $t \in \tau$ ,  $Z_t$  é uma variável aleatória.

Nessas condições, um processo estocástico é uma família de variáveis aleatórias, que supomos definidas num mesmo espaço de probabilidade  $(\Omega, \mathcal{A}, \mathcal{P})$ . Cada  $t \in \mathbb{Z}$ , tal que  $Z_t$

seja uma variável aleatória. Na Figura a seguir, podemos ilustrar um processo estocástico como uma família de trajetórias.

**Observação:** Uma série temporal é uma realização de um processo estocástico, isto é,  $Z_t^{(j)}$  para algum  $j = 1, \dots, n$ .



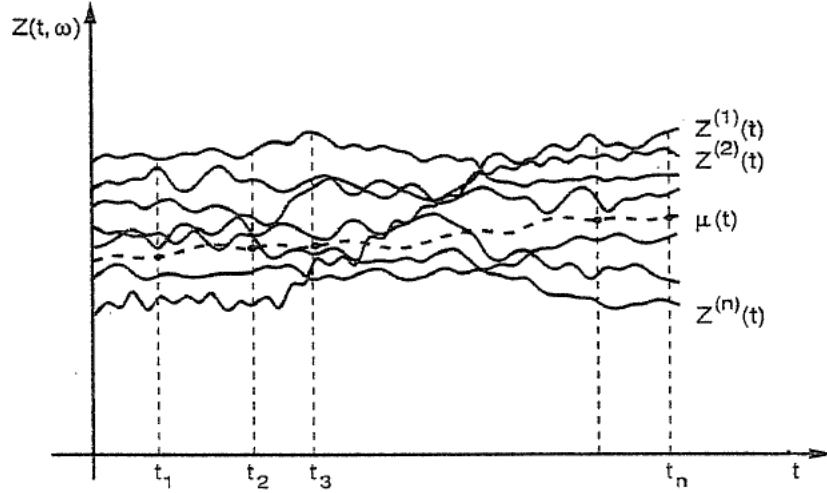


Figure 1: Fonte: Morettin & Toloi (2018)

### Processos estocásticos estacionários

**Definição:** Um processo estocástico  $Z = \{Z(t), t \in \tau\}$  diz-se estritamente estacionário se todas as distribuições finito-dimensionais permanecem as mesmas sob translações no tempo, ou seja,

$$F(z_1, \dots, z_n; t_1 + \tau, \dots, t_n + \tau) = F(z_1, \dots, z_n; t_1, \dots, t_n)$$

para quaisquer  $t_1, \dots, t_n, \tau$ .

Isto significa, em particular, que todas as distribuições unidimensionais são invariantes sob translações do tempo, logo a média  $\mu(t)$  e a variância  $V(t)$  são constantes, isto é,

$$\mu(t) = \mu, \quad V(t) = \sigma^2,$$

para todo  $t \in \tau$ . Sem perda de generalidade, podemos supor que  $\mu = 0$ ; caso contrário, considere o processo  $\{Z(t) - \mu\}$ .

**Definição:** Um processo estocástico  $Z = \{Z(t), t \in T\}$  diz-se fracamente estacionário ou estacionário de segunda ordem (ou em sentido amplo) se e somente se

- (i)  $E\{Z(t)\} = \mu(t) = \mu$ , constante, para todo  $t \in T$ ;
- (ii)  $E\{Z^2(t)\} < \infty$ , para todo  $t \in T$ ;
- (iii)  $\gamma(t_1, t_2) = \text{Cov}\{Z(t_1), Z(t_2)\}$  é uma função de  $|t_1 - t_2|$ .

### Função de autocovariância

Seja  $\{Z_t, t \in \mathbb{Z}\}$  um processo estocástico real com tempo discreto, de média zero e função de autocovariância (FACV)  $\gamma_r = \mathbf{E}(Z_t Z_{t-r})$ .

**Proposição:** A FACV  $\gamma_r$  satisfaz as seguintes propriedades:

- (i)  $\gamma_0 > 0$
- (ii)  $\gamma_{-r} = \gamma_r$
- (iii)  $|\gamma_r| \leq \gamma_0$

### Modelo de decomposição

Toda série temporal pode ser decomposta em três componentes, como segue,

$$Z_t = T_t + S_t + \epsilon_t,$$

em que  $T_t$  é a tendência,  $S_t$  é a sazonalidade e  $\epsilon_t$  é um componente aleatório, de média zero e variância  $\sigma_\epsilon^2$ .

O interesse principal em considerar um modelo de decomposição será o de estimar a tendência  $\hat{T}_t$  ou  $\hat{S}_t$  e construir séries livres de tendência ou sazonalidade. Isto é,

$$Z_t^* = Z_t - \hat{S}_t \quad (\text{Sem sazonalidade})$$

ou

$$Z_t^* = Z_t - \hat{T}_t \quad (\text{Sem tendência})$$

### Tendência Polinomial

Um procedimento bastante comum para estimar a tendência, é o modelo de tendência polinomial,

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \cdots + \beta_p t^p,$$

em que  $p$  é o grau do polinômio e tem que ser menor que o número de observações. Para estimar os parâmetros  $\beta_j$  utiliza-se o método dos mínimos quadrados, ou seja, se resume a um problema de regressão linear.

## Modelos de suavização exponencial

Os métodos de previsões em séries temporais, em geral, baseia-se na ideia de que observações passadas contêm informações sobre o padrão de comportamento da série temporal. Uma técnica utilizada para entender esse comportamento da série ao longo do tempo são os modelos de suavização exponencial.

Essas técnicas buscam suavizar os valores de uma série baseado no comportamento mais recente da série.

## Modelos para séries localmente constantes

Vamos considerar inicialmente o caso de uma série temporal  $Z_1, \dots, Z_N$ , localmente composta de seu nível mais um ruído aleatório, isto é,

$$Z_t = \mu_t + a_t, \quad t = 1, \dots, N,$$

onde  $E(a_t) = 0$ ,  $\text{Var}(a_t) = \sigma_a^2$  e  $\mu_t$  é um parâmetro desconhecido, que pode variar lentamente com o tempo.

## Médias móveis simples

- **Procedimento**

A técnica de média móvel consiste em calcular a média aritmética das  $r$  observações mais recentes, isto é.

$$M_t = \frac{Z_t + Z_{t-1} + \dots + Z_{t-r+1}}{r}$$

ou

$$M_t = M_{t-1} + \frac{Z_t - Z_{t-r}}{r}$$

Assim,  $M_t$  é uma estimativa do nível  $\mu_t$  que não leva em conta (ou não pondera) as observações mais antigas, o que é razoável devido ao fato do parâmetro variar suavemente com o tempo.

O nome média móvel é utilizado porque, a cada período, a observação mais antiga é substituída pela mais recente, calculando-se uma média nova.

- **Previsão**

A previsão de todos os valores futuros é dada pela última média móvel calculada, isto é,

$$\hat{Z}_t(h) = \hat{Z}_{t-1}(h+1) + \frac{Z_t - Z_{t-r}}{r}, \quad h > 0.$$

A equação acima pode ser interpretada como um mecanismo de atualização de previsão, pois a cada instante (ou a cada nova observação) corrige a estimativa prévia de  $Z_{t+h}$ .

- **Determinação de  $r$  (tamanho do janelamento)**

As propriedades do método dependem do número de observações utilizadas na média (valor de  $r$ ). Um valor grande de  $r$  faz com que a previsão acompanhe lentamente as mudanças do parâmetro  $\mu_t$ : um valor pequeno implica numa reação mais rápida. Existem dois casos extremos:

- (i) se  $r = 1$ , então o valor mais recente da série é utilizado como previsão de todos os valores futuros (este é o tipo de previsão mais simples que existe e é denominado “método ingênuo”);
- (ii) se  $r = N$ , então a previsão será igual à média aritmética de todos os dados observados.

Entretanto, como podemos encontrar um  $r$  ‘ótimo’ para o nosso problema? Não existe uma resposta exata para esta pergunta, porém o que podemos fazer é escolher, por exemplo, o valor de  $r$  que forneça a melhor previsão.

- **Vantagens e desvantagens do método**

As principais **vantagens** são:

- (i) simples aplicação;
- (ii) é aplicável quando se tem um número pequeno de observações;
- (iii) permite uma flexibilidade grande devido à variação de  $r$  de acordo com o padrão da série:

e as **desvantagens** são:

- (i) deve ser utilizado somente para prever séries estacionárias, caso contrário a precisão das previsões obtidas será muito pequena, pois os pesos atribuídos às  $r$  observações são todos iguais e nenhum peso é dado às observações anteriores a esse período;
- (ii) necessidade de armazenar pelo menos  $(r - 1)$  observações; e
- (iv) dificuldade em determinar o valor de  $r$ .

Na prática, o método de médias móveis não é utilizado frequentemente, pois o Método de Suavização Exponencial Simples, que veremos logo a seguir, possui todas as vantagens anteriores e mais algumas, que o tornam mais atraente.

## Suavização exponencial simples (SES)

Em séries temporais é razoável pensar que as informações mais recentes tenham um peso maior que as informações mais antigas. Este é exatamente o conceito por trás da suavização exponencial simples. As previsões são calculadas usando médias ponderadas, onde os pesos diminuem à medida que as observações ficam mais longe do tempo atual.

- **Procedimento**

Suavização exponencial simples pode ser descrita matematicamente por

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1 - \alpha) y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + \dots,$$

Onde  $0 \leq \alpha \leq 1$  é o parâmetro de suavização.. A previsão um passo à frente para o tempo  $T + 1$  é uma média ponderada de todas as observações na série  $y_1, \dots, y_T$ . A taxa na qual os pesos diminuem é controlada pelo parâmetro  $\alpha$ .

A tabela abaixo mostra os pesos anexados às observações para quatro valores diferentes de  $\alpha$  ao prever usando suavização exponencial simples. Observe que a soma dos pesos, mesmo para um pequeno valor de  $\alpha$  será aproximadamente um para qualquer tamanho de amostra razoável.

	$\alpha = 0,2$	$\alpha = 0,4$	$\alpha = 0,6$	$\alpha = 0,8$
$y_T$	0,2000	0,4000	0,6000	0,8000
$y_{T-1}$	0,1600	0,2400	0,2400	0,1600
$y_{T-2}$	0,1280	0,1440	0,0960	0,0320
$y_{T-3}$	0,1024	0,0864	0,0384	0,0064
$y_{T-4}$	0,0819	0,0518	0,0154	0,0013
$y_{T-5}$	0,0655	0,0311	0,0061	0,0003

Para qualquer  $\alpha$  entre 0 e 1, os pesos atribuídos às observações diminuem exponencialmente à medida que voltamos no tempo, daí o nome “suavização exponencial”. Se  $\alpha$  é pequeno (isto é, próximo de 0), mais peso é dado às observações do passado mais distante. Se  $\alpha$  é grande (ou seja, próximo de 1), mais peso é dado às observações mais recentes.

- **Determinação da constante  $\alpha$**

Quanto menor for o valor de  $\alpha$  mais estáveis serão as previsões finais, uma vez que a utilização de baixo valor de  $\alpha$  implica que pesos maiores serão dados às observações passadas e, conseqüentemente, qualquer flutuação aleatória, no presente, exercerá um peso menor no cálculo da previsão.

Brown (1962) faz alguns comentários sobre a determinação dessa constante, de acordo com alguns critérios, tais como tipo de autocorrelação entre os dados e custo de previsão. Um