MBA USP ESALQ

SÉRIES TEMPORAIS

Prof. Fabiano Guasti Lima

*A responsabilidade pela idoneidade, originalidade e licitude dos conteúdos didáticos apresentados é do professor.

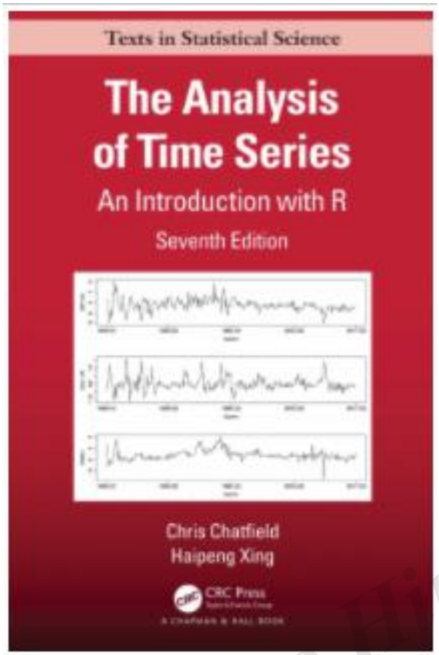
MBA USP ESALQ

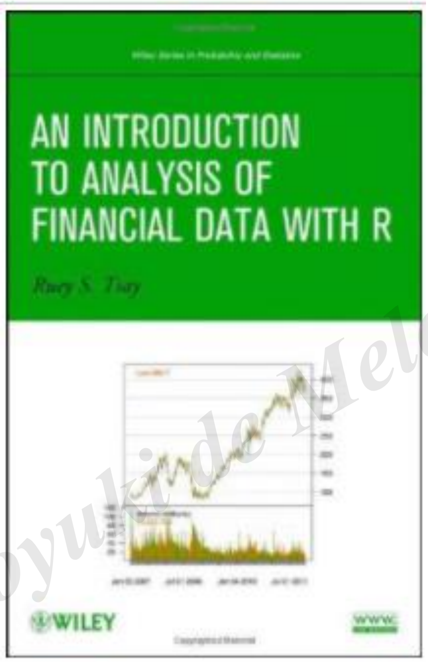
SÉRIES TEMPORAIS

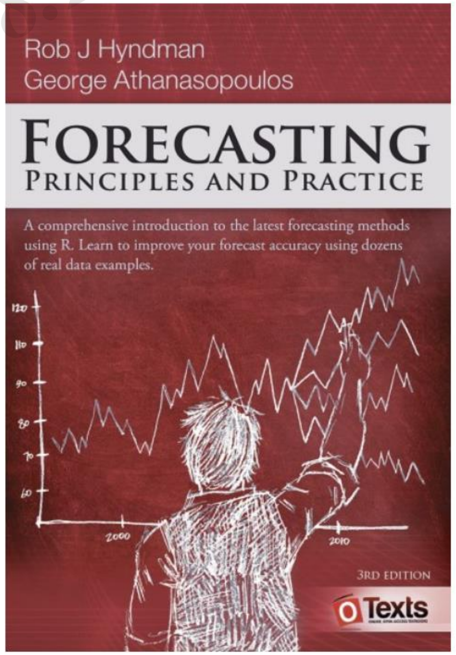
Tópicos

Leitura de dados em série temporal; Plotagem e decomposição de séries temporais; Método de Holt-Winters; Decomposição de séries não sazonais; Alisamento exponencial (exponential smoothing); Decomposição de séries sazonais e ajustamento sazonal; Forecast; Modelos ARIMA; Diferenciação de séries temporais; Seleção de modelos ARIMA; Forecast para modelos ARIMA; Exemplos e exercícios adicionais dos modelos estudados.

Referências

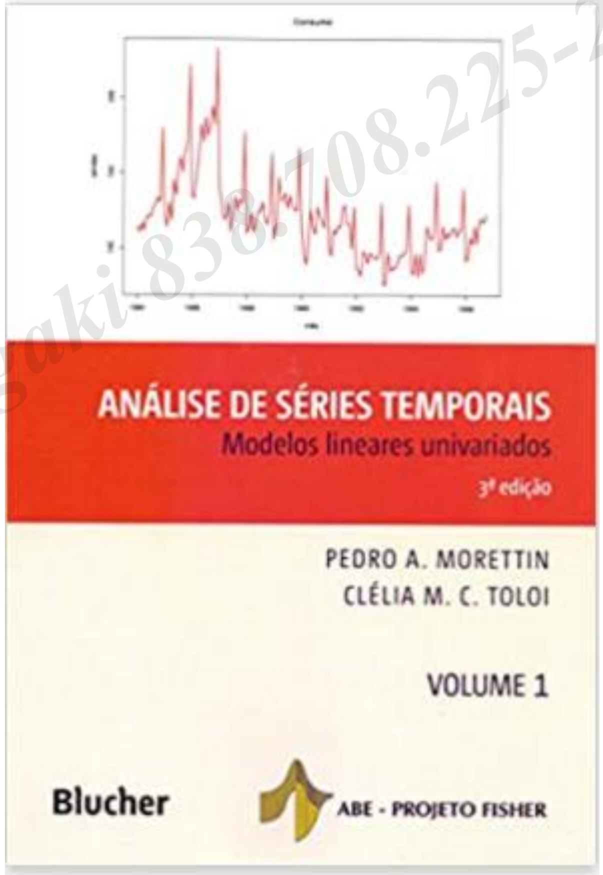






https://www.msperlin.com/adfeR/ https://otexts.com/fpp3/

Referências



SÉRIES TEMPORAIS

Conjunto de observações ordenadas no tempo.

$$\{X_t\}_{t=1}^{n} = \{X_1, X_2, \cdots, X_n\}$$

Se n>=50, chama-se sucessão cronológica.
- Ordem: dependência de ordem!

- Índice IBOVESPA diário;
- Retorno das ações da Petrobrás mensal;
- Índices Mensais da Inflação no Brasil;
- Taxas de Câmbio Real/US$ diário

ELEMENTOS DE UMA SÉRIE TEMPORAL

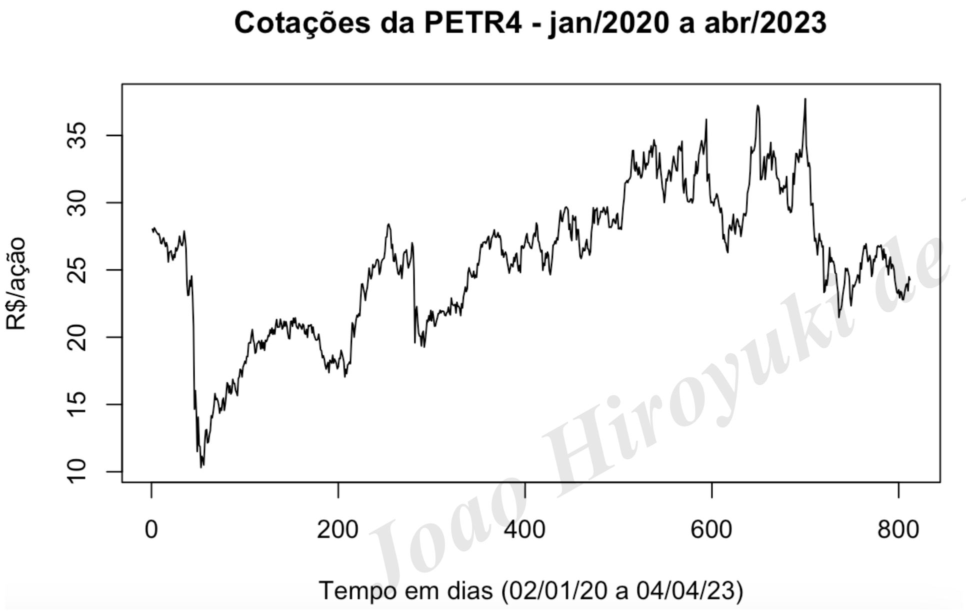
Índice IBOVESPA (B3) diário

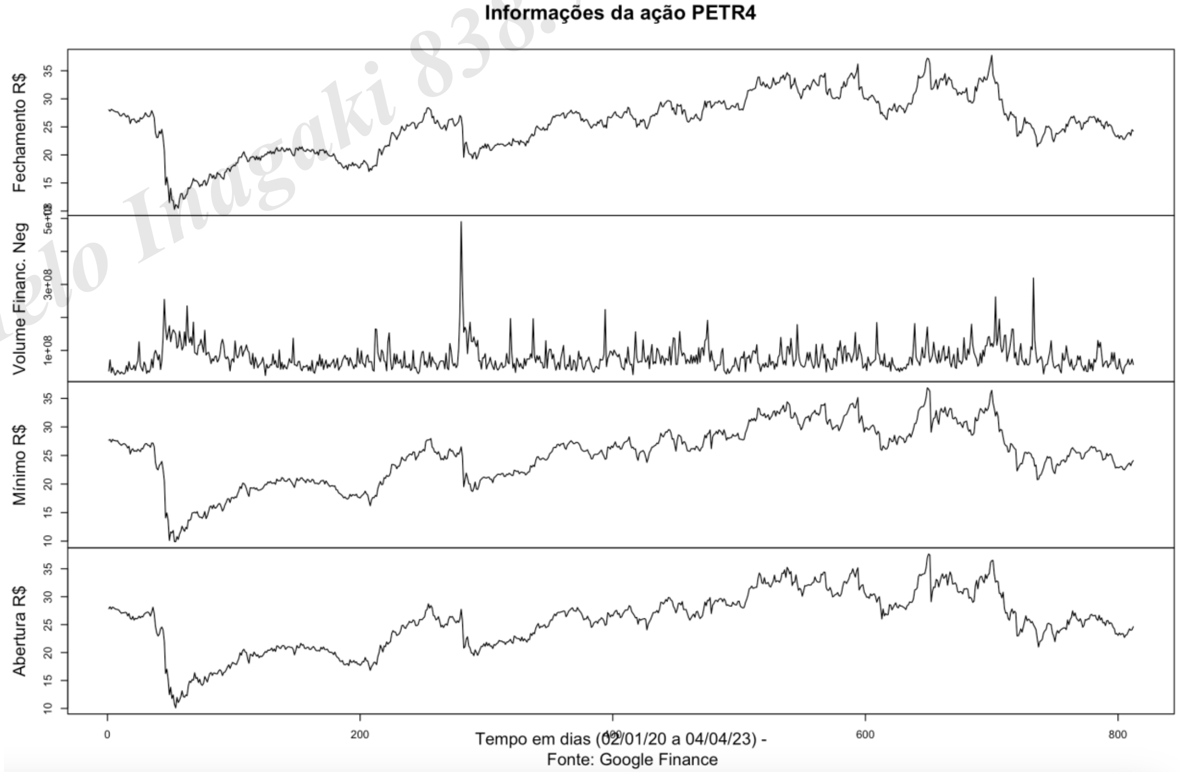
MEDIDA FATO Unidade de tempo

Fonte: B3, abril,2023 – Acesso em 05/04/2023

Univariadas e Multivariadas

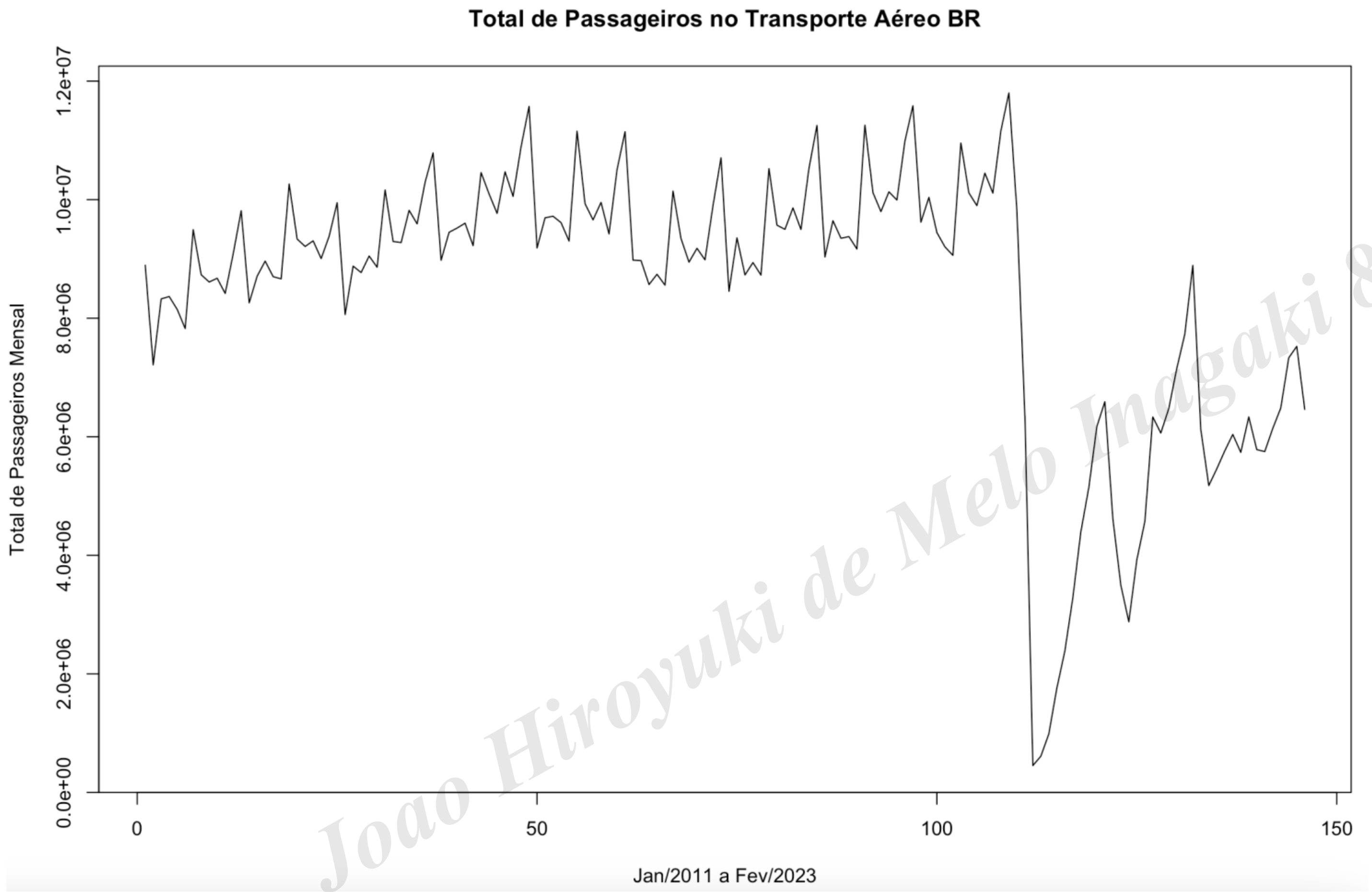
- **Univariadas:** apenas uma variável conectada ao tempo
- **Multivariadas:** duas ou mais variáveis conectadas ao tempo





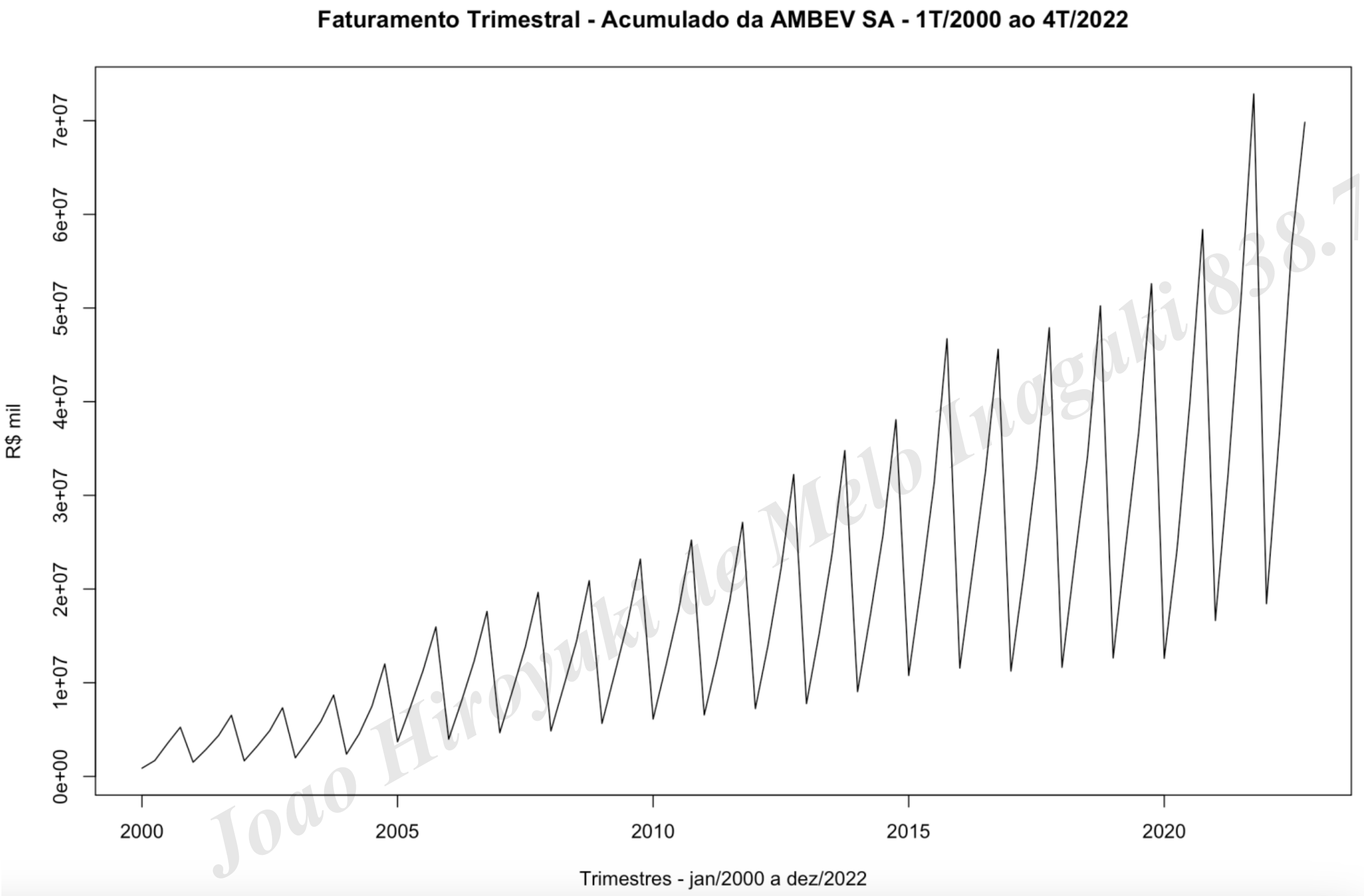
Fonte: googlefinance, abril/2023

Exemplos Séries Temporais



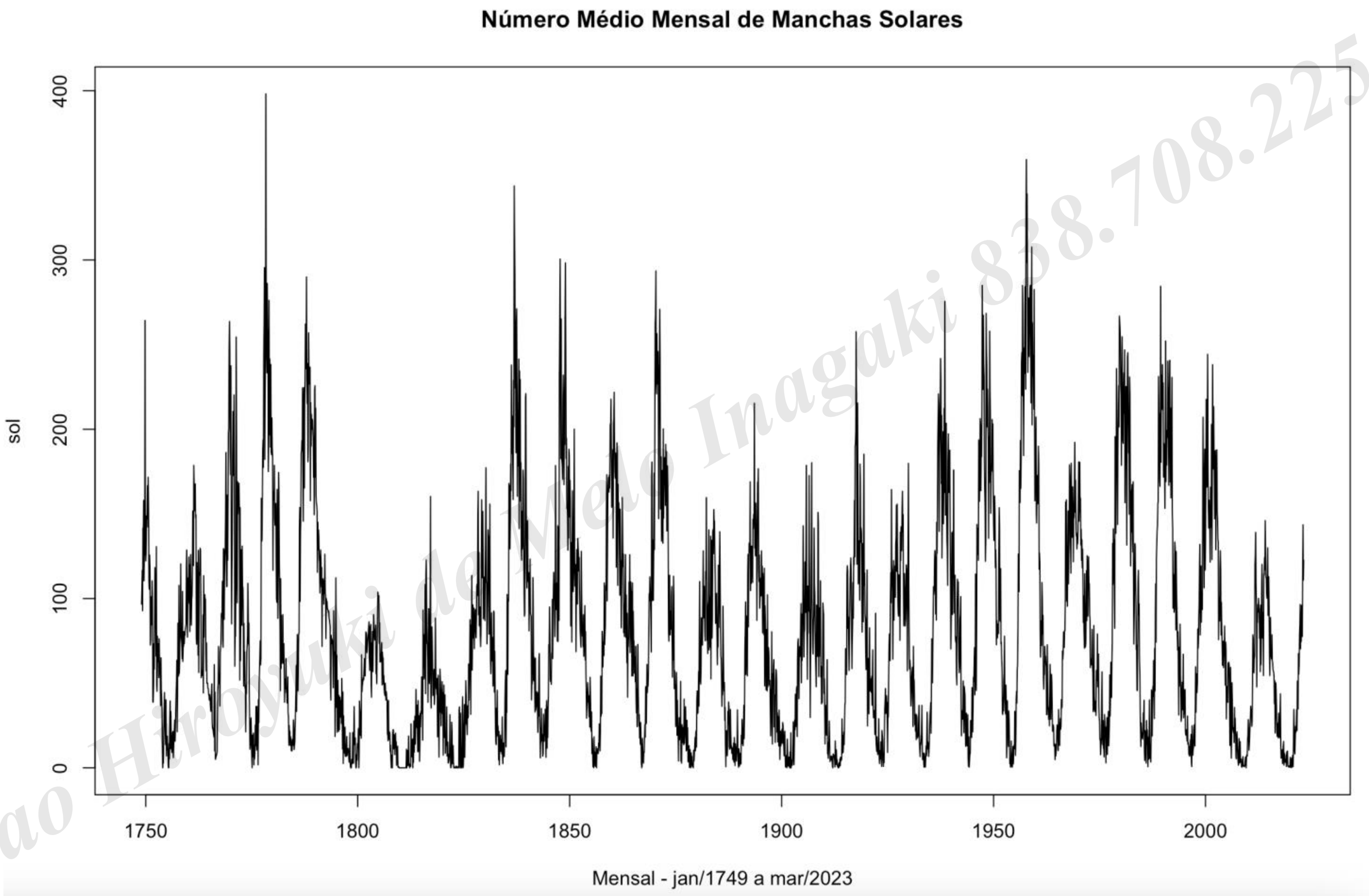
Fonte: https://www.gov.br/anac/pt-br/assuntos/dados-e-estatisticas/dados-estatisticos/dados-estatisticos

Exemplos Séries Temporais



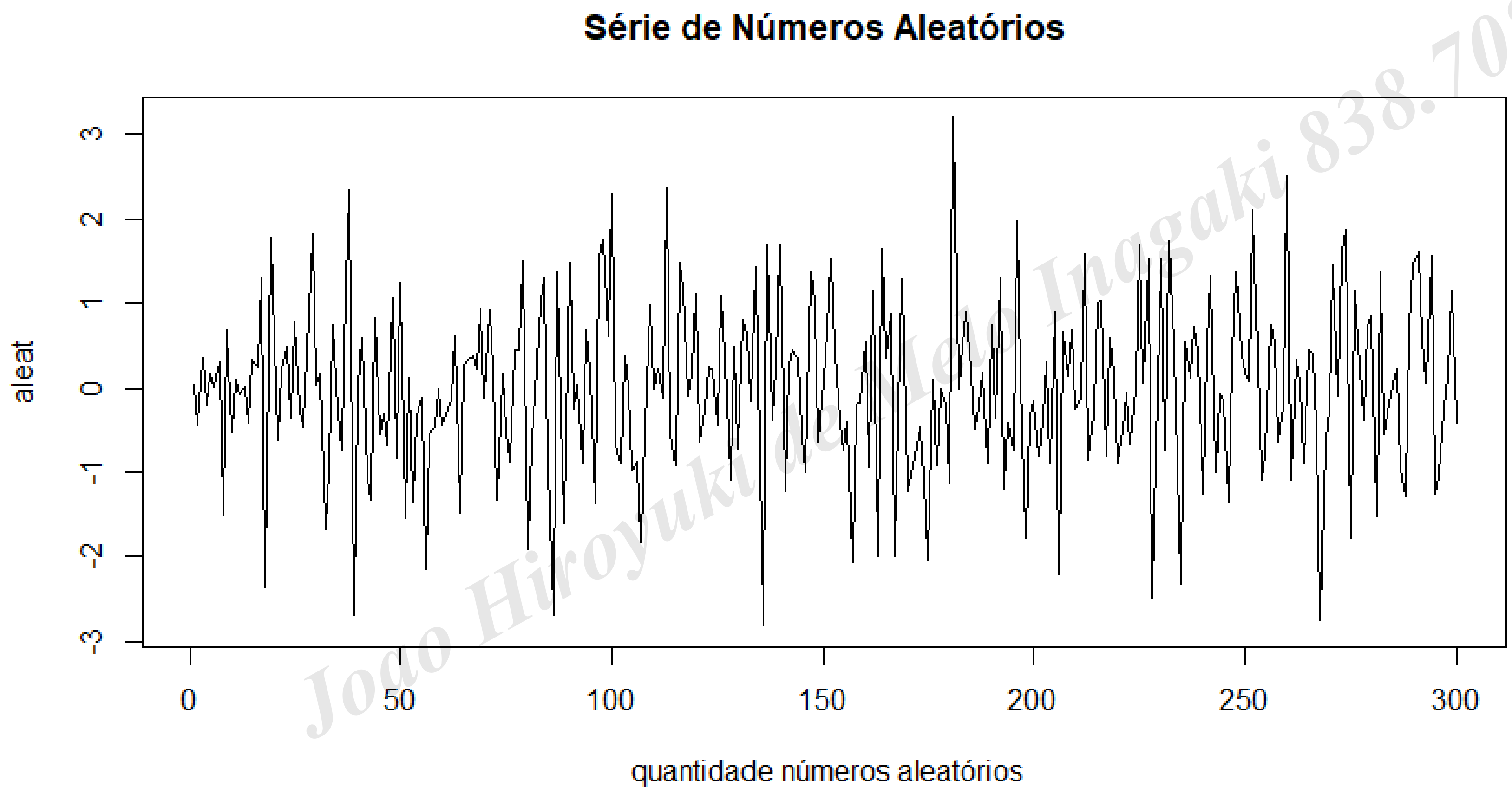
Fonte: CVM, dez/2022

Exemplos Séries Temporais

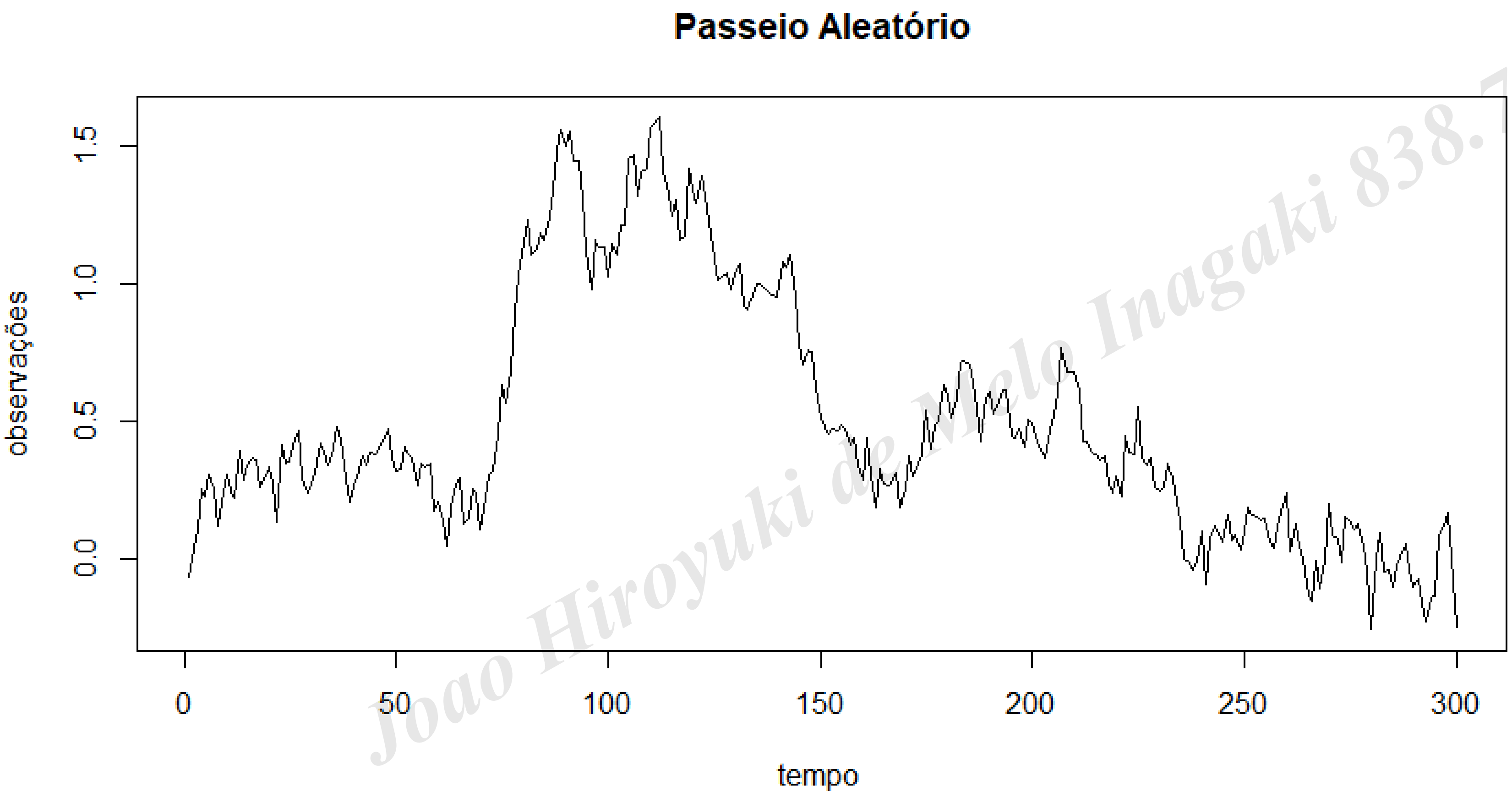


Fonte: http://sidc.be/silso/infosnmtot

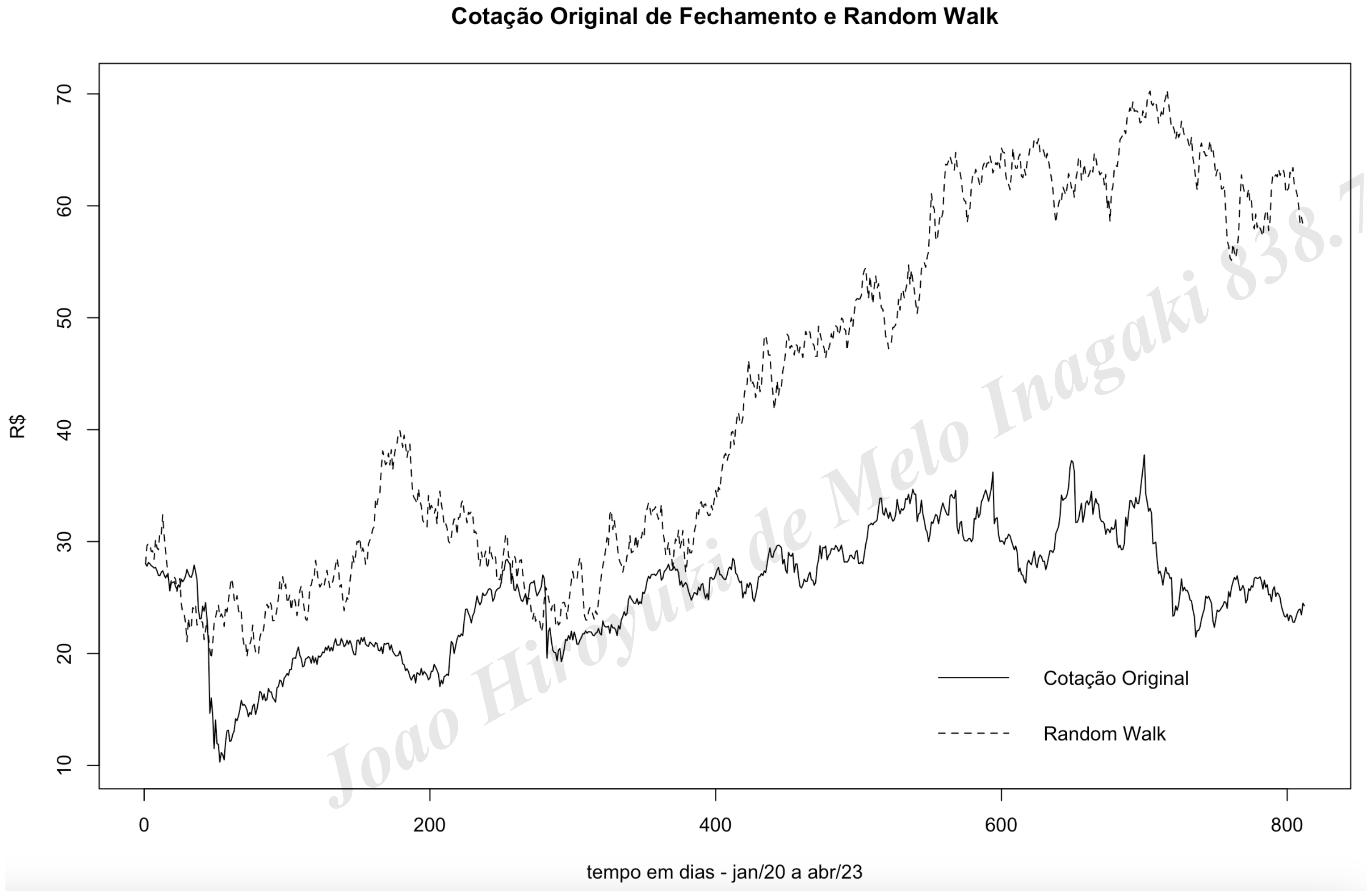
Exemplo Séries Temporais



Exemplos Séries Temporais



Exemplos Séries Temporais



Objetivos do Estudo de Séries de Tempo

Objetivos das análises de séries temporais

- Investigação do mecanismo que gera a série temporal
- Fazer previsões de valores futuros da série
- Descrever apenas o comportamento da série
- Procurar periodicidades relevantes nos dados

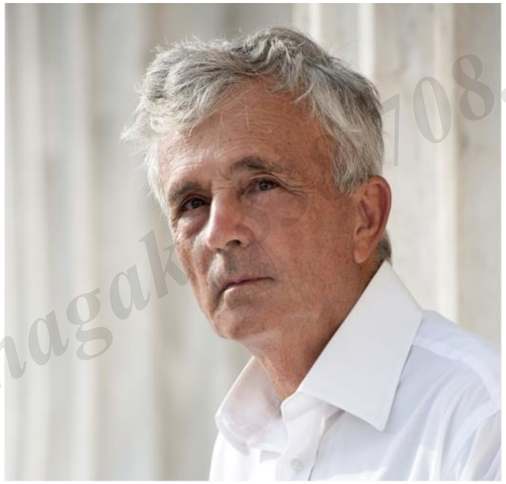
Histórico...

- **Stigler (1699)** – primeiro esquema "empírico" da demanda publicado por Charles Davenant;
- **Rodulfo Enini (1907)** – primeiros estudos;
- **1930** – Econometric Society;
- **Antes de 1955** – Modelos Clássicos de Decomposição;
- **1957 – 1962** – Modelos de Alisamento Exponencial (Holt-Winters e Brown);
- **Décadas de 60/70** – Modelos de Box-Jenskins (ARIMA);
- **Década de 80** – Modelos estruturais clássicos e bayesianos (Filtro de Kalman);
- **Década de 80/90** – Cointegração e econometria de Séries Temporais.

Séries Temporais

George Box (1919-2013)

Fonte: Google Imagens



Gwilym Jenkins (1932-1982)

Todos os modelos
estão errados, mas
alguns são úteis.

MBA USP ESALQ

Classificação das Séries Temporais

- **Discretas**: são séries em que o intervalo de observações (t) pertence a um conjunto discreto. Ou seja, as observações são feitas em intervalos de tempo fixos.

- **Contínuas**: são séries em que as observações são obtidas continuamente através de algum intervalo no tempo.

Classificação das Séries Temporais

- **Determinística**: quando pode ser descrita por uma função matemática para estabelecer exatamente os valores futuros da série.

- **Estocástica**: quando os valores futuros da série somente podem ser estabelecidos em termos probabilísticos, pois o modelo compõe-se também de um termo aleatório.

Métodos Estudados

SIMPLES	CLÁSSICA	OUTRAS
Naive	Suavização Exponencial	Regressão
Mean	ARIMA	Redes Neurais
Drift		Decomposição

Métodos SIMPLES

- **NAIVE**: Projeta o último valor para o futuro
- **NAIVE SAZONAL**: Considera o último valor no mesmo período de tempo (para séries com sazonalidade)
- **Média**: usa a média histórica como previsão para o futuro
- **Drift**: faz uma previsão que acompanha a tendência da série (equivale a traçar uma renta entre o primeiro e o último ponto)

Métodos SIMPLES

- **NAIVE**: Projeta o último valor para o futuro

$$\{X_t\}_{t=1}^{n} = \{X_1, X_2, \cdots, X_n\}$$

$$X_{t+1} = X_t$$

$$X_{t+h} = X_t \pm Z_{\alpha/2}\% \,.\, \sigma_{resíduos} \,.\, \sqrt{h}$$

Métodos SIMPLES

- **NAIVE SAZONAL**: Projeta o último período sazonal

$$\{X_t\}_{t=1}^{n} = \{X_1, X_2, \cdots, X_n\}$$

$$X_{t+1} = X_{t+m}$$

$$X_{t+h} = X_{t+m} \pm Z_{\alpha/2}\% \, . \, \sigma_{resíduos} \, . \, \sqrt{k+1}$$

$$k = (\text{parte inteira}) \left(\frac{h-1}{m} \right)$$

Métodos SIMPLES

- **MÉDIA**: média de toda a série

$$\{X_t\}_{t=1}^{n} = \{X_1, X_2, \cdots, X_n\}$$

$$X_{t+1} = Média = \frac{1}{n}.\sum_{t=1}^{n} X_t$$

$$X_{t+h} = Média \pm t_{\alpha/_2\%} .\ \sigma_{resíduos} .\ \sqrt{1+\frac{1}{n}}$$

Métodos SIMPLES

- **DRIFT**: equivale a traçar uma reta entre o primeiro e o último ponto

$$\{X_t\}_{t=1}^{n} = \{X_1, X_2, \cdots, X_n\}$$

$$X_{t+h} = X_t + h.\frac{X_t - X_1}{t - t_0}$$

$$X_{t+h} = X_t \pm Z_{\alpha/2}\% \cdot \sigma_{resíduos} \cdot \sqrt{h.\left(1 + \frac{h}{n_{resíduos} - 1}\right)}$$

Estatísticas de Erro das Previsões

- **ME**: Mean Error – É a média da diferença entre realizado e o previsto.

$$erro_t = X_t - \hat{X}_t$$

$$ME = \frac{\sum_{t=1}^{h} erro_t}{h}$$

- **MAE**: Mean Absolute Error – É a média da diferença absoluta entre realizado e previsto

$$MAE = \frac{\sum_{t=1}^{h} |erro_t|}{h}$$

Estatísticas de Erro das Previsões

- **RMSE**: Root Mean Square Error – É o desvio padrão total da amostra da diferença entre o previsto e o realizado.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{h}(erro_t)^2}{h}}$$

Estatísticas de Erro das Previsões

- **MPE**: Mean Percentage Error – É a diferença percentual do erro.

$$MPE = \frac{\sum_{t=1}^{h} \frac{erro_t}{X_t}}{h} \times 100\%$$

- **MAPE**: Mean Absolute Percentage Error – É a diferença absoluta percentual do erro.

$$MAPE = \frac{\sum_{t=1}^{h} \frac{|erro_t|}{X_t}}{h} \times 100\%$$

Estatísticas de Erro das Previsões

- **TIC**: Theil Inequality Coefficient – **Theil´s U**

É o grau de ajuste da previsão. Quanto menor, melhor. Zero ideal.

$$Theil's\ U = \sqrt{\frac{\sum_{t=1}^{h}\left(\frac{\hat{X}_{t+1} - X_{t+1}}{X_t}\right)^2}{\sum_{t=1}^{h}\left(\frac{X_{t+1} - X_t}{X_t}\right)^2}}$$

Estatísticas de Erro das Previsões

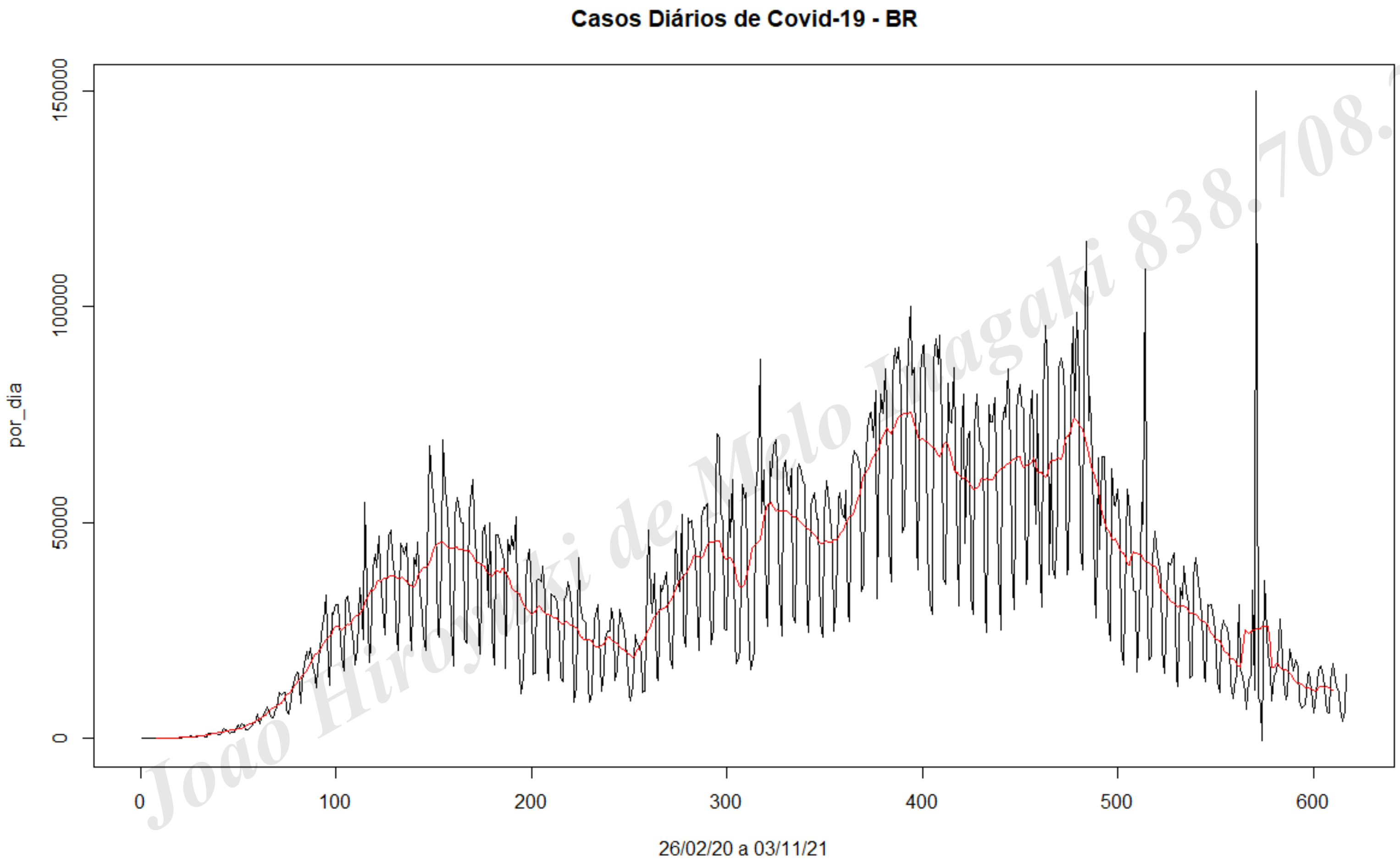
- **ACF1**: First-Order Autocorrelation Function – Aucorrelação dos resíduos.

$$ACF_k = \frac{cov(R_{it}, R_{i,t-k})}{variância(R_{it})}$$

Métodos CLÁSSICOS

- **DECOMPOSIÇÃO**: projeções por decomposição da série
- **SUAVIZAÇÃO EXPONENCIAL**: método de amortecimento (suavização), ideal para tendências e inclui variação sazonal
  - **Aditivo**: para variação sazonal constante
  - **Multiplicativo**: variação sazonal varia na série

Médias Móveis



Fonte: https://covid.saude.gov.br/

Componentes de uma série temporal

TENDÊNCIA

Movimento oculto no dados, seguindo uma direção - crescente, decrescente ou estacionária

sazonal

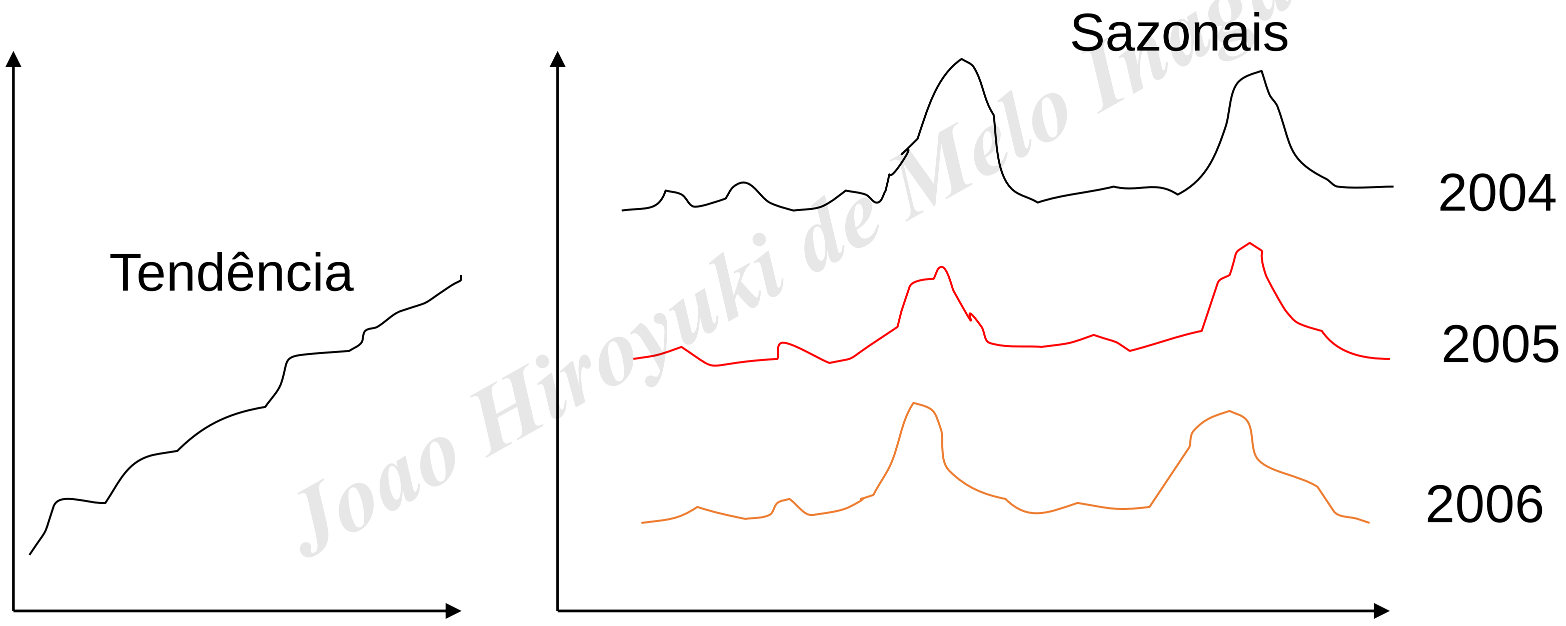
Flutuações regulares dentro de um período completo de tempo(dia, semana, mês, etc.)
- Representam um tipo de padrão que se repete. (picos, depressões) normalmente dentro de um ano

CÍCLICA

Flutuações de **longo** prazo nos dados e são similares aos fatores sazonais. Padrão que se repete com regularidade mas sem período fixo

Série com tendência e sazonalidade

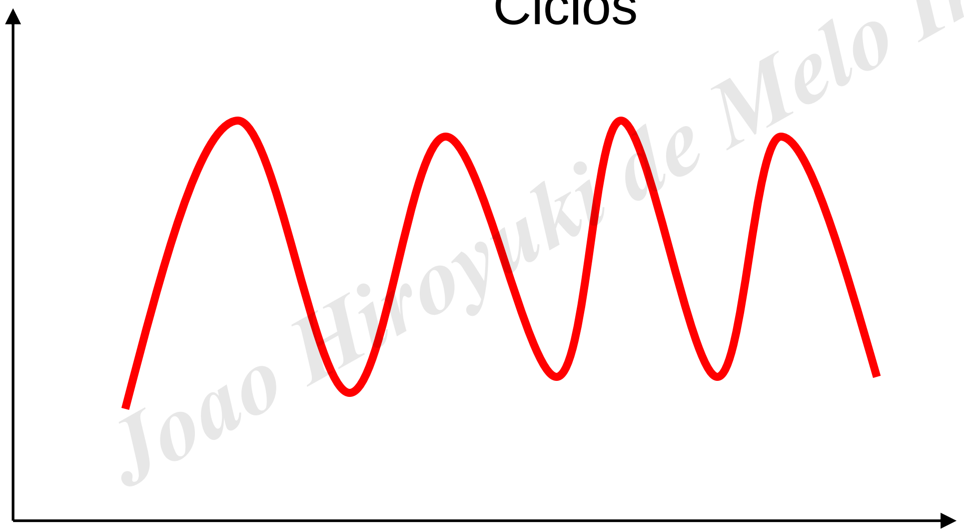
Sazonalidade são as flutuações regulares dentro de um período completo de tempo (um dia, uma semana, um mês, etc).
O importante sobre fatores sazonais é que eles representam um tipo de padrão que se repete.



Série com ciclo

Ciclos São flutuações a longo prazo nos dados e são similares aos fatores sazonais. Eles podem ser difíceis de serem identificados a menos que uma série de dados longa esteja disponível.

Ciclos



Série com tendência, sazonalidade e variações cíclicas

- Muitas séries apresentam junto a uma tendência variações cíclicas e sazonais;
- Estas variações aparecem devido a clima, fatores econômicos, hora, etc;
- Se estas variações podem ser observadas, a sua consideração pode ajudar a melhorar as previsões.
- São usados o método da decomposição:

- MULTIPLICATIVO
- ADITIVO

Aditivo

$$Y = T+C+S+E$$

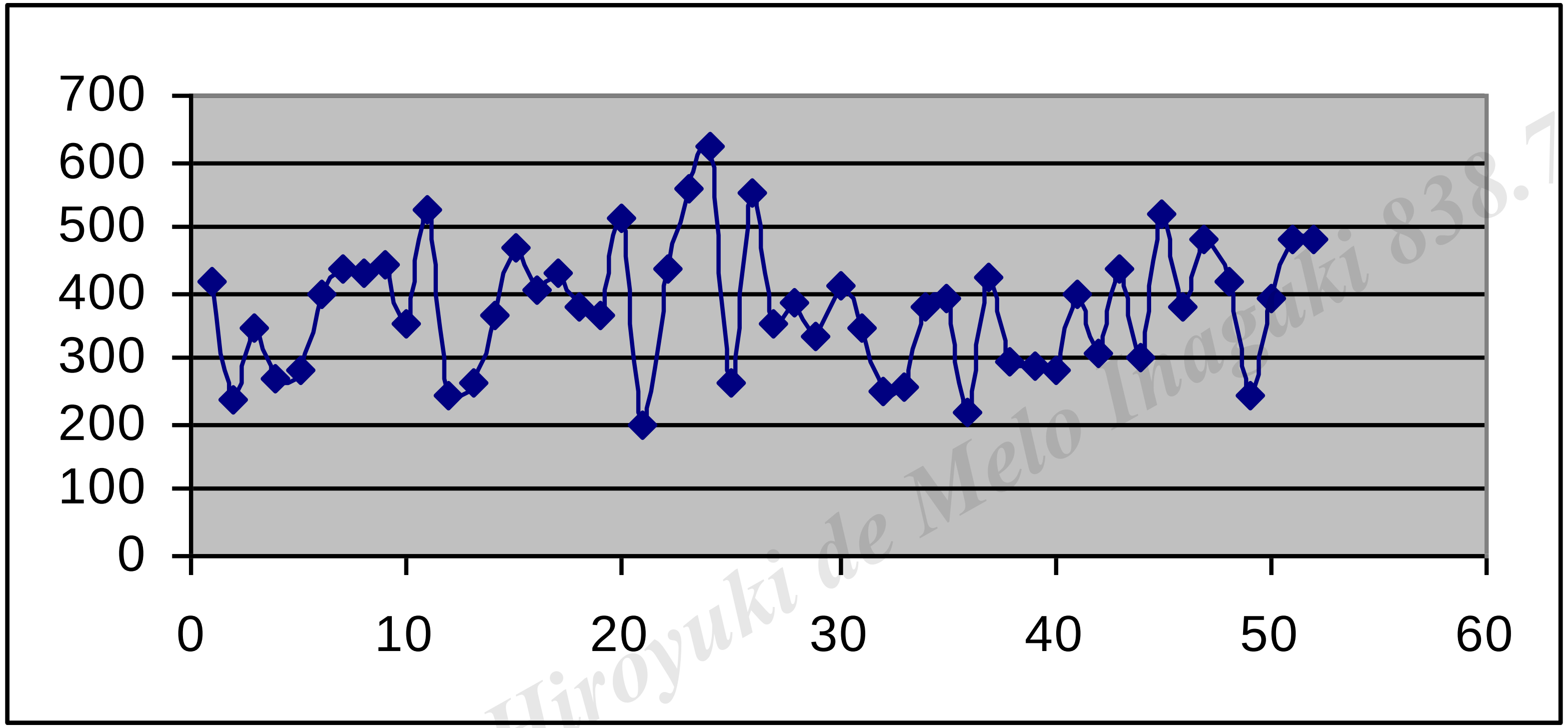
- Y - valor da série no instante t
- T - componente de tendência para o instante t
- C - componente cíclica para o instante t
- S - componente sazonal para o instante t
- E - componente aleatória para o instante t

Multiplicativo

$$Y = T*C*S*E$$

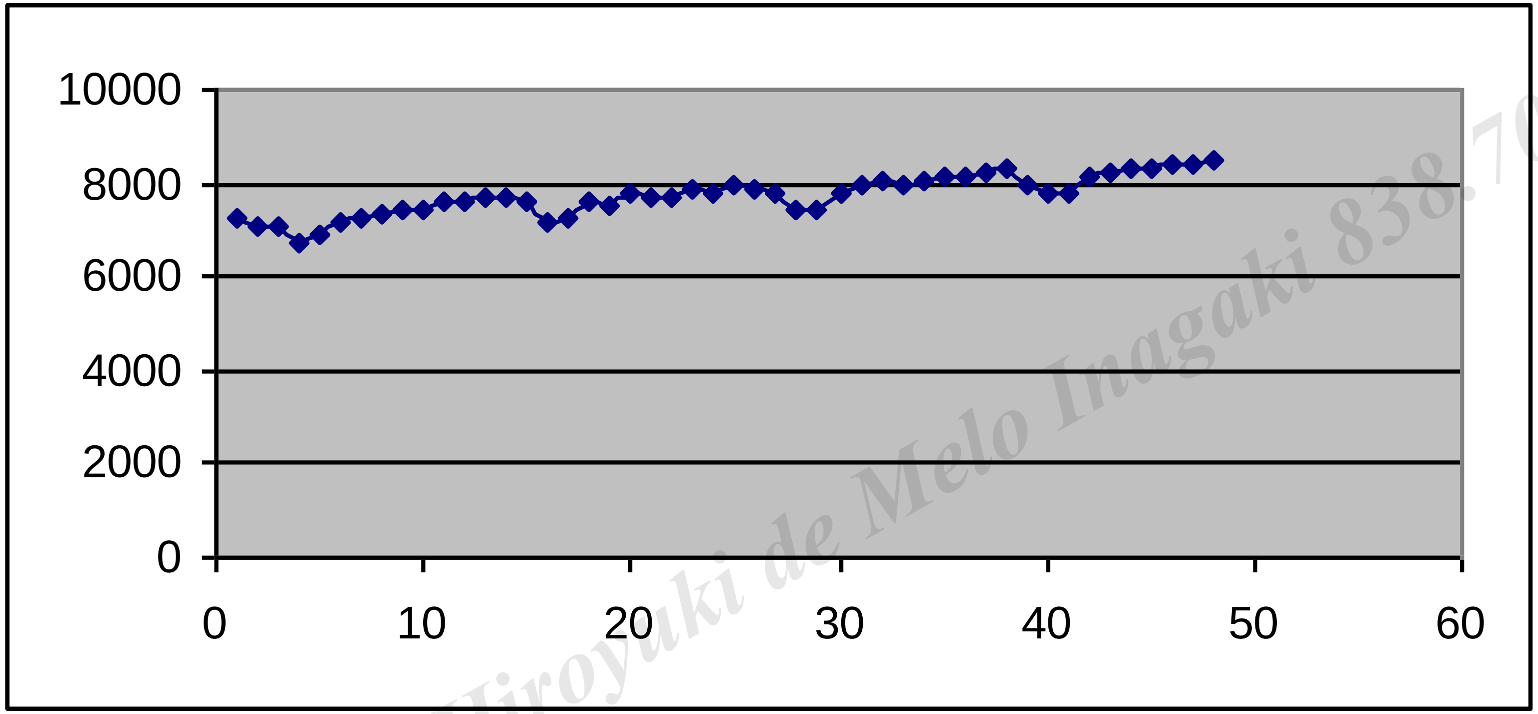
- Y - valor da série no instante t
- T - componente de tendência para o instante t
- C - componente cíclica para o instante t
- S - componente sazonal para o instante t
- E - componente aleatória para o instante t

Modelo Multiplicativo



O modelo multiplicativo é normalmente aplicado a dados em que o tamanho dos efeitos sazonais aumentam.

Modelo Aditivo



O modelo aditivo geralmente é considerado mais adequado para dados em que as flutuações sazonais permanecem aproximadamente do mesmo tamanho com o tempo.

Modelos de Suavização Exponencial (SES)

- Série temporal que não apresenta tendência e nem sazonalidade;
- Série temporal com Tendência mas sem sazonalidade – Suavização Exponencial de Holt (SEH)
- Série temporal com Tendência e Sazonalidade – Suavização Exponencial de Holt-Winters

Suavização Exponencial Simples (SES)

- Dá pesos maiores às observações mais recentes captando melhor as mudanças de comportamento.
- Previsão é igual ao último valor exponencial suavizado.

$$\{X_t\}_{t=1}^{n} = \{X_1, X_2, \cdots, X_n\}$$

$$\hat{X}_{t+1} = \alpha X_t + (1-\alpha)\hat{X}_t$$

$$\widehat{X_1} = X_1, \qquad 0 \le \alpha \le 1, \qquad t = 1, \dots, n$$

Suavização Exponencial Simples (SES) – Previsão

$$\hat{X}_t(h) = \hat{X}_t \pm Z\alpha_{/2\%}.SE$$

$$SE = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n-1}\left(\hat{X}_t - X_t\right)^2}$$ Erro para h = 1

$$SE_h = SE \times \sqrt{1 + (h-1).\alpha^2}$$

Suavização Exponencial de Holt (SEH)

- Para Séries temporais com tendência linear

$$\{X_t\}_{t=1}^{n} = \{X_1, X_2, \cdots, X_n\}$$

$$\hat{u}_t = \alpha X_t + (1 - \alpha)(\hat{u}_{t-1} + \hat{T}_{t-1})$$

$$\hat{T}_t = \beta(\hat{u}_t - \hat{u}_{t-1}) + (1 - \beta)(\hat{T}_{t-1})$$

$$\hat{X}_{t+1} = u_t + \hat{T}_t$$

$$\hat{u}_1 = X_1, \qquad \hat{T}_1 = 0, \qquad 0 \leq \alpha \leq 1, \quad 0 \leq \beta \leq 1 \qquad t = 1, \ldots, n$$

Suavização Exponencial de Holt (SEH) – Previsão

$$\hat{X}_t(h) = \hat{X}_t \pm Z\alpha_{/2\%}.SE$$

$$SE = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n-1}(\hat{X}_t - X_t)^2}$$

Suavização Exponencial de Holt (SEH) – Previsão

$$\hat{X}_t(h) = \hat{X}_t \pm Z_{\alpha/2} . SE(h)$$

$$SE(h) = \mathrm{SE}.\sqrt{1 + k.\alpha^2}$$

$$k = \sum_{i=2}^{h} (1 + \beta.(i-1))^2$$

Suavização Exponencial de Holt-Winters (HW - Aditivo)

- Para Séries temporais com comportamento sazonal (c=período sazonal).
  - **Modelo Aditivo**

$$\hat{X}_t = \hat{L}_{t-1} + \hat{T}_{t-1} + \hat{S}_{t-c} \qquad \{X_t\}_{t=1}^{n} = \{X_1, X_2, \cdots, X_n\}$$

$$\hat{L}_t = \alpha\left(\hat{X}_t - \hat{S}_{t-c}\right) + (1-\alpha).\,(L_{t-1} + T_{t-1})$$

$$\hat{T}_t = \beta\left(\hat{L}_t - \hat{L}_{t-1}\right) + (1-\beta).\,T_{t-1}$$

$$\hat{S}_t = \gamma\left(\hat{X}_t - \hat{L}_{t-1}\right) + (1-\gamma).\,S_{t-c}$$

$$0 \leq \alpha \leq 1$$
$$0 \leq \beta \leq 1$$
$$0 \leq \gamma \leq 1 - \alpha$$

$$\hat{X}_{t+h} = \hat{L}_t + h\hat{T}_t + \hat{S}_{t+h-ch'}$$

$$h' = INT\left({}^{(h-1)}/_{c}\right) + 1$$

Suavização Exponencial de Holt-Winters (HW - Aditivo) – Previsão

$$\hat{X}_t(h) = \hat{X}_t \pm Z_{\alpha/_2}\,.SE(h)$$

$$SE(h) = \mathrm{SE}.\sqrt{1 + \sum_{i=1}^{h-1} \varphi_i^2}$$

$$\varphi_i = \begin{cases} \alpha(1+\beta i), c \neq i-1 \\ \alpha(1+\beta i) + \gamma(1-\alpha), c = i-1 \end{cases}$$

Suavização Exponencial de Holt-Winters (HW - Multiplicativo)

- Para Séries temporais com comportamento sazonal (c=período sazonal).
  - **Modelo Multiplicativo**

$$\hat{X}_t = (\hat{L}_{t-1} + \hat{T}_{t-1}).\hat{S}_{t-c} \qquad \{X_t\}_{t=1}^{n} = \{X_1, X_2, \cdots, X_n\}$$

$$\hat{L}_t = \alpha\left(\hat{X}_t / \hat{S}_{t-c}\right) + (1-\alpha).(L_{t-1} + T_{t-1})$$

$$\hat{T}_t = \beta\left(\hat{L}_t - \hat{L}_{t-1}\right) + (1-\beta).T_{t-1}$$

$$\hat{S}_t = \gamma\left(\hat{X}_t / \hat{L}_{t-1}\right) + (1-\gamma).S_{t-c}$$

$$0 \leq \alpha \leq 1$$
$$0 \leq \beta \leq 1$$
$$0 \leq \gamma \leq 1-\alpha$$

$$\hat{X}_{t+h} = (\hat{L}_t + \hat{T}_t).\hat{S}_{t+h-ch'}$$

$$h' = INT\left({}^{(h-1)}/_{c}\right) + 1$$

Previsão
Modelo ETS (Error, Trend, Seasonal)

- Definirá o melhor modelo

Erro	Tendência	Sazonalidade
A	N	N
M	A	A
Z	M	M
	Z	Z

model = “AAA”

Legenda	
A	Aditivo
M	Multiplicativo
N	Nenhum
Z	Automático

Tendência	
Ad	Aditivo Amortecido
Md	Multiplicativo Amortecido

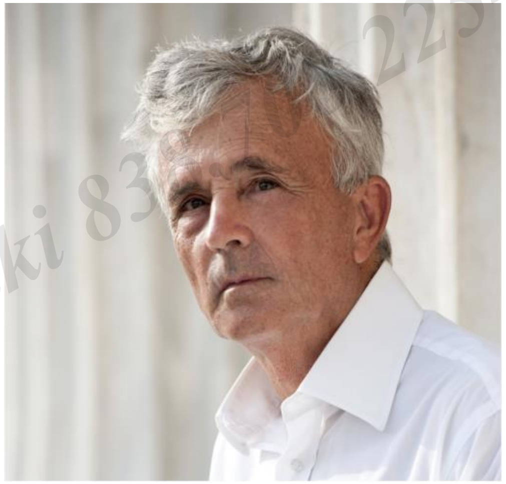
Modelos de Séries Temporais

- **Série Estacionária**: movimento de tendência não é significativo ao longo do tempo.

Série Estacionária	Série Não-Estacionária
○ Média móvel ○ Média móvel ponderada ○ Alisamento exponencial	○ Tendência linear ○ Método de Holt

MODELOS A R I M A

George Box (1919-2013)



Gwilym Jenkins (1932-1982)

Fonte: Google Imagens

“ Nos modelos ARIMA os dados falam por si mesmo”

Modelos ARIMA (Box-Jenkins)

- Robusto: pode ser aplicado em praticamente qualquer tipo de série temporal
- Funciona melhor com dados estáveis, com poucos *outliers* (embora podemos removê-los) - tsclean
- **Requer dados estacionários**
  - Pode ser transformada usando diferenciação: remove tendências
  - Diferenciação: subtrai a observação atual da anterior
  - Diferenciação pode ser feita 1x: diferenciação de primeira ordem
  - Diferenciação 2x: diferenciação de segunda ordem (mais raro)

Tipos de Modelos – ARIMA não Sazonal

- **Modelos auto-regressivos (AR)**: avalia a relação entre os períodos (*lags*): autocorrelação – extrai a influência;
- **Integrado (I)**: Aplicado à diferenciação, quando necessário
- **Modelos médias móveis (MA)**: avalia erros entre períodos e extrai esses erros;
- **Modelos auto-regressivos e de médias móveis (ARMA)**
- **Modelos auto-regressivos integrados e de médias móveis (ARIMA)**

Modelos ARIMA

ARIMA

(p,d,q)

q: ordem da média móvel

d: grau de diferenciação

p: ordem da parte autorregressiva

Modelos ARIMA

- p = 1, significa que uma determinada observação pode ser explicada pela observação prévio + erro
- p = 2, significa que uma determinada observação pode ser explicada por duas observação prévias + erro

- d = 0, significa que não é aplicada diferenciação
- d = 1, significa que será aplicada diferenciação de primeira ordem
- d = 2, significa que será aplicada diferenciação de segunda ordem

- q = 1, significa que uma determinada observação pode ser explicada pelo erro da observação prévia
- q = 2, significa que uma determinada observação pode ser explicada pelo erro de duas observações prévias.

Modelos ARIMA

- AR(1) ou ARIMA(1,0,0) – Apenas elemento auto-regressivo de 1ª ordem
- AR(2) ou ARIMA(2,0,0) – Apenas elemento auto-regressivo de 2ª ordem

- MA(1) ou ARIMA(0,0,1) – Apenas média móvel

- ARMA(1,1) ou ARIMA(1,0,1) – Auto-regressão e média móvel de 1ª ordem

Modelos ARIMA – Como definir os valores de p, d, e q

- p: ordem da parte autorregressiva: PACF
- d: grau de diferenciação – Teste de Estacionariedade
- q: ordem da média móvel: ACF

Escolha dos Modelos

- Critério de AIC – Critério de Informação de Akaike
  - O AIC estima a quantidade relativa de informações perdidas por um determinado modelo: quanto menos informações um modelo perde, maior a qualidade desse modelo e menor a pontuação AIC.

- Critério de BIC – Critério de Informação Bayesiano
  - BIC mais baixo implica em melhor ajuste.

Modelos ARIMA

- Série Estacionária: média e a autocovariância são constantes no tempo.
- Formas detectar estacionariedade:
  - Gráfico ACF da série (processo não estacionário apresenta lento decaimento da sua função de autocorrelação.
    - Série com tendência é o motivo mais comum para não estacionariedade

Modelos Simulados

$$AR(1)$$

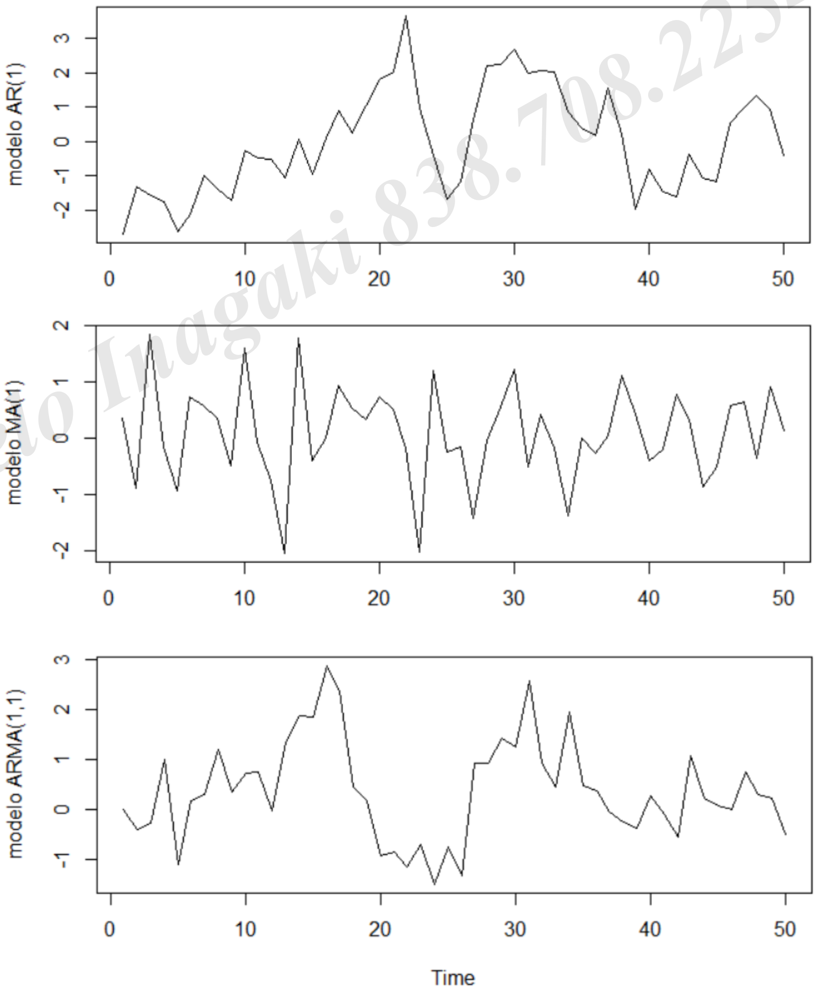
$$X_t = 0{,}8X_{t-1} + \varepsilon_t$$

$$MA(1)$$

$$X_t = -0{,}3\varepsilon_{t-1} + \varepsilon_t$$

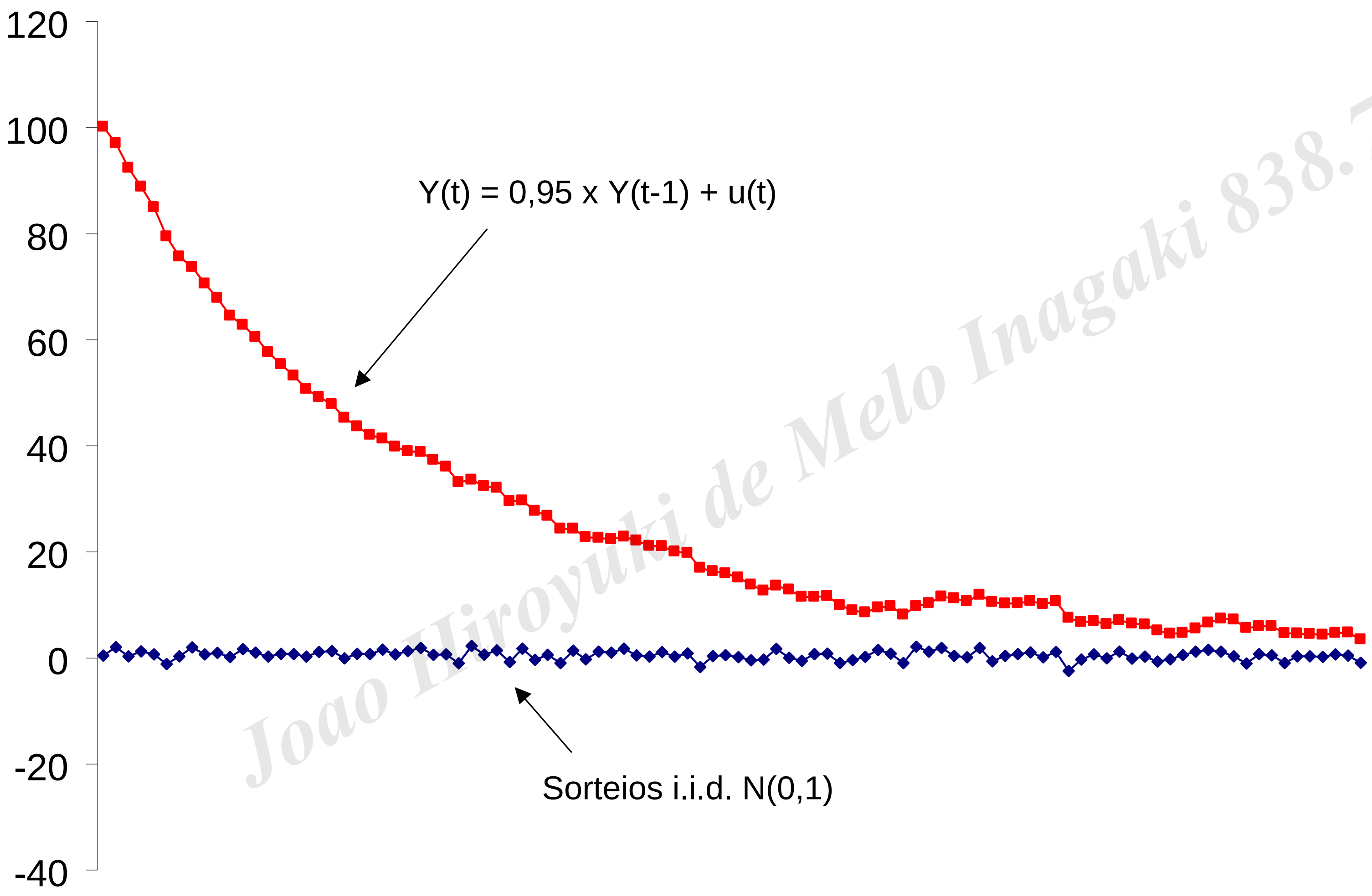
$$ARMA(1{,}1)$$

$$X_t = -0{,}8X_{t-1} - 0{,}3\varepsilon_{t-1} + \varepsilon_t$$



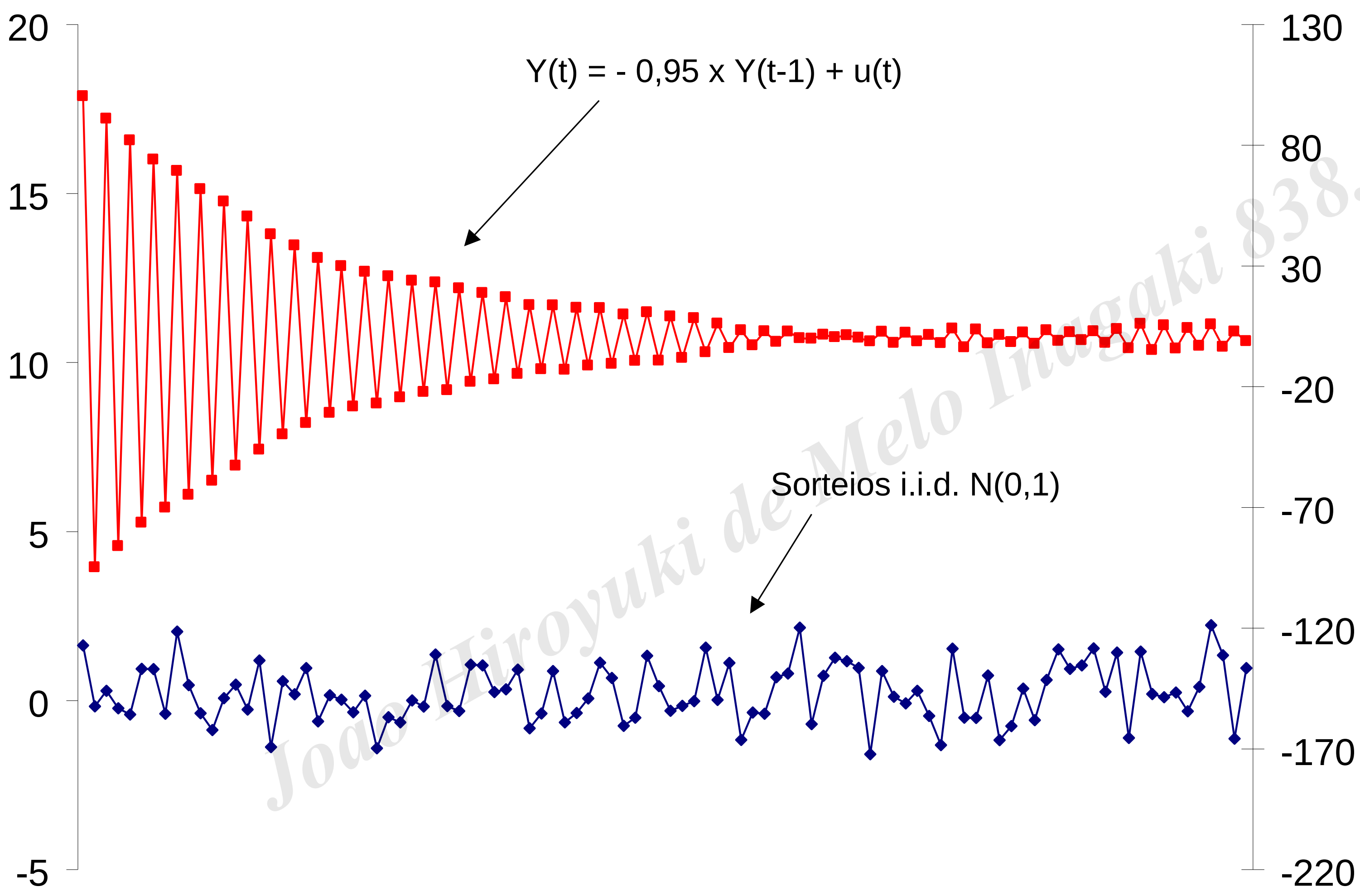
Exemplo: Processo AR(1) de baixa frequência

$$Y_t = 0{,}95 \times Y_{t-1} + u_t, \quad Y_0 = 100, \quad u_t \sim N(0{,}1)$$



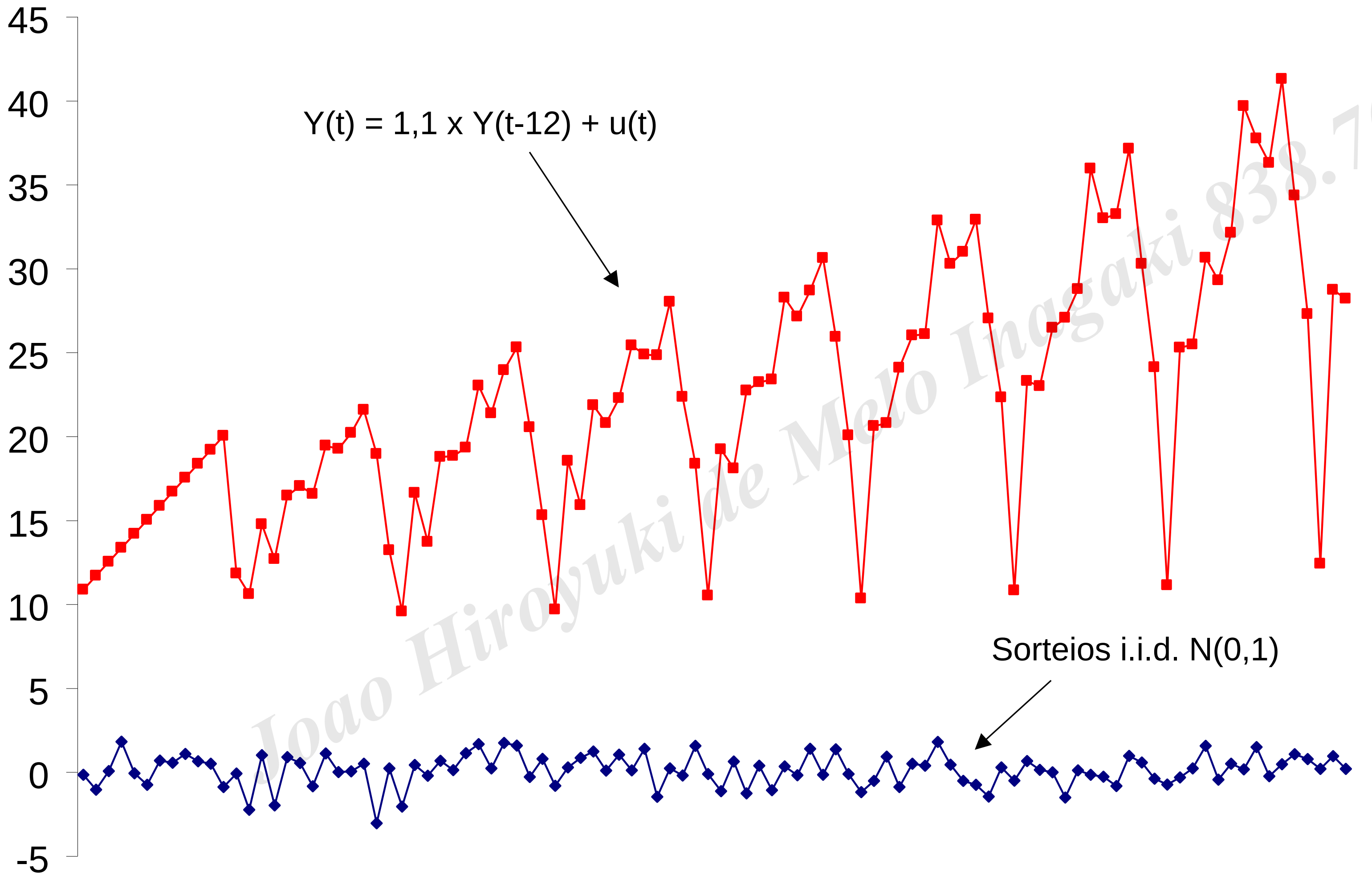
Exemplo: Processo AR(1) de baixa frequência

$$Y_t = -0{,}95 \times Y_{t-1} + u_t, \quad Y_0 = 100, \quad u_t \sim N(0{,}1)$$



Exemplo: Processo AR(12):

$$Y_t = 1{,}1 \times Y_{t-12} + u_t\,, \quad Y_0 = 10, \quad u_t \sim N(0{,}1)$$



Modelos Auto-regressivos (AR)

- Os valores correntes de uma série Y_t dependem apenas de seus valores passados e dos erros aleatórios.

- Exemplo AR(p): p é o número de defasagens

Modelo AR(1)

$Y_t = \phi_1 Y_{t-1} + \varepsilon_t$

Teste de Estacionariedade

[...] um processo estocástico é estacionário se suas média e variância forem constantes ao longo do tempo e o valor da covariância entre dois períodos de tempo depender apenas da distância ou defasagem entre os dois períodos, e não do período de tempo efetivo em que a covariância é calculada (ENDERS, 2003).

Teste KPSS – Kwiatkowski-Phillps-Schmidt - Shin

H_0 : a série é estacionária (Não apresenta raiz unitária)

H_1 : a série Não é estacionária (Possui raiz unitária)

Teste de Estacionariedade
Teste de Dickey-Fuller

Teste feito porque não se sabe se a série temporal possui mais de uma raiz unitária, pois o número de termos de diferenças defasadas é, muitas vezes, determinado empiricamente.

$$H_0: \text{a série não é estacionária (apresenta raiz unitária)}$$
$$H_1: \text{a série é estacionária (sem raiz unitária)}$$

Teste PP – Phillips-Perron

$$H_0: \text{a série não é estacionária (apresenta raiz unitária)}$$
$$H_1: \text{a série é estacionária (sem raiz unitária)}$$

Autocorrelação

- Coeficiente de Correlação de Perason (r)
- Coeficiente de Determinação (R^2)
- Autocorrelação
  - Mede se existe uma relação matemática entre os intervalos da série temporal
  - Também deve estar entre – 1 e +1, sendo zero a ausência de autocorrelação.
  - Medida em intervalos (*lags*)
    - 1 intervalo: mede como os valores de 1 período (vizinhos) distante estão relacionados
    - 2 intervalos: mede com os valores de 2 períodos distantes estão relacionados

Autocorrelação

É o coeficiente de correlação entre observações defasadas no tempo:

$$r_1 = \frac{\sum_{t=1}^{n-1}(x_t - \overline{x}_1)(x_{t+1} - \overline{x}_2)}{\sqrt{\sum_{t=1}^{n-1}(x_t - \overline{x}_1)^2(x_{t+1} - \overline{x}_2)^2}}$$

onde as médias amostrais são:

$$\overline{x}_1 = \sum_{i=1}^{n-1} x_t/(n-1)$$ e $$\overline{x}_2 = \sum_{i=2}^{n} x_t/(n-1)$$

Função de Autocorrelação FAC(k)

A expressão anterior pode ser generalizada para k períodos de tempo (defasagem):

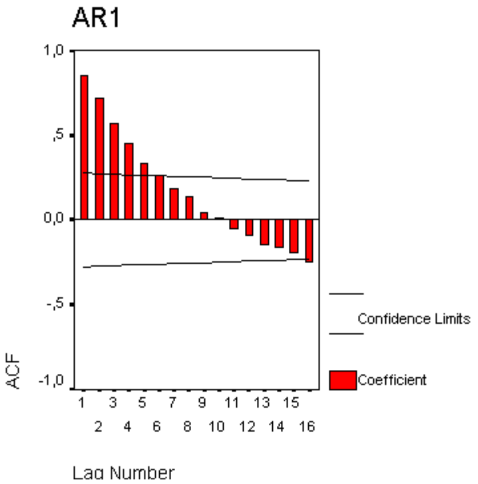
$$r_k = FAC(k) = \frac{\sum_{t=1}^{n-k}(x_t - \overline{x})(x_{t+k} - \overline{x})}{\sum_{t=1}^{n-1}(x_t - \overline{x})^2}$$

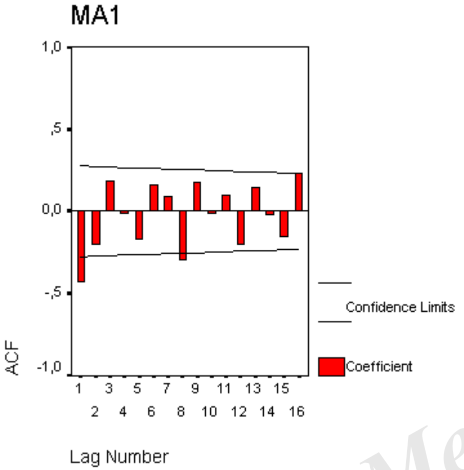
Como, tanto a covariância como a variância apresentam as mesmas unidades, não tem unidade. Oscila entre –1 e 1 e o gráfico feito colocando contra k é chamado de **correlograma amostral** da FAC.

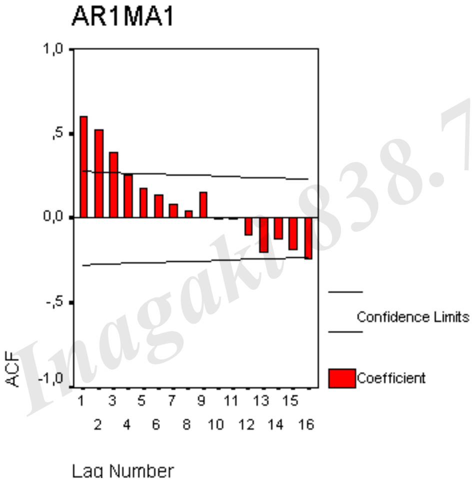
Correlogramas

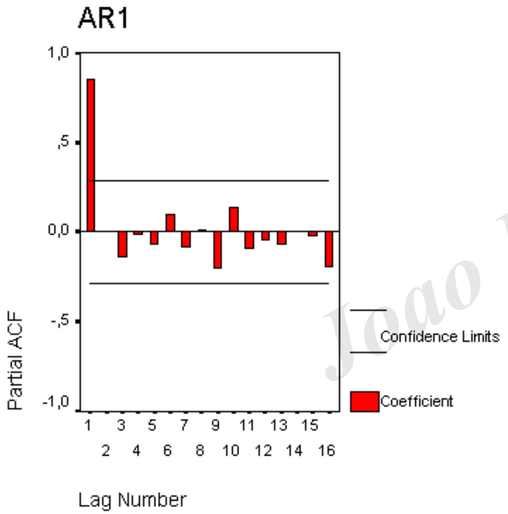
- FAC (ACF) – Função de Autocorrelação
  - Mostra as autocorrelações em uma série temporal
  - Linhas mostram significância (intervalo de confiança)
  - A 1ª autocorrelação é igual a 1. Cada traço do gráfico mostra uma defasagem e uma correlação (autocorrelação).
- FACP (PACF) – Função de Autocorrelação Parcial
  - Mede a autocorrelação não entre *lags* mas entre diferentes intervalos.

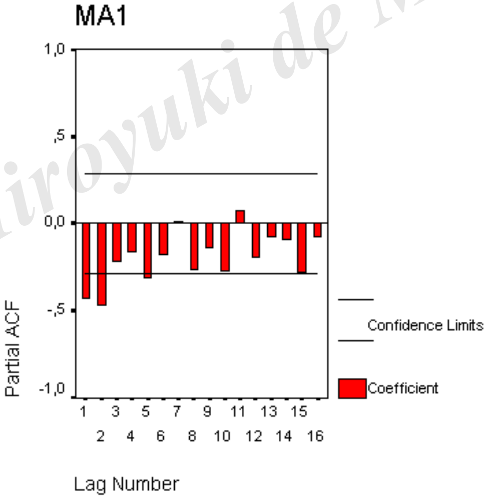
FAC dos Modelos Simulados

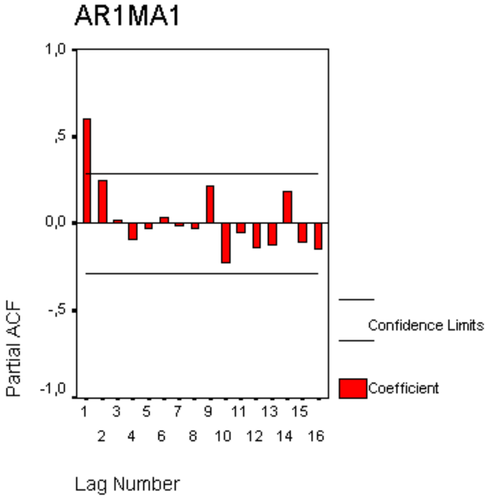












Características da FAC

	Padrão típico da FAC	Padrão típico da FACP
AR(p)	Decai exponencialmente para zero ou com padrão de onda senoidal amortecida, ou ambos.	Valores significativos, ou seja, não nulos, até a defasagem p
MA(q)	Valores significativos, ou seja, não nulos, até a defasagem q	Decai exponencialmente para zero.
ARMA(p,q)	Decai exponencialmente para zero.	Decai exponencialmente para zero.

Metodologia de Box-Jenkins

Etapas	Processo
Identificação	descobrir os valores apropriados de p e q. Para isso usamos o correlograma e o correlograma parcial para perceber em que períodos de defasagem existe mais correlação com a variável dependente ou de correlação entre as observações com k períodos de defasagem.
Estimação	estimar os parâmetros dos termos auto-regressivo e de média móvel incluídos no modelo
Checagem	Um teste simples do modelo escolhido é ver se os resíduos estimados desse modelo são ruídos brancos; se são, podemos aceitar o ajuste específico; se não são, devemos começar tudo de novo.

Modelo ARIMA(p,d,q)

Pode-se pensar num modelo ARIMA como uma função de regressão populacional para Y_t em que há apenas 2 tipos de "variáveis explicativas":

(1) Valores passados de $Y_t \rightarrow$ A parte "autorregressiva".

(2) Valores presente e passados do distúrbio normal u_t (ou "inovação") $\rightarrow$ A parte de "médias móveis".

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + u_t - \theta_1 u_{t-1} - \ldots - \theta_q u_{t-q}$$

- Hiperparâmetro *p:* a defasagem máxima de Y_t presente na equação.
- Hiperparâmetro *q:* a defasagem máxima de u_t presente na equação.
- Hiperparâmetro *d:* ordem de integração, se o processo for não-estacionário

Modelo AR(p) - Fundamentos

- Um processo linear **estacionário** auto-regressivo de ordem p, ou simplesmente AR(p), é definido como:

$$Y_t = \rho Y_{t-1} + \rho Y_{t-2} + \ldots + \rho Y_{t-p} + \varepsilon_t$$

Ruído Branco

Uma sequência ε_t é um processo ruído branco se, para qualquer t:

1) $E(\varepsilon_t) = 0$ → **média constante e nula**
2) $Var(\varepsilon_t) = E(\varepsilon_{t2})$ → **variância constante**
3) $Cov(\varepsilon_t, \varepsilon_{t-1}) = E(\varepsilon_t \varepsilon_{t-1}) = 0$ → **ausência de autocorrelação**

Se $\varepsilon_t \sim N$, então será Ruído Branco Gaussiano

Classes de modelos

Exemplos de modelos da classe ARIMA:

- **Modelo AR(1):** $Y_t = \phi_1 Y_{t-1} + u_t$
- **Modelo AR(2):** $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + u_t$
- **Modelo MA(1):** $Y_t = u_t - \theta_1 u_{t-1}$
- **Modelo ARMA(1,1):** $Y_t = \phi_1 Y_{t-1} + u_t - \theta_1 u_{t-1}$

Modelo ARIMA SAZONAL
Modelo SARIMA(p,d,q)(P,D,Q)

- Notação: SARIMA (0,0,0) $(1,0,0)_{12}$

$$Y_t = \alpha + \phi_{12} Y_{t-12} + \epsilon_t$$

- Notação: SARIMA (0,0,0) $(0,0,1)_{12}$

$$Y_t = \alpha + \theta_{12} e_{t-12} + \epsilon_t$$

- Generalizando: SARIMA$(p,d,q)(P,D,Q)_s$
- P - número de termos auto regressivos sazonais (defasagens no lado direito da equação)
- d – número de diferenças sazonais
- q – número de médias móveis sazonais (erros defasados no lado direito da equação)
- s – ciclo sazonal

Teste de Normalidade

Assimetria $$A(X) = E\left[\frac{(X-\mu)^3}{\sigma^3}\right]$$

Curtose $$K(X) = E\left[\frac{(X-\mu)^4}{\sigma^4}\right]$$

Normal: A = 0 e K = 3

Teste de Jarque Bera (1981) $$JB = \left(\frac{n}{6}\right)A^2 + \left(\frac{n}{24}\right)(K-3)^2$$

H_0 : série é normal

$$JB \sim \chi^2(2)$$

OBRIGADO!

linkedin.com/in/fabiano-guasti-lima-b9830282