

Séries Temporais

EELT7055 Modelagem e Avaliação de Desempenho

Prof. Carlos Marcelo Pedroso

Pós Graduação em Engenharia Elétrica - PPGE

2025

- Considere um processo onde o valor presente de uma série depende dos valores passados.
- Uma série temporal é a amostragem sequencial de uma variável durante um intervalo de tempo, via de regra longo, para possibilitar a identificação de padrões.
- Pode ser obtida analiticamente, ou também por simulação numérica, ou ainda, por medição experimental.

Séries Temporais: Notação

- Observação de uma série temporal discreta em instantes $\tau_1, \tau_2, \dots, \tau_t, \dots, \tau_n$: $X(\tau_1), X(\tau_2), \dots, X(\tau_t), \dots, X(\tau_n)$.
- Para N valores sucessivos com intervalo fixo h : $X_1, X_2, \dots, X_t, \dots, X_N$.
- **Esperança**: $E(X) = \mu$.
- **Variância**: $V[X] = E[(X - \mu)^2] = \sigma^2$.
- **Auto-covariância com defasagem k** :

$$\gamma_k = E[(X_t - \mu)(X_{t-k} - \mu)]$$

- **Auto-correlação com defasagem k** :

$$\rho_k = \gamma_k / \gamma_0$$

- ▶ X_t é estritamente estacionário se $[X_{t_1}, X_{t_2}, \dots, X_{t_n}]$ e $[X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k}]$ possuem a mesma distribuição conjunta para todo $n \in \mathbb{N}$.
- ▶ O processo deslocado por k , X_{t+k} , e o primeiro, X_t , devem ser equivalentes.
- ▶ A função de auto-covariância $\gamma(r, s) = E[(X_r - \mu)(X_s - \mu)]$ deve satisfazer a relação de invariância $\gamma(r, s) = \gamma(r + k, s + k)$ para $r, s, k \in \mathbb{Z}$.

Teste de Estacionariedade: Dickey-Fuller Aumentado (ADF)

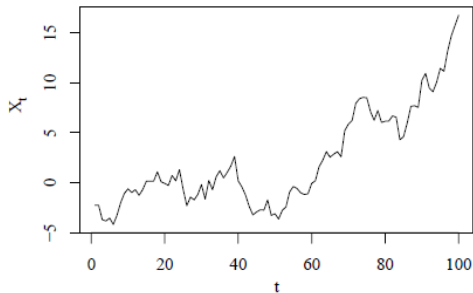
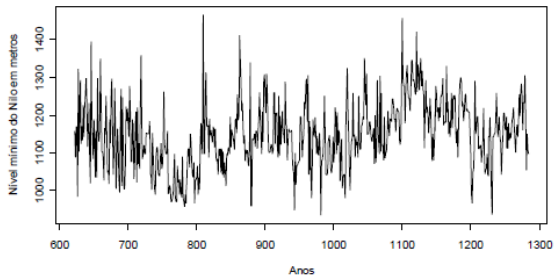
- ▶ **Objetivo:** Determinar se uma série temporal (Y_t) é estacionária ou se possui uma *raiz unitária*.
- ▶ **Hipóteses:** O teste ADF é um teste de raiz unitária.
 - **Hipótese Nula (H_0):** A série Y_t é **NÃO ESTACIONÁRIA** (Possui raiz unitária, $\gamma = 0$).
 - **Hipótese Alternativa (H_a):** A série Y_t é **ESTACIONÁRIA** ($\gamma < 0$).
- ▶ **Equação de Teste Geral:**

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^p \delta_i \Delta Y_{t-i} + \epsilon_t$$

- ▶ **Interpretação (no R):**
 - ✓ Se $p\text{-value} \leq \alpha$ (ex: 0.05): **Rejeitamos H_0** . Conclusão: a série é **ESTACIONÁRIA**.
 - ✓ Se $p\text{-value} > \alpha$: **Não Rejeitamos H_0** . Conclusão: a série é **NÃO ESTACIONÁRIA**.

A inclusão dos termos α (intercepto) e βt (tendência) depende do modelo escolhido.

Exemplos de Estacionariedade



Coeficiente de Correlação ($\rho_{X,Y}$)

- **Função de Correlação (População):** Mede a **dependência linear** entre duas variáveis aleatórias X e Y .
- **Definição:** O Coeficiente de Correlação de Pearson, $\rho_{X,Y}$, é dado por:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

➤ Propriedades:

- ✓ Varia no intervalo: $-1 \leq \rho_{X,Y} \leq 1$.
- ✓ $\rho_{X,Y} = 1$: Correlação linear positiva perfeita.
- ✓ $\rho_{X,Y} = -1$: Correlação linear negativa perfeita.
- ✓ $\rho_{X,Y} = 0$: Ausência de correlação linear.

Estimador Amostral do Coeficiente de Correlação

$(\hat{\rho}_{X,Y})$

- **Amostra:** Considere uma amostra de n pares de observações: $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$.
- **Estimador Amostral:** O estimador de correlação (amostral) $\hat{\rho}_{X,Y}$ é dado por:

$$\hat{\rho}_{X,Y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

- **Onde:**
 - ✓ $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ é a média amostral de X .
 - ✓ $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ é a média amostral de Y .
 - ✓ O numerador é a **Covariância Amostral** (não-viesada, sem o $\frac{1}{n-1}$ ou $\frac{1}{n}$ no denominador, mas a fórmula acima é a forma padrão do coeficiente).
- **Uso:** $\hat{\rho}_{X,Y}$ é uma estatística que estima o verdadeiro parâmetro populacional $\rho_{X,Y}$.

Exemplo de Cálculo: Estimador de Correlação ($\hat{\rho}$)

Estimador Amostral ($\hat{\rho}_{X,Y}$):

$$\hat{\rho}_{X,Y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Dados Amostrais (10 Pares):

i	X _i	Y _i	(X _i - $\bar{x}$)	(Y _i - $\bar{y}$)	(X _i - $\bar{x}$)(Y _i - $\bar{y}$)
1	2	8	-4.5	3.0	-13.5
2	3	7	-3.5	2.0	-7.0
3	4	7	-2.5	2.0	-5.0
4	5	5	-1.5	0.0	0.0
5	6	6	-0.5	1.0	-0.5
6	7	4	0.5	-1.0	-0.5
7	8	3	1.5	-2.0	-3.0
8	9	4	2.5	-1.0	-2.5
9	10	2	3.5	-3.0	-10.5
10	11	4	4.5	-1.0	-4.5
Soma dos Produtos dos Desvios (SPD):					-47.0

- SPD: $\sum (X_i - \bar{x})(Y_i - \bar{y}) = -47.0$
- SQD_X: $\sum (X_i - \bar{x})^2 = 82.5$
- SQD_Y: $\sum (Y_i - \bar{y})^2 = 34.0$
- $\hat{\rho}_{X,Y} = -0.8874$

Função de Auto-Correlação

- A análise da Função de Auto-Correlação (FAC) de uma série é muito importante.
- A FAC estabelece o coeficiente de correlação entre um valor X_i da série e o valor defasado X_{i-k} .
- O índice k é o deslocamento (ou *lag*).
- A análise da FAC tem um papel central na análise de desempenho de sistemas de comunicações.

➤ **Auto-Covariância:**

$$\gamma_k = E[(X_t - \mu)(X_{t-k} - \mu)]$$

➤ **Auto-Correlação:**

$$\rho_k = \gamma_k / \gamma_0$$

➤ **Estimador da Auto-Correlação:**

$$\bar{\gamma}_k = \frac{\sum_{i=1}^{N-k} (X_i - \bar{X})(X_{i+k} - \bar{X})}{\sum_{i=1}^N (X_i - \bar{X})^2}$$

Modelo Auto-Regressivo (AR)

- O modelo $AR(p)$, ou auto-regressivo, é dado por:

$$\tilde{X}_t = \phi_1 \tilde{X}_{t-1} + \cdots + \phi_p \tilde{X}_{t-p} + a_t$$

onde $\tilde{X}_t = X_t - \mu$, a_t é um ruído branco e $\phi_1, \dots, \phi_p$ são parâmetros do modelo.

- A função de auto-correlação decai exponencialmente, o que faz com que o modelo possa ser representado aproximadamente por um modelo de Markov Modulado.
- Para identificação, o analista deve examinar a estrutura da função de auto-correlação e estimar o valor de p .

Modelo Auto-Regressivo (AR)

Aqui uma lista de comandos, anotar o significado na aula!

Comandos R

```
serie_ar1 <- arima.sim(n = 1000, model = list(ar = 0.9), sd = 3) + 10
tseries::adf.test(serie_ar1)
plot(serie_ar1)
acf(serie_ar1)
pacf(serie_ar1)
modelo <- arima(serie_ar1, order = c(1, 0, 0))
modelo
cat("\n--- Criterios de Informacao: Quanto menor o AIC/BIC, melhor o modelo ---\n")
cat("Critério de Informacao de Akaike (AIC) modelo AR(1):", AIC(modelo), "\n")
cat("Critério de Informacao Bayesiano (BIC) modelo AR(1):", BIC(modelo), "\n")
str(modelo)
residual<-na.omit(modelo$resid)
par(mfrow = c(2, 2))
plot(residual)
hist(residual,freq=FALSE)
lines(x,dnorm(x,mean=mean(residual),sd=sd(residual)),col=2)
acf(residual)
plot(serie_ar1)
lines(residual, col=2)
```

Interpretação do Teste Dickey-Fuller Aumentado

Resultado do Teste

Dickey-Fuller = -6.6239, Lag order = 9, **p-value = 0.01**
alternative hypothesis: stationary

- ▶ **Hipótese Nula (H_0):** A série é **NÃO ESTACIONÁRIA**.
- ▶ **Hipótese Alternativa (H_a):** A série é **ESTACIONÁRIA**.

Decisão Estatística ($\alpha = 5\%$)

- ▶ **Nível de Significância (α):** Adota-se 0.05 (5%).
- ▶ **Valor-p Reportado:** 0.01.

Comparação: p-value (0.01) $\leq$ α (0.05)

Conclusão

Como o **p-value** é **menor** que o nível de significância de 5%, **rejeitamos a Hipótese Nula (H_0)**.

A série serie_ar1 é ESTACIONÁRIA. ✓

Modelo Média Móvel (MA)

- O modelo $MA(q)$, ou médias móveis, é dado por:

$$\tilde{X}_t = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

- O valor atual de X_t é formado pela soma dos choques ponderados de ruídos aleatórios passados.
- $\theta_1, \dots, \theta_q$ são os parâmetros do modelo.

Modelo Média Móvel (MA)

Aqui uma lista de comandos, anotar o significado na aula!

Comandos R

```
serie_mal <- arima.sim(n = 1000, model = list(ma = -0.5), sd = 3) + 10
plot(serie_mal)
acf(serie_mal)
pacf(serie_mal)
modelo <- arima(serie_mal, order = c(0, 0, 1))
modelo
cat("\n--- Critérios de Informacao: Quanto menor o AIC/BIC, melhor o modelo ---\n")
cat("Critério de Informacao de Akaike (AIC) modelo AR(1):", AIC(modelo), "\n")
cat("Critério de Informacao Bayesiano (BIC) modelo AR(1):", BIC(modelo), "\n")
str(modelo)
residual<-na.omit(modelo$resid)
par(mfrow = c(2, 2))
plot(residual)
hist(residual,freq=FALSE)
lines(x,dnorm(x,mean=mean(residual),sd=sd(residual)),col=2)
acf(residual)
plot(serie_mal)
lines(residual, col=2)
```

Modelo ARMA

- O modelo $\text{ARMA}(p, q)$ combina os modelos AR e MA.
- Os valores de p e q indicam, respectivamente, o número de parâmetros no modelo AR e MA.
- O modelo é dado por:

$$\tilde{X}_t = \phi_1 \tilde{X}_{t-1} + \cdots + \phi_p \tilde{X}_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

- A estimação de parâmetros é mais difícil que no modelo AR, envolvendo a resolução de equações não-lineares.

Modelo ARMA

Aqui uma lista de comandos, anotar o significado na aula!

Comandos R

```
serie_armall <- arima.sim(n = 1000, model = list(ar = 0.7, ma = -0.5))
plot(serie_armall)
acf(serie_armall)
pacf(serie_armall)
modelo <- arima(serie_armall, order = c(1, 0, 1))
modelo
cat("\n--- Critérios de Informacao: Quanto menor o AIC/BIC, melhor o modelo ---\n")
cat("Critério de Informacao de Akaike (AIC) modelo AR(1):", AIC(modelo), "\n")
cat("Critério de Informacao Bayesiano (BIC) modelo AR(1):", BIC(modelo), "\n")
str(modelo)
residual<-na.omit(modelo$resid)
par(mfrow = c(2, 2))
plot(residual)
hist(residual,freq=FALSE)
lines(x,dnorm(x,mean=mean(residual),sd=sd(residual)),col=2)
acf(residual)
plot(serie_armall)
lines(residual, col=2)
```

Modelo ARIMA

- O modelo ARIMA (Auto Regressive Integrated Moving Average) é uma extensão do modelo ARMA(p, q).
- É dado por ARIMA(p, d, q), onde ∇^d é um operador de diferenças.
- $\nabla^d X_t$ é definido como $\nabla^1 X_t = X_t - X_{t-1}$.
- A diferenciação da série ($\nabla^d X_t$) é normalmente aplicada na tentativa de torná-la **estacionária**.
- Após a diferenciação, aplica-se um modelo ARMA sobre a série diferenciada.
- Desta maneira, o modelo pode ser utilizado em séries **não estacionárias**.

Modelo ARIMA

Aqui uma lista de comandos, anotar o significado na aula!

Comandos R

```
serie_arima111 <- arima.sim(n = 1000, model = list(order = c(1, 1, 1), ar = 0.7, ma =  
  -0.5))  
tseries::adf.test(serie_arima111)  
plot(serie_arima111)  
acf(serie_arima111)  
pacf(serie_arima111)  
serie_diferenciada<-diff(serie_arima111, lag = 1, differences = 1)  
plot(serie_diferenciada)  
acf(serie_diferenciada)  
tseries::adf.test(serie_diferenciada)  
modelo <- arima(serie_arima111, order = c(1, 1, 1))  
modelo  
cat("\n--- Critérios de Informação: Quanto menor o AIC/BIC, melhor o modelo ---\n")  
cat("Critério de Informação de Akaike (AIC) modelo AR(1):", AIC(modelo), "\n")  
cat("Critério de Informação Bayesiano (BIC) modelo AR(1):", BIC(modelo), "\n")  
str(modelo)  
residual<-na.omit(modelo$resid)  
par(mfrow = c(2, 2))  
plot(residual)  
hist(residual,freq=FALSE)  
lines(x,dnorm(x,mean=mean(residual),sd=sd(residual)),col=2)  
acf(residual)  
plot(serie_arima111)  
lines(residual, col=2)
```

Teste de Estacionariedade no exemplo

Resultado do Teste

Dickey-Fuller = -1.958, Lag order = 9, p-value = 0.5961
alternative hypothesis: stationary

- Para o Teste de Dickey-Fuller Aumentado (ADF):
 - ▶ **Hipótese Nula (H_0):** A série é **NÃO ESTACIONÁRIA**.
 - ▶ **Hipótese Alternativa (H_a):** A série é **ESTACIONÁRIA**.
- Decisão Estatística:
 - ✓ Valor-p: 0.3276. Nível de Signifância Padrão (α): 0.05 (5%)
 - ✓ Regra de Decisão: Compare o p -value com α .
 - ✓ p -value (**0.3276**) > α (**0.05**)
 - ✓ Como o p -value é maior que 0.05, você **NÃO REJEITA** a Hipótese Nula (H_0).
- **A série é NÃO ESTACIONÁRIA**
- Verificar após a aplicação do operador `nabla`!

O que é uma Série Temporal Sazonal?

- ▶ Uma **Série Temporal** (Y_t) é uma sequência de observações ordenadas no tempo e registradas em intervalos regulares.
- ▶ O componente de **Sazonalidade** (S_t) refere-se a um padrão repetitivo e previsível que ocorre em períodos fixos (ex: mensal, trimestral, semanal).
- ▶ **Exemplo Clássico:** Vendas no varejo, consumo de energia, e a famosa série *AirPassengers* (Passageiros Aéreos).
- ▶ **Desafio:** O modelo ARIMA tradicional (não-sazonal) não consegue capturar essa periodicidade de longo prazo de forma eficaz.

Componentes de uma Série Temporal

A maioria das séries temporais não-estacionárias é composta por 3 ou 4 elementos:

- ❶ **Tendência (T_t):** Comportamento de longo prazo (crescimento ou decrescimento).
- ❷ **Sazonalidade (S_t):** Variações regulares e periódicas (ex: picos em Dezembro).
- ❸ **Ciclo (C_t):** Flutuações de longo prazo, mas irregulares e de duração variável (ligadas a ciclos econômicos, por exemplo).
- ❹ **Irregular/Ruído (ϵ_t):** Variações aleatórias e não explicadas.

Decomposição (Modelo Aditivo):

$$Y_t = T_t + S_t + C_t + \epsilon_t$$

Decomposição (Modelo Multiplicativo):

$$Y_t = T_t \times S_t \times C_t \times \epsilon_t$$

Revisão: O Modelo ARIMA (p, d, q)

O ARIMA (AutoRegressive Integrated Moving Average) é base para a modelagem:

➤ **AR(p): AutoRegressive**

- ✓ Usa observações passadas de Y_t para prever o valor atual.
- ✓ p = número de termos autorregressivos.

➤ **I(d): Integrated**

- ✓ **Diferenciação** para tornar a série **Estacionária**.
- ✓ d = número de diferenciações necessárias (remove Tendência).

➤ **MA(q): Moving Average**

- ✓ Usa os erros (ϵ_t) das previsões passadas para prever o valor atual.
- ✓ q = número de termos de média móvel.

O Modelo SARIMA (ARIMA Sazonal)

O SARIMA é uma extensão do ARIMA que adiciona termos para a sazonalidade.

Notação do SARIMA:

$$\text{SARIMA}(p, d, q) \times (P, D, Q)_s$$

- (p, d, q) (**Parte Não-Sazonal**): Descreve a estrutura de curto prazo (não sazonal).
- $(P, D, Q)_s$ (**Parte Sazonal**): Descreve a estrutura sazonal.
 - ✓ **P** (AR Sazonal): Termos AR na defasagem s .
 - ✓ **D** (Diferenciação Sazonal): Remove a sazonalidade.
 - ✓ **Q** (MA Sazonal): Termos MA na defasagem s .
 - ✓ **s**: Período de sazonalidade (ex: $s = 12$ para dados mensais).

O Coração do SARIMA: Diferenciação Sazonal (D)

- O parâmetro **D** (Diferenciação Sazonal) é crucial para séries como a *AirPassengers*.
- **Função:** Eliminar o padrão sazonal, garantindo que as propriedades estatísticas sejam constantes em múltiplos do período s .
- **Exemplo ($s = 12$):**
 - ✓ A diferenciação compara uma observação (Y_t) com a observação do **mesmo período no ano anterior** (Y_{t-12}).
 - ✓ **Operador Diferença Sazonal:** $\nabla_s Y_t = Y_t - Y_{t-s}$
- Na série *AirPassengers*, geralmente precisamos de $d = 1$ (Tendência) e $D = 1$ (Sazonalidade) para atingir a estacionariedade.

Exemplo: Aplicação do SARIMA na Série AirPassengers

- **Série Temporal:** Número de passageiros aéreos internacionais (mensal, $s = 12$).
- **Pré-Processamento (R-Cran):**
 - ✓ → Aplicação de $\log()$ para estabilizar a variância.
 - ✓ → Aplicação da diferenciação não-sazonal ($d = 1$).
 - ✓ → Aplicação da diferenciação sazonal ($D = 1$).
- **Modelo Comum (Clássico Box-Jenkins):**

$$\text{SARIMA}(0, 1, 1) \times (0, 1, 1)_{12}$$

- **Interpretação:**
 - ✓ $(0, 1, 1)$: Um termo de Média Móvel (MA) e uma Diferenciação ($d = 1$).
 - ✓ $(0, 1, 1)_{12}$: Um termo de Média Móvel Sazonal ($Q = 1$) e uma Diferenciação Sazonal ($D = 1$) no período 12.

Modelo SARIMA

Aqui uma lista de comandos, anotar o significado na aula!

Comandos R

```
dados_passagens <- AirPassengers
plot(dados_passagens, main = "AirPassengers", ylab = "Passageiros (milhares)", xlab = "Ano")
log_passagens <- log(dados_passagens)
par(mfrow = c(2, 2))
plot(log_passagens, main = "LogAirPassengers", ylab = "Ano")
acf(log_passagens, main = "ACF - LogAirPassengers", lag.max=50)
pacf(log_passagens, main = "PACF - LogAirPassengers", lag.max=50)

# Diferenciacao Nao-Sazonal (d=1)
diff_log_passagens_d1 <- diff(log_passagens, differences = 1)
par(mfrow = c(2, 2))
plot(diff_log_passagens_d1)
acf(diff_log_passagens_d1, main = "ACF - Diferenca d=1")
pacf(diff_log_passagens_d1, main = "PACF - Diferenca d=1")

# Diferenciacao Sazonal (D=1, periodo s=12)
diff_sazonal <- diff(diff_log_passagens_d1, lag = 12)
par(mfrow = c(2, 2))
plot(diff_sazonal)
acf(diff_sazonal, lag.max = 50, main = "ACF - Diferenca Sazonal")
pacf(diff_sazonal, lag.max = 50, main = "PACF - Diferenca Sazonal")
modelo <- Arima(log_passagens, order = c(1, 1, 1), seasonal = list(order = c(1, 1, 1),
  period = 12))
summary(modelo)
```

Modelo SARIMA

Aqui uma lista de comandos, anotar o significado na aula!

Comandos R

```
# Ajuste Automatico
modelo_auto <- auto.arima(log_passagens,
                           trace = TRUE, # mostra os modelos testados
                           stepwise = FALSE, # testa mais combinacoes
                           approximation = FALSE) # nao usa aproximacao

summary(modelo_auto)

checkresiduals(modelo_auto)

# Previsao para os proximos 3 anos (h=3*12)
previsao <- forecast(modelo_auto, h = 36)

# Plotar a previsao
plot(previsao, main = "Previsao AirPassengers com SARIMA")

# Exibir os valores previstos (revertendo o log)
previsao_revertida <- exp(previsao$mean)
previsao_revertida
```

Conclusão: Modelagem com SARIMA

O SARIMA é o modelo ideal para séries sazonais porque:

- Ele **integra** a estrutura não-sazonal (curto prazo) e a estrutura sazonal (longo prazo) em um único modelo.
- Usa a **Diferenciação** (não-sazonal e sazonal) para remover tendências e sazonalidade e garantir a **Estacionariedade**.
- É uma ferramenta **robusta** e amplamente utilizada para previsão.

Próxima Etapa: Diagnóstico dos Resíduos e Previsão.

Outros Modelos: Média Móvel

- O método da Média Móvel considera como previsão para o período futuro a média das observações passadas recentes.
- O termo é usado porque a média é recalculada com a próxima observação, incluindo-a e desprezando a observação mais antiga.
- **Fórmula** (Média Móvel de N períodos):

$$\bar{X}_t = \frac{X_{t-1} + X_{t-2} + \cdots + X_{t-N}}{N}$$

EWMA?

Média Móvel Exponencialmente Ponderada

- O **EWMA** é uma técnica de **Suavização Exponencial** utilizada para calcular uma média móvel de uma série temporal.
- **Diferencial:** Diferente da Média Móvel Simples (SMA), o EWMA atribui **pesos decrescentes exponencialmente** às observações à medida que elas ficam mais distantes no passado.
- **Conceito Chave:** Os dados **mais recentes** têm uma importância significativamente **maior** na determinação do valor atual do EWMA.
- **Aplicações Principais:**
 - ➊ **Previsão de Séries Temporais** (modelos de suavização).
 - ➋ **Estimativa de Volatilidade** (Mercado Financeiro - RiskMetrics).
 - ➌ **Controle Estatístico de Processo (CEP)** (Gráficos de Controle EWMA).

Cálculo da Média Ponderada

O EWMA (S_t) no tempo t é calculado como uma combinação linear da observação atual (X_t) e do valor EWMA do período anterior (S_{t-1}).

$$S_t = \lambda X_t + (1 - \lambda)S_{t-1}, \quad 0 \leq \lambda \leq 1$$

- S_t : Valor do EWMA no tempo t .
- X_t : Observação atual da série temporal no tempo t .
- S_{t-1} : Valor do EWMA no período anterior.
- λ : **Fator de Suavização (Peso) ou Decaimento.**

O Fator de Suavização (λ)

Ajustando a Memória do Modelo

- O λ (Lambda) é o único parâmetro do modelo, e $0 < \lambda \leq 1$.
- Ele controla o peso dado à observação mais recente (Y_t) versus o histórico (S_{t-1}).
- **EWMA com λ Alto (Perto de 1):**
 - ✓ O modelo é **mais sensível** a mudanças recentes.
 - ✓ Tem **pouca memória** do passado distante (acompanha mais de perto o ruído).
- **EWMA com λ Baixo (Perto de 0):**
 - ✓ O modelo é **mais suave** (menos volátil).
 - ✓ Tem **muita memória** do passado (lento para reagir a choques recentes).
- **Exemplos:** Para estimativa de volatilidade financeira, EWMA com $\lambda = 0.94$ é frequentemente utilizado para dados diários. Para estimativa do Round Trip Time do TCP, EWMA com $\lambda = 0.8$ a $\lambda = 0.9$ é comum. No escalonador PF do 4G/5G, é usado o EWMA para estimar o consumo de banda dos dispositivos com $\lambda = 0.8$.

Pesos Decrescentes Exponencialmente

Expandindo a fórmula recorrente, vemos a ponderação exponencial:

$$S_t = \lambda X_t + \lambda(1 - \lambda)X_{t-1} + \lambda(1 - \lambda)^2 X_{t-2} + \cdots + \lambda(1 - \lambda)^k X_{t-k} + \dots$$

- ▶ O peso dado à observação k períodos no passado (Y_{t-k}) é $\lambda(1 - \lambda)^k$.
- ▶ Como $1 - \lambda$ é um valor entre 0 e 1, o termo $(1 - \lambda)^k$ **diminui rapidamente** à medida que k aumenta.
- ▶ Isso formaliza o conceito: **o passado recente é mais relevante que o passado remoto.**

Aplicação em Volatilidade (EWMA-Vol)

Modelo utilizado pelo J.P. Morgan RiskMetrics (1994)

O EWMA é frequentemente utilizado para estimar a **Volatilidade** (σ_t^2) dos retornos de ativos financeiros (r_t^2).

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2$$

- ▶ σ_t^2 : Previsão da Volatilidade (Variância) no tempo t .
- ▶ σ_{t-1}^2 : Volatilidade (Variância) prevista no período anterior.
- ▶ r_{t-1}^2 : Retorno ao Quadrado (choque/inovação) no período anterior.

Vantagem: Simples, fácil de implementar e captura a característica de clustering de volatilidade (a volatilidade muda ao longo do tempo).

EWMA vs. GARCH(1,1)

- O modelo EWMA de volatilidade pode ser visto como um **caso especial** do modelo GARCH(1,1).
- **Fórmula GARCH(1,1) para Volatilidade:**

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$

- **Relação com EWMA:** O EWMA ocorre quando:
 - ✓ $\omega = 0$ (o termo de longo prazo/variância incondicional é zero).
 - ✓ $\alpha + \beta = 1$ (a persistência é máxima, $\alpha = 1 - \lambda$ e $\beta = \lambda$).
- **Implicação:** No EWMA, a média de longo prazo da volatilidade é zero, o que pode ser uma limitação em séries financeiras de longo prazo.

Modelo EWMA

Aqui uma lista de comandos, anotar o significado na aula!

Comandos R

```
library(forecast)
library(fpp2)

# Serie Temporal de Exemplo (dados anuais de producao de cerveja australiana)
dados_cerveja <- window(ausbeer, start = 1990)

# Visualizar a Serie
plot(dados_cerveja,
     main = "Producao de Cerveja Australiana (desde 1990)",
     ylab = "Milhoes de Litros")

# Aplicar a EWMA
# A funcao 'ses' ajusta o modelo SES, otimizando o parametro alpha (lambda)
# h = 8: Previsao para os proximos 8 periodos (trimestres)
modelo_ses <- ses(dados_cerveja, h = 8)

# Visualizar o Resultado do Ajuste e Previsao
plot(modelo_ses,
     main = "Previsca com EWMA",
     ylab = "Milhoes de Litros")

# Exibir o modelo
summary(modelo_ses)
```

Exemplo: MACD

(Moving Average Convergence Divergence)

- **Objetivo:** Identificar a força, direção e momentum de uma tendência de preços.
- **Método:** Baseado na **diferença** entre duas Médias Móveis Exponenciais (MME) de períodos distintos (rápida e lenta).
- **Pacotes R:** TTR e quantmod.

Composição do MACD (Padrão 12, 26, 9)

- ➊ **Linha MACD:** $MME_{12} - MME_{26}$
- ➋ **Linha de Sinal:** MME_9 da Linha MACD
- ➌ **Histograma:** Linha MACD – Linha de Sinal

Comandos R-CRAN para MACD

Análise da Ação AAPL (Apple)

Comandos R

```
# 1. Carregar dados (usando quantmod)
library(quantmod)

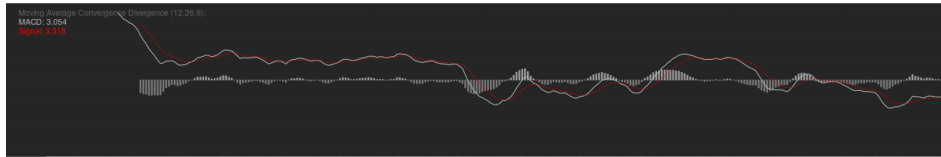
getSymbols("AAPL", src = "yahoo", from = "2023-01-01", to = "2025-01-01")

chartSeries(AAPL['/2024-06'], # Mostra dados ate Junho de 2024
            name = "AAPL com MACD - Foco ate Junho/2024")

addMACD(fast = 12, slow = 26, signal = 9)
```

- MME Rápida (fast = 12): o λ implícito é relativamente alto (próximo a $1/12 \approx 0.083$).
- MME Lenta (slow = 26): O λ implícito é menor (próximo a $1/26 \approx 0.038$).
- Linha de Sinal (signal = 9): Cálculo: A Média Móvel Exponencial é aplicada não aos preços, mas sim à própria Linha MACD (a diferença entre a EMA de 12 e a EMA de 26), usando 9 períodos. O λ implícito é moderado (próximo a $1/9 \approx 0.11$). Esta linha atua como um filtro. Ela suaviza o movimento rápido da Linha MACD, tornando os cruzamentos (os sinais de negociação) menos frequentes e, teoricamente, mais confiáveis.

Interpretação e Sinais do MACD



Sinal	Condição	Implicação (Previsão)
Compra	MACD cruza a Linha de Sinal <u>acima</u>	Momentum de alta acelerando.
Venda	MACD cruza a Linha de Sinal <u>abaixo</u>	Momentum de baixa acelerando.
Zero	Linha MACD cruza o eixo zero	Forte mudança na média de longo prazo.

Divergência

Se o preço faz nova máxima, mas o MACD não, isso é um sinal forte de que a **tendência está perdendo força** e pode reverter.

O Desafio da Amostragem Assíncrona

Por Que o EMA Clássico Não Funciona?

- **Amostragem Síncrona (Clássica):** Assume um intervalo de tempo (Δt) constante entre as observações.
- **Fórmula EMA (Síncrona):**

$$\mu_t = \alpha Y_t + (1 - \alpha)\mu_{t-1} \quad \text{onde } \alpha \text{ é constante.}$$

- **Amostragem Assíncrona:** Ocorre quando o intervalo $t_i = t_i - t_{i-1}$ é **variável** ou irregular (ex: sensores de rede, dados de mercado de alta frequência, registro de eventos).
- **Problema:** Se Δt é maior, o peso do histórico (μ_{t-1}) deveria decair mais rapidamente. O uso de um α constante introduz **viés**.

Adaptação: Constante de Tempo (τ)

Relacionando o α com o Intervalo de Tempo

- No EMA/EWMA, o parâmetro α está fundamentalmente ligado a uma **Constante de Tempo** (τ).
- τ representa o "horizonte de memória" do filtro, ou o tempo que leva para o peso decair a um certo nível ($e^{-1} \approx 36.8\%$).
- **Relação Geral (Δt constante):**

$$\alpha = 1 - e^{-\Delta t / \tau}$$

A Solução para o Caso Assíncrono

Para lidar com t_i variável, devemos usar um fator de suavização α_i que **também é variável**.

O Modelo EMA para t Variável

Fórmula com Fator de Suavização Dinâmico

- ❶ **Passo 1: Calcular o Fator de Decaimento (β_i)** O termo $(1 - \alpha)$ na fórmula clássica é substituído pelo fator de decaimento:

$$\beta_i = e^{-\Delta t_i / \tau}$$

- ❷ **Passo 2: Definir o Novo α Variável (α_i)**

$$\alpha_i = 1 - \beta_i = 1 - e^{-\Delta t_i / \tau}$$

- ❸ **Passo 3: Aplicar a Fórmula EMA Dinâmica** O valor suavizado (μ_i) é calculado no novo tempo t_i :

$$\mu_i = \alpha_i Y_i + (1 - \alpha_i) \mu_{i-1}$$

Interpretação e Vantagens

Garantindo a Consistência do Filtro

- **Consistência Garantida:** O filtro mantém a mesma **Constante de Tempo** (τ) (o "horizonte de memória") independentemente da frequência de amostragem.
- **Reação ao t :**
 - ✓ Se t_i é **pequeno** $\rightarrow \alpha_i$ é pequeno $\rightarrow$ O novo dado Y_i tem pouco peso.
 - ✓ Se t_i é **grande** $\rightarrow \alpha_i$ é grande $\rightarrow$ O novo dado Y_i tem um peso maior (pois a média anterior, μ_{i-1} , está "velha").
- **Desafio Prático:** A principal dificuldade é **estimar ou escolher** a Constante de Tempo (τ) de forma apropriada para o processo.

Aplicações em Sistemas Assíncronos

- ❶ **Sistemas de Controle em Rede (NCS):** Sensores e atuadores que se comunicam através de redes com atrasos e perdas de pacotes variáveis.
- ❷ **Séries Temporais de Alta Frequência (Finanças):** Dados *tick-by-tick* ou *event-based*, onde o tempo entre as transações é irregular.
- ❸ **Internet das Coisas (IoT):** Dispositivos que enviam dados apenas quando um limite é cruzado (*event-triggered sampling*), gerando intervalos de tempo irregulares.

Comandos R

```
# 1. Dados de Exemplo (Simulacao Assincrona)
# Definir tempos e valores com espacamento irregular

# Tempo (em dias) - Assincrono
tempos <- c(0, 1, 3, 4, 8, 9, 10, 15, 17, 20)

# Valores (Ex: leitura de um sensor)
valores <- c(53.0, 52.0, 58.0, 57.5, 60.0, 61.5, 59.0, 67.0, 65.5, 64.0)

# 2. Parametro Chave: Constante de Tempo (tau)
# Tau (tau): Representa a "memoria" do filtro em dias.
# Um Tau menor (ex: 2 dias) torna o filtro mais reativo.
# Um Tau maior (ex: 10 dias) torna o filtro mais suave.
tau <- 5 # 5 dias de horizonte de memoria

# 3. Calcular os Intervalos de Tempo (Delta_t)
# diff(tempos) calcula a diferenca entre um tempo e o tempo anterior
delta_t <- c(NA, diff(tempos)) # 0 primeiro Delta_t e NA

# 4. Inicializar o Vetor de EWMA (mu)
ewma_assincrono <- numeric(length(valores))
ewma_assincrono[1] <- valores[1] # Inicializa o primeiro valor com a 1a observacao
```

Comandos R

```
# 5. Loop para o Calculo Dinamico (onde a logica "Next" e aplicada)
for (i in 2:length(valores)) {
  # 5a. Calcular o alpha (alpha_i) dinamico
  #  $\alpha_i = 1 - \exp(-\Delta t_i / \tau)$ 
  alpha_i <- 1 - exp(-delta_t[i] / tau)
  # 5b. Aplicar a formula EMA (EWMA)
  #  $\mu_i = \alpha_i * Y_i + (1 - \alpha_i) * \mu_{i-1}$ 
  ewma_assincrono[i] <- alpha_i * valores[i] + (1 - alpha_i) * ewma_assincrono[i-1]
}
# 6. Reunir os Resultados no Data Frame
resultado_cran <- data.frame(
  Tempo_dias = tempos, Valor_Y = valores, Delta_t = delta_t,
  Alpha_i = round(c(NA, 1 - exp(-delta_t[-1] / tau)), 4), # Calculando e arredondando
  ewma_alpha
  EWMA_Assincrono = round(ewma_assincrono, 4)
)
# 7. Exibir o Resultado
print("--- Tabela de Resultados do EWMA Assincrono (tau = 5 dias) ---")
print(resultado_cran)
```

Comandos R

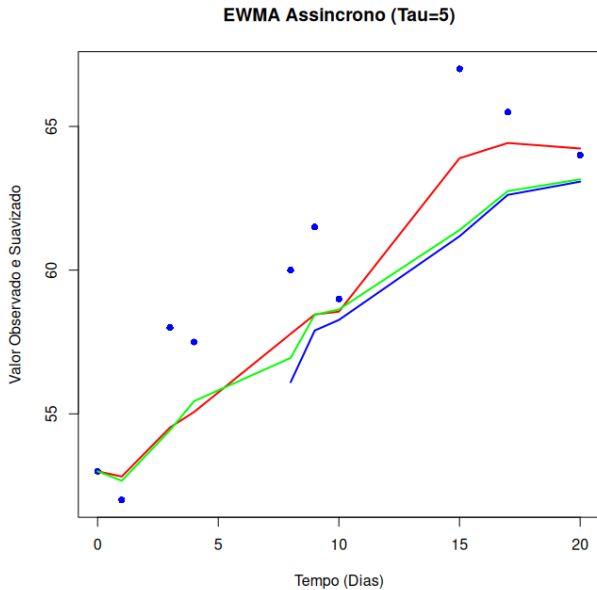
```
# 8. Visualizacao
plot(tempos, valores, type = 'p', pch = 16, col = 'blue',
     main = "EWMA Assincrono (Tau=5)",
     xlab = "Tempo (Dias)",
     ylab = "Valor Observado e Suavizado")

lines(tempos, ewma_assincrono, col = 'red', lwd = 2)

# 9. Comparação de resultados
# Calcular o EMA (EWMA) tradicional com n=5, corresponde a um alpha (peso) de  $2/(5+1) = 0.333$ 
ewma_resultado <- EMA(valores, n = 5)
lines(tempos, ewma_resultado, col = 'blue', lwd = 2)

# Outra forma
library(qcc)
ewma_smooth_resultado <- ewmaSmooth(x = tempos, y = valores, lambda = 0.33)
lines(ewma_smooth_resultado$x, ewma_smooth_resultado$y, col = 'green', lwd = 2)
```

EWMA Assincrono comparado



Outros Modelos: Alisamento Exponencial Linear

- Utilizado para evitar erros sistemáticos (superestimação ou subestimação) quando o Alisamento Exponencial Simples é aplicado em séries com **tendência**.
- Procura reconhecer a presença de tendência na série de dados.
- **Fórmulas:**

$$F_{t+m} = S_t + mT_t$$

$$S_t = \alpha X_t + (1 - \alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1}$$

- α é o peso atribuído à observação ($0 < \alpha < 1$) e β é o coeficiente de alisamento da tendência ($0 < \beta < 1$).

Outros Modelos: Alisamento Exponencial Sazonal e Linear de Winter

► Fórmulas:

$$S_t = \alpha \frac{X_t}{L_{t-l}} + (1 - \alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1}$$

$$L_t = \gamma \frac{X_t}{S_t} + (1 - \gamma)L_{t-l}$$

$$F_{t+m} = (S_t + mT_t)L_{t-l+m}$$

- L_t corresponde ao alisamento do fator de sazonalidade $\frac{X_t}{S_t}$.
- l é o intervalo da sazonalidade.
- γ corresponde ao peso atribuído ao fator de sazonalidade.

- ❶ Modele a série representada pelo valor de fechamento diário das ações da Petrobras (PETR4).
 - ✓ Utilize o modelo ARIMA
 - ✓ Explique o resultado. Qual outro modelo esse resultado indica?
 - ✓ Analise o comportamento com o método MACD e explique o significado.
- ❷ Modele a série representada pelo valor do IPCA mensal no Brasil.
- ❸ Modele a série Producao de Cerveja Australiana (desde 1990) com o modelo ARIMA.
- ❹ Modele a série de Vendas ao Varejo no Brasil com o modelo ARIMA.

Exercício PETR4

Comandos R

```
library(quantmod)
library(forecast)

ticker <- "PETR4.SA"
data_inicio <- "2023-01-01" # Exemplo: 1º de janeiro de 2023 até hoje

# Baixa os dados históricos do Yahoo Finance
getSymbols(ticker,
           src = "yahoo",
           from = data_inicio,
           to = Sys.Date(),
           auto.assign = TRUE)

# O objeto criado é chamado "PETR4.SA"
# Extraí o preço de fechamento ajustado (ideal para análise de séries temporais)
# 'Cl' é um atalho do quantmod para a coluna de fechamento ajustado
fechamento_ajustado <- Cl(get(ticker))

# Estuda a série
print(tail(fechamento_ajustado))
plot(fechamento_ajustado)
acf(fechamento_ajustado)
pacf(fechamento_ajustado)
hist(fechamento_ajustado)
```

Quais as hipóteses?

Exercício PETR4

Comandos R

```
# AJUSTE DO MODELO ARIMA AUTOMÁTICO (auto.arima)
modelo_arima <- auto.arima(fechamento_ajustado,
                           trace = TRUE,           # Imprime o processo de busca
                           stepwise = FALSE,        # Busca mais exaustiva (lenta, mas melhor)
                           approximation = FALSE)    # Sem aproximação (mais preciso)

print(summary(modelo_arima))

# DIAGNÓSTICO DOS RESÍDUOS
checkresiduals(modelo_arima)

# PREVISÃO (FORECAST)
horizonte_previsao <- 21
previsao <- forecast(modelo_arima, h = horizonte_previsao)
print(previsao)
plot(previsao,
     main = "Previsão ARIMA para PETR4.SA (Fechamento Ajustado)",
     xlab = "Data",
     ylab = "Preço (R$)")
```

Exercício PETR4

Comandos R

```
# =====  
# MACD  
library(TTR)  
library(quantmod)  
  
# O MACD usa os preços de Fechamento.  
precos_fechamento <- Cl(PETR4.SA)  
  
# Para uma EMA de 12 períodos ( $n=12$ ), o multiplicador  $\alpha$  será:  $2 / (12 + 1) \approx 0.1538$ , ou cerca de  $15.38\%$ .  
EMA_rapida <- EMA(precos_fechamento, n = 12)  
  
# Calcula a EMA de 26 períodos  
EMA_lenta <- EMA(precos_fechamento, n = 26)  
  
# Plota o gráfico principal (OHLC) sem indicadores (TA=NULL)  
chartSeries(PETR4.SA,  
            name = "PETR4.SA - EWMA para MACD",  
            theme = "white",  
            TA = NULL)  
  
addTA(EMA_rapida, on = 1, col = "red", lty = 1)  
addTA(EMA_lenta, on = 1, col = "blue", lty = 1)  
  
# Usando o pacote já implementado  
addMACD(fast = 12, slow = 26, signal = 9)
```

Exercício IPCA

Comandos R

```
library(rbcB)
library(forecast)

# Baixando dados mensais do IPCA (Código BCB 433) com rbcB
ipca <- rbcB::get_series(
  code = 433,
  start_date = "2010-01-01"
)

str(ipca)
ipca      #observe os dados
ipca[["433"]]

# Verificação da série:")
par(mfrow = c(2, 2))
plot(ipca, type="l")
hist(ipca[["433"]])
acf(ipca[["433"]])
pacf(ipca[["433"]])
```

Análise as Hipóteses

Comandos R

```
# Ajuste do Modelo; ARIMA? SARIMA?

modelo_sarima_ipca <- auto.arima(ipca_ts,
                                trace = TRUE,
                                stepwise = FALSE,
                                approximation = FALSE,
                                seasonal = TRUE)

print(summary(modelo_sarima_ipca))

checkresiduals(modelo_sarima_ipca)

horizonte_previsao <- 12
previsao <- forecast(modelo_sarima_ipca, h = horizonte_previsao)
print(previsao)
```

Analise e proponha alternativas'

Comandos R

```
library(forecast)
library(fpp2)

# Serie producao de cerveja australiana)
cerveja <- window(ausbeer, start = 1990)

cerveja      # observe que são dados trimestrais!!
str(cerveja)

# Verificação da série:)
par(mfrow = c(2, 2))

plot(cerveja,
      main = "Producao de Cerveja Australiana (desde 1990)",
      ylab = "Milhoes de Litros")

hist(cerveja)
acf(cerveja)
pacf(cerveja)
```

Analise as Hipóteses e proponha alternativas

Exercício Vendas no Varejo no Brasil

Comandos R

```
library(rbcbr)
library(forecast)

# Código 20108: Índice de Volume de Vendas no Varejo (Anual)
vendas <- rbcbr::get_series(
  code = 20108,
  start = "1990-01-01", # Geralmente, séries anuais têm um histórico mais longo
)

vendas # observe os dados
str(vendas)
# Verificação da série:")
par(mfrow = c(2, 2))
plot(vendas, ,type="l")
hist(vendas[["20108"]])
acf(vendas[["20108"]],lag.max=100)
pacf(vendas[["20108"]],lag.max=24)
```

Analise as Hipóteses e proponha alternativas