

CIÊNCIA DE DADOS (BIG DATA)

ANÁLISE ESTATÍSTICA

Professor curador: Mário Olímpio de Menezes



Mackenzie



TRILHA 7

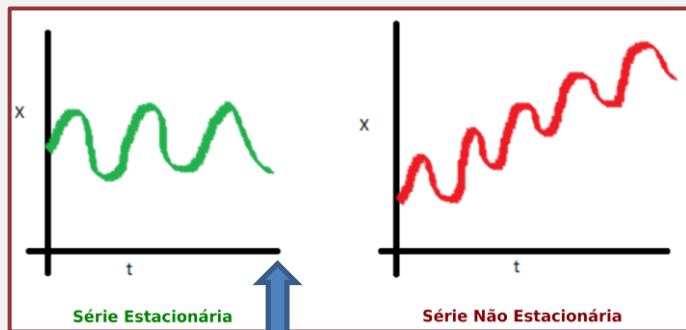
PARTE B – INTRODUÇÃO A SÉRIES TEMPORAIS

PARTE B – INTRODUÇÃO A SÉRIES TEMPORAIS

MODELAGEM DE SÉRIES TEMPORAIS

MODELOS AR E MA

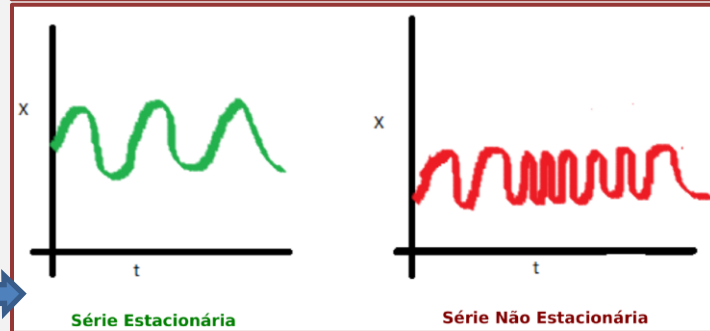
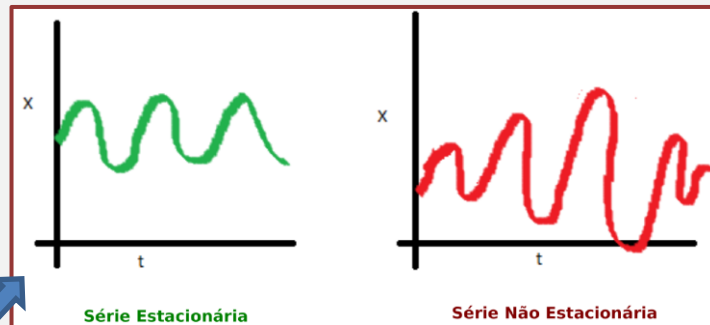
SÉRIES ESTACIONÁRIAS



Média constante

Variância não
deve ser função
do tempo

Covariância não
deve ser função
do tempo



POR QUE SÉRIES ESTACIONÁRIAS?

- Alguns modelos precisam que a série seja estacionária para serem aplicados.
- Quando o critério estacionário é violado, o primeiro requisito é *estacionarizar* a série temporal e, então, tentar algum modelo estocástico para prever a série temporal.
- Duas das várias maneiras de estacionarizar uma série são: *Detrending* e Diferenciando.

MODELOS ARMA DE SÉRIES TEMPORAIS

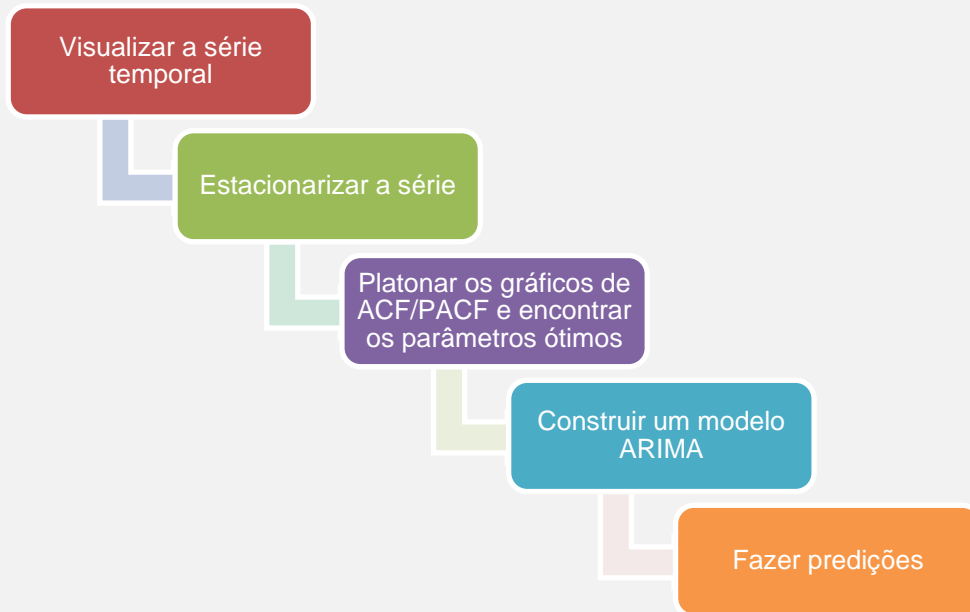
- AR – autorregressão.
- MA – *moving average* (média móvel).
- Aplicáveis a séries estacionárias.

AUTO-CORRELAÇÃO E AUTO-CORRELAÇÃO PARCIAL

Uma vez que se tem séries estacionárias, precisamos saber:

- É um processo AR ou MA?
- Qual ordem do processo AR ou MA precisaremos usar?

MODELAGEM DE SÉRIES TEMPORAIS COM ARIMA



MODELAGEM DE SÉRIES TEMPORAIS COM ARIMA

Depois da visualização da série, passamos à etapa de estacionarização. **Três técnicas** podem ser empregadas:

1. Remover a tendência.
2. Diferenciar, isto é, explorar as diferenças entre os termos.
3. Tratar a sazonalidade.

MODELAGEM DE SÉRIES TEMPORAIS COM ARIMA

- A diferenciação, ou a modelagem das diferenças dos termos, é dada por:

$$x(t) - x(t - 1) = ARMA(p, q)$$

- Esta diferenciação representa a **Integração** dos dois modelos, ou seja, AR(I)MA.

MODELAGEM DE SÉRIES TEMPORAIS COM ARIMA

TEMOS TRÊS PARÂMETROS

- **p** – para o modelo AR.
- **d** – para a Integração.
- **q** – para o modelo MA.

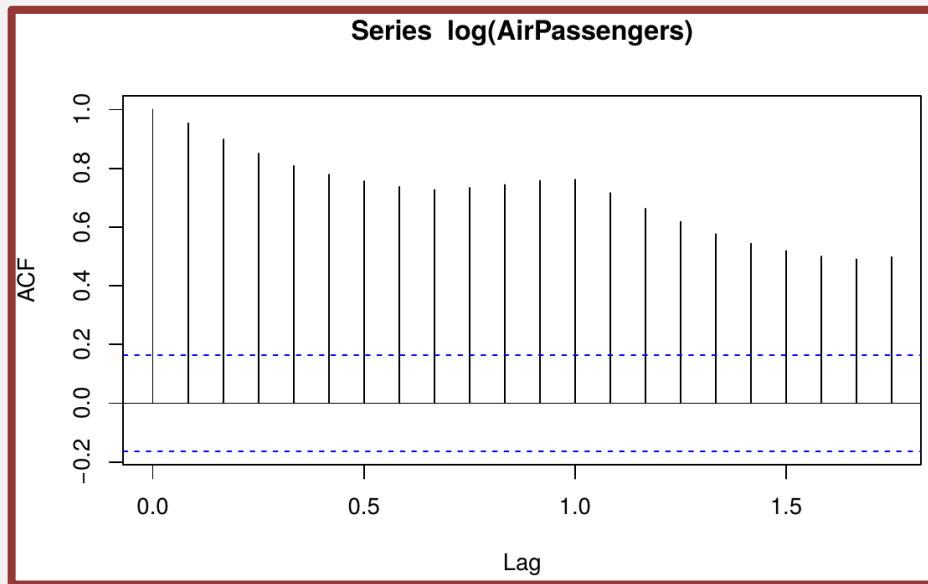
Estes parâmetros são encontrados pelos gráficos ACF e PACF.

APLICANDO O MODELO ARIMA

- Para remover a desigualdade da variância de uma série temporal, uma técnica utilizada é fazer o **log** da série.
- Para remover o componente de tendência de uma série temporal, uma técnica utilizada é pegar a **diferença** da série.
- Faremos isso na série **AirPassengers**.

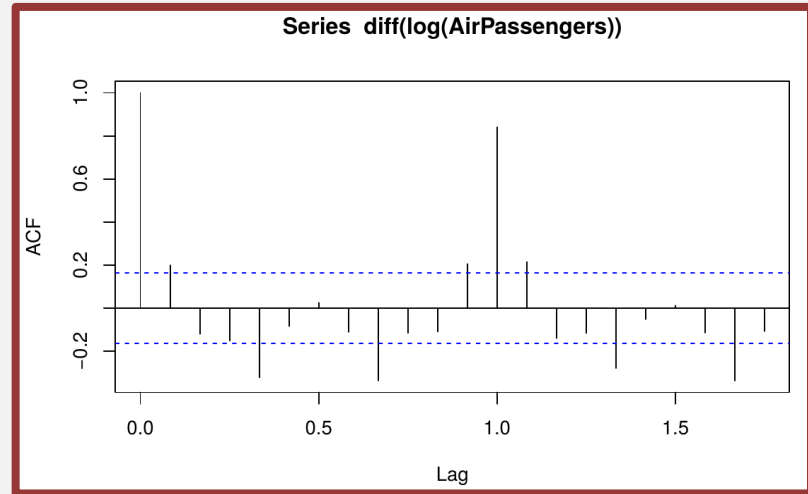
APLICANDO O MODELO ARIMA

- Decaimento muito lento de ACF.
- Somente o **log** não deixou a série estacionária.



APLICANDO O MODELO ARIMA

- Fazendo a diferença dos **logs**, vemos que **ACF** corta após o primeiro lag.
- O valor de **q** (processo **MA**) poderia ser 1 ou 2; o valor do segundo lag está próximo do limite de significância.



APLICANDO O MODELO ARIMA

Comparando os modelos com $q = 1$ e $q = 2$, verifica-se que **1** é o melhor valor de q . Este valor foi utilizado, então.

