

## PERGUNTAS REALIZADAS DURANTE A AULA E RESPONDIDAS POR E-MAIL

**Curso:** MBA em Data Science e Analytics

**Disciplina:** Séries Temporais II

**Data:** 19/10/2021

### Sem nome

Professor, no exemplo 5 a função Arima é diferente da mostrada anteriormente, inclusive com novos parâmetros, pode nos falar o motivo?

Marcos, o exemplo 5 está nos mostrando o caminho para se chegar ao melhor modelo, que pode ser um dentre os inúmeros modelos possíveis de séries temporais. A Família ARIMA possui algumas vertentes, como modelos que permitem identificar e considerar a sazonalidade (SARIMA), modelos AR, MA, ARMA, dentre outros.

### Janara Tatjana Llanque Sousa

Eu me perdi no Box-Ljung e o p-value é superior a 0.05, não tem correlação. Isso no modelo é bom ou não? Precisa de ajuste?

Janara, estatística Ljung-Box testa a hipótese nula de que as autocorrelações até lag  $k$  são iguais a zero (ou seja, os valores dos dados são aleatórios e independentes até um certo número de lags). Um pequeno valor  $p$  é uma evidência contra a independência.

### Marina Bendit Szacher 22:49

ainda não ficou claro como definir o  $p$  e  $q$ . O "d" deu para entender pois está relacionado à deixar a os dados estacionários

Marina, vamos tentar explicar de uma outra forma. Cada letra de  $p$ ,  $d$  e  $q$  representa uma parte da sigla ARIMA. Desta forma, o "p" representa o "AR" do nome do modelo, enquanto "d" representa o "I" e, por fim, "q" representa o "MA". Estes componentes podem ser definidos, grosso modo, como segue:

- $p$ : termo autoregressivo;
- $d$ : número de diferenciações; e
- $q$ : termo da média móvel.

Marina, caso ainda não tenha ficado claro, peço que, por gentileza, nos acione pelo e-mail da monitoria que iremos prontamente lhe auxiliar.

### Marcos Henrique Da Silva 22:13

professor, existe um indicador que mostra a "estacionaridade" da base? sem a necessidade de fazermos uma análise visual, suscetível a falhas?

Marcos, os testes utilizados para esta verificação são os seguintes: o teste ADF (Dickey-Fuller Aumentado), o teste PP (Phillips e Perron), o teste KPSS (Kwiatkowski, Phillips, Schmidt e Shin) e o teste LMC (Leybourne e McCabe), cada qual com suas aplicações e peculiaridades.

### André Araújo 22:04

É possível fazer uma análise de séries temporais numa perspectiva multinível?

André, creio que o tipo de modelo adequado para o caso são os modelos lineares de efeitos mistos multinível. Estes modelos permitem explorar e modelar adequadamente a variabilidade de dados com erros aleatórios heterocedásticos e/ou correlacionados,

agrupados, considerando ainda mais do que um fator – ou seja, agrupando os dados conforme uma perspectiva hierárquica.

**Patricia Barbosa Pacheco 21:57**

Professor teria um livro recomendado? Na aula do Favero confesso que além de rever a aula várias vezes acompanhei o Livro de análise de Dados dele e da Patricia. Teríamos uma leitura recomendada? obg

Patrícia, não há um livro texto para o módulo de séries temporais. Na bibliografia recomendada há alguns materiais. Se ainda assim não for suficiente, peço que nos acione pelo e-mail da monitoria que iremos prontamente lhe auxiliar.

**Maria Leidiane Kelly Dionísio Barros Sales 21:27**

Quando eu sei que a sazonalidade é ou não constante?

Maria, a maneira de determinar o período do fenômeno sazonal em uma série temporal é através da própria inspeção visual do gráfico. Uma outra maneira de determinar o período do fenômeno sazonal é através da análise da disposição dos resíduos oriundos da Regressão Linear para os valores de um gráfico do tipo Scatter. Neste caso estaríamos fazendo uma inspeção acerca dos componentes de erro vistos ao longo do tempo, e teríamos uma ideia sobre a sazonalidade inerente aos dados.

**Vitor Carvalho de Oliveira 20:40**

O que seria o parâmetro  $h$ , da função forecast?

Vítor, o parâmetro  $h$  diz respeito ao número de período de forecasting.

**Alisson Tavares Corrêa 20:28**

Boa noite, Prof Poderia pontuar com mais clareza sobre os conceitos dos componentes das ST e porque o ciclo sai do contexto dessa aula

Alisson, o componente determinístico diz respeito ao componente da série temporal que pode ser escrito através de uma função matemática determinada por uma ou mais variáveis, diz-se que ela contém apenas o componente determinístico. Ao passo que o componente aleatório, diz respeito aos componentes não captados pela função matemática descrita pelo componente determinístico.

A forma matemática pode ser descrita conforme segue:

$$y_t = \underbrace{a + b \cdot x_t}_{\text{determinístico}} + \underbrace{\epsilon_t}_{\text{estocástico}}$$

**Alisson Tavares Corrêa 20:27**

Boa noite professor Fvaor poderia explicar com mais clareza sobre o ARIMA ? E sobre a forma  $p, d$  e  $q$ ?

ARIMA é o acrônimo em inglês para AutoRegressive Integrated Moving Average. Em séries temporais, designa um modelo auto-regressivo integrado de médias móveis. Pode ser considerado ainda como uma generalização de um modelo auto-regressivo de médias móveis, se constituindo em uma família de modelos possíveis a depender do contexto e das

características de cada modelagem proposta com base nos termos  $p$ ,  $d$  e  $q$ . Os quais podem ser definidos, grosso modo, como segue:

- $p$ : termo autoregressivo;
- $d$ : número de diferenciações; e
- $q$ : termo da média móvel.

#### Gabriela Alves De Almeida 20:14

prof. tanto na aula quanto no material, não ficou claro se estocástico e determinístico são componentes de séries temporais, junto com tendência, sazonalidade, ciclo e aleatório.

Gabriela, o componente determinístico diz respeito ao componente da série temporal que pode ser escrito através de uma função matemática determinada por uma ou mais variáveis, diz-se que ela contém apenas o componente determinístico. Ao passo que o componente aleatório, diz respeito aos componentes não captados pela função matemática descrita pelo componente determinístico.

#### André Araújo 20:05

O  $ndiffs$  entrega pra gente a quantidade de diferenciações que devem ser feitas?

Andre, exatamente. O “ $ndiffs$ ” nos fornece uma medida para a quantidade de diferenciações de uma série temporal. Uma explicação sobre os parâmetros da função pode ser obtida por meio do seguinte link:

<https://search.r-project.org/CRAN/refmans/forecast/html/ndiffs.html>

Caso não tenha ficado claro, André, peço que, por gentileza, nos acione no e-mail da monitoria que iremos prontamente lhe auxiliar.

#### Marina Bendit Szacher 19:49

Olá, mas esses valores de  $p$ ,  $d$  e  $q$ , é escolhido aleatoriamente ou o modelo vai dar esses valores? Não entendi da onde eles vêm?

Marina, estes valores são obtidos por intermédio da observação dos componentes da série temporal, conforme segue:

- $p$ : termo autoregressivo;
- $d$ : número de diferenciações; e
- $q$ : termo da média móvel.

#### André Zacharias 21:26

Em alguns materiais vemos referência a 3 componentes (sazonalidade, tendência e erro) mas em outras partes vemos um quarto componente chamado ciclo... ciclo é de fato diferente da sazonalidade?

André, os componentes de uma série temporal dão os seguintes: tendência, ciclo, sazonalidade e componente aleatório.

#### Samya de Lara Lins de Araujo Pinheiro 21:24

Professor, boa noite. Poderia por gentileza explorar mais o significado dos parâmetros  $\alpha$ ,  $\beta$  e  $\gamma$ ? Devemos forçar  $\beta$  e  $\gamma$  zero baseado em análise visual? Ou aplicamos e avaliamos cada um?

Samya, alfa, beta e gama são as constantes de suavização para o nível, a tendência e a sazonalidade, respectivamente.

**Vanessa Hoffmann de Quadros 21:10**

Professor, qual a razão de ter aplicado o logaritmo aos dados no exemplo da linha 232?

Vanessa, a transformação logarítmica pode ser utilizada para dados que aparentam não serem aderentes a uma distribuição normal padrão.

Hernandes Matias Junior 124.799.116-40