**Professor:** Wesley Augusto de Freitas Borges

**Alunos:**

EDUARDO FLEISCHMANN

GUSTAVO JOSE MAZZOTTI MARTINS

RAFAEL DA SILVA

**Motivação:**

Realizar testes de estacionaridade, analisar os resultados e construir modelos ARIMA e SARIMA para melhor compreensão dos dados do indicador IBC-Br.

**Material:**

Para execução da atividade iremos considerar a série mensal IBC-Br entre o período de 2015 e 2024.

**ATIVIDADE EM GRUPO:**

Considerando a série mensal do IBC-Br entre 2015 e 2024 (código 24363 do Sistema Gerenciador de Séries Temporais do Banco Central do Brasil - BCB), realize as atividades descritas a seguir.

Tabela

Descrição gerada automaticamente

Campos que constam na série de dados:

ref.date, value, id.num e series.name (padrão do SGS)

1. Agora que você acessou e visualizou a série do IBC-Br, prossiga com as seguintes análises:
2. Identifique e descreva as principais características do IBC-Br no período analisado, focando em aspectos de tendência e sazonalidade.

Verifique se a série é estacionária utilizando tanto a Função de Autocorrelação (FAC) com o acf() quanto o teste de Dickey-Fuller Aumentado (ADF) com o adf.test().

• OBS.: Lembre-se de que uma das técnicas para estacionarizar uma série temporal é utilizar a transformação da taxa de crescimento, seja ela simples ou contínua.

**Resposta:**

Abaixo vamos plotar a série temporal e tentar capturar algum insight visual no gráfico:

Gráfico

Descrição gerada automaticamente

Avaliando o gráfico do índice IBC-Br podemos identificar claramente um padrão possivelmente associado a sazonalidade. Visualmente a série não aparenta ser estacionária e aparenta possuir alguma tendência. Há uma quebra estrutural em abril/2020 associado a pandemia, e podemos identificar dois padrões distintos a partir deste ponto, antes, aparentemente estacionário e após, aparentemente não estacionária e com tendência.

Gráfico, Gráfico de barras

Descrição gerada automaticamente

Analisando o subplot acima, conseguimos identificar alguns padrões associados aos meses do ano, sobretudo verificando o mês de março por exemplo. Indicando um padrão de sazonalidade.

Aplicando as funções de autocorrelação e autocorrelação parcial na série IBC-Br:

Gráfico, Histograma

Descrição gerada automaticamente

Analisando o gráfico de retorno da função de autocorrelação, podemos ver que há um decaimento lento ao logo dos lags mais distantes, o que pode indicar que a série é não estacionária.

Gráfico, Histograma

Descrição gerada automaticamente

Analisando a função de autocorrelação parcial vemos que existem picos em lags específicos o que pode indicar a presença de sazonalidade. O resultado apresentado neste teste é um tanto quanto inconclusivo. É necessário aplicarmos testes mais robustos para confirmar esta inferência que realizamos.

Seguindo nossas análises, realizamos a decomposição da série IBC-Br de modo a identificarmos os padrões de tendência e sazonalidade.

Interface gráfica do usuário, Gráfico

Descrição gerada automaticamente

Como visto na aula 3, apenas aplicar a função de autocorrelação (FAC) talvez não seja um método suficiente para capturar a estacionaridade de séries temporais mais complexas. Existem testes adicionais mais robustos que auxiliam na identificação da existência de uma raiz unitária na série temporal que estamos analisando. Um teste complementar para identificarmos se a série é ou não estacionária é o teste **ADF (Dickey-Fuller Aumentado)** apresentado a seguir:

**Augmented Dickey-Fuller Test**

data: na.omit(ibc\_ts)

Dickey-Fuller = -3.2049, Lag order = 4, p-value = 0.09056

alternative hypothesis: stationary

**Hipótese nula (H0):** A série tem raiz unitária (não estacionária).

**Hipótese alternativa (H1):** A série é estacionária.

Considerando os resultados do teste ADF, podemos obter duas análises distintas se considerarmos níveis de significância diferentes. Por exemplo, se considerarmos 5% de nível de significância, em função do P-VALOR ser de 0.09056 podemos não rejeitar a hipótese nula, indicando que a série é não estacionária. Porém, se considerarmos um nível de significância de 10%, a análise muda de figura. Neste caso, em função do P-VALOR de 0.09056 ser menor que o nível de significância, podemos rejeitar a hipótese nula de que a série é não estacionária, e existe a possibilidade de não haver uma raiz unitária. Como nosso P-VALOR ficou muito próximo de 10% gerando um alto nível de incerteza, iremos considerar um nível de significância de 5%, e **não rejeitaremos a hipótese nula** de que a **série é não estacionária e possui raiz unitária.**

Para complementar a evidência de que a série IBC-Br em nível é não estacionária, vamos aplicar o **teste KPSS** e avaliar seus resultados:

* **Hipótese nula (H0):** A série é estacionária.
* **Hipótese alternativa (H1):** A série não é estacionária.

**#######################**

**# KPSS Unit Root Test #**

**#######################**

Test is of type: mu with 4 lags.

Value of test-statistic is: 1.061

Critical value for a significance level of:

10pct 5pct 2.5pct 1pct

critical values 0.347 0.463 0.574 0.739

Em função do valor do teste estatístico ser 1.061 e este valor ser maior que qualquer um dos valores críticos, podemos rejeitar a hipótese nula e concluir que a série IBC-Br é não estacionária.

Como identificamos através das funções de Autocorrelação, e dos testes ADF e KPSS que a série IBC-Br é não estacionária, iremos aplicar a diferenciação de modo a provocar a estacionariedade nesta série:

Uma imagem contendo Gráfico

Descrição gerada automaticamenteGráfico

Descrição gerada automaticamente

**Augmented Dickey-Fuller Test**

data: na.omit(ibc\_diff)

Dickey-Fuller = -4.9295, Lag order = 4, **p-value = 0.01**

alternative hypothesis: stationary

**#######################**

**# KPSS Unit Root Test #**

**#######################**

Test is of type: mu with 4 lags.

**Value of test-statistic is: 0.0558**

Critical value for a significance level of:

10pct 5pct 2.5pct 1pct

critical values 0.347 0.463 0.574 0.739

Como podemos perceber, ao aplicar a diferenciação na série IBC-Br, induzimos a estacionariedade na série, deixando-a mais adequada para construção dos modelos de predição.

1. Após verificar a estacionariedade da série e tendo realizado as transformações necessárias, construa um modelo ARIMA para o IBC-Br utilizando a função Arima() ou auto.arima(). Apresente os resultados dos modelos, comparando os critérios de AIC e BIC, e interprete os coeficientes dos parâmetros estimados.

Construindo o modelo Arima utilizando a função auto.arima() com base na série IBC-Br em nível:

Series: ibc\_ts **(IBC-Br em nível)**

ARIMA(1,1,2)

Coefficients:

ar1 ma1 ma2

-0.4311 0.0411 -0.6477

s.e. 0.1544 0.1312 0.0834

sigma^2 = 21.15: log likelihood = -331.74

**AIC=671.47 AICc=671.84 BIC=682.38**

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set 0.2885022 4.517265 3.286808 0.1070083 2.368131 0.8001108 0.01797069

Construindo o modelo Arima utilizando a função auto.arima() com base na série IBC-Br Diferenciada:

Series: ibc\_diff (IBC-Br diferenciado)

ARIMA(1,0,2) with zero mean

Coefficients:

ar1 ma1 ma2

-0.4311 0.0411 -0.6477

s.e. 0.1544 0.1312 0.0834

sigma^2 = 21.15: log likelihood = -331.74

AIC=671.47 AICc=671.84 BIC=682.38

Training set error measures: ME RMSE MAE MPE MAPE MASE ACF1

Training set: 0.289825 4.537189 3.314666 56.17213 169.0413 1.10053 0.01783476

ARIMA(1,0,2) with zero mean

Coefficients:

ar1      ma1     ma2

-0.4311  0.0411  -0.6477

s.e.  0.1544  0.1312   0.0834

O modelo analisado é autorregressivo de primeira ordem, sem diferenciação e com dois termos de média móvel.

O ar1 de -0.4311 indica que o valor atual da série é negativamente influenciado pelo número imediatamente anterior. Mais especificamente em 43% da variação anterior.

O ma1 de 0.041 expressa a dependência do valor atual ao erro de previsão do período anterior, logo se o modelo errou para cima na previsão anterior, o número atual vai receber um pequeno acréscimo.

O ma2 é similar, porém muito mais relevante, e leva em consideração o erro de dois períodos atrás. Nesse caso, se o erro de t-2 foi negativo por exemplo, o modelo irá gerar uma correção positiva de 64% desse valor para o lado contrário, ou seja, positiva.

1. Avalie a adequação do modelo ARIMA analisando a FAC dos resíduos e aplicando o teste de Ljung-Box com Box.test() para verificar se os resíduos são ruídos brancos. Com base nessas análises, discuta a qualidade do modelo construído para modelar o IBC-Br. Indique se o modelo é apropriado para previsões futuras e quaisquer limitações que possam existir.

Gráfico, Histograma

Descrição gerada automaticamenteGráfico

Descrição gerada automaticamente

Box-Pierce test

data: residuals(modelo\_arima\_nivel)

X-squared = 0.036816, df = 1, p-value = 0.8478

Box-Pierce test

data: residuals(modelo\_arima\_diff)

X-squared = 0.035943, df = 1, p-value = 0.8496

Avaliando os resultados da função de autocorrelação, podemos verificar que em alguns lags existem pontos que fogem dos intervalos de confiança o que indica que esses modelos não capturaram algum fator que influencia no comportamento da série IBC-Br. Observando com um pouco mais de atenção, podemos ver que há um padrão nesses pontos o que pode significar que nosso modelo não capturou a sazonalidade contida na série.

Por outro lado, os resultados apresentados pelo Ljung Test nos indicam que os resíduos do modelo se comportam como ruido branco, portanto, não podemos rejeitar a hipótese nula deste teste.

Com base nos resultados acima, podemos considerar que a qualidade do modelo não é a ideal para previsões futuras em função do modelo não capturar a sazonalidade contida na série IBC-Br.

1. Considerando os aspectos cíclicos da economia e sua influência sobre a atividade econômica, realize as seguintes análises para explorar a sazonalidade e como isso pode influenciar a modelagem do IBC-Br:
2. Avalie e argumente a presença de sazonalidade na série utilizando a FAC, a FACP com pacf() e o processo de decomposição com decompose() ou stl().

Resposta: Durante a exploração dos dados na questão 1, nós já evidenciamos a presença de sazonalidade e tendência, ao plotarmos o gráfico decomposto da série IBC-Br. Vamos analisar o resultado desta decomposição.

Interface gráfica do usuário, Gráfico

Descrição gerada automaticamente

Como podemos verificar, existe claramente um padrão sazonal, sobretudo nos meses de março e abril e existe uma tendência de crescimento, que fica mais acentuada após a quebra estrutural que ocorreu entre os meses de março e abril de 2020.

1. Construa um modelo SARIMA para a série utilizando a função Arima() ou auto.arima() com parâmetros sazonais. Apresente os resultados e interprete os coeficientes dos parâmetros estimados.

Series: ibc\_ts

ARIMA(1,1,1)(2,0,0)[12]

Coefficients:

ar1 ma1 sar1 sar2

0.6983 -0.9478 0.4955 0.2748

s.e. 0.1122 0.0571 0.0906 0.0982

sigma^2 = 11.98: log likelihood = -303.31

**AIC=616.62 AICc=617.18 BIC=630.26**

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set 0.3078386 3.384455 2.438155 0.166174 1.763446 0.5935223 -0.04718798

Series: ibc\_diff

ARIMA(1,0,1)(2,1,0)[12]

Coefficients:

ar1 ma1 sar1 sar2

0.7105 -0.9246 -0.5171 -0.2618

s.e. 0.1049 0.0630 0.1000 0.0948

sigma^2 = 12.03: log likelihood = -269.29

**AIC=548.57 AICc=549.2 BIC=561.65**

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set 0.5490089 3.213223 2.304845 44.11862 151.9563 0.7652509 -0.03247928

Ao utilizarmos a função auto.arima() obtivemos um modelo SARIMA com a seguinte estrutura:

Considerando a série IBC-Br diferenciada, ARIMA(1,0,1)(2,1,0)[12]

Vamos interpretar estas informações:

Este modelo possui um termo autorregressivo de ordem 1, não possui um termo de integração, pois nós já aplicamos a diferenciação na série anteriormente, possui um termo de média móvel de ordem 1, possui termo Autorregressivo Sazonal de ordem 2, possui termo de diferenciação sazonal de ordem 1 e não possui nenhum termo de média móvel sazonal.

Ao construirmos o modelo acima obtivemos os seguintes coeficientes:

**ar1: 0.7105** – Coeficiente do termo autorregressivo positivo, indicando que há relação positiva entre o valor atual e o valor anterior.

**ma1: - 0.9246** – Coeficiente do termo de média móvel de ordem 1, negativo, indicando que o erro do passo anterior tem influencia negativa no valor atual. Este valor indica que os erros passados têm peso relevante na correção do valor atual.

**sar1: -0.5171 e sar2: -0.2618** – Esses termos se referem aos coeficientes dos termos autorregressivos sazonais de ordem 1 e 2. Seus valores negativos indicam que a relação entre os valores observados em lags sazonais de 12 lags é inverso.

1. Compare a performance dos modelos ARIMA e SARIMA em termos de critérios de seleção de modelos (AIC e BIC) e análise dos resíduos. Justifique qual modelo seria mais adequado para prever a atividade econômica e explique as implicações de cada modelo para a tomada de decisões na sua instituição.

Comparando os critérios de seleção de modelos (AIC e BIC) temos o seguinte cenário:

**Modelos Arima sem componente Sazonal:**

**AIC=**671.47 **AICc=**671.84 **BIC=**682.38

Teste Ljung Box para os modelos Arima (em nível e diferenciado)

**Box-Pierce test - data: residuals(modelo\_arima\_nivel)**

X-squared = 0.036816, df = 1, p-value = 0.8478

**Box-Pierce test - data: residuals(modelo\_arima\_diff)**

X-squared = 0.035943, df = 1, p-value = 0.8496

Modelos Arima com componente Sazonal (SARIMA):

Valores para série em nível (não estacionária)

**AIC=**616.62 **AICc=**617.18 **BIC=**630.26

Valores para a série diferenciada (com estacionariedade induzida)

**AIC=**548.57 **AICc=**549.2 **BIC=**561.65

**Box-Pierce test - data: residuals(modelo\_sarima\_nivel)**

X-squared = 0.25384, df = 1, p-value = 0.6144

**Box-Pierce test - data: residuals(modelo\_sarima\_diff)**

X-squared = 0.1192, df = 1, p-value = 0.7299

Com os resultados acima, podemos concluir que, utilizando os critérios de seleção de modelos, o modelo SARIMA aplicado a série IBC-Br com estacionaridade induzida é um modelo mais robusto e adequado a se utilizar para tomada de decisões. Apesar de ter mais coeficientes inclusos, a redução nos valores dos critérios de seleção é expressiva em comparação com os demais modelos. Para evidenciar, podemos plotar uma previsão para os próximos 6 meses considerando os modelos e visualizando seu comportamento:

Forecast dos modelos Arima utilizando as séries em Nível e diferenciada

Gráfico

Descrição gerada automaticamenteInterface gráfica do usuário

Descrição gerada automaticamente

Forecast dos modelos SARIMA utilizando as séries em Nível e diferenciada

Gráfico

Descrição gerada automaticamente com confiança médiaInterface gráfica do usuário

Descrição gerada automaticamente

Como podemos evidenciar nos gráficos, as previsões fornecidas pelo modelo SARIMA, são mais coerentes do que as previsões do modelo ARIMA.