



## **EAE0327 – Econometria III**

Prof. André Luis Squarize Chagas

Prof. Assistente: Felipe Guimarães dos Santos

### **Trabalho da Disciplina**

Partes I e II

São Paulo

2023

## **1. Orientações**

O trabalho é dividido em Parte 1 (modelos univariados) e Parte 2 (modelos multivariados). Em ambas as partes, vocês devem usar os tópicos levantados para elaborar um texto em formato de artigo, problematizando as perguntas, discorrendo sobre os modelos e anexando figuras e tabelas que auxiliem seus resultados.

Os artigos podem ser escritos separadamente (um para cada parte), embora explorar uma relação entre as partes seja proveitoso. Além da solução das perguntas descritas neste trabalho, o uso de citações, a coesão do texto e a discussão metodológica serão pontos avaliados.

A base que deve ser utilizada já está inserida no Moodle, e seus dados foram obtidos no Ipeadata, com periodicidade mensal. O trabalho deve ser feito em grupos (até 6 integrantes), não sendo aceitos trabalhos feitos de forma individual.

A cada entrega parcial devem ser enviados o artigo em formato PDF e a rotina (script) em R, devidamente comentada e separada em tópicos conforme apresentado em aula. As datas das entregas parciais estão disponíveis no programa do curso, sendo que apenas um componente do grupo deve realizar o envio do trabalho, contendo uma capa com o nome e número USP de todos os integrantes do grupo.

## 2. Parte I – Modelos Univariados

Como um dos fatores de maior impacto sobre o IDH e a distribuição de renda, o emprego no Brasil é um das variáveis mais relevantes para a elevação do bem-estar social. Não é de se surpreender, portanto, tamanha preocupação dada à permanência do emprego informal e sobre a elevação do desemprego, como mostrada pela PNAD Covid-19, fruto do esforço do IBGE para avaliar essa dimensão durante o isolamento social.

A falta de trabalho na localidade associada às restrições do período levou 16,8 milhões de brasileiros à desocupação. Tal subutilização da força de trabalho é uma preocupação que é discutida em diferentes contextos pelas correntes econômicas, que propõe soluções para a redução do desemprego e das desigualdades que dele advém (Peggy Lowell, 1995).

Assim, a proposta da primeira parte deste trabalho é modelar uma série de ocupação, realizar sua análise e previsão, para que se possam fazer apontamentos sobre o comportamento esperado do mercado de trabalho brasileiro antes da pandemia.

Para os exercícios abaixo considere a série ‘pessoas ocupadas - Pessoa (mil)’.

**Questão 01.** No código, introduza a base em um diretório, e defina as variáveis como séries temporais.

Verifique os dados disponíveis procurando por eventuais duplos registros, erros tipográficos e outliers. Teste o uso de um *boxplot* para encontrar outliers, e compare com o uso de uma função específica para séries temporais. Se for necessário (considere a tendência da série e seus conhecimentos econômicos) empregue os valores sugeridos pela função *tsclean()* para substituir outliers.

Apresente os dados formalmente em seu artigo, explicitando o método de coleta, instituição responsável, periodicidade e períodos disponíveis. Descreva sinteticamente as eventuais correções que realizaram.

**Questão 02.** Apresente os gráficos da série de pessoas ocupadas entre março de 2012 a junho de 2020, com legendas e títulos. Realize a decomposição da série entre tendência, sazonalidade e erro.

Explicita quais fatores econômicos que podem gerar sazonalidade e erro na formação dessa série. Explicita a diferença entre a série empregada e outras fornecidas pelo IBGE tais como PIA, Pessoas na Força de Trabalho e Força de trabalho potencial.

**Questão 03.** Qual é o comportamento das funções de autocorrelação e autocorrelação parcial dessas séries? Sem um teste de raiz unitária qual conclusão seria induzida?

Realize e reporte os resultados de um teste de raiz unitária. Com uma série em primeira diferença, proceda novamente a análise gráfica da FAC e FACP. Quais são as conclusões induzidas pelos gráficos?

Explique a diferença entre a diferenciação em nível e a sazonal. Determine se nesse modelo convém realizar a diferenciação sazonal.

**Questão 04.** Utilizando a função *auto.arima()* identifique um modelo (SARIMA). Discuta se o modelo proposto condiz com o esperado a partir dos testes realizados nas questões anteriores.

Estime outras duas opções de modelos com ordens diferentes, escolhidas a partir dos resultados obtidos nas questões anteriores. Ao menos um modelo dentre estes deve apresentar, ao final do trabalho, resíduos homocedásticos e não autocorrelacionados.

**Questão 05.** Proceda com testes de Ljung-Box para 4, 12 e 24 meses para cada um dos modelos propostos na questão anterior. O que os resultados implicam? Ao realizar uma análise gráfica dos resíduos do modelo obtém-se o mesmo resultado?

Realize uma análise gráfica para examinar a normalidade dos resíduos de cada modelo. Ao se empregar o teste de Shapiro ou o teste de Jarque Bera poderíamos contestar o resultado da análise gráfica?

Realize o teste de heterocedasticidade ARCH LM para os modelos. O que os resultados implicam?

Apresente os resultados dos testes realizados para cada modelo, sintetizados em uma única tabela.

**Questão 06.** Dentre os modelos com resíduos homocedásticos e não correlacionados exiba uma previsão para a ocupação de 24 meses, conjuntamente a um intervalo de confiança de 95%.

Realize o teste de acurácia dentro da amostra para cada modelo com resíduos homocedásticos e não autocorrelacionados. Qual modelo apresenta o menor RMSE?

**Questão 07.** Realize o teste de acurácia fora da amostra para cada modelo, considerando como treino os dados de março de 2012 até dezembro de 2018 e como teste os dados de janeiro de 2019 até junho de 2020.

Explique qual é o papel deste teste e como o resultado projetado de cada modelo se comportou quando comparado aos dados originais. Explique as limitações dos modelos para prever os valores observados a partir de janeiro de 2020.

**Questão 08.** Considerando os resultados de 2020, o valor projetado pelo seu modelo com menor RSME fora da amostra e o valor apresentado para o salário mínimo real (presente na base de dados) informe qual foi a perda real estimada devido à queda do número de ocupações frente ao que era previsto (atenção para com as das unidades de medida das variáveis).

### 3. Parte II – Modelos Multivariados

Desde a formulação da Curva de Phillips Ampliada a elaboração de um complexo ferramental teórico permitiu a discussão profunda sobre as causas e efeitos da inflação e do desemprego. A inserção de variáveis e postulados novo-keynesianos e a formação da equação de Tobin, permitiram demonstrar o papel ativo dos Ministérios da Economia e dos Bancos Centrais na determinação do nível de emprego e redução da inflação.

No cenário posto em 2020 ambas as variáveis assumiram a atenção pública, onde 1/3 dos brasileiros veem o desemprego como o pior problema da economia (JOTA, 2020) e dois em cada três brasileiros acreditam que a inflação irá aumentar (Datafolha, 2020). Este processo de incerteza e queda da produção já levou mais de 16,8 milhões de brasileiros à desocupação, por conta da pandemia ou falta de trabalho na localidade (Pnad 2020). O principal custo desse cenário é a produção perdida, que reduz o produto interno e, consequentemente, o bem-estar de muitos brasileiros.

Os principais afetados neste contexto são os indivíduos dos grupos de menor renda, mais expostos à desocupação, e que desde a formação do plano real viram uma elevação do seu poder de compra e renda. É neste contexto que se faz relevante, realizar neste trabalho, uma consulta à literatura existente, para que por sua sumarização e formação de modelos econométricos, se possam levantar algumas relações entre ocupação, produção e inflação esperada.

Para as questões abaixo considere as séries indo de março de 2012 até junho de 2020.

**Questão 01.** Escolha ao menos duas variáveis dentro do banco de dados com as quais acreditam que a ocupação possa ser parcialmente explicada. Defina as variáveis escolhidas como séries temporais e realize a correção dos dados, conforme realizado na parte I. Retire o componente sazonal das variáveis usando a função do filtro X-13: seas().

Apresente os dados formalmente em seu artigo, explicitando o método de coleta, instituição responsável, periodicidade e períodos disponíveis. Descreva sinteticamente as eventuais correções que realizaram e apresente os gráficos das séries.

**Questão 02.** Ao realizar a análise gráfica da FAC e da FACP de cada série qual é a conclusão? Realize um teste de raiz unitária e determine se é necessário tomar diferenciações para tornar as séries estacionárias.

**Questão 03.** Realize a identificação do modelo VAR por meio da função VARselect(), considerando o critério Bayesiano para a tomada de defasagens. Explícite o motivo de escolha da ordem de inserção das variáveis.

**Questão 04.** Proceda com os testes de estabilidade, autocorrelação serial e heterocedasticidade. Apresente os resultados dos testes realizados, sintetizados em uma única tabela. Explícite brevemente as diferenças entre os testes e seus resultados.

Realize uma análise gráfica para examinar a normalidade dos resíduos de cada variável do modelo. Ao se empregar o teste de Shapiro ou o teste de Jarque Bera poderíamos contestar o resultado da análise gráfica?

**Questão 05.** Realize o teste de acurácia dentro da amostra e posteriormente teste fora da amostra para cada modelo, considerando como teste os dados de janeiro de 2019 até junho de 2020.

**Questão 06.** Obtenha as funções de resposta ao impulso das variáveis e realize a decomposição do erro de previsão do modelo proposto.