

## ***Trabalho e sazonalidade: econometria como ferramenta aplicável a vagas de emprego no Brasil***

O desenvolvimento econômico nacional encontra-se intimamente ligado à quantidade de empregos disponíveis. Assim, para estudar esse desenvolvimento, tornou-se necessário analisar variáveis relacionadas a ele, tais como admissões e demissões. Para tanto, o objetivo deste estudo detém-se em encontrar modelos matemáticos capazes de explicar o comportamento dessas variáveis por meio de modelos de previsão lineares (ARIMA), tendo sido possível com a aplicação da metodologia de Box e Jenkins. A análise refere-se à quantidade de admissões e demissões, no período de maio de 1999 até outubro de 2022. Os modelos matemáticos encontrados para a variável de admissões foi um SARIMA (1,1,1)(1,1,0)<sub>12</sub>, bem como para demissões, também um SARIMA(1,1,0)(2,0,0)<sub>12</sub>. Esses modelos são capazes de prever o comportamento das séries, acompanhando a movimentação econômica do país. Dessa forma, pode-se proporcionar um estudo mais aprofundado sobre o desenvolvimento econômico, por meio dessas modelagens, visando identificar as melhores épocas do ano para a troca de emprego, quando ocorre maior quantidade de vagas livres no mercado, assim como, se utilizar desses modelos para aliar a séries de índices econômicos e embasar o comportamento dessas variações.

**Palavras-chave:** Modelagem ARIMA; Metodologia Box e Jenkins; Emprego; Admissões; Demissões.

## ***Work and seasonality: econometrics as a tool applicable to job vacancies in Brazil***

National economic development is closely linked to the number of available jobs. Thus, in order to study this development, it became necessary to analyze variables related to it, such as admissions and dismissals. Therefore, the objective of this study focuses on finding mathematical models capable of explaining the behavior of these variables through linear forecasting models (ARIMA), which was possible with the application of the methodology of Box and Jenkins. The analysis refers to the number of admissions and dismissals, from May 1999 to October 2022. The mathematical models found for the admission variable were a SARIMA (1,1,1)(1,1,0)<sub>12</sub>, as well as for dismissals, also a SARIMA(1,1,0)(2,0,0)<sub>12</sub>. These models are able to predict the behavior of the series, following the economic movement of the country. In this way, a more in-depth study of economic development can be provided through these models, in order to identify the best times of the year for changing jobs, when there is a greater number of vacancies in the market, as well as, if using these models to combine series of economic indices and support the behavior of these variations.


**Keywords:** ARIMA Modeling; Box and Jenkins Methodology; Employment; Admissions; Dismissals.


Topic: **Pesquisa Operacional**


Received: **01/12/2022**

Approved: **25/11/2022**

Reviewed anonymously in the process of blind peer.

**Renata Rocha de Oliveira**   
Universidade Federal de Santa Maria, Brasil  
<http://lattes.cnpq.br/8846571240091795>  
<https://orcid.org/0000-0002-2057-0217>  
[renata.oliveira@ufsm.br](mailto:renata.oliveira@ufsm.br)

**Lorena Vicini**   
Universidade Federal de Santa Maria, Brasil  
<http://lattes.cnpq.br/8676745244195312>  
<https://orcid.org/0000-0002-9959-6902>  
[vicini22@gmail.com](mailto:vicini22@gmail.com)

**Luis Felipe Dias Lopes**   
Universidade Federal de Santa Maria, Brasil  
<http://lattes.cnpq.br/1074372911061770>  
<https://orcid.org/0000-0002-2438-0226>  
[lflopes67@yahoo.com](mailto:lflopes67@yahoo.com)



DOI: 10.6008/CBPC2179-684X.2022.004.0011

### **Referencing this:**

OLIVEIRA, R. R.; VICINI, L.; LOPES, L. F. D.. Trabalho e sazonalidade: econometria como ferramenta aplicável a vagas de emprego no Brasil. **Revista Brasileira de Administração Científica**, v.13, n.4, p.147-163, 2022. DOI: <http://doi.org/10.6008/CBPC2179-684X.2022.004.0011>

## INTRODUÇÃO

A palavra emprego tem sua origem na língua inglesa, aparecendo em meados de 1400 D.C. até o início do século XVIII, a qual descrevia alguma atividade ou determinado labor; não estava conectada a cargos ou posição dentro de uma organização. No século XIX, passou a ser entendida como o trabalho realizado nas fábricas ou nas nações em que o capitalismo estava em ascensão na fase de industrialização (WOLEK, 2008).

As temáticas do emprego são sempre muito discutidas, pois são uma das maiores influências para que uma economia seja desenvolvida ou esteja em processo desenvolvimento. A quantidade de admissões descreve o andamento do mercado, abertura de empresas e novas oportunidades econômicas. Por outro lado as demissões revelam estagnação no mercado financeiro e perda econômica para o país.

Para analisar a quantidade de admissões e demissões os dados foram organizados em observações mensais, a fim de estudar e descrever seu comportamento. Dessa forma, esse trabalho objetiva analisar por meio de séries temporais, as variáveis admissão e demissão. De acordo com a teoria de Box et al. (1970) e Reinsel (1994), a análise de séries temporais tem os objetivos de analisar as séries, modelar e a realizar a previsão. Essa metodologia permite que valores futuros de uma série sejam previstos tendo como base seus valores passados. A previsão é realizada por meio da correlação temporal entre os valores existentes na série.

O mercado de trabalho no Brasil tem apresentado, nas últimas décadas, intensa variação na taxa de participação de pessoas economicamente ativas (de 15 a 64 anos que trabalha ou procura emprego), em relação à criação de novos postos de trabalho, no perfil dos trabalhadores (em termos de sexo, idade, escolaridade e qualificação) e em diversos aspectos relacionados a qualidade dos empregos e condições de contratação (CEPAL et al., 2013) (IPEA, 2014) (IPEA, 2019).

Algumas variáveis afetam o mercado de trabalho de forma direta, tal como o desemprego, que é um problema macroeconômico. Para a maioria das pessoas, a perda de um emprego significa um padrão de vida reduzido e uma angústia psicológica (BURGARD, 2009). Não causa surpresa o fato de que o desemprego seja um tópico frequente em debates políticos, e que frequentemente declarem que as políticas econômicas propostas por eles auxiliam na criação de novos postos de trabalho, e de exclusão de postos, conforme políticas internacionais (MANKIW, 2011) (MTP, 2021).

Dada à importância da quantidade de empregos disponíveis em um país, para seu desenvolvimento econômico, justifica-se analisar as variáveis relacionadas com este tema, incluindo a pandemia de Covid-19 que abateu o mundo em 2020. Com base na economia brasileira e suas oscilações em relação aos postos de trabalho, analisa-se as séries de admissões e demissões, a fim de compreender o comportamento e prever essas variações ao longo do tempo e do contexto econômico, para que se necessário, sejam tomadas novas estratégias na criação de novos postos de trabalho.

A problemática se encontra em analisar as variáveis referentes a admissões e demissões, e prever a quantidade de empregos, no Brasil, no período de maio de 1999 até outubro de 2022, utilizando para isso a metodologia Box e Jenkins, para modelos lineares.

## DISCUSSÃO TEÓRICA

### Uma breve definição de empregos

O relatório da CEPA et al. (2013) e CEPA et al. (2022), declaram que o mercado de trabalho no Brasil, entre o início dos anos 90 e 2006, caracterizou-se pela heterogeneidade e por um elevado déficit de trabalho condizente aos padrões considerados pela Organização Internacional do Trabalho (OIT), e tem como principais razões as altas taxas de desemprego e de informalidade, resultando em baixo grau de proteção social e inadequada inserção dos trabalhadores no mercado de trabalho.

Dessa forma, não apenas o desemprego é um grave problema no Brasil, mas também a insuficiente qualidade das ocupações. Assim, a taxa de desocupação não é capaz de indicar, isoladamente, todo o grau de precariedade da inserção dos trabalhadores no mercado (CEPAL et al., 2022). Seguem conceitos principais desse estudo:

**Admissões** - O conceito de admissões vem de um compromisso firmado por meio do contrato de trabalho, estabelecido entre empregado e empregador, no qual o colaborador compromete-se a prestar serviço não eventual para o empregador mediante recebimento de salário (VIANNA, 2014).

**Demissões** - Demissões é um conceito que se refere à renúncia ou ao abandono de um emprego, de um cargo, de uma comissão etc. Conforme Nascimento (2009) a demissão ocorre por meio da rescisão do contrato de trabalho, sendo esse o termo utilizado para indicar o fim da relação de emprego. Cabe ao empregador contratar e demitir o funcionário. A demissão não é necessariamente uma punição, nem mesmo depende da aceitação do empregado (NASCIMENTO, 2009).

### Análise de Séries Temporais

Uma série temporal é um conjunto de observações em uma sequência de tempo se referindo a um fenômeno aleatório (BAYER, 2008), sendo composta por quatro elementos: a tendência (evidencia o sentido do deslocamento da série ao longo do tempo); o ciclo (periodicidade do movimento da série ao longo dos anos); a sazonalidade (picos repetitivos em intervalos iguais de tempo dentro da série, geralmente inferior a um ano); e o erro aleatório (corresponde a variabilidade intrínseca dos dados, não podendo ser modelado). Em suma, o conceito de série temporal está relacionado a um conjunto de observações de uma determinada variável feita em períodos sucessivos ao longo de um determinado intervalo.

De acordo com Morettin et al. (2006), os objetivos da análise de uma série temporal, observada nos instantes  $t$  de tempo, são: a) Investigar o mecanismo gerador da série temporal; como por exemplo, analisar uma série de altura de ondas, surgindo o questionamento de como as ondas foram geradas; b) Fazer previsões de valores futuros de uma série; estas previsões podem ser de curto prazo, como uma série de vendas, de produção de estoque, ou de longo prazo, como séries populacionais, etc.; c) Descrever o comportamento da série; aqui, a construção do gráfico, a verificação da existência de tendências, ciclos ou variações sazonais, construção de histogramas e diagramas de dispersão, etc.; podem ser ferramentas úteis a serem analisadas; d) Procurar periodicidades relevantes nos dados; neste caso, pode ser de grande valia o

uso da análise espectral.

Existem dois principais aspectos a serem considerados no estudo de séries temporais: análise e modelagem. A primeira condição para a análise e modelagem de séries temporais é a estacionariedade. A detecção da existência ou não de estacionariedade em uma série pode ser feita a partir da análise gráfica, em busca de padrões, ou realizando testes estatísticos de raiz unitária. Séries de tempo não estacionárias geram resultados distorcidos, mas podem ser transformadas em estacionárias através do uso de diferenças (MATOS, 2000).

Conforme Matos (2000), um processo estocástico é estacionário, quando sua média e sua variância são constantes ao longo do tempo e quando o valor da covariância entre os dois períodos depende apenas da defasagem entre eles. Assim, de acordo com as equações escritas a seguir, pode-se observar:

- (i)  $E[y_t] = [y_{t-s}] = \mu$ , com  $t \neq s$ , média constante ao longo do tempo;  
 $E[(y_t - \mu)^2] = E[(y_{t-s} - \mu)^2] = \sigma_y^2$ , variância constante;
- (ii)  $E[(y_t - \mu)(y_{t-s} - \mu)] = E[(y_{t-s} - \mu)(y_{t-j-s} - \mu)] = \gamma_s$ , covariância constante.

Dessa maneira, usam-se testes específicos para se testar a estacionariedade de uma série, tais como, Teste Dickey-Fuller (DF) (1979), Teste Dickey-Fuller Aumentado (ADF) (1984), Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS) (1992), Teste Philips-Perron (PP) (1998). Para fins deste trabalho, será utilizado o teste ADF. E como um complemento ao teste será utilizado o teste KPSS.

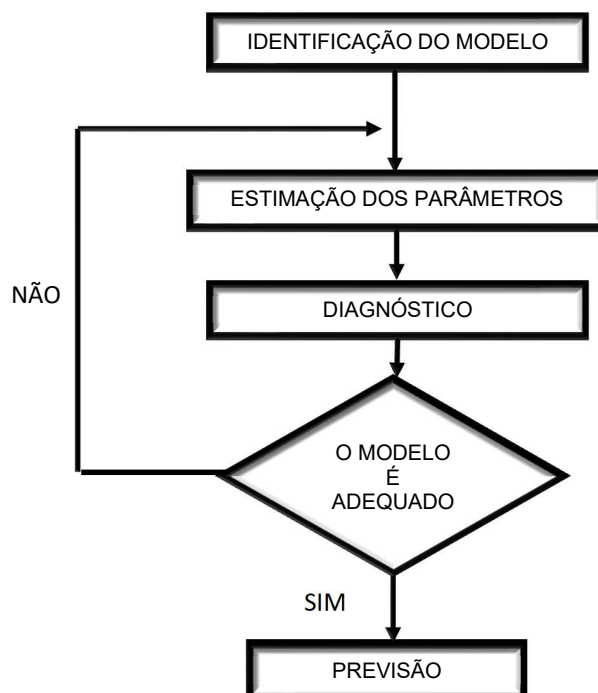
### Metodologia Box e Jenkins

Uma das técnicas quantitativas mais difundidas na década de 70 é a metodologia de Box e Jenkins, que parte da ideia de que cada valor da série (temporal) pode ser explicado por valores prévios, a partir do uso da estrutura de correlação temporal que, geralmente, há entre os valores da série (BOX et al., 1970). Estes modelos têm sido amplamente utilizados para a modelagem e previsão em aplicações médicas, ambientais, financeiras e de engenharia (PEREIRA et al., 1986) (FAVA, 2000) (WERNER et al., 2003) (MORETTIN et al., 2006).

A metodologia de Box e Jenkins tem sua modelagem baseada em um ciclo iterativo, em que a escolha da estrutura do modelo é a partir dos dados e considera as seguintes etapas: 1) Na especificação, avalia-se uma classe geral de modelos; 2) A fase de identificação é realizada com base na análise de autocorrelações, autocorrelações parciais e outros critérios; 3) Após, identificado o modelo, estima-se os parâmetros dos modelos; 4) E por fim, a verificação ou diagnóstico do modelo ajustado, por meio de uma análise de resíduos, para certificar-se de que o modelo é adequado para a finalidade da técnica, a previsão. Caso o modelo não seja adequado, repete-se o ciclo, voltando para a fase inicial de identificação.

A Modelagem de Box e Jenkins é uma das técnicas mais utilizadas para a previsão comportamental de variáveis, empregando modelos univariados na econometria contemporânea. A construção desses modelos baseia-se na teoria de que existe uma grande quantidade de informação presente na série de dados,

sendo esses dados capazes de fornecer estimativas sobre o comportamento futuro da variável, uma previsão. Assim, será apenas o comportamento da variável que responderá pela sua dinâmica futura. Este é conhecido, na literatura, como o modelo autorregressivo integrado e de médias móveis (ARIMA), ou definido simplesmente por modelo ARIMA (BOX et al., 1970).



**Figura 1:** Organograma adaptado da metodologia Box e Jenkins. **Fonte:** Souza (2011).

De acordo com o organograma apresentado, observa-se o detalhamento dos passos da metodologia Box e Jenkins. Essa metodologia só é aplicada após identificação de alguns requisitos fundamentais na série para que seja possível aplicar a modelagem. A primeira atividade a se realizar é a identificação do modelo, seguindo com a estimação dos parâmetros desse modelo. Após os parâmetros serem estimados, busca-se identificar se os resíduos do modelo possuem comportamento ruído branco (RB) para que o modelo seja adequado. Caso não seja adequado, volta-se para a etapa de estimação onde uma nova equação e novos parâmetros serão encontrados para a série. Essa sequência se repete até a obtenção de modelos que estejam dentro dos pressupostos exigidos pela metodologia, só assim, após escolher o melhor modelo, pode-se realizar a previsão.

### Modelos de séries temporais

Os modelos de séries temporais são dados pelas especificações de processos autorregressivos (AR), processos de médias móveis (MA), processos autorregressivos e médias móveis (ARMA) ou ainda modelos com processos mistos (ARIMA) (JUNIOR et al., 2013).

### Modelos autorregressivos AR(p)

O modelo AR(p), em que  $Z_t$  é descrito por seus valores passados e por seu RB  $a_t$ , sendo que  $a_t \sim RB(0, \sigma^2)$ . Em que sua versão mais simples é aquela em que  $Z_t$  depende somente de  $Z_{t-1}$  e de  $a_t$ .

Assim caracteriza-se o modelo autorregressivo de ordem 1, o qual é indicado compactamente por AR(1).

Denotam-se os modelos autorregressivos de ordem  $p$ :

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t, \quad (1)$$

em que  $p$  representa o número de defasagens de  $Z_t$ . Sendo que  $\phi$  é o coeficiente autorregressivo que acompanha o instante de  $Z_t$ ; e  $a_t$  é o erro aleatório do modelo. A forma mais simples de um processo autorregressivo é a que contém apenas uma defasagem, ou seja,  $p = 1$ . Uma vez especificado o número de defasagens, pode-se usar o método de estimação de mínimos quadrados para estimar os parâmetros do modelo AR (MATOS, 2000).

### Modelos de médias móveis MA(q)

Conforme Matos (2000), nos modelos de médias móveis, os valores de  $Z_t$  são resultados de uma soma ponderada, semelhante a uma média, dos valores mais recentes dos erros. O modelo de médias móveis, expressa os valores correntes da série  $Z_t$  como função linear dos erros passados não correlacionados até um número limitado de defasagens. O modelo MA com  $q$  defasagens, é definido como MA(q), é escrito da seguinte forma:

$$Z_t = \theta_0 + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}, \quad (2)$$

em que  $a_t$  são os erros ou resíduos aleatórios não correlacionados, não observáveis, que possuem média zero e variância constante. E  $\theta$  é o coeficiente de médias móveis.

Os modelos AR e MA são fáceis de identificar, pois suas funções de autocorrelações e autocorrelações parciais apresentam características definidas como estacionariedade e invertibilidade dos modelos Morettin et al., (2006).

### Modelos Autorregressivos e de Médias Móveis ARMA(p,q)

Alguns tipos de séries apresentam a necessidade de utilizar uma modelagem composta, em que se utilizam os dois modelos em conjunto, o AR e o MA, pois em separado eles acabam não se tornando adequados para analisar as mesmas. Dependendo do comportamento dessas séries, os modelos compostos, podem resultar em coeficientes com valores mais baixos, tendo também uma previsão mais próxima do real. Esses modelos são denominados modelos ARMA(p,q).

Conforme Matos (2000), o modelo ARMA é tratado como uma série  $Z_t$  e considera suas duas primeiras diferenças  $Z_t = Z_t - Z_{t-1}$ ,  $p$  defasagens para essa variável e  $q$ , para os erros aleatórios, tem-se o modelo ARMA(p,q). Esse modelo pode ser escrito da seguinte forma:

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}. \quad (3)$$

## Modelos Autorregressivos Integrado e de Médias Móveis ARIMA (p,d,q)

Os modelos autorregressivos integrados de médias móveis são modelos compostos pelos AR e MA, possuindo ainda um componente de integração I, que mostra o número de diferenças necessárias para que a série em estudo se torne estacionária. Para Morettin e Tolo (2006), se  $W_t = \Delta^d Z_t$  ( $W_t$ ) representa a série diferenciada entre dois instantes de tempo de  $Z_t$ , caso essa série seja estacionária, pode-se representar  $W_t$  por um modelo ARMA(p,q), ou seja,

$$\phi(B)W_t = \theta(B)a_t \quad (4)$$

Se  $W_t$  for uma diferença de  $Z_t$ , então  $Z_t$  é uma integral de  $W_t$ , daí diz-se que  $Z_t$  segue um modelo autorregressivo, integrado, de médias móveis, ou modelo ARIMA,

$$\phi(B)\Delta^d Z_t = \theta(B)a_t \quad (5)$$

de ordem (p,d,q) e escreve-se ARIMA(p,d,q), sendo p e q as ordens de  $\phi(B)$  e  $\theta(B)$ ,  $d = 1$ , respectivamente (MORETTIN; TOLOI, 2006).

Na equação 7 todas as raízes de  $\phi(B)$  devem estar fora do círculo unitário, sendo esta a condição para que o modelo encontrado seja não explosivo.

$$\phi(B)Z_t = \theta(B)a_t \quad (6)$$

em que  $\phi(B)$  é um operador autorregressivo, de ordem  $p + d$ , com  $d$  raízes iguais a um (sobre o círculo unitário) e as restantes  $p$  fora do círculo unitário, ou seja,

$$\phi(B) = \phi(B)\Delta^d = \phi(B)(1 - B)^d \quad (7)$$

Desta forma, o modelo da equação 8, supõe que a d-ésima diferença da série  $Z_t$  pode ser representada por um modelo ARIMA. Em grande parte dos casos usuais,  $d = 1$  ou  $d = 2$ , que correspondem a dois casos comuns de não-estacionariedade (MORETTIN et al., 2006): a) Séries não-estacionárias quanto ao nível: oscilam ao redor de um nível médio durante algum tempo e depois passam para outro nível temporário. Para que se tornem estacionárias é suficiente tomar apenas uma diferença; b) Séries não-estacionárias quanto à inclinação: oscilam em uma direção por um tempo e depois mudam para outra direção, também temporária. Para estacionarizá-las é necessária uma segunda diferença. Pressupostos da Análise de Séries Temporais

### Autocorrelação: Teste de Ljung-Box

A estatística de Ljung et al. (1978) tem o objetivo de identificar a existência de autocorrelação entre os resíduos da série analisada. A não existência dessa autocorrelação, mostra que os resíduos tem comportamento de RB e que possuem média zero e variância constante. Para séries temporais os resíduos devem ter comportamento de RB, do contrário, o modelo que não possuir essa característica é descartado.

Teste proposto por Ljung et al. (1978) em que sua estatística é dada pela equação:

$$Q(K) = n(n + 2) \sum_{k=1}^h \frac{\hat{r}_k^2}{(n-K)}, \quad (10)$$

em que  $\hat{r}_k^2$  significa a autocorrelação no *lag*  $K$ ,  $n$  representa o número de observações e  $h$  significa o número de *lag*'s que serão testados, os quais seguem uma distribuição de Qui-Quadrado com  $p$  graus de liberdade,  $\chi_{2,p}^2$ . As hipóteses do teste são:  $H_0$ : os resíduos são i.i.d.  $H_1$ : os resíduos não são i.i.d.

### Estacionariedade (Raiz Unitária) - Teste Dickey-Fuller Aumentado (ADF)

Os testes de raiz unitária consistem em avaliar a presença de raízes dentro do círculo unitário. Existe um procedimento específico para a realização desse teste em uma série  $Z_t$ , que se baseia em regredir a série  $Z_t$  contra seus valores defasados de um período  $Z_{t-1}$ . Na sequência da metodologia, testa-se a significância do parâmetro de  $Z_{t-1}$ .

Existem três tipos de relação entre  $Z_t$  e  $Z_{t-1}$ , que são: a) Equação sem intercepto e sem tendência ( $Z_t = \phi Z_{t-1} + a_t$ ); b) Equação com intercepto e sem tendência ( $Z_t = \alpha + \phi Z_{t-1} + a_t$ ); e c) Equação com intercepto e com tendência ( $Z_t = \beta + \delta T + a_t$ ). Nas quais  $a_t$  é RB, o que significa que possui média zero, variância constante e ausência de autocorrelação serial.

Considerando que  $\phi = 0$ , as equações acabam não dependendo de  $Z_{t-1}$ , tornando-as apenas dependente da tendência ( $T$ ), o que indica que  $Z_t$  segue um passeio aleatório (NOTINI, 2009). Uma série é não estacionária se possui uma raiz unitária, ou seja, quando a hipótese de que  $\phi = 1$  for aceita. As hipóteses pertencentes ao Teste ADF são:  $H_0$ : tem raiz unitária (não é estacionária)  $H_1$ : não tem raiz unitária (é estacionária).

### Teste KPSS

Este teste também serve para identificar presença de estacionariedade na série, corroborando com o Teste ADF. As hipóteses do teste KPSS são as seguintes:  $H_0$ : A série é estacionária.  $H_1$ : A série não é estacionária.

Com a aplicação de ambos os testes, KPSS e ADF, é possível uma conclusão mais precisa, uma vez que se pode distinguir entre séries com comportamento estacionário, séries com comportamento não estacionário, e séries em que os dados não são suficientemente informativos para saber sobre a estacionariedade.

### Heterocedasticidade Condicional Residual

A heterocedasticidade condicional acontece quando a variância dos resíduos, ou erros, não é constante. Para que se teste a homocedasticidade dos resíduos se fez necessária a utilização de um teste ARCH que se baseia na Estatística F.



## Estatística F

Morettin et al. (2006), explicam que o Teste F é utilizado para pequenas amostras e baseia-se na seguinte estatística:

$$F = \frac{(SQR_0 - SQR_1)/m}{SQR_1/(N-2m-1)} \sim F(m, N - 2m - 1) \quad (11)$$

Em que  $SQR_0 = \sum_{n=m+1}^N (r^2 - \bar{r})^2$  e  $SQR_1 = \sum_{n=m+1}^N \hat{u}_n^2$ , em que  $\bar{r}$  é a média amostral dos valores de  $r_n^2$  e  $\hat{u}_n$ , que são resíduos dos Mínimos Quadrados Ordinários (MQO) referente à regressão da equação 11. Se o valor de F for significativo, diz que há heterocedasticidade condicional na série.  $H_0$ : Todos os coeficientes de regressão são iguais à zero.  $H_1$ : Os coeficientes de regressão são diferentes de zero.

## Crítérios de Informação - Critério de Informação Akaike (AIC) e Critério de Informação Bayesiano (BIC)

Para Bueno (2008), o critério de informação é uma forma de encontrar o número ideal de parâmetros para um modelo e a qualidade do ajuste realizado. O critério de informação associa uma penalidade para balancear a redução de erros, devido ao aumento de regressores (parâmetros) que causam a diminuição à soma dos resíduos. Se a penalidade for menor que a redução da soma dos resíduos, o regressor deve ser adicionado ao modelo. Em caso de a penalidade ser maior que a diminuição da soma, esse regressor deve ser descartado, pois prejudicará o modelo (BUENO, 2008). A estatística de Akaike, e a estatística de Schwartz ou critério bayesiano são dadas da seguinte maneira:

$$AIC(k) = n \log \sigma^2 + 2k \quad (12)$$

$$BIC(k) = n \log \sigma^2 + k \log n \quad (13)$$

Dessa forma,  $k$  corresponde ao número de parâmetros estimados, o  $n$  é o tamanho da amostra e o  $\sigma^2$  é a variância (ALMEIDA, 2010). Assim, apresenta-se o método, descreve-se etapa por etapa do estudo e sua principal colaboração de pesquisa.

## Método

O conjunto de dados considerado corresponde ao período de maio de 1999 a outubro de 2022, sendo que as observações que compõem a amostra correspondem ao número total de pessoas que foram admitidas e que foram demitidas em cada um dos 281 meses de coleta, para cada série analisada. Estes dados foram obtidos a partir do site oficial<sup>1</sup> do Instituto de Pesquisa Econômica Aplicada (IPEA), na aba regional de tema emprego.

A metodologia proposta neste trabalho foi desenvolvida nas seguintes etapas: a) Gerou-se o gráfico com cada série original, de modo a verificar se elas possuíam tendência, sazonalidade, ciclo e oscilações abruptas, com isso é possível verificar se elas são estacionárias. Caso não sejam, aplicam-se diferenciações

<sup>1</sup> <http://www.ipeadata.gov.br/default.aspx>

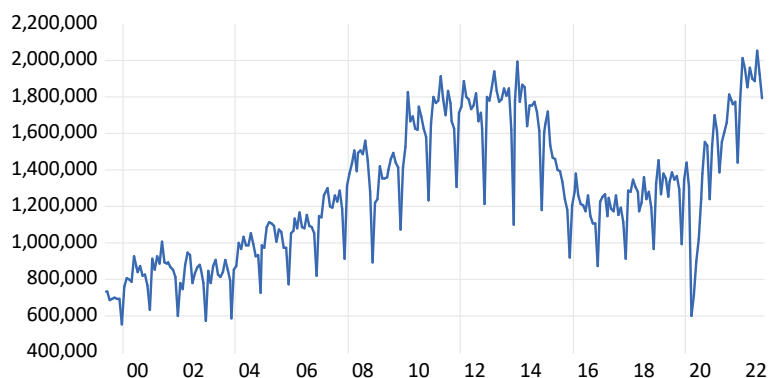
buscando estacionarizá-las; b) Em seguida foi gerado o correlograma com a FAC e FACP da série original a fim de identificar a categoria de um provável modelo e também se realizou testes de hipótese para confirmar a estacionariedade ou não-estacionariedade da série; c) Definiu-se o modelo a ser utilizado para cada série, se estimou os parâmetros, realizou-se a análise dos resíduos, para verificar se os mesmos são RB; após, também se analisou a homocedasticidade dos resíduos quadráticos. Em se tratando das séries deste estudo, não houve série com heterocedasticidade;

A escolha do modelo que melhor ajuste a série para a previsão, depende de uma prévia seleção de três modelos que as representem. Estes modelos são denominados modelos concorrentes, e a partir dos critérios de penalização, define-se o melhor, sendo os critérios escolhidos para esse estudo, o Critério de Informação Akaike (AIC) e o Critério de Informação Bayesiano (BIC). A partir do modelo escolhido realizaram-se as previsões. Para que fossem cumpridos todos os passos da modelagem, utilizaram-se os softwares Statistica 10 Enterprise, Microsoft Office Excel 2010 e EViews 12.

## RESULTADOS

Apresentam-se os resultados da análise do comportamento dos empregos, mais especificamente da quantidade de admissões e demissões, assim como as discussões dessa pesquisa por meio de autores e relatórios que corroboram os resultados das análises. Por meio da modelagem ARIMA, verificou-se a necessidade de utilização de modelos lineares ou não lineares nas séries. Assim, visando observar o comportamento dessas variáveis, em escala nacional dessas séries.

Série de admissões - Conforme a figura 2 observa-se o gráfico temporal da série de admissões. O gráfico representa o comportamento da série original em que se verifica a não estacionariedade, havendo um crescimento com pontos de queda no decorrer do período o que caracteriza a presença de uma possível sazonalidade.



**Figura 2:** Série de admissões no período de 1999 a 2022.

Importante salientar que no período pré-pandêmico, entre 2014 e 2018, existe uma queda significativa nas contratações de funcionários formais, ainda, no início da pandemia foi a pior situação de admissões desde os anos 2000 a 2004. Entretanto, hoje se pode considerar a melhor fase de admissões desde o início das medições, pois nunca houve tantas admissões como em 2022, corroborado pela CEPAL e OIT (2022). Assim, para identificar a estacionariedade da série, utilizaram-se as Funções de Autocorrelação (FAC)

e Autocorrelação Parcial (FACP) da série de admissões, as quais são apresentadas na Figura 3, que também fornecem o possível modelo a ser estimado.

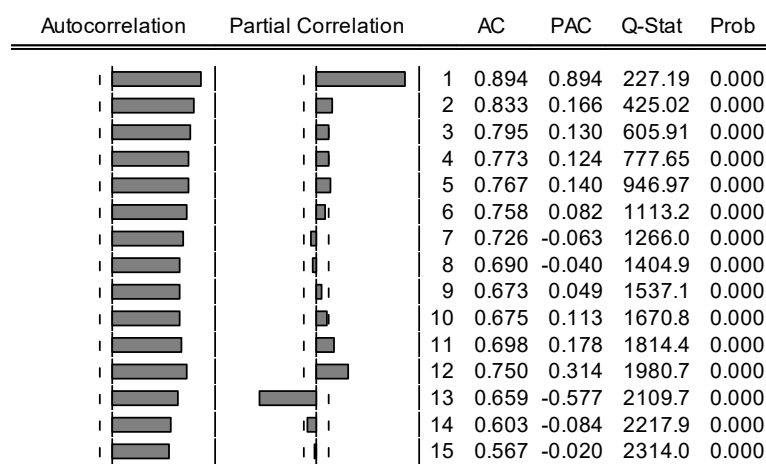


Figura 3: FAC, FACP e estatística Ljung-Box da série de admissões.

Observa-se que a FAC amostral apresenta um decaimento lento entre janeiro e fevereiro, apresentando uma longa dependência serial, o que pode ser uma indicação de um modelo fracionário. A FACP, neste caso, indica que a ordem do modelo seja AR(1), pois apenas o primeiro *lag* está fora dos limites, assim como, também identifica-se o *lag* 13 com decaimento contrário ao *lag* 1 indicando um modelo MA(1) com sazonalidade. Ainda pode-se observar que conforme o teste de Ljung-Box, seus respectivos p-valores são menores que o nível de 5%, o que evidencia a presença de autocorrelação para todos os 15 *lag*'s. Pode-se considerar, dessa forma, que a série de admissões é não estacionária.

A fim de verificar presença ou não de estacionariedade da série de admissões, realizaram-se os seguintes testes de hipótese, ADF e KPSS na série original e primeira diferença, sendo que ambos são apresentados na Tabela 1.

Tabela 1: Testes de Estacionariedade ADF e KPSS da série original e primeira diferença da série de admissões.

Série	ADF Estatística do Teste	p-valor ADF	Valor Crítico ADF (5%)	KPSS Estatística do Teste	Valor Crítico KPSS (5%)
Original	-1,183475	0,6822	-2,872121	1,050078	0,463000
1ª Diferença	-4,668068	0,0001		0,221759	

Analisando os resultados da Tabela 2, temos que pelo teste KPSS não se rejeita a hipótese nula de estacionariedade, na série diferenciada. Pelo teste ADF, o p-valor esteve um pouco acima de 0,05, ou seja, a série diferenciada foi considerada estacionária, pois nessa série a modelagem obteve um melhor comportamento e considerando que ambos os testes se complementam. Análise do melhor modelo e quadro de modelos concorrentes – Os modelos determinados para a série foram escolhidos seguindo os preceitos obrigatórios da metodologia proposta. A seguir serão apresentados, exemplos aplicados da metodologia Box e Jenkins, e a escolha do modelo que melhor explica esse conjunto de dados. Levando em conta a escolha de um melhor modelo para representar cada série a ser analisada, foram estimados diversos modelos concorrentes, conforme Quadro 1, e dentre esses, escolheu-se o que obteve os menores valores de AIC e

BIC.

**Quadro 1:** Quadro de modelos concorrentes da série de admissões.

Modelo	Coefficientes	AIC	BIC
SARIMA(1,1,0)(1,1,0) <sub>1</sub>	$\phi_1: -0,028194$ $\theta: +0,937902$	26,78474	26,83653
SARIMA(1,1,1)(1,1,0) <sub>1</sub>	$\phi: +0,404874$ $\theta: +0,993060$ $\Theta: -0,799151$	26,70212	26,76686
<b>SARIMA(1,1,1)(1,1,0)<sub>12</sub></b>	<b><math>\phi: +0,923026</math></b> <b><math>\theta: +0,757852</math></b> <b><math>\Theta: -0,245240</math></b>	<b>25,94776</b>	<b>26,01250</b>

O modelo que melhor ajustou a série de admissões, foi o SARIMA(1,1,1)(1,1,0)<sub>12</sub>, pois pelos critérios de informação, esse possui um menor valor do AIC e de BIC, corroborado por Reimbold et.al. (2017). Após a estimação do modelo, é realizada a análise de resíduos, a qual deve possuir comportamento de RB, dessa forma, indicando comportamento estacionário. A Figura 8 apresenta a FAC e FACP dos resíduos, bem como a estatística de teste Ljung-Box dos resíduos e dos resíduos quadráticos.

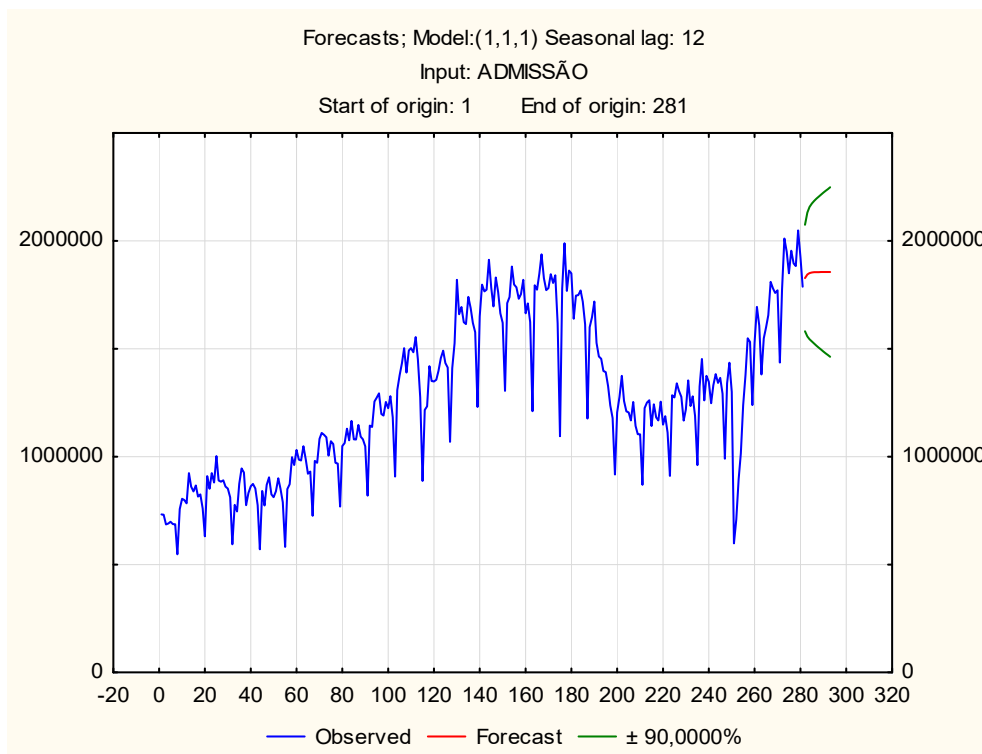
Na Figura 4, pode-se verificar que se tem um processo estacionário, devido ao fato que a FAC e a FACP dos resíduos apresentam as autocorrelações de seus *lag's* dentro dos limites de confiabilidade, considerando um nível de significância de 5%. Para confirmar que os resíduos não são autocorrelacionados fez-se o teste Ljung-Box(Q). Conforme os p-valores observa-se que seus valores foram superiores aos níveis de significância adotados, indicando assim, que não se rejeita a hipótese nula, ou seja, existe estacionariedade nos resíduos. Assim, realizou-se as previsões. Na Figura 5, apresenta-se as previsões para 16 meses a frente, para a série de admissões, utilizando-se o modelo SARIMA(1,1,1)(1,1,0)<sub>12</sub>.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.004	0.004	0.0057	
		2	-0.067	-0.067	1.2953	
		3	0.073	0.074	2.8131	
		4	-0.018	-0.024	2.9036	0.088
		5	-0.036	-0.026	3.2834	0.194
		6	0.028	0.020	3.5058	0.320
		7	-0.032	-0.034	3.7964	0.434
		8	-0.004	0.004	3.8021	0.578
		9	0.012	0.003	3.8425	0.698
		10	0.034	0.039	4.1836	0.758
		11	-0.006	-0.006	4.1943	0.839
		12	-0.114	-0.115	8.0521	0.529
		13	0.028	0.028	8.2864	0.601
		14	-0.062	-0.079	9.4462	0.581
		15	-0.043	-0.020	10.003	0.616

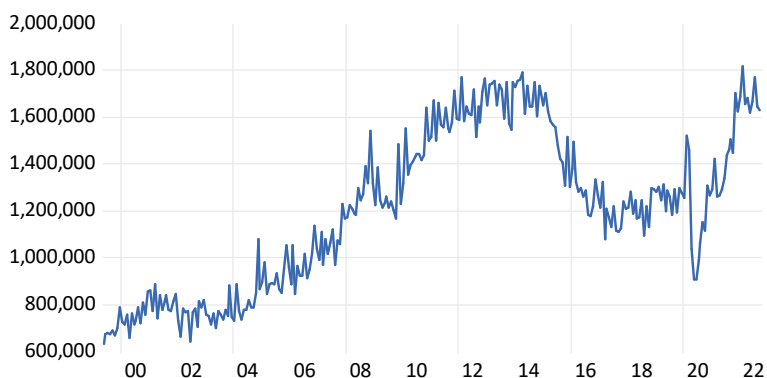
**Figura 4:** FAC, FACP dos resíduos do modelo SARIMA(1,1,1)(1,1,0)<sub>12</sub> e estatística Ljung-Box da série de admissões.

Conforme a Figura 5 pode-se verificar que os resultados da previsão se encontram dentro dos intervalos de confiança. Observa-se também que no eixo horizontal têm-se os meses correspondentes às observações, sendo que, no mês 293 ocorreu um pico nas previsões, correspondendo este ao mês de setembro, no qual este apresenta uma característica diferente dos demais meses, conforme Figura 2. Nesse modelo de admissões encontrado, apresentando sazonalidade coincide com um estudo realizado por Gonçalves Júnior, Fehr e Rodrigues (2021) em que estudaram a previsão do Imposto sobre Circulação de

Mercadorias e Serviços (ICMS) em estados da região sudeste do Brasil e deles, no Espírito Santo, pôde ser modelado por um sarima. Série de demissões - Teste de estacionariedade - Com o intuito também de verificar o comportamento da série de demissões, faz-se aqui uma análise descritiva. Dessa forma, pode-se verificar se esta série possui oscilações ou se comportou de maneira estável no período de maio de 1999 até outubro de 2022. A Figura 6 apresenta-se a série de demissões.



**Figura 5:** Previsão do modelo SARIMA(2,1,0)(1,0,0)<sub>12</sub>.



**Figura 6:** Série de demissões no período de 1999 a 2022.

Conforme Mueller (2013), em 2004, houve demissões em massa nas indústrias de diversos setores no país, tais como automobilística, moveleira etc. No ano de 2008, assim como nas admissões, as demissões aconteceram, grande parte, devido à crise internacional. Já em 2012, contava-se que o país não seria mais afetado pela crise, mas essa variabilidade ocorreu pela estagnação da economia brasileira, vinda do ano de 2011, e um fraco crescimento econômico no ano que se seguiu. Considerando a estabilidade e redução de demissões entre 2016 e 2019, nota-se um pico em fevereiro de 2020, com posterior queda, quando da notificação mundial da OMS, em virtude da epidemia de Covid-19. Da mesma forma que as admissões, identificam-se um pico de demissões em fevereiro de 2022.

Pela Figura 7, observa-se a série diferenciada, na qual se verifica um comportamento de estacionariedade. Sendo assim, ela será utilizada para estimar os parâmetros do modelo em estudo. Apresenta-se a FAC e FACP da série original de demissões.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.946	0.946	254.20	0.000
		2	0.935	0.380	503.32	0.000
		3	0.924	0.175	747.32	0.000
		4	0.893	-0.134	976.51	0.000
		5	0.890	0.140	1204.8	0.000
		6	0.864	-0.103	1420.4	0.000
		7	0.856	0.116	1632.9	0.000
		8	0.839	-0.064	1837.9	0.000
		9	0.832	0.165	2040.4	0.000
		10	0.810	-0.203	2233.1	0.000
		11	0.796	0.078	2419.6	0.000
		12	0.803	0.165	2610.2	0.000
		13	0.758	-0.328	2780.8	0.000
		14	0.752	0.027	2949.1	0.000
		15	0.735	0.013	3110.5	0.000

Figura 7: FAC, FACP e estatística Ljung-Box da Série de demissões.

Conforme Figura 17, observa-se que a FAC amostral apresenta decaimento lento com longa dependência serial. A FACP indica que a ordem do modelo seja um AR(2), pois possui dois lag's fora dos limites de confiabilidade. Mas, no período de doze meses, é possível visualizar que há uma incidência de um filtro MA no mês de janeiro. Essa variação indica a necessidade de um estudo de possível sazonalidade na série.

Assim como na série de admissões, a série de demissões apresentou decaimento lento, o que poderia indicar um modelo fracionário, mas não foi encontrado nenhum modelo que se comporte melhor com fracionariedade. Na Tabela 2 constam os valores dos testes de estacionariedade ADF e KPSS, juntamente com o p-valor respectivo.

Tabela 2: Testes de Estacionariedade ADF e KPSS da série original e primeira diferença da série de demissões.

Série Demissões	ADF Estatística do Teste	p-valor ADF	Valor Crítico ADF (5%)	KPSS Estatística do Teste	Valor Crítico KPSS (5%)
Original	-1,286081	0,6366	-2,872121	1,208295	0,46300
1ª Diferença	-3,996276	0,0017		0,101945	

Os testes da Tabela 2 apresentaram resultados de estacionariedade para a série de demissões em primeira diferença, Série essa que será realizada a modelagem. Observa-se que pelo teste KPSS não se rejeita a hipótese nula de estacionariedade, na série diferenciada de demissões. Assim como, pelo teste ADF, em que o p-valor foi significativo ao nível de 5%, ou seja, a série diferenciada foi considerada estacionária.

Quadro 2: Quadro de modelos concorrentes da série de demissões.

Modelo	Coefficientes	AIC	BIC
SARIMA(1,1,0)(2,0,0) <sub>1,12</sub>	$\phi_1: +0,974074$ $\Phi_1: -0,330843$ $\Phi_2: +0,519534$	25,31818	25,38292
SARIMA(1,1,0)(1,1,0) <sub>12</sub>	$\phi_1: -0,665672$ $\Phi: +0,622324$	25,46775	25,51954
SARIMA(1,1,1)(1,1,0) <sub>12</sub>	$\phi: +0,980306$ $\Phi: +0,541679$ $\theta_1: -0,440312$	25,32692	25,39166

Considerando o melhor modelo selecionado para representar a série de demissões, estimou-se 3 modelos concorrentes, dentre os quais os que tiveram melhor resultado estão dispostos no Quadro 2, que segue.

O modelo que melhor ajustou a série de demissões foi um SARIMA(1,1,0)(2,0,0)<sub>1,12</sub>, pois apresentou melhor modelagem na série de demissões e menores valores dos Critérios de Informação Akaike e Bayesiano, também utilizado por Raimbold et. al. (2017) em sua pesquisa. Após a escolha do modelo, analisou-se se os resíduos desse modelo têm comportamento de RB, em que se busca validar o modelo escolhido mediante a estatística de Ljung-Box dos resíduos e dos resíduos quadráticos. Na Figura 8, constam a FAC e FACP da série dos resíduos do modelo SARIMA(1,1,0)(2,0,0)<sub>1,12</sub>.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.096	-0.096	2.6148	
		2 -0.142	-0.153	8.3868	
		3 0.100	0.072	11.234	
		4 -0.121	-0.129	15.404	0.000
		5 0.048	0.052	16.067	0.000
		6 0.031	-0.005	16.337	0.001
		7 -0.084	-0.049	18.384	0.001
		8 0.072	0.045	19.909	0.001
		9 0.143	0.151	25.918	0.000
		10 -0.064	-0.012	27.126	0.000
		11 -0.030	-0.018	27.393	0.001
		12 0.003	-0.015	27.396	0.001
		13 -0.060	-0.038	28.472	0.002
		14 0.019	-0.016	28.583	0.003
		15 0.056	0.049	29.508	0.003

Figura 8: FAC, FACP dos resíduos do modelo SARIMA(1,1,0)(2,0,0)<sub>1,12</sub> e estatística Ljung-Box(Q).

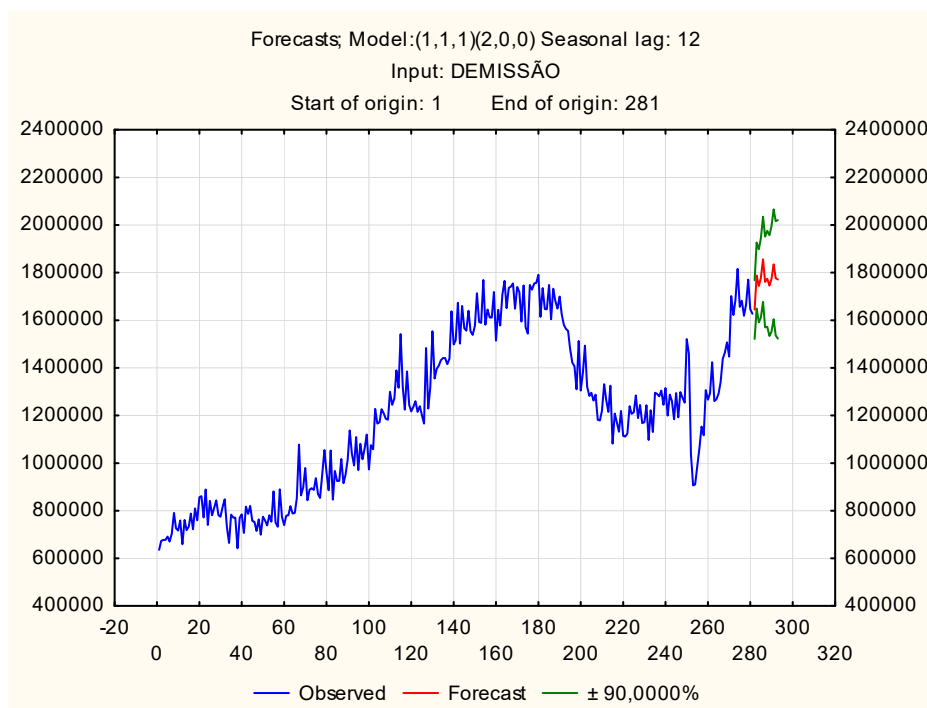


Figura 10: Previsão do modelo SARIMA(1,1,0)(2,1,0)<sub>1,12</sub>.

Dessa forma, o comportamento é estacionário, podendo ser verificado pela FAC e FACP, em que seus lag's estão dentro dos limites de confiabilidade e os p-valores foram superiores aos níveis de significância

adotados. Na Figura 9, apresenta-se as previsões para 12 meses a frente para a série de demissões, utilizando-se o modelo SARIMA(1,1,0)(2,0,0)<sub>1,12</sub>.

Na Figura 10 pode-se verificar que os resultados da previsão para a série de demissões encontram-se dentro dos limites de confiança. Verifica-se também, que a série de demissões prevista apresenta certa estabilidade com alguns picos.

## CONCLUSÕES

Com o modelo proposto foi possível analisar e prever o comportamento de cada série, no período de maio de 1999 a outubro de 2022, sendo que a série de admissões apresentou algumas tendências, sugerindo assim a utilização de um modelo sazonal devido seu caimento lento na FAC da série original. Pela análise também foi possível identificar, na série de demissões, que ela também apresentou um comportamento sazonal, a qual teve um decaimento lento com longa dependência serial.

A utilização dos modelos propostos SARIMA(1,1,1)(1,1,0)<sub>12</sub> e SARIMA(1,1,0)(2,0,0)<sub>1,12</sub>, justifica-se pelo fato de que as séries em estudo apresentaram efeito sazonal que se repete de doze em doze meses, com longa dependência serial, com isso tornou-se necessário a aplicação de modelos que levem em consideração esta característica, captando assim o verdadeiro processo gerador da série. As admissões tiveram como melhor modelo um SARIMA(1,1,0)(1,1,0)<sub>12</sub>, o que significa que admissões possuem sazonalidade intensa no mês de dezembro, propiciada pela época de Natal e virada de ano. De acordo com uma análise específica das demissões, escolheu-se o melhor modelo um SARIMA(1,1,0)(2,0,0)<sub>1,12</sub>, com sazonalidade pois em todos os janeiros e dezembros, levantando a hipótese de pessoas trocarem de empregos por possuírem melhores propostas nessas mesmas épocas ou pelos empregos temporários que surgem em dezembro e grandes saídas em janeiro.

Quando observado se a série possuía volatilidade, os resíduos não apresentaram esta característica, portanto não há flutuação muito grande em termos de variabilidade de um período para outro. O que ocorre é um incremento de valor que a sazonalidade consegue explicar. Há de se observar que os modelos SARIMA refletem o comportamento do nível da série e os modelos de volatilidade o comportamento de variabilidade. Sugere-se como estudo futuro, continuar periodicamente as análises e testar outras variáveis bem como outros modelos, de forma a trazer opções mais robustas de modelagem e metodologia.

## REFERÊNCIAS

ALMEIDA, A. F. M.. **Análise comparativa da aplicação de modelos para imputação do volume médio diário de Séries Históricas de volume de tráfego**. Dissertação (Mestrado em Engenharia de Transportes) – Universidade Federal do Ceará, Fortaleza, 2010.

BAYER, F. M.. **Previsão do Preço e da Volatilidade de Commodities Agrícolas, por meio de Modelos Arfima-Garch**. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Maria, Santa Maria, 2008.

BOX, G.; JENKINS, G. M.. **Time series analysis: Forecasting and control**. San Francisco: Holden-Day, 1970.

BOX, G.; JENKINS, G. M.; REINSEL, G.. **Time series analysis: Forecasting and control**. 3 ed. New York. Prentice Hall, 1994.

BUENO, R. L. S.. **Econometria de séries temporais**. FGV – Fundação Getúlio Vargas. São Paulo: Cengage Learning, 2008.

CEPAL/PNUD/OIT. **Emprego, desenvolvimento humano e trabalho decente: a experiência brasileira recente**. Brasília: CEPAL/ PNUD/OIT, 2013.

CEPAL/OIT. **Coyuntura Laboral en América Latina y el**



**Caribe.** Dinâmica da produtividade do trabalho na América Latina. Brasília: CEPAL/OIT, 2022.

FAVA, V. L.. **Manual de econometria.** São Paulo: Atlas, 2000.

GONÇALVES JÚNIOR, G. A.; FEHR, L. C. F. A.; RODRIGUES, R. S.. Previsão do ICMS nos estados da região Sudeste: utilização do modelo sazonal simples e Holt-Winter. **Revista Brasileira de Administração Científica**, v.12, n.4, p.432-440, 2021. DOI: <http://doi.org/10.6008/CBPC2179-684X.2021.004.0030>

MANKIW, N. G.. **Macroeconomia.** Rio de Janeiro: LTC, 2011. MATOS, O. C.. **Econometria Básica: Teoria e Aplicações.** 3 ed. São Paulo: Atlas, 2000.

MTP. MINISTÉRIO DO TRABALHO E PREVIDÊNCIA. **Grupo de Altos Estudos do Trabalho.** Relatórios dos Grupos de Estudos Temáticos. Brasília. 2021

MORETTIN, P. A.; TOLOI, C. M. C.. **Análise de séries temporais.** Associação Brasileira de Estatística ABE – Projeto Fisher, 2006.

NASCIMENTO, A. M.. **Curso de Direito do Trabalho.** 24 ed. São Paulo: Saraiva, 2009.

OIT. ORGANIZAÇÃO INTERNACIONAL DO TRABALHO. **Crise Financeira Mundial – Impactos Sociais e no Mercado de Trabalho.** ANFIP - Associação Nacional dos Auditores Fiscais

da Receita Federal do Brasil. Brasília, 2009.

OIT. ORGANIZAÇÃO INTERNACIONAL DO TRABALHO. **Crise Financeira Mundial – Impactos Sociais e no Mercado de Trabalho.** ANFIP - Associação Nacional dos Auditores Fiscais da Receita Federal do Brasil. Brasília: 2014.

PADILHA JUNIOR, J. B.. Modelagem Box-Jenkins (Arima) para gestão de risco de preço na Pecuária de Corte do Estado do Paraná. In: CONGRESSO VIRTUAL BRASILEIRO DE ADMINISTRAÇÃO, 10. **Anais.** 2013.

PEREIRA, B. B.; PAIS, M. B. Z.; SALES, P. R. H.. **Análise Espectral de séries temporais: Uma introdução para Engenharia, Economia e Estatística.** Rio de Janeiro: Arte Final, 1986.

VIANNA, C. S. V.. **Manual Prático das Relações Trabalhistas.** 12 ed. São Paulo: LTr, 2014.

WERNER, L.; RIBEIRO, J. L. D.. Previsão de demanda: uma aplicação dos Modelos Box-Jenkins na área de assistência técnica de computadores pessoais. **Revista Gestão e Produção**, v.10, n.1, p.47-67, 2003. DOI: <https://doi.org/10.1590/S0104-530X2003000100005>

WOLEK, A.. **O trabalho, a ocupação e o emprego: uma perspectiva histórica.** Instituto Catarinense de Pós-Graduação, 2008.

Os autores detêm os direitos autorais de sua obra publicada. A CBPC – Companhia Brasileira de Produção Científica (CNPJ: 11.221.422/0001-03) detêm os direitos materiais dos trabalhos publicados (obras, artigos etc.). Os direitos referem-se à publicação do trabalho em qualquer parte do mundo, incluindo os direitos às renovações, expansões e disseminações da contribuição, bem como outros direitos subsidiários. Todos os trabalhos publicados eletronicamente poderão posteriormente ser publicados em coletâneas impressas ou digitais sob coordenação da Companhia Brasileira de Produção Científica e seus parceiros autorizados. Os (as) autores (as) preservam os direitos autorais, mas não têm permissão para a publicação da contribuição em outro meio, impresso ou digital, em português ou em tradução.