

## **Análise do Valor- $p$ Determinado pela Estatística $\tau$ na Aplicação do Teste de Dickey-Fuller Aumentado**

A. G. SILVEIRA<sup>1\*</sup>, V. L. D. MATTOS<sup>2</sup>, L. R. NAKAMURA<sup>3</sup>, M. C. AMARAL<sup>4</sup>,  
A. C. KONRATH<sup>5</sup> e A. C. BORNIA<sup>6</sup>

Recebido em 28 de junho de 2021 / Aceito em 18 de outubro de 2021

**RESUMO.** O presente artigo avaliou a interferência da quantidade de defasagens utilizadas no resultado do valor- $p$  associado à estatística utilizada no teste de Dickey-Fuller aumentado (ADF), bem como identificou algumas propriedades das séries estudadas que interferem em seu resultado. Foram realizados três experimentos com séries de diferentes amplitudes, considerando estrutura do modelo em relação à presença ou não de constante e/ou tendência, e quantidade de defasagens como fatores. Modelos autorregressivos do tipo AR(1) foram considerados para a geração de dados pelo método de Monte Carlo, que poderiam apresentar diferentes intensidades para o coeficiente associado à primeira defasagem. Depois da aplicação do teste ADF, foram determinadas as proporções de rejeição da hipótese nula em cada condição experimental, sendo utilizada uma análise de variância com a estatística qui-quadrado para verificar a influência da quantidade de defasagens no valor- $p$ . Os resultados mostram que se houver raiz unitária, o teste apresenta bons resultados, independentemente da quantidade de defasagens considerada. Entretanto, o mesmo não foi observado nos casos em que a série temporal não apresenta raiz unitária.

**Palavras-chave:** estacionariedade, raiz unitária, séries temporais, teste ADF.

### **1 INTRODUÇÃO**

Nas últimas décadas, vários métodos têm sido utilizados para modelar séries temporais, que se caracterizam como uma sequência de observações ordenadas no tempo. Estes métodos podem

---

\*Autor correspondente: Anderson Garcia Silveira – E-mail: a.garcia.ee@gmail.com

<sup>1</sup>Programa de Pós-Graduação em Modelagem Computacional, Universidade Federal do Rio Grande, Av. Itália, s/n, km 8, 69203-900, Rio Grande, RS, Brasil – E-mail: a.garcia.ee@gmail.com <https://orcid.org/0000-0002-3954-6145>

<sup>2</sup>Instituto de Matemática, Estatística e Física, IMEF, Universidade Federal do Rio Grande, Av. Itália, s/n, km 8, 69203-900, Rio Grande, RS, Brasil – E-mail: viviane.leite.mattos@gmail.com <https://orcid.org/0000-0002-3512-6290>

<sup>3</sup>Departamento de Informática e Estatística, INE, Universidade Federal de Santa Catarina, R. Delfino Conti, s/n, 88040-900, Florianópolis, SC, Brasil – E-mail: luiz.nakamura@ufsc.br <https://orcid.org/0000-0002-7312-2717>

<sup>4</sup>Programa de Pós-Graduação em Modelagem Computacional, Universidade Federal do Rio Grande, Av. Itália, s/n, km 8, 69203-900, Rio Grande, RS, Brasil – E-mail: marianecamaral@gmail.com <https://orcid.org/0000-0002-6286-3077>

<sup>5</sup>Departamento de Informática e Estatística, INE, Universidade Federal de Santa Catarina, R. Delfino Conti, s/n, 88040-900, Florianópolis, SC, Brasil – E-mail: andreack@gmail.com <https://orcid.org/0000-0002-3742-5032>

<sup>6</sup>Departamento de Engenharia de Produção e Sistemas, EPS, Universidade Federal de Santa Catarina, R. Delfino Conti, s/n, 88040-900, Florianópolis, SC, Brasil – E-mail: cesar.bornia@ufsc.br <https://orcid.org/0000-0003-3468-7536>

estar fundamentados em conceitos estatísticos, matemáticos ou computacionais, sendo alguns deles resultado da integração entre conceitos destas áreas. Na modelagem da demanda, resultados satisfatórios têm sido encontrados por meio de métodos estatísticos, entre os quais aqueles propostos por Box-Jenkins [4]: os modelos autorregressivos de médias móveis (ARMA) e suas variações. Estes modelos, assim como a maioria dos modelos estatísticos, supõe que as séries temporais são estacionárias [17].

Na presença de estacionariedade, a função amostral que descreve o processo gerador dos dados tem a mesma forma em todos os instantes, facilitando a identificação de estimativas dos parâmetros desconhecidos dos modelos especificados. O trabalho de Nelson e Plosser [18], considerado por alguns estudiosos como ponto de partida para estudos nesta área, mostrou que as flutuações econômicas são melhor explicadas por mudanças na componente tendência do que nos componentes cíclicos, evidenciando a necessidade da avaliação da estacionariedade como uma importante propriedade do processo gerador dos dados de uma série temporal.

Alguns procedimentos gráficos relativamente simples, como gráficos em linhas e gráficos da função de autocorrelação e da função de autocorrelação parcial, podem ser utilizados por pessoas experientes para o reconhecimento da existência desta propriedade. Entretanto, técnicas analíticas mais formais apresentam melhor desempenho, além de não dependerem da experiência do analista [2].

Nas últimas décadas vários procedimentos formais para identificar a presença de raiz unitária em uma série temporal foram apresentados, entre os quais o teste de Dickey-Fuller aumentado (ADF) e suas extensões, que consideram a sua presença (ausência de estacionariedade) como hipótese nula [8]. Ao não rejeitar esta hipótese nula, não é encontrada evidência de estacionariedade, razão pela qual é mais plausível considerar a sua ausência. Neste caso, uma prática frequentemente adotada consiste na diferenciação sucessiva da série até atingi-la, sendo a última diferenciação determinante de sua ordem de integração.

Entretanto, possivelmente existem alguns problemas sérios na interpretação destes testes como salientado em [5]. Embora o teste ADF seja, talvez, o mais popular entre os testes de raiz unitária, apresenta baixo poder, ou seja, tende a não rejeitar a hipótese nula com mais frequência que o desejado [3]. Ainda com base em [3], o resultado destes testes está associado à quantidade de defasagens considerada, enquanto que [2] evidenciam a relação de seus resultados com a amplitude da série estudada e a intensidade do parâmetro autorregressivo, principalmente, se este último estiver próximo da unidade [5]. Finalmente, o desempenho destes testes também estão associados à presença de tendência determinística, porém alguns dos problemas apresentados podem ser contornados se a série analisada sofre uma transformação via mínimos quadrados generalizados (GLS) antes da aplicação do teste ADF, identificado como ADF-GLS, na avaliação da estacionariedade da parte não sazonal da série [2].

O presente estudo foi realizado para avaliar a interferência da quantidade de defasagens utilizadas no cálculo do valor- $p$ , determinado a partir da estatística  $\tau$  no teste ADF, identificando algumas propriedades da série estudada que interferem em seu resultado.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 Estacionariedade e ergodicidade

Um processo estocástico que gerou uma série temporal  $\{y_t\}$  é dito estacionário quando for invariante no tempo, sendo denominado de processo fracamente estacionário se satisfizer as seguintes condições: (i)  $E(y_t) = E(y_{t+k}) = \mu, \forall t, k$ ; (ii)  $E[(y_t - \mu)^2] = \sigma_t^2 = \sigma^2$ , sendo  $\sigma < \infty, \forall t$ ; e (iii)  $\gamma_k = \text{Cov}[y_t, y_{t+k}] = \text{Cov}[y_{t+m}, y_{t+m+k}], \forall m, t, k$ . Ademais, uma série temporal é dita ergódica se  $[y_{t+1}, y_{t+2}, \dots, y_{t+k}]$  e  $[y_{t+m+1}, y_{t+m+2}, \dots, y_{t+m+k}]$  forem independentes, com  $m$  suficientemente grande.

Como consequência, se uma série temporal  $\{y_t\}$ , em que  $E(y_t) = \mu$ , for estacionária e ergódica, a média amostral de qualquer amostra de tamanho  $T, \bar{y}$ , converge assintoticamente para média populacional  $\mu$ .

### 2.2 Processo estocástico com raiz unitária

Considere um processo estocástico que possui um modelo autorregressivo de ordem um – AR(1) –, isto é,

$$y_t = \phi y_{t-1} + u_t, \quad (2.1)$$

como mecanismo gerador dos dados, em que  $\phi$  é um parâmetro a ser estimado e  $u_t$  é um processo de ruído branco, que se caracteriza como uma sequência de variáveis aleatórias independentes e normalmente distribuídas (iid), com média zero e variância constante, ou seja,  $u_t \sim RB(0, \sigma^2)$ .

O comportamento deste processo estocástico apresenta diferentes propriedades, para diferentes valores de  $\phi$ . Se  $\phi = \pm 1$ , a equação (2.1) se reduz a  $y_t = y_{t-1} + u_t$ , ou ainda  $y_t = y_0 + \sum_{i=1}^T u_i$ , no qual o processo estocástico é dito não estacionário (passeio aleatório). Se  $|\phi| < 1$ , então é possível considerar  $y_t = \sum_{i=1}^{\infty} \phi^i u_{t-i}$ , no qual o processo estocástico é dito estacionário. Esta representação pode ser usada para evidenciar as diferenças entre os dois processos estocásticos [7, 10] conforme Quadro 1 (observe que as mesmas propriedades são válidas para um processo autorregressivo de médias móveis – ARMA – estacionário e invertível).

Finalmente, observe que o modelo da equação (2.1) não apresenta elementos referentes a constante e tendência, o que raramente acontece em séries estacionárias [7].

Quadro 1: Propriedades de processos estocásticos, conforme valor de  $\phi$ .

Propriedade	$ \phi  < 1$	$ \phi  = 1$ ou $ \phi  > 1$
Variabilidade	$V(y_t) = \frac{\sigma_u^2}{1-\phi^2}$ $V(y_t) < V(\sigma_u^2)$	$V(y_t) = \infty$
Inovação	Perde gradualmente seu efeito sobre os dados à medida que o processo vai acontecendo	Tem efeito permanente sobre os dados gerados pelo processo estocástico
Densidade espectral	Finita	Infinita
Relação com a média do processo	O tempo esperado entre cruzamentos de $\{y_t\}$ é finito, logo o processo se movimenta em torno da média e tende para ela	O tempo esperado entre cruzamentos de $\{y_t\}$ é infinito, logo o processo não apresenta a tendência de retornar para sua média teórica
Autocorrelação ( $\gamma$ )	Diminui à medida que aumenta a quantidade de defasagens	$\lim_{t \rightarrow \infty} \gamma_k = 1, \forall k$
Variância da previsão	$V(y_{t+h}) = \sigma_u^2(1 + \phi^2 + \dots + \phi^{2(h-1)})$	$V(y_{t+h}) = h\sigma_u^2$

Fonte: Engle e Granger [10] e Choi [7].

### 2.3 Teste DF e teste ADF

O teste de Dickey-Fuller (DF) [8], além de poder ser utilizado no modelo da equação (2.1), ainda pode identificar a presença de raiz unitária em dois outros modelos que utilizam o AR(1) como mecanismo de geração dos dados, a saber

$$y_t = \mu + \phi y_{t-1} + u_t, \quad (2.2)$$

e

$$y_t = \mu + \beta t + \phi y_{t-1} + u_t, \quad (2.3)$$

em que  $\mu$  é uma constante,  $\beta$  é um efeito de tendência.

Nos três modelos, descritos nas equações (2.1), (2.2) e (2.3), testa-se a hipótese nula  $H_0 : \phi = 1$  (há presença de raiz unitária), contra a alternativa  $H_1 : |\phi| < 1$  (não há presença de raiz unitária), podendo o processo ser considerado estacionário. Chama-se a atenção para o fato de que ao não rejeitar a hipótese nula, não se encontram evidências de que não há presença de raiz unitária, sendo mais plausível considerar a sua presença, o que conduz a um processo estocástico não estacionário.

Na equação (2.2) há a presença de uma constante, então inicialmente testam-se as hipóteses de que o processo estacionário possui média ( $\mu$ ) igual ou diferente zero, para, posteriormente,

avaliar a estacionariedade da série como um todo. Na equação (2.3) além da constante, há ainda o parâmetro relacionado à tendência  $\beta$ , então testa-se, inicialmente, se este parâmetro é ou não nulo. Em seguida, se a hipótese nula ( $\beta = 0$ ) for rejeitada, avalia-se a estacionariedade que poderá identificar um processo estocástico com tendência estacionária. Neste último caso, a presença de constante também deve ser avaliada.

Na modelagem do processo representado em (2.1), assumindo que  $y_0 = 0$  e que  $u_t \sim N(0, \sigma^2)$ , a significância da estimativa de  $\phi$  obtida pelo método dos mínimos quadrados ordinários poderia ser testada comparando-a à estatística  $t = (\hat{\phi} - \mu(\hat{\phi}))/EP(\hat{\phi})$ , proveniente de uma distribuição  $t$  de Student [17]. Entretanto, na análise de raízes unitárias, a distribuição normal não se aplica e, logo, a distribuição  $t$  de Student não pode ser utilizada para a realização do teste. Neste caso, pode-se utilizar a estatística  $\tau$  na avaliação [7, 9, 16], cuja distribuição assintótica é representada por

$$Q(\hat{\phi} - 1) = \frac{\int_0^1 W(r)dW(r)}{\int_0^1 W^2(r)dr} = \hat{\tau},$$

independentemente da distribuição de  $\{u_t\}$ , em que  $\{W(r), 0 \leq r \leq 1\}$  é definido como um processo contínuo (um movimento Browniano padrão), tal que: i)  $W(0) = 0$ ; ii) as variáveis aleatórias  $W(r_k) - W(r_{k-1})$  são independentes e  $\{W(s) - W(r)\} \sim N(0; s - r)$ ; iii)  $W(s) - W(t)$  e  $W(s + r) - W(t + r)$  tem a mesma distribuição para  $s, t$  e  $r$  assumindo valores no intervalo  $[0; 1]$ ; iv)  $W(r) \sim N(0; r), \forall 0 \leq r \leq 1$ ; e v) as trajetórias de  $W(r)$  são contínuas com probabilidade igual a um.

Ainda, na realização do teste ADF [9], é usual trabalhar com o operador diferença  $\Delta$ , isto é, podemos reescrever a equação (2.1) como

$$\begin{aligned} y_t - y_{t-1} &= \phi y_{t-1} - y_{t-1} + u_t, \\ \Delta y_t &= (\phi - 1)y_{t-1} + u_t, \\ \Delta y_t &= \delta y_{t-1} + u_t. \end{aligned}$$

Neste caso,  $\delta = \phi - 1$  e as hipóteses testadas se transformam em  $H_0 : \delta = 0$  e  $H_1 : \delta \neq 0$ . Se  $\delta = 0$ , então  $\phi = 1$ , o que indica a presença de raiz unitária. A estatística do teste é dada por  $\hat{\tau} = \hat{\delta}/EP(\hat{\delta})$ , sendo o erro padrão e o estimador da variância do erro determinados por

$$EP(\hat{\delta}) = \frac{s}{\sqrt{\sum_{t=1}^T y_{t-1}^2}} \quad \text{e} \quad s^2 = \frac{\sum_{t=2}^T (\Delta y_t - \hat{\delta} y_{t-1})^2}{T - 2},$$

respectivamente, sendo  $s$ , o desvio padrão.

Alguns valores críticos desta estatística de teste associados a determinadas probabilidades de ocorrência, que representam possíveis níveis de significância, foram determinados por

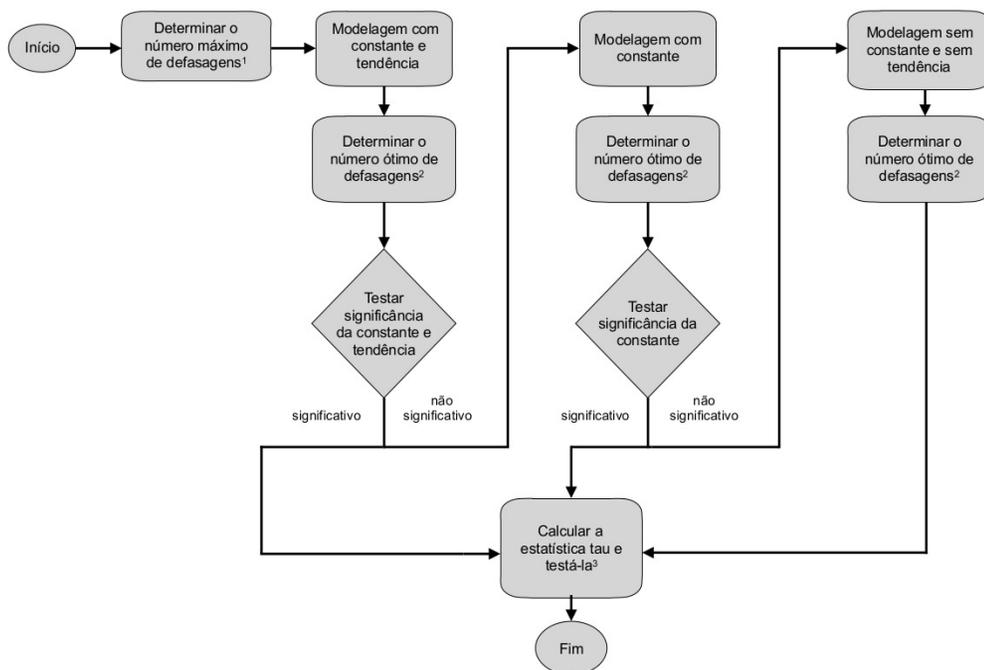


Figura 1: Dinâmica de aplicação do teste de raiz unitária ADF. Nota: 1. o número máximo de defasagens deve ser determinado por algum critério; 2. o número ótimo de defasagens pode ser determinado pelo teste  $t$  ou pelos critérios de informação; 3. ao rejeitar a hipótese nula, assume-se estacionariedade.

simulações de Monte Carlo. As estatísticas  $t$  e  $\tau$  diferem em probabilidade no infinito [7], ao passo que a estatística  $\tau$  torna o teste mais poderoso, embora seu poder diminua a medida que forem acrescentados mais termos ao polinômio que descreve o processo estocástico [7, 8]. O procedimento supracitado pode ser usado para modelos AR(1) com constante, assim como também para modelos com constante e tendência.

O teste ADF (observe o fluxograma na Figura 1), entretanto, é sensível à quantidade de defasagens considerada ( $k$ ), sendo este valor inversamente proporcional ao seu poder [3]. Na prática, utiliza-se um processo iterativo, partindo de uma defasagem máxima ( $k_{\text{máx}}$ ) até encontrar uma ótima ( $k_{\text{ótimo}}$ ) [19]. Isto é, inicialmente, indica-se uma defasagem máxima, que deve ser previamente indicada, como, por exemplo,  $k_{\text{máx}} = \text{int} \left[ 12 \left( \frac{T}{100} \right)^{1/4} \right]$  [24] ou  $k_{\text{máx}} = \text{int} \left[ 4 \left( \frac{T}{100} \right)^{2/9} \right]$  [13]. Em etapa posterior, ocorre o ajuste de sucessivos modelos, até que se obtenha uma defasagem ótima [13, 24]. Tal defasagem pode ser selecionada, por exemplo, por meio de critérios de qualidade de ajuste, tais como o critério de informação de Akaike (AIC) [1], o critério de informação bayesiano (BIC) [23] ou o critério de informação de Hannan-Quinn (HQIC) [12]. Entretanto, não existe um real consenso sobre a utilização de tais critérios, uma vez que, por exemplo, eles ten-

dem a identificar um valor muito pequeno de defasagens se a série apresentar valor inicial muito grande e/ou presença de quebra estrutural na tendência [11]. Todavia, a quantidade selecionada de defasagens por meio de tais critérios apresentam melhor desempenho quando comparados ao teste  $t$  [6].

### 3 METODOLOGIA

O presente estudo foi realizado seguindo a metodologia de simulações de Monte Carlo, sendo seus resultados analisados pela metodologia de análise de experimentos. Na simulação os dados foram gerados de acordo com os modelos AR(1), descritos pelas equações (2.1), (2.2) e (2.3), considerando  $u_t \sim N(0, 1)$ ,  $t = 1, \dots, T$ ,  $\beta_0 = 10$  e  $\beta_1 = 1, 0$ . Foram geradas 1.000 séries de amplitudes 30, 50 e 100 para cada um dos três modelos, sendo testada a  $H_0 : \phi = 1$  (equivalente à  $H_0 : \delta = 0$ ), ao nível de 5% de significância, a partir da estatística de teste  $\tau$ , para todos os valores de defasagem pertencentes ao intervalo  $[0, k_{\max} + 2]$ , em que  $k_{\max} = \text{int} \left[ 12 \left( \frac{T}{100} \right)^{1/4} \right]$  [24]. Os valores obtidos para a defasagem máxima foram 8, 10 e 12, para as séries de amplitudes 30, 50 e 100, respectivamente.

No primeiro cenário foi considerado  $\phi = 1$ , representando a presença de raiz unitária. Logo, nas testagens espera-se que o teste rejeite a hipótese de nulidade em, no máximo, 5% dos casos. Foram considerados três experimentos de acordo com a amplitude das séries com um fator, componentes do modelo, ensaiado em três diferentes níveis: modelo com constante e tendência (CT), modelo só com constante (C), modelo sem constante nem tendência (N). As defasagens foram consideradas como blocos, sendo computada, para cada condição experimental, a proporção de casos ( $\hat{p}$ ) em que a hipótese nula foi rejeitada.

Vale ressaltar que proporções entre zero e um podem se afastar muito de uma distribuição normal, o que não seria desejável, principalmente se os valores pertencerem aos intervalos  $[0; 0, 3]$  ou  $[0, 7; 1]$  [26]. Se a  $\hat{p}$  for transformada em  $\text{arc sen} \sqrt{\hat{p}}$ , a variância se estabiliza e os resultados tendem a apresentar uma distribuição bastante próxima da normal, razão pela qual foi utilizada esta transformação matemática. Para analisar o efeito de cada fator nas proporções transformadas foi realizada uma análise de variância, que faz uso da estatística  $F$ , determinada a partir da estimativa da variância do erro. Não obstante, a transformação matemática utilizada estabiliza a variância do erro em  $1/(4N)$ , sendo  $N$  a quantidade de simulações. Logo, uma vez adotado  $N = 1.000$ , é possível considerar a variância do erro conhecida é igual a 0,00025 e, sendo conhecida, a significância do resultado pode ser avaliada por meio da estatística de teste  $\chi^2 = \frac{\text{SQ}_{\text{fonte de variação}}}{\sigma_{\text{erro}}^2}$ , que apresenta distribuição qui-quadrado com  $gl_{\text{fonte de variação}}$  graus de liberdade.

Posteriormente, outros cenários são considerados para diferentes valores de  $\phi < 1$ , gerando séries com ausência de raiz unitária, sendo esta característica considerada como um outro fator ensaiado em três níveis (0,95; 0,75; 0,55).

Os procedimentos descritos foram todos realizados com o auxílio do software R [22], utilizando, principalmente, os pacotes `forecast` [14], `urca` [21], `lmtest` [27] e `agricolae` [15].

## 4 RESULTADOS E DISCUSSÕES

### 4.1 Cenário 1: $\phi = 1$

Neste contexto espera-se que a proporção de rejeições da hipótese nula seja baixa. Como o nível de significância considerado foi 0,05, esta proporção deve ser de, no máximo, 5%, o que parece acontecer para a maioria dos casos. As medidas resumo obtidas a partir das séries geradas são apresentadas na Tabela 1, onde é possível constatar que, como esperado, para cada estrutura considerada o valor- $p$  tende a diminuir com o aumento da amplitude da série. Comportamento um pouco diferenciado foi apresentado pelas séries com  $T = 100$ , sem constante e sem tendência, estrutura que, na prática, dificilmente acontece.

Tabela 1: Média (e desvio padrão) da proporção de casos em que a hipótese nula é rejeitada, no primeiro cenário com simulações considerando  $\phi = 1$ .

$T$	Constante	Constante e tendência	Sem constante e sem tendência
30	0,04011 (0,00411)	0,04091 (0,00630)	0,05555 (0,00446)
50	0,3085 (0,00428)	0,00862 (0,00409)	0,04723 (0,00588)
100	0,01506 (0,00282)	0,00435 (0,00154)	0,05918 (0,00484)

A Figura 2 apresenta a distribuição dos resultados das proporções de casos em que  $H_0$  é rejeitada obtidas a partir de séries com diferentes amplitudes e com diferentes estruturas. Os gráficos sugerem que a proporção parece diminuir à medida que a amplitude da série aumenta, existindo uma aparente estabilidade de sua tendência central em relação ao número de defasagens. O mesmo acontece em relação à sua estrutura, evidenciando a superioridade dos resultados obtidos em estruturas com menos componentes (sem constante e sem tendência), principalmente para  $T = 100$ .

A Tabela 2 apresenta os resultados obtidos a partir da análise de variância. Considerando nível de significância de 0,05, não foram encontradas evidências de que a quantidade de defasagens interfere na proporção de casos em que  $H_0$  é rejeitada em séries menos amplas  $T = 30$  e  $T = 50$ , o mesmo não acontecendo quando esta amplitude foi aumentada  $T = 100$ , embora o valor- $p$  tenha ficado muito próximo do nível de significância adotado. Também foram encontradas evidências de que a estrutura do modelo interfere na proporção de casos em que a hipótese nula é rejeitada, sendo aqueles que não tem constante nem tendência os que mais ultrapassaram 5%. Ainda, modelos sem constante e sem tendência são os que mais se afastam deste limite, sendo este fato mais evidente em séries mais longas.

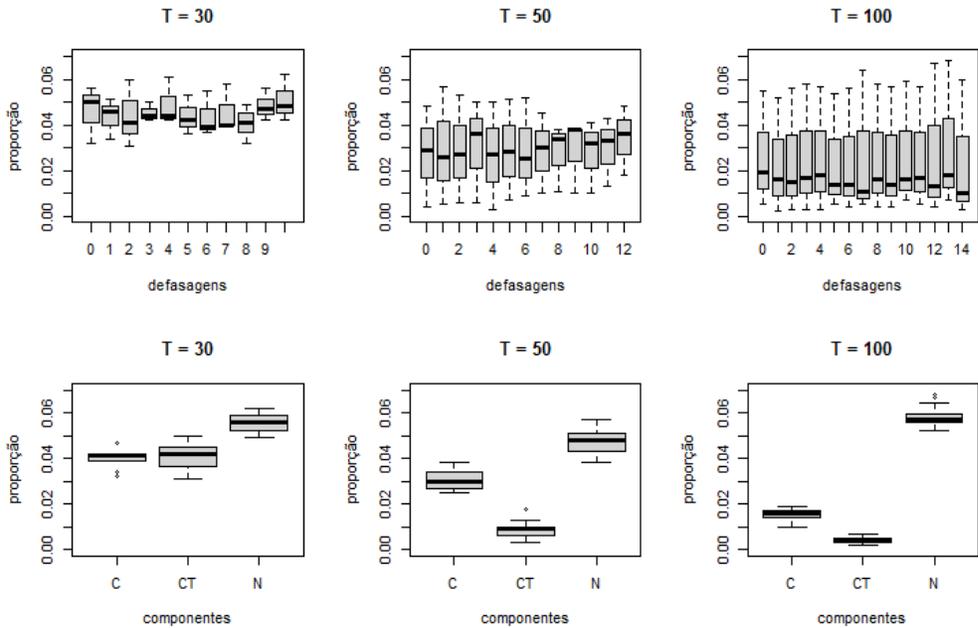


Figura 2: Gráficos de caixa de acordo com a quantidade de defasagens e componentes do modelo no primeiro cenário com simulações considerando  $\phi = 1$ .

Tabela 2: Resultados da análise de variância no primeiro cenário com simulações considerando  $\phi = 1$ .

$T$	Fonte de variação	gl	SQ	$X^2_{\text{calc}}$	valor- $p$
30	Defasagens	10	0,00146	5,84	0,0817
	Componentes	2	0,00927	37,08	<0,0001
50	Defasagens	12	0,00235	9,40	0,0869
	Componentes	2	0,11075	443,00	<0,0001
100	Defasagens	14	0,00192	7,68	0,0478
	Componentes	2	0,25189	1.007,56	<0,0001

#### 4.2 Cenário 2: $\phi < 1$

Neste contexto, diferentemente do anterior, espera-se que a proporção de rejeições da hipótese nula seja elevada, pois as séries geradas não apresentam raiz unitária, caracterizando-se como estacionárias. As medidas resumo das proporções encontradas para as séries geradas são apresentadas na Tabela 3, onde é possível constatar que, para cada estrutura considerada, a proporção de rejeições da  $H_0$  tende a aumentar com o aumento da amplitude da série, com a diminuição do coeficiente e com a simplificação da estrutura utilizada em sua geração.

Tabela 3: Média (e desvio padrão) da proporção de casos em que a hipótese nula é rejeitada, no segundo cenário com simulações considerando  $\phi < 1$ .

$T$	$\phi$	Sem constante			Total
		Constante	Constante e tendência	e sem tendência	
30	0,95	0,05736 (0,00642)	0,04554 (0,00706)	0,28318 (0,10915)	0,12870 (0,12681)
	0,75	0,27339 (0,27055)	0,07782 (0,04468)	0,53309 (0,33477)	0,29509 (0,30720)
	0,55	0,33618 (0,35251)	0,21045 (0,25953)	0,54191 (0,35732)	0,36285 (0,35401)
	Total	0,22248 (0,27648)	0,11127 (0,16413)	0,45273 (0,30574)	0,26216 (0,29104)
50	0,95	0,06431 (0,01212)	0,08477 (0,02396)	0,35623 (0,12262)	0,16844 (0,15214)
	0,75	0,37669 (0,32106)	0,26400 (0,23523)	0,67008 (0,28933)	0,4369 (0,32639)
	0,55	0,43977 (0,36722)	0,41300 (0,35827)	0,70185 (0,29068)	0,51821 (0,35669)
	Total	0,29359 (0,32068)	0,25392 (0,27689)	0,57605 (0,28781)	0,37452 (0,32663)
100	0,95	0,10607 (0,02086)	0,21727 (0,06903)	0,54647 (0,11764)	0,28993 (0,20813)
	0,75	0,63773 (0,28566)	0,56573 (0,27417)	0,89433 (0,13363)	0,69927 (0,28273)
	0,55	0,71113 (0,28211)	0,69440 (0,28655)	0,92780 (0,10584)	0,77778 (0,25899)
	Total	0,58753 (0,31349)	0,56950 (0,28203)	0,88103 (0,12968)	0,59694 (0,32952)

Os gráficos da Figura 3 apresentam a distribuição dos resultados referentes às proporções obtidas no que tange a rejeição da hipótese nula em séries de diferentes amplitudes, para diferentes defasagens, diferentes estruturas e diferentes intensidades de  $\phi$ . Nestes gráficos, observa-se que esta proporção diminui à medida que aumenta o número de defasagens, o que ocorre mais rapidamente em séries de menor amplitude. Esta proporção tende a ser maior em séries de estrutura mais simples, sendo esta diferença mais evidente em séries de menor amplitude, e menor em séries geradas com  $\phi$  próximos da unidade. Esta diferença se acentua em séries de maior amplitude.

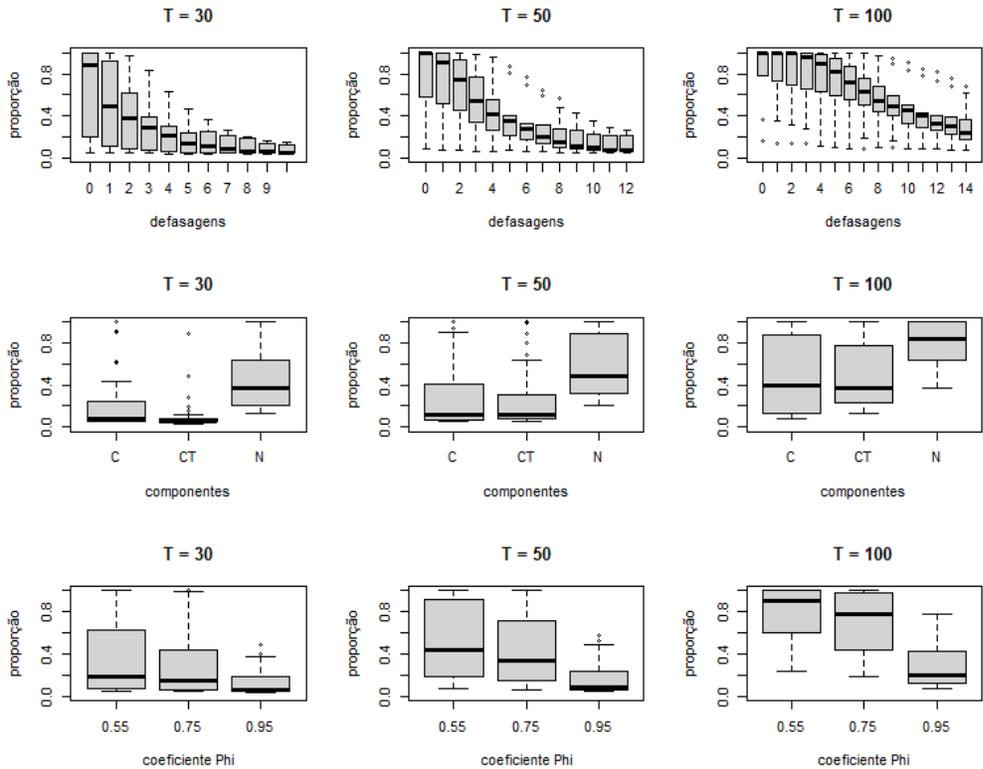


Figura 3: Gráficos de caixa de acordo com a quantidade de defasagens e componentes do modelo no segundo cenário com simulações considerando  $\phi < 1$ .

A Tabela 4 apresenta os resultados obtidos na análise de variância, mostrando que para as três amplitudes foram encontradas evidências de que a quantidade de defasagens, a estrutura do modelo e a intensidade do coeficiente interferem na proporção de casos em que a hipótese nula é rejeitada, concordando com a análise gráfica. Ademais, existem evidências de interação entre quantidade de componentes e defasagens e entre intensidade do coeficiente  $\phi$  e defasagens, representadas na Figura 4.

No que se refere à interação entre o número de componentes e de defasagens, os dados representados nos gráficos das Figuras 4 (a)-(c) mostram que valores mais elevados da proporção de rejeições da hipótese nula são encontrados para modelos sem constante e nem tendência. Estes valores são bastante próximos para os dois outros tipos de estrutura do modelo, tornando-se iguais a partir da sétima e oitava defasagem em séries de amplitude 30 e 50, respectivamente, e a partir da terceira defasagem em séries de amplitude 100.

Ainda, no que se refere à interação entre o número de defasagens e a magnitude do parâmetro  $\phi$ , os dados representados nos gráficos das Figuras 4 (d)-(f) mostram que valores mais elevados da

Tabela 4: Resultados da análise de variância no primeiro cenário com simulações considerando  $\phi = 1$ .

$T$	Fonte de variação	gl	SQ	$X^2_{\text{calc}}$	valor- $p$
30	Defasagens	10	5,068	20.272	<0,0001
	Componente	2	3,549	14.196	<0,0001
	Coefficiente	2	1,640	6.560	<0,0001
	Componente $\times$ Coeficiente	4	0,212	848	<0,0001
	Defasagens $\times$ Componente	20	0,860	3.440	<0,0001
	Defasagens $\times$ Coeficiente	20	1,912	7.648	<0,0001
50	Defasagens	12	8,585	34.340	<0,0001
	Componente	2	3,874	15.496	<0,0001
	Coefficiente	2	4,328	17.312	<0,0001
	Componente $\times$ Coeficiente	4	0,157	628	<0,0001
	Defasagens $\times$ Componente	20	0,258	1.032	<0,0001
	Defasagens $\times$ Coeficiente	20	2,767	11.068	<0,0001
100	Defasagens	14	7,546	30.184	<0,0001
	Componente	2	4,172	16.688	<0,0001
	Coefficiente	2	10,179	40.716	<0,0001
	Componente $\times$ Coeficiente	4	0,328	1.312	<0,0001
	Defasagens $\times$ Componente	20	0,298	1.192	<0,0001
	Defasagens $\times$ Coeficiente	20	1,649	6.596	<0,0001

proporção de rejeições de  $H_0$  são encontrados para coeficientes de menor intensidade, sendo esta diferença maior para menores defasagens. Entretanto, como a diminuição destes valores é mais acentuada para menores valores deste coeficiente, estas proporções se igualam a partir da oitava defasagem em séries de amplitude 30, valor da defasagem máxima ( $k_{\text{máx}}$ ), obtido pelo critério de Schwert [24], e a partir da 12ª defasagem em séries de amplitude 50, duas unidades acima do  $k_{\text{máx}}$ . Nas séries de amplitude 100, estes valores se aproximariam para resultados acima de  $k_{\text{máx}} + 2$ .

Não obstante, os resultados obtidos ao relação ao parâmetro  $\phi$ , estão de acordo com os encontrados em [2], em que, para séries com 50 observações, o poder do teste é muito baixo para  $\phi > 0,5$ , enquanto que para séries de 100 observações o poder do teste decresce rapidamente quando  $\phi > 0,8$ . Além disso, como pode-se observar, para pequenas amostras, o poder do teste decresce com o aumento do número de defasagens selecionado, corroborando assim com [3, 20].

Em relação aos termos determinísticos, em [25], para o caso do teste em que há a inclusão de tendência linear os resultados indicam que há um aumento na probabilidade de rejeitar a hipótese nula quando esta é verdadeira, associando isso ao fato de que a inclusão da tendência leva a subestimar o número de defasagens a serem incluídas na regressão do teste ADF. Estes achados não estão de acordo com o que foi encontrado neste trabalho, uma vez que aqui os indícios

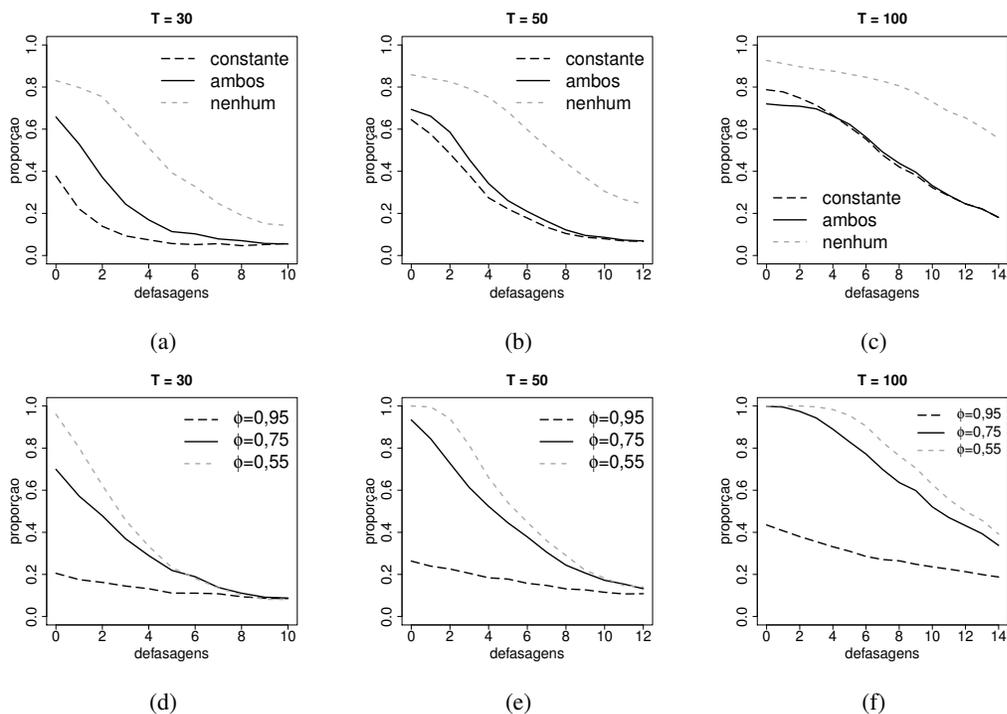


Figura 4: Representação da proporção de casos em que a hipótese nula é rejeitada e as interações entre as defasagens e: (a), (b), (c) número de componentes (com constante; ambos – com constante e tendência; nenhum – sem constante, nem tendência); (d), (e), (f) intensidade de  $\phi$ .

mostram que a inclusão de termos determinísticos reduz ligeiramente a proporção de rejeição da hipótese nula na presença da raiz unitária. Entretanto, cabe salientar que os dados estudados em [25] seguiam um processo ARIMA, diferentemente dos dados deste trabalho.

## 5 CONSIDERAÇÕES FINAIS

Este estudo apresenta os resultados de algumas simulações realizadas com o objetivo de avaliar a relação entre o valor- $p$ , determinado a partir da estatística  $\tau$  encontrada na aplicação do teste de Dickey-Fuller aumentado (ADF), e a quantidade de defasagens utilizada no respectivo cálculo, considerando modelos com diferentes estruturas no que se refere a componentes e intensidade do coeficiente  $\phi$ , fundamentando-se na proporção de rejeições da hipótese de nulidade, utilizando-se um nível de 5% de significância.

Os resultados obtidos mostraram que, se a série temporal possui raiz unitária, a quantidade de defasagens pouco interfere no resultado, independentemente da estrutura do modelo. Entretanto, se a série temporal não possui raiz unitária, a quantidade de defasagens interfere no resultado desta proporção, e consequentemente no valor- $p$ , diminuindo com o aumento do número de

defasagens, o que indica uma diminuição do poder do teste. Menores proporções são encontradas em modelos com constante e tendência, assim como também em modelos com coeficiente  $\phi$  próximo de 1,0.

Para trabalhos futuros, outras simulações podem ser realizadas com o objetivo de avaliar o desempenho das técnicas utilizadas na determinação da defasagem máxima e da defasagem ótima a serem utilizadas no teste ADF, empregado na identificação de raiz unitária em uma série temporal.

**ABSTRACT.** In this paper we evaluate how much influence the number of selected lags has in the p-value obtained from the statistic considered in an augmented Dickey-Fuller test (ADF). Three experiments with different series sizes were carried out considering the presence (or not) of a constant and/or trend, and also different number of lags. Autoregressive models (AR1) are used in the data generating process through the Monte Carlo method, considering different values for the coefficient associated with the first lag. After the application of the ADF test, we calculate the proportion of times when the null hypothesis was rejected in each scenario, and perform an analysis of variance (considering chi-square statistic) to verify the influence of the number of selected lags in the p-value. Results show that if a unit root is verified, then the test presents a good performance, regardless of the number of lags. However, the same was not observed in cases where the time series does not present a unit root.

**Keywords:** stationarity, unit root, time series, ADF test.

## REFERÊNCIAS

- [1] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**(6) (1974), 716–723.
- [2] M. Arltová & D. Fedorová. Selection of unit root test on the basis of length of time series and value of AR(1) parameter. *Statistika: Statistics and Economy Journal*, **96**(3) (2016), 47–64.
- [3] E. Aylar, S. Smeekes & J. Westerlund. Lag truncation and the local asymptotic distribution of the ADF test for a unit root. *Statistical Papers*, **60** (2019), 2109–2118.
- [4] G.E.P. Box, G.M. Jenkins & G.C. Reinsel. “Time Series Analysis: Forecasting and Control”. Wiley (2015).
- [5] M. Caner & L. Kilian. Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate. *Journal of International Money and Finance*, **20**(5) (2001), 639–657.
- [6] G. Cavaliere, P.C.B. Phillips, S. Smeekes & A.M.R. Taylor. Lag Length Selection for Unit Root Tests in the Presence of Nonstationary Volatility. *Econometric Reviews*, **34**(4) (2015), 512–536.
- [7] I. Choi. “Almost all about unit roots: Foundations, Developments, and Applications”. Cambridge University Press (2015).

- [8] D.A. Dickey & W.A. Fuller. Distribution of the estimators for autoregressive time series with unit root. *Journal of the American Statistical Association*, **74**(366) (1979), 427–431.
- [9] D.A. Dickey & W.A. Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, **49**(4) (1981), 1057–1072.
- [10] R.F. Engle & C.W.J. Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica*, **55**(2) (1987), 251–276.
- [11] H. Ferrer-Pérez, M.I. Ayuda & A. Aznar. The sensitivity of unit root tests to the initial condition and to the lag length selection: A Monte Carlo Simulation Study. *Communications in Statistics - Simulation and Computation*, (2019). doi:10.1080/03610918.2019.1577967.
- [12] E.J. Hannan & B.G. Quinn. The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, **41**(2) (1979), 190–195.
- [13] R.J. Hyndman & G. Athanasopoulos. *Forecasting: principles and practice* (2020). URL <http://otexts.com/fpp2/>.
- [14] R.J. Hyndman & Y. Khandakar. Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, **26**(3) (2008), 1–22.
- [15] F. Mendiburu & M. Yaseen. “agricolae: Statistical Procedures for Agricultural Research” (2020). R package version 1.4.0.
- [16] P.A. Morettin. “Econometria Financeira: um Curso em Séries temporais Financeiras”. Blucher (2011).
- [17] P.A. Morettin & C.M.C. Toloi. “Análise de Séries Temporais”. Blucher (2018).
- [18] C.R. Nelson & C.R. Plosser. Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics*, **10**(2) (1982), 139–162.
- [19] S. Ng & P. Perron. Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, **90**(429) (1995), 268–281.
- [20] E. Paparoditis & D.N. Politis. The asymptotic size and power of the augmented Dickey–Fuller test for a unit root. *Econometric Reviews*, **37**(9) (2018), 955–973.
- [21] B. Pfaff. “Analysis of Integrated and Cointegrated Time Series with R”. Springer (2008).
- [22] R Core Team. R: a language and environment for statistical computing (2020). URL <http://R-project.org/>.
- [23] G. Schwarz. Estimating the Dimension of a Model. *Annals of Statistics*, **6**(2) (1978), 461–464.
- [24] G.W. Schwert. Tests for unit roots: a Monte Carlo investigation. *Journal of Business & Economic Statistics*, **7**(2) (1989), 147–159.
- [25] A.M.R. Taylor. The Finite Sample Effects of Deterministic Variables on Conventional Methods of Lag-selection in Unit Root Tests. *Oxford Bulletin of Economics and Statistics*, **62**(2) (2000), 293–304.
- [26] J.H. Zar. “Biostatistical Analysis”. Pearson (1999).

- [27] A. Zeileis & T. Hothorn. Diagnostic Checking in Regression Relationships. *R News*, **2**(3) (2002), 7–10.

