

Dissertação de Mestrado:

**“COMPARAÇÃO DE TESTES DE RAIZ UNITÁRIA E
COINTEGRAÇÃO EM MODELOS DE LONGA DEPENDÊNCIA”**

Por

Fabiana de Assis Alves

Orientadora

Prof^a Glaura da Conceição Franco

Co-Orientador

Prof. Valdério Anselmo Reisen

Dezembro 2008

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE ESTATÍSTICA

Dissertação de Mestrado

**COMPARAÇÃO DE TESTES DE RAIZ UNITÁRIA E COINTEGRAÇÃO
EM MODELOS DE LONGA DEPENDÊNCIA**

ALUNA: FABIANA DE ASSIS ALVES
ORIENTADORA: PROF. DRA. GLAURA DA CONCEIÇÃO FRANCO
CO-ORIENTADOR: PROF. VALDÉRIO ANSELMO REISEN

BELO HORIZONTE, 01 DE DEZEMBRO DE 2008.

*Aos meus queridos pais,
aos meus irmãos, à Erica, ao Marcus
e aos meus pequenos Lucas e Arthur.*

AGRADECIMENTOS

Foi uma longa jornada até mais esta conquista, mas agora, olhando para trás, vejo aqueles que me ajudaram a transpor todos os obstáculos e gostaria muito de agradecê-los:

A Deus, por estar ao meu lado e por ter me dado forças para conseguir superar todas as dificuldades que surgiram durante essa árdua etapa. Obrigada Senhor, por também colocar pessoas maravilhosas no meu caminho e que certamente me ajudaram a chegar até aqui.

À Glaura, que durante todos esses anos, foi muito mais do que uma orientadora. Obrigada pelos conselhos, pela amizade, pelo profissionalismo e principalmente pela paciência que teve com as minhas dúvidas acadêmicas, profissionais e até mesmo pessoais.

Aos meus queridos Pais, exemplos de serenidade e sabedoria, que mesmo à distância se fizeram presentes através de incentivo, amor e dedicação.

Aos meus irmãos e cunhados, por estarem sempre ao meu lado e por suportarem uma boa dose de uma estudante estressada.

Aos meus queridos sobrinhos Lucas e Arthur, duas preciosidades na minha vida, que certamente tornaram essa jornada mais suave, alegre e colorida.

Ao Marcus, pelo amor, compreensão, companheirismo e apoio emocional.

Aos Professores Valdério, Dani e Salvato, por partilharem seus conhecimentos e pelas grandes contribuições que deram a este trabalho.

Aos amigos de Corredor, aos amigos da Graduação, aos amigos da Pós, enfim a todas as amizades que conquistei durante todos esses anos no Departamento de Estatística. Gostaria muito de citar nomes, mas por medo de ser injusta e me esquecer de alguém, agradeço a todos de uma maneira geral, pelos momentos de descontração, pela força dada nos momentos difíceis, pelas dicas, pelo apoio e pela amizade que construímos nesta etapa.

Ao Departamento de Estatística da UFMG, por oferecer as condições necessárias para mais esta conquista. Obrigada aos professores pelo conhecimento compartilhado e aos funcionários pela atenção dispensada.

SUMÁRIO

1 INTRODUÇÃO	10
2. LONGA DEPENDÊNCIA	13
2.1 O MODELO ARFIMA.....	13
2.2 ESTIMAÇÃO DO PARÂMETRO d	15
3 RAIZ UNITÁRIA	19
3.1 DEFINIÇÕES.....	19
3.2 TESTES DE RAIZ UNITÁRIA.....	21
3.2.1 Teste de Dickey-Fuller Aumentado (ADF).....	21
3.2.2 Teste de raiz unitária usando a Distribuição Assintótica do GPH.....	24
3.2.3 Testes de raiz unitária baseado na técnica <i>bootstrap</i>	24
4 COINTEGRAÇÃO.....	27
4.1 TESTES DE COINTEGRAÇÃO.....	29
4.1.1 Método de Engle Granger (EG)	29
4.1.2 Procedimento de Johansen.....	30
4.1.3 Teste de cointegração usando a distribuição assintótica do GPH	32
4.1.4 Testes de cointegração baseados na técnica <i>bootstrap</i>	33
5 RESULTADOS DE SIMULAÇÃO	34
5.1 PODER E TAMANHO DOS TESTES DE RAIZ UNITÁRIA.....	34
5.2 PODER E TAMANHO DOS TESTES DE COINTEGRAÇÃO	38
6 APLICAÇÃO A DADOS REAIS	43
6.1 DESCRITIVA PARA AS SÉRIES IBOVESPA E DOW JONES.....	44
6.2 VERIFICAÇÃO DA EXISTÊNCIA DE COINTEGRAÇÃO	45
7 CONCLUSÕES	48
REFERÊNCIAS BIBLIOGRÁFICAS	50
Anexo 1 - Programa para os Testes de Raiz Unitária.....	54
Anexo 2 - Programa para os testes de Cointegração	56

LISTA DE FIGURAS

Figura 5.1 Curva de Poder dos testes de raiz unitária para $n=100$	36
Figura 5.2 Curva de Poder dos testes de raiz unitária para $n=300$	37
Figura 5.3 Curva de Poder dos testes de raiz unitária para $n=500$	38
Figura 5.4 Curva de Poder dos testes de cointegração para $n=100$	40
Figura 5.5 Curva de Poder dos testes de cointegração para $n=300$	41
Figura 5.6 Curva de Poder dos testes de cointegração para $n=500$	42
Figura 6.1 Séries Ibovespa e Dow Jones	44
Figura 6.2 Box Plot das séries Ibovespa e Dow Jones	44
Figura 6.3 Plot Ibovespa vs Dow Jones.....	45
Figura 6.3 Gráfico para os resíduos do modelo ajustado na Tabela 6.4.....	47

LISTA DE TABELAS

Tabela 5.1 Resultados de tamanho e poder para os testes de raiz unitária para séries de tamanho 100	35
Tabela 5.2 Resultados de tamanho e poder para os testes de raiz unitária para séries de tamanho 300	36
Tabela 5.3 Resultados de tamanho e poder para os testes de raiz unitária para séries de tamanho 500	37
Tabela 5.4 Resultados de tamanho e poder para os testes de cointegração para séries de tamanho 100	40
Tabela 5.5 Resultados de tamanho e poder para os testes de cointegração para séries de tamanho 300	41
Tabela 5.6 Resultados de tamanho e poder para os testes de cointegração para séries de tamanho 500	42
Tabela 6.1 Testes de raiz unitária para as séries	45
Tabela 6.2 Modelo de regressão ajustado para as séries Ibovesa e Dow Jones	46
Tabela 6.3 Testes de cointegração estudados	47

RESUMO

Neste trabalho são propostos dois testes baseados na técnica *bootstrap* para testar a presença de raiz unitária e relações de cointegração em processos de longa dependência. Estes testes utilizam o estimador GPH proposto por Geweke e Porter-Hudak (1983) para estimação do parâmetro d fracionário de modelos ARFIMA(p,d,q). Foram utilizadas simulações Monte Carlo para comparar as performances dos testes sugeridos com as de outros testes amplamente divulgados na literatura. Para raiz unitária, foram considerados no estudo os testes ADF (Dickey e Fuller, 1981) e assintótico para o estimador GPH. Na avaliação de cointegração foram usados os testes de Engle e Granger (1987), Johansen (1990) e o baseado na distribuição assintótica do GPH. Os resultados mostram que a técnica *bootstrap* é uma boa alternativa para a construção de testes de hipótese sobre o parâmetro fracionário d de modelos de longa dependência, pois de uma maneira geral, os testes sugeridos apresentaram resultados satisfatórios frente aos testes comumente aplicados nos mais diversos estudos econométricos.

ABSTRACT

In this work two tests based on the bootstrap technique are considered to test the presence of unit root and relations of cointegration in long memory processes. The proposed tests use the GPH estimator considered by Geweke and Porter-Hudak (1983) for estimating the memory parameter d of ARFIMA(p,d,q) models. Monte Carlo simulations have been used to compare the performances of the suggested tests with other widely used tests in the literature. For unit root, the ADF test (Dickey and Fuller, 1981) and the test based on the asymptotic distribution of GPH have been considered. In the cointegration evaluation, the Engle and Granger test (1987), the Johansen procedure (1990) and the test based on the asymptotic distribution of the GPH have been used. The results show that the bootstrap technique is a good alternative for the construction of hypothesis tests on the parameter of long memory models. In general, the suggested tests presented satisfactory results if compared to the usually applied tests in econometrical studies.

1. INTRODUÇÃO

Uma das suposições mais freqüentes que se faz a respeito de uma série temporal é a de que ela é estacionária, ou seja, se desenvolve no tempo aleatoriamente ao redor de uma média e variância constantes. Todavia, na prática, a maior parte das séries encontradas apresenta alguma forma de não estacionaridade.

Durante a década de 1980, economistas aplicados utilizavam modelos econométricos de séries temporais sem estarem cientes de certas dificuldades que resultam quando há presença de raiz unitária, ou seja, processos estocásticos não estacionários, induzidos por acúmulo persistente de efeitos passados. Ao ignorar este fato na estimação de um modelo de regressão contendo variáveis não estacionárias, por exemplo, na melhor das hipóteses são ignoradas informações importantes dos processos que geram os dados e, na pior das hipóteses, resultados espúrios são encontrados. Por esta razão é importante a aplicação de um teste para verificar a presença de raízes unitárias e, caso elas sejam encontradas (o que ocorre na maioria das vezes), utilizar um procedimento apropriado para a modelagem dos dados.

Existem vários testes para verificar a presença de raiz unitária, sendo que os mais utilizados na prática são o teste DF proposto por Dickey e Fuller (1979), o teste ADF que é uma extensão do testes DF e que também foi proposto por Dickey e Fuller (1981) e o teste não paramétrico de Phillips e Perron (1988).

Para estudar relações entre variáveis econômicas, o conceito de cointegração, introduzido por Granger (1981), tem sido amplamente discutido, pois ao utilizar modelos de regressão com variáveis não estacionárias, pode-se obter resultados espúrios. A chamada regressão espúria é um problema reportado com freqüência na literatura econométrica e que consiste na tendência de aceitar a existência de uma relação entre séries quando ela não existe. Assim, a idéia básica de cointegração é a de que cada uma das séries que compõem um vetor Y_t é não estacionária, mas alguma combinação linear delas, digamos $u_t = \beta'Y_t$, é estacionária. Desde o trabalho de Granger (1981), vários estudos sobre este assunto foram desenvolvidos, sendo que o mais referenciado na literatura é o apresentado por Engle e Granger (1987). Neste estudo os autores apresentaram um procedimento para testar cointegração e, apesar de considerarem a possibilidade do parâmetro de diferenciação ser fracionário, trabalharam com valores inteiros. Ainda avaliando procedimentos clássicos para testar cointegração, ou seja, aqueles que consideram ordem de integração inteira, outra metodologia bastante citada é o procedimento de Johansen proposto no artigo de Johansen

(1988). Este teste, apesar de ser mais complexo em termos teóricos e computacionais, tem como vantagem poder identificar, nos casos em que se tenha mais do que duas séries, quantos vetores de cointegração existem entre as variáveis, enquanto o teste de Engle e Granger (1987) apenas indica se há ou não cointegração.

Novas abordagens surgiram e as pesquisas relacionadas à cointegração fracionária têm se tornado cada vez mais expressivas. Diversas metodologias estão sendo implementadas na estimação e na construção de teste de processos fracionalmente cointegrados. O modelo ARFIMA (p,d,q) proposto por Granger & Joyeux (1980) e Hosking (1981) que consiste numa generalização do modelo ARIMA de Box e Jenkins (1976) tem sido bastante utilizado para este fim, pois o parâmetro d fracionário consegue capturar a memória de uma série, e seus diversos estimadores podem ser usados na construção de métodos de estimação e testes de processos fracionalmente cointegrados. Os artigos de Tse *et al.* (1999), Dittmann (2000) e Santander *et al.* (2003) são exemplos de trabalhos teóricos e de aplicação em cointegração que utilizam estimadores do parâmetro d do modelo ARFIMA (p,d,q) .

A principal contribuição do presente estudo é a utilização da técnica de reamostragem *bootstrap* como alternativa para construir testes de hipóteses para o parâmetro d fracionário, tanto na parte de raiz unitária, quanto na parte de cointegração. O *bootstrap* é uma técnica de reamostragem que permite aproximar a distribuição de uma função das observações pela distribuição empírica dos dados e consiste basicamente em gerar repetidas amostras com reposição, um número suficientemente grande de vezes (Efron, 1979). No contexto de séries temporais, a reamostragem não deve ser aplicada diretamente nas observações, uma vez que estas são correlacionadas. Uma alternativa que tem sido bastante empregada em diversos estudos é aplicar a reamostragem nos resíduos do modelo ajustado à série.

Franco e Reisen (2004, 2007) e Arteché e Orbe (2005) em seus trabalhos, além de utilizarem diversas técnicas *bootstrap* para fazer inferência sobre o parâmetro de memória d do modelo ARFIMA (p,d,q) também fazem comparações entre as metodologias sugeridas. Andrews e Liberman (2002) provam propriedades assintóticas do *bootstrap* paramétrico. Peretti (2002) e Grau-Charles (2004) também empregam a técnica *bootstrap* na construção de testes para processos de longa dependência. Andersson e Grendenholf (2000) utilizam o *bootstrap* para testar integração fracionária. Nos artigos de Davidson (2002,2004), o *bootstrap* também é considerado como alternativa para testar a existência de relação de cointegração em processos fracionalmente integrados. Em seu estudo de 2002, Davidson ainda conta com uma aplicação a dados reais, onde os testes propostos foram empregados na avaliação de cointegração entre as séries de consumo e rendimento líquido no Reino Unido.

Neste estudo são propostas duas abordagens baseadas na técnica *bootstrap* para averiguar a presença de raiz unitária e a existência de cointegração em séries de longa dependência. Além disso, uma outra contribuição interessante deste trabalho é que serão realizados estudos de simulação Monte Carlo visando comparar empiricamente o poder e o tamanho dos testes propostos com o das principais metodologias existentes nos contextos clássico e fracionário. Dentre estas metodologias são considerados, no caso de raiz unitária, os testes Dickey e Fuller aumentado (ADF) e o teste baseado na distribuição assintótica do estimador GPH de Geweke e Porter-Hudak (1983). Para a cointegração são usados o teste de Engle e Granger, o procedimento de Johansen e o teste assintótico para o estimador GPH. Os testes *bootstraps* sugeridos neste estudo, bem como o teste de Engle Granger e o teste que utiliza a distribuição assintótica do GPH, são aplicáveis quando se deseja testar cointegração entre duas séries, uma vez que estas metodologias consideram apenas uma equação de cointegração. Já o procedimento de Johansen é um teste mais geral e pode ser aplicado quando se trabalha com mais de duas séries.

Este trabalho conta ainda com uma etapa de aplicação à dados reais, onde serão aplicados os testes de raiz unitária e cointegração descritos no estudo, com intuito de averiguar se os índices Ibovespa e Dow Jones apresentam raiz unitária e são cointegrados.

Este relatório está estruturado da seguinte maneira: O Capítulo 2 é dedicado à característica de Longa Dependência com uma breve descrição do modelo ARFIMA (p,d,q) e do estimador GPH para o parâmetro d . No Capítulo 3 são apresentados os conceitos de estacionaridade e raiz unitária e os testes utilizados neste estudo para detectar a presença de raiz unitária em uma série temporal. A cointegração é tratada no capítulo 4, bem como os testes utilizados neste estudo para detectar a presença de cointegração entre duas ou mais séries temporais. Um estudo de simulações Monte Carlo que compara os testes de raiz unitária e os testes de cointegração é apresentado no Capítulo 5. No Capítulo 6 é mostrada uma aplicação das técnicas estudadas neste trabalho para um conjunto de dados reais e as conclusões são dadas no Capítulo 7.

2. LONGA DEPENDÊNCIA

A característica de longa dependência é observada em séries que apresentam um ajuste lento e de longo prazo. Suas observações, mesmo separadas por um intervalo longo de tempo, são significativamente dependentes. Esta característica tem sido observada em diferentes áreas de estudo, como exemplo, pode-se citar as áreas de meteorologia, astronomia, hidrologia e economia.

Segundo Murteira et al. (1993), em geral, as séries temporais não se encontram estacionarizadas e, como os modelos ARMA propostos por Box e Jenkins (1976) só se aplicam a processos estacionários, faz-se necessário operar sobre os dados originais um certo número de diferenciações, de modo que a série resultante possa ter a variância estabilizada, a tendência neutralizada e os movimentos de caráter estritamente periódico possam ser eliminados. Quando uma série não estacionária é diferenciada d vezes até que se torne estacionária, o modelo de Box e Jenkins apropriado para descrevê-la é o modelo ARIMA (p,d,q) (Autorregressivo Integrado Média Móvel).

Se o parâmetro d (grau de diferenciação) de um modelo ARIMA (p,d,q) , assumir valores não inteiros, o modelo é chamado processo geral com diferenciação fracionária. Então quando d assumir tais valores estaremos nos referindo ao modelo ARFIMA (p,d,q) (*General Fractional Differenced Process*) e este é bastante utilizado para descrever as séries com característica de curta e longa dependência dependendo do valor que d assumir.

O modelo ARFIMA (p,d,q) e suas principais características são apresentados na Seção 2.1. Além do modelo ARFIMA, neste capítulo também é apresentado um estimador para o parâmetro d fracionário. Existem vários métodos propostos na literatura para estimação deste parâmetro, sendo que neste trabalho considerou-se o método de estimação semiparamétrica proposto por Geweke & Porter-Hudak (1983), denominado GPH, que será visto na Seção 2.2

2.1 O MODELO ARFIMA

Se y_t é um processo linear satisfazendo:

$$\Phi_p(B)y_t = \Theta_q(B)(1-B)^{-d} e_t \quad d \in (-0.5,0.5) \quad (2.1)$$

onde e_t é um ruído branco com $\mu = 0$ e variância σ_e^2 , B é o operador de retardo da forma $By_t = y_{t-1}$, $\Phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ e $\Theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ são os polinômios

autorregressivo e média móvel, respectivamente, com raízes não comum e fora do círculo unitário. Então, y_t é um processo *geral com diferenciação fracionária* (*General Fractional Differenced*) e é denominado processo ARFIMA (p,d,q) , onde d é o grau de diferenciação fracionário (Granger & Joyeux, 1980 e Hosking, 1981). O termo $(1-B)^d$ na expressão (2.1) pode ser descrito através da expansão binomial

$$(1-B)^d = 1 - dB + \frac{d(d-1)B^2}{2!} - \frac{d(d-1)(d-2)B^3}{3!} + \dots \quad (2.2)$$

A função densidade espectral de y_t , $f(\omega)$, é definida por

$$f(\omega) = f_U(\omega) \left[2 \sin\left(\frac{\omega}{2}\right) \right]^{-2d} \quad (2.3)$$

onde $\omega, \omega \in [-\pi, \pi]$, são as freqüências de Fourier e $f_U(\omega)$ é a função densidade espectral de $U_t = (1-B)^d y_t$:

$$f_U(\omega) = \sigma_{a_t}^2 \frac{|\theta_q(e^{-i\omega})|^2}{|\phi_p(e^{-i\omega})|^2}, \quad (2.4)$$

onde $\sigma_{a_t}^2$ é a variância do processo U_t .

As principais características do modelo ARFIMA (p,d,q) são:

- Se $d < 0.5$ e todas as raízes de $\phi(B) = 0$ estiverem fora de círculo unitário o processo é estacionário.
- Se $d > -0.5$ e todas as raízes de $\theta(B) = 0$ estiverem fora do círculo unitário o processo é invertível.
- Para $d \in (-0.5, 0.0)$ todas as autocorrelações são negativas e tendem hiperbolicamente para zero e o processo é tido como antipersistente ou de memória intermediária.
- Para $d=0$ o processo é considerado de memória curta.
- Para $d \in (0.0, 0.5)$ as autocorrelações são positivas tendendo hiperbolicamente para 0 de uma forma muito lenta. Este processo é dito como persistente ou de longa dependência estacionário, apresentando dependência entre observações distantes da série.
- Quando $0.5 < d < 1.0$ o processo é não estacionário, pois sua variância não é finita, porém é reversível na média.
- Se $d > 1$ o processo é não estacionário e não reversível na média.

2.2 ESTIMAÇÃO DO PARÂMETRO d

Existem vários métodos na literatura para estimar o parâmetro d do modelo ARFIMA (p,d,q) . Basicamente, pode-se dizer que eles se dividem em dois métodos principais: a estimação paramétrica e a semiparamétrica. Os métodos paramétricos são baseados na função de verossimilhança e os semiparamétricos são construídos a partir de uma equação de regressão estimada via mínimos quadrados. Neste estudo, optou-se pelo uso de um estimador semiparamétrico, uma vez que os métodos paramétricos, apesar de apresentarem maior eficiência, demandam um grande esforço computacional. Robinson, em seu artigo de 2003 apresenta uma discussão detalhada sobre estimadores paramétricos e não paramétricos do parâmetro de memória d .

Os estimadores de Geweke e Porter-Hudak (1983), Reisen (1994) e Hurvich e Ray (1995) são alguns dos estimadores semiparamétricos mais utilizados na estimação do parâmetro d fracionário de um modelo ARFIMA (p,d,q) . Para esta pesquisa optou-se pelo estimador GPH proposto Geweke e Porter-Hudak (1983), pois se verificou que para o intervalo a ser considerado neste estudo, este apresentou melhor comportamento. Estudos de simulação recentes, comparando diferentes procedimentos de estimação para processos de longa dependência podem ser encontrados em Bisaglia e Guégan (1998), Reisen *et al.* (2001) e Franco e Reisen (2007).

O método de Geweke e Porter-Hudak consiste em tomar o logaritmo da função densidade espectral $f(\omega)$ dada em (2.3), e estimar o parâmetro d por mínimos quadrados através da equação de regressão obtida.

Logo tirando o logaritmo da função espectral temos:

$$\ln f(\omega) = \ln f_u(\omega) - d \ln \left\{ 2 \sin \left(\frac{\omega}{2} \right) \right\}^2,$$

que pode ser rescrita da seguinte maneira:

$$\ln f(\omega) = \ln f_u(0) - d \ln \left\{ 2 \sin \left(\frac{\omega}{2} \right) \right\}^2 + \ln \left\{ \frac{f_u(\omega)}{f_u(0)} \right\}. \quad (2.5)$$

Como $f(\omega)$ é desconhecida, é necessário substituí-la então pelo seu estimador, a função Periodograma, definida a seguir.

Definição 2.1 Sejam y_1, y_2, \dots, y_n observações do processo y_t . A função $I(\omega)$, chamada periodograma, é definida por

$$I(\omega) = 2 \left[R(0) + 2 \sum_{j=1}^{n-1} R(j) \cos(\omega_j) \right], \quad \omega \in [-\pi, \pi]. \quad (2.6)$$

onde $R(j)$, a autocovariância amostral (para uma amostra de tamanho n) é dada por

$$R(j) = \frac{1}{n} \sum_{i=1}^{n-j} (y_i - \bar{y})(y_{i+j} - \bar{y}), \quad j = 0, \pm 1, \dots, \pm(n-1) \quad (2.7)$$

e $R(0)$, a variância amostral, é um caso particular de (2.7)

Tomando $I_a(\omega) = \frac{I(\omega)}{4\pi}$ temos:

$$I_a(\omega) = \frac{1}{2\pi} \left[R(0) + 2 \sum_{s=1}^{n-1} R(s) \cos(\omega_s) \right], \quad \omega \in (-\pi, \pi). \quad (2.8)$$

Verifica-se que $I_a(\omega)$ é um estimador assintoticamente não viciado de $f(\omega)$, pois:

$$E[I_a(\omega)] = \frac{1}{2\pi} \left[R(0) + 2 \sum_{j=1}^{n-1} R(j) \cos(\omega_j) \right] \rightarrow f(\omega), \quad n \rightarrow \infty.$$

Mas (2.8) não é consistente, pois sua variância não tende a zero à medida que $n \rightarrow \infty$. $I_a(\omega)$ é conhecido como estimado “natural” da função espectral $f(\omega)$.

Substituindo-se ω pelas frequências harmônicas, $\omega_j = 2\pi j / n$ com $j = 1, 2, \dots, g(n)$ e $f(\omega)$ por seu estimador assintoticamente não viciado, a função periodograma $I_a(\omega_j)$, na equação (2.5) temos:

$$\ln I_a(\omega_j) = \ln f_u(0) + \ln \left\{ f_u(\omega_j) / f_u(0) \right\} - d \ln (2 \operatorname{sen}(\omega_j / 2))^2 + \ln (I_a(\omega_j) / f(\omega_j))$$

O limite superior de j é dado por uma função $g(n)$, a qual deve satisfazer:

$$\frac{g(n)}{n} \rightarrow 0 \text{ quando } n \rightarrow \infty \text{ e para } \omega_j \text{ próximo de zero, } \omega_j \leq \omega_{g(n)} \text{ onde } \omega_{g(n)} \text{ é pequeno.}$$

Quanto à escolha de $g(n)$, o mais indicado é escolher $g(n) = n^\alpha$, $\alpha \in (0, 1)$. O valor de α

deve ser definido de acordo com o estudo. Quando se trabalha com o modelo ARFIMA (p,d,q) é usual utilizar $\alpha = 0,5$, já quando se trabalha com o modelo ARFIMA $(0,d,0)$, que é o caso deste estudo, é recomendado a utilização de maiores valores para o α (Reisen et al., 2001).

O termo $\ln(f_u(\omega_j)/f_u(0))$ é considerado desprezível comparado aos demais termos.

Portanto temos:

$$\ln I_a(\omega_j) \cong \ln f_u(0) - d \ln(2\text{sen}(\omega_j/2))^2 + \ln(I_a(\omega_j)/f(\omega_j)) \quad (2.9)$$

Esta equação resulta numa equação de regressão simples da forma $z_j = a + bx_j + e_j$ com $j= 1,2,\dots, g(n)$, onde:

$$z_j = \ln I_a(\omega_j),$$

$$x_j = \ln[2\text{sen}(\omega_j/2)]^2,$$

$$a = \ln f_u(0) - c,$$

$$b=-d,$$

$$e_j = \ln \left[\frac{I_a(\omega_j)}{f(\omega_j)} \right] + c.$$

Como as variáveis $\ln(I_a(\omega_j)/f(\omega_j))$, $j=1,2,\dots, g(n)$, são aproximadamente independentes com distribuição de Gumbel com média $-0,577216$ e variância $\pi^2/6$ quando $d \in (-0.5,0.0)$, então as variáveis e_j também são aproximadamente independentes com distribuição Gumbel $(0, \pi^2/6)$.

Este resultado sugere o estimador de d pelo método de mínimos quadrados da regressão. Assim:

$$\hat{d}_{GPH} = -\hat{b} = -\frac{\sum_{i=1}^{g(n)} (x_i - \bar{x})(z_i - \bar{z})}{\sum_{i=1}^{g(n)} (x_i - \bar{x})^2} \quad (2.10)$$

onde \hat{d}_{GPH} é o estimador de d usando a função periodograma como estimadora da função espectral.

Segundo Geweke e Porter_Hudak (1983), este estimador tem valor esperado $E(\hat{d}_{GPH})=d$ e variância $\sigma_{GPH}^2 = \frac{\pi^2}{6 \sum_{i=1}^{g(n)} (x_i - \bar{x})^2}$. Sua distribuição assintótica é tal que

$$\frac{\hat{d}_{GPH} - d}{\sqrt{\sigma_{GPH}^2}} \rightarrow N(0,1) \text{ onde } \sigma_{GPH}^2 \text{ é a variância de } d_{GPH}.$$

A distribuição assintótica do estimador GPH será utilizada na construção de testes para verificar a presença de raiz unitária e detectar existência de cointegração entre duas séries temporais como será visto mais adiante.

3. RAIZ UNITÁRIA

Métodos clássicos de estimação que habitualmente são empregados em trabalhos econométricos aplicados são baseados nas suposições de que médias e variâncias são constantes bem definidas e independentes do tempo. Entretanto, é fácil constatar a partir de testes de raízes unitárias que estas suposições não estão satisfeitas para um grande número de séries temporais.

Atualmente na literatura econométrica, existem diversos testes que verificam a presença de raiz unitária em uma série temporal. Estes testes foram inicialmente propostos por Fuller (1976) e Dickey e Fuller (1979). Mais tarde, foram publicados outros trabalhos sobre este assunto nos artigos de Said e Dickey (1984) e Phillips e Perron (1988).

Este capítulo apresenta alguns testes de raiz unitária, a saber, o teste ADF proposto por Dickey e Fuller (1981) e que constitui-se em um dos mais populares e utilizados na prática, o teste baseado na distribuição assintótica do estimador GPH proposto por Geweke e Porter-Hudak (1983) e dois outros testes baseados na técnica de reamostragem *Bootstrap* e que também utilizam o estimador GPH serão sugeridos neste estudo. Antes de apresentar a teoria de cada teste, serão definidos alguns conceitos importantes a respeito de estacionaridade e integração clássica e fracionária.

3.1 DEFINIÇÕES

A presença de raiz unitária em uma série está intimamente relacionada com o conceito de estacionaridade. Intuitivamente diz-se que uma série é estacionária quando esta se desenvolve no tempo de forma constante.

Considere um conjunto finito de variáveis aleatórias $\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$ de um processo estocástico $\{X_t\}_{t \in N}$ e seja $F_{X_{t_1}, \dots, X_{t_n}}(X_1, \dots, X_n)$ a distribuição de probabilidade conjunta de dimensão n do vetor $\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$, isto é:

$$F_{X_{t_1}, \dots, X_{t_n}}(X_1, \dots, X_n) = P(X_{t_1} \leq X_1, X_{t_2} \leq X_2, \dots, X_{t_n} \leq X_n). \quad (3.1)$$

Definição 3.1 (*Processo Estritamente Estacionário*): Um processo $\{X_t\}_{t \in N}$ é dito ser estritamente estacionário, ou fortemente estacionário, se todas as distribuições finito-dimensionais (3.1) permanecem invariantes no tempo, ou seja,

$$F_{X_{t_1}, \dots, X_{t_n}}(X_1, \dots, X_n) \equiv F_{X_{t_1+k}, \dots, X_{t_n+k}}(X_{1+k}, \dots, X_{n+k}) \quad (3.2)$$

para qualquer vetor (t_1, t_2, \dots, t_n) de número inteiros e qualquer k inteiro.

Este tipo de estacionariedade é muito difícil de ser observado na prática, dado que geralmente a função F não é conhecida. Assim é possível se trabalhar bem com processos que atendem ao conceito dado a seguir.

Definição 3.2 (*Processo Fracamente Estacionário*): Um processo $\{X_t\}_{t \in \mathbb{N}}$ é dito ser fracamente estacionário, ou simplesmente estacionário, se e somente se:

- i. $E(X_t) = E(X_{t+k}) = \mu$;
- ii. $E(X_t^2) = \sigma_Y^2 < \infty$;
- iii. $Cov(X_t, X_{t+k})$ é uma função que só depende da distância entre t e $t+k$.

A característica de não estacionariedade é frequentemente observada em estudos de séries temporais. Um procedimento bastante comum, para se prosseguir com a análise deste tipo de série é a diferenciação, que tem por objetivo torná-la estacionária e assim aplicar modelagens conhecidas na análise de séries estacionárias, como por exemplo, a utilização dos modelos ARMA(p, q) de Box e Jenkins.

Um conceito bastante ligado ao de diferenciação e muito utilizado na literatura é o de integração. Este conceito foi introduzido por Engle e Granger(1987) e é dado a seguir:

Definição 3.3 Uma série sem componente determinística, que tem uma representação ARMA(p, q) estacionária e invertível depois de ter sido diferenciada d vezes, é chamada série integrada de ordem d e denotada $y_t \sim I(d)$.

Neste contexto, um processo $I(d)$, $y_t, t = 1, \dots, n$, é definido por

$$(1 - B)^d y_t = e_t, \quad t = 1, 2, \dots, n$$

onde e_t é $I(0)$, ou seja, é um processo ruído branco com média zero e variância constante.

Quando d é um inteiro maior ou igual a 1 se diz que y_t tem “raízes unitárias”, mas quando d é real a série é dita “fracionariamente integrada”.

A teoria acerca da integração fracionária foi introduzida por Granger e Joyeux (1980) e Hosking (1981) e é considerada na modelagem de séries que apresentam a característica de longa dependência. O lento decaimento dos efeitos dos choques e o vagaroso, mas eventual ajustamento para o equilíbrio é um dos atrativos deste tipo de processo. Ao permitir que a ordem de integração das séries assuma qualquer valor real conduz-se aos modelos de longa dependência, que apresentam como ponto de interesse, entre outros, o fato de implicarem

diferentes previsões de longo prazo e efeitos dos choques em relação às abordagens convencionais. Uma das características destes tipos de processos é que estes aparentam ser não estacionários, mas quando diferenciados percebe-se que houve uma sobrediferenciação.

3.2 TESTES DE RAIZ UNITÁRIA

O objetivo dos testes de raiz unitária é essencialmente detectar se uma dada série foi suficientemente diferenciada para se tornar estacionária. Para tanto, testa-se a hipótese nula de que a série não é estacionária, ou seja, possui raiz unitária contra a alternativa de que a série é estacionária. Na prática, há várias maneiras de testar a presença de raiz unitária em uma série temporal. O modelo ARFIMA(p,d,q), discutido no capítulo anterior, tem recebido considerável atenção em Econometria, pois é a partir da distribuição assintótica de alguns estimadores do parâmetro d fracionário que desenvolveu-se algumas técnicas para testar a presença de raiz unitária, como será visto adiante. Aqui serão utilizadas três metodologias distintas. A primeira, o teste de Dickey-Fuller Aumentado – ADF (Dickey e Fuller, 1981), que consiste num método clássico que utiliza uma extensão do testes de Dickey-Fuller (1979). A segunda é construída a partir do modelo ARFIMA (p,d,q) e é baseada na distribuição assintótica do estimador GPH de Geweke e Porter-Hudak (1983). A terceira metodologia é também construída a partir do estimador GPH, mas se utilizando da técnica *bootstrap* que será vista mais adiante.

3.2.1 Teste de Dickey-Fuller Aumentado (ADF)

A forma mais simples de testar estacionaridade baseia-se em um processo AR(1), que pode ser apresentado de uma maneira simplificada como:

$$y_t = \rho y_{t-1} + e_t \quad (3.3)$$

onde e_t é o termo do erro aleatório conhecido como ruído branco, que apresenta média zero, variância σ^2 constante e é não auto-correlacionado.

O teste de Dickey-Fuller (1979) consiste em testar:

$$H_0: \rho=1 \quad \textit{versus} \quad H_1: \rho<1.$$

Sob a hipótese nula, y_t não é um processo estacionário, sendo, portanto um caminho aleatório com mudança de nível. Sob a hipótese alternativa, y_t é um processo estacionário AR(1).

O estimador de máxima verossimilhança para ρ é o estimador de mínimos quadrados:

$$\hat{\rho} = \left[\sum_{t=2}^n y_{t-1}^2 \right]^{-1} \sum_{t=2}^n y_t y_{t-1} \quad (3.5)$$

Rubin (1950) mostrou que $\hat{\rho}$ é um estimador consistente para todos os valores de ρ .

O teste de razão de verossimilhança para $\rho = 1$ é baseado na seguinte estatística

$$t = (\hat{\rho} - 1) S^{-1} \left(\sum_{t=2}^n y_{t-1}^2 \right)^{-1/2} \quad (3.6)$$

onde a estimativa da variância do erro é dada por $S^2 = (n-2)^{-1} \sum_{t=2}^n (y_t - \hat{\rho} y_{t-1})^2$. A razão- t apresentada em (3.6) não tem distribuição limite normal. Ela é assimétrica à esquerda e a maior parte de sua massa está abaixo de zero, refletindo o fato de que os valores de ρ maiores que 1 poderiam gerar um particular conjunto de dados com probabilidade muito baixa.

Se um simples modelo AR (1) é usado quando de fato y_t segue um processo AR(p), então o termo de erro apresentado em (3.3) poderá ser autocorrelacionado para compensar a perda de especificação da estrutura do modelo dinâmico de y_t . Então, a presença de autocorrelação nos resíduos poderia invalidar o uso do teste de Dickey-Fuller, pois este é baseado na suposição de que e_t é um ruído branco. Contudo, Dickey e Fuller (1981) demonstraram que a distribuição limite e os valores críticos que eles obtiveram sob a suposição de que e_t é um ruído branco, também são válidos quando e_t apresenta-se autocorrelacionado. Assim, supondo-se que os dados são gerados de acordo com a equação (3.3) com $\rho = 1$ e que e_t é um processo estacionário autorregressivo de ordem p :

$$e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \dots + \phi_p e_{t-p} + u_t \quad (3.7)$$

onde u_t , agora, define um processo IID (Independente e identicamente distribuído).

Considere a versão reparametrizada de (3.3) mostrada abaixo,

$$\Delta y_t = \delta y_{t-1} + e_t \quad (3.8)$$

tal que $\delta = \rho - 1$ e $H_0: \delta = 0$ está sendo testada contra $H_1: \delta < 0$. Dada a equação para e_t em (3.7), (3.8) pode ser escrita como:

$$\Delta y_t = \delta y_{t-1} + \phi_1 e_{t-1} + \phi_2 e_{t-2} + \dots + \phi_p e_{t-p} + u_t \quad (3.9)$$

que pode ser escrita, desde que (3.3) com $\rho = 1$ fornece $e_t = \Delta y_{t-1} = y_t - y_{t-1}$, como:

$$\Delta y_t = \delta y_{t-1} + \phi_1 \Delta y_{t-1} + \phi_2 \Delta y_{t-2} + \dots + \phi_p \Delta y_{t-p} + u_t$$

implicando em uma regressão de Δy_t sobre $y_{t-1}, \Delta y_{t-1}, \Delta y_{t-2}, \dots, \Delta y_{t-p}$. Dickey e Fuller (1981) demonstraram que a estatística- t para $\delta = 0$ ($\rho = 1$ se y_t é a variável dependente), proveniente desta regressão, isto é, a estatística $ADF(p)$, tem a mesma distribuição limite não normal da estatística apresentada em (3.6) se $\rho = 1$. Isto implica que os valores críticos para um teste baseado na estatística $ADF(p)$ são idênticos àqueles associados à estatística do teste DF.

A estatística ADF é dada por $ADF = (C_{11} \hat{\sigma}^2)^{-1/2} (\hat{\rho} - 1)$, onde C_{11} é o primeiro elemento de R_0^{-1} , que é a matriz invertida de soma de quadrados e produtos cruzados da regressão, e $\hat{\sigma}^2$ é o estimador da variância do erro da regressão apresentada acima. Os valores críticos para o teste ADF podem ser encontrados em Dickey e Fuller (1981).

Said e Dickey (1984) fornecem uma generalização deste resultado. Além disto, eles demonstram que o procedimento ADF é válido assintoticamente, contanto que o valor de p (isto é, o número de *lags* de Δy_t introduzidos como regressores), usado na estimação, aumente com o tamanho da amostra n em uma razão controlada ($n^{1/3}$).

É bem provável o desconhecimento do valor correto de p na utilização da regressão ADF. A escolha do número de *lags* (p) deve ser feita com base nos dados avaliados. Na prática usa-se incluir tantos termos na variável dependente defasada, quanto é necessário para que se encontre um ruído branco residual.

Perron (1997) sugere um procedimento denominado *t-sig* para que a escolha do número de *lags* seja adequada. Este procedimento recursivo é baseado na estatística- t do coeficiente associado ao último *lag* da auto-regressão estimada. Mais especificamente, este procedimento seleciona um valor k , tal que o coeficiente do último *lag* da auto-regressão de ordem k é significativa, e o coeficiente do último *lag* da auto-regressão de ordem maior que k não é significativa.

Outros procedimentos muito utilizados na escolha do número de *lags* (p) são os critérios de informação de Akaike (AIC) e Schwarz (BIC) definidos abaixo, em combinação com o valor do coeficiente de determinação (R^2). Uma alternativa seria a fórmula $l_{12} = \text{int}\{12(n/100)^{1/4}\}$ encontrada por Schwert (1989) que permite que a ordem da auto-regressão aumente com o tamanho da amostra n .

$$AIC = \ln \hat{\sigma}^2 + \left(\frac{2}{n}\right) (\text{número de parâmetros})$$

$$BIC = \ln \hat{\sigma}^2 + \left(\frac{\ln n}{n} \right) (\text{número de parâmetros})$$

onde $\hat{\sigma}^2$ é a soma dos quadrados dos resíduos do processo auto-regressivo de ordem p e n refere-se ao número de observações da amostra.

3.2.2 Teste de raiz unitária usando a Distribuição Assintótica do GPH

Uma maneira simples e bastante utilizada no contexto de longa dependência para se testar a presença de raiz unitária é baseado no estimador GPH do parâmetro d fracionário de processos ARFIMA (p, d, q) . Considere uma série temporal y_t . Para averiguar se esta possui raiz unitária, ou seja, testar as hipóteses $H_0: y_t \sim I(1)$ contra $H_1: Y_t \sim I(d)$, $d < 1$, baseando-se na distribuição assintótica de \hat{d}_{GPH} , a estatística de teste utilizada é dada por:

$$t = \frac{\hat{d}_{GPH} - 1}{\sigma_{\hat{d}_{GPH}}}. \quad (3.10)$$

Os valores críticos para o percentil de ordem 5% para a estatística de teste acima podem ser encontrados em Santander *et al.* (2003).

3.2.3 Testes de raiz unitária baseado na técnica *bootstrap*

O *bootstrap* é uma técnica introduzida por Efron (1979) baseada numa reamostragem que permite aproximar a distribuição de uma função das observações pela distribuição empírica dos dados, com base em uma amostra finita.

Existem dois tipos distintos de *bootstrap*, o paramétrico e o não paramétrico. Neste estudo será utilizado o *bootstrap* não paramétrico, onde a reamostragem é feita a partir da amostra original e supõe-se que as observações são obtidas da função de distribuição empírica que designa uma massa de probabilidade igual a $\frac{1}{n}$ para cada um dos pontos da amostra.

Em séries temporais, a aplicação desta técnica requer vários cuidados, devido à não independência das observações. Neste caso a série deve ser reamostrada “indiretamente”, pois caso contrário sua estrutura original pode ser perdida nas séries reamostradas. Maiores detalhes a respeito de aplicação de *bootstrap* em séries temporais podem ser encontradas em Efron e Tibshirani (1993) e Franco e Reisen (2004 e 2007).

No caso deste trabalho, será utilizado o *bootstrap* não paramétrico nos resíduos do modelo ARFIMA ajustado.

Considere y_t uma série de n observações modelada por um ARFIMA(p,d,q),

$$\phi_p(B)(1-B)^d y_t = \theta_q(B)e_t, \quad d \in (-1.0,1.0).$$

Após estimar os parâmetros ϕ_p, θ_q e d , os resíduos amostrais são obtidos através de

$$\hat{e}_t = \hat{\theta}_q^{-1}(B)\hat{\phi}_p(B)(1-B)^{\hat{d}} y_t, \quad (3.11)$$

que são supostos *IID*.

É conveniente centralizar os resíduos, pois eles tendem a ser menores que os erros verdadeiros (Franco, 1998):

$$\tilde{e}_t = \left(\hat{e}_t - \frac{\sum_{t=1}^n \hat{e}_t}{n} \right). \quad (3.12)$$

Em seguida reamostra-se \tilde{e}_t com reposição, atribuindo a cada $\tilde{e}_t, t=1,2,\dots,n$, uma massa de probabilidade igual a $\frac{1}{n}$, obtendo assim os resíduos *bootstrap* e_t^* .

Então é possível construir recursivamente a série *bootstrap* y_t^* a partir de e_t^* da seguinte maneira:

$$y_t^* = \hat{\theta}_q(B)\hat{\phi}_p^{-1}(B)(1-B)^{-\hat{d}} e_t^*.$$

Note que $\hat{d}, \hat{\phi}_p$ e $\hat{\theta}_q$ são mantidos fixos durante as replicações *bootstrap*, ou seja, eles são tratados como parâmetros populacionais verdadeiros. Esta é considerada uma abordagem não-paramétrica do método de reamostragem *bootstrap* por não utilizar nenhuma suposição sobre a distribuição dos resíduos para reamostrá-los.

Em testes de hipótese, a técnica *bootstrap* pode ser aplicada tanto no cômputo de valores críticos para a estatística de teste quanto para calcular a probabilidade de significância, ou valor-p do teste.

Aqui, a técnica *bootstrap* será empregada para obter a distribuição empírica de duas estatísticas de teste utilizadas para testar a hipótese nula de existência de uma raiz unitária. A primeira estatística foi apresentada originalmente no artigo de Franco *et al.* (2006) e é baseada no estimador GPH para o parâmetro d que será obtido em cada replicação *bootstrap*, \hat{d}_{GPH}^* .

A segunda estatística de teste, proposta neste trabalho, também é baseada no estimador GPH, e é dada por:

$$t^* = \frac{\hat{d}_{GPH}^* - 1}{\sigma_{\hat{d}_{GPH}^*}} \quad (3.13)$$

onde $\sigma_{\hat{d}_{GPH}^*}$ é o erro padrão de \hat{d}_{GPH}^* estimado na regressão.

Neste estudo, o primeiro teste *bootstrap* será denominado BOOTd e o segundo BOOTReg. Para facilitar o entendimento da técnica, abaixo são descritos os passos do teste de uma maneira genérica:

- 1º) Para uma dada série y são reamostradas B amostras, $y_1^*, y_2^*, \dots, y_B^*$, independentes. Lembrando que no contexto de séries temporais, a reamostragem é feita nos resíduos e, a partir deles, obtém-se as amostras $y_1^*, y_2^*, \dots, y_B^*$, conforme descrito anteriormente.
- 2º) Para cada reamostragem *bootstrap* o valor da estatística de teste é calculado. Sendo que neste estudo a estatística de teste pode ser qualquer uma das duas apresentadas acima.
- 3º) As B 's estatísticas de teste são ordenadas e é calculado o percentil 5% que é o valor crítico do teste *bootstrap*.
- 4º) Compara-se o valor da estatística de teste da amostra com o valor crítico.
- 5º) Conclusão: se o valor da estatística de teste for menor do que o valor crítico a hipótese nula é rejeitada, ou seja, a série não apresenta raiz unitária.

Maiores detalhes sobre testes *bootstrap* podem ser encontrados em Davison e Hinkley (1997).

4. COINTEGRAÇÃO

A idéia básica da cointegração é que cada um dos processos considerados deve ser não estacionário, mas alguma combinação linear dos mesmos produz um processo estacionário. Harris, em seu artigo de 1995, apresenta a seguinte interpretação econômica para a cointegração: se duas (ou mais) séries possuem uma relação de equilíbrio de longo prazo, então mesmo que as séries possam conter tendências estocásticas (isto é, serem não estacionárias), elas irão mover-se juntas no tempo e a diferença entre elas será estável (isto é, estacionária). Em suma, o conceito de cointegração indica a existência de um equilíbrio de longo prazo, para o qual o sistema econômico converge no tempo.

Engle e Granger (1987) introduziram uma definição formal de cointegração: seja Y_t o vetor de séries temporais. Os componentes deste vetor Y_t são ditos cointegrados de ordem d , b , denotado por $Y_t \sim CI(d, b)$, se:

- (i) todos os componentes de $Y_t; y_{1t}, \dots, y_{kt}$ são $I(d)$, onde k é o número de séries de Y_t ;
- (ii) existe um vetor $\beta (\neq 0)$ tal que $e_t = \beta' Y_t \sim I(d-b)$, $0 < b \leq d$. O vetor β é denominado vetor cointegrante.

Interpretando $E(\beta' Y_t) = 0$ como o equilíbrio de longo prazo, a cointegração implica que o vetor de desvios do equilíbrio, e_t , geralmente chamado de erro cointegrante, é estacionário com variância finita, mesmo que as séries sejam não estacionárias e com variância infinita (maiores detalhes podem ser encontrados em Engle e Granger, 1987). Portanto, considere um exemplo com duas séries, y_t e x_t , cointegradas de ordem $CI(1,1)$, então estas são $I(1)$ e $e_t \sim I(0)$. Assim para estimar a relação de equilíbrio de longo prazo entre y_t e x_t , é necessário apenas estimar o modelo estático:

$$y_t = \beta x_t + e_t \quad (4.1)$$

Uma estimativa consistente dessa relação pode ser obtida utilizando-se o método de mínimos quadrados ordinários (MQO) e todas as questões dinâmicas e de endogeneidade podem ser ignoradas assintoticamente. Isto ocorre devido à propriedade de “super consistência” (Harris, 1995) do estimador MQO, quando as séries são cointegradas.

A definição de cointegração de Engle e Granger (1987), dada anteriormente, pode ser estendida para o caso de cointegração fracionária. Neste caso, os componentes do vetor Y_t são

fracionariamente cointegrados se o erro de cointegração é dito fracionalmente integrado e de ordem menor que as séries que compõe o vetor Y_t .

Definição 4.1 Seja Y_t um vetor composto de k séries Y_{it} , tal que $Y_{it} \sim I(d_i)$, $d_i > 0$, $i = 1, \dots, k$, $t = 1, \dots, n$. Diz-se que Y_t é fracionariamente cointegrado, e denota-se $Y_t \sim FCI(d_1, \dots, d_k; d_u)$, se existe um vetor $\beta \neq 0$ tal que $u_t = \beta Y_t \sim I(d_u)$, onde $0 \leq d_u < \min_{1 \leq i \leq k} d_i$. (Dittmann, 2000)

Esta definição é válida se e somente se $d_i = d_j$ para algum $i \neq j$. Além disso, uma condição necessária para β ser vetor de cointegração é que seu i -ésimo componente seja zero, se $d_i > d_j$ para todo $i \neq j$. No caso em que $d_1 = d_2 = \dots = d_k = d$, é usual escrever $Y_t \sim CI(d, b)$, onde $b = d_j - d_u$ mede a força da relação de cointegração. É importante destacar que a possibilidade de d ou d_u ser um valor fracionário foi mencionada no artigo original de Engle e Granger (1987), entretanto o artigo focalizou o estudo de séries $C(I, I)$.

O caso relevante neste estudo de cointegração fracionário é quando a hipótese de que os componentes do vetor Y_t são $I(1)$, mas o termo de correção de erro ou erro cointegrante é $I(1-b)$, com $b > 0$, ou seja, a ordem de integração pode assumir um valor não inteiro. Neste contexto, o termo de correção de erro é reversível na média, embora, não necessariamente estacionário em covariância, uma vez que um choque no sistema persiste, mas eventualmente desaparece. Existe um equilíbrio de longo prazo entre as variáveis do sistema mesmo quando ajustes para o equilíbrio podem levar um longo tempo para se completar.

Para que uma série temporal seja reversível na média, não necessariamente ela precisa ser integrada de ordem zero, mas sim de ordem menor ou igual a um. Se este for o caso os testes convencionais de raízes unitárias podem levar a conclusões equivocadas porque, como demonstrado por Diebold e Rudebusch (1991) e Hassler e Wolters (1994), eles possuem menor poder quando as séries são reversíveis na média, mas não são $I(0)$. Por esta razão uma maior flexibilidade pode ser encontrada na análise de cointegração fracionária.

Assim o objetivo principal deste capítulo é apresentar os testes de cointegração a serem utilizados neste estudo. Serão considerados para efeito de comparação, dois testes clássicos, a saber, o teste de Engle Granger e o Procedimento de Johansen e um teste fracionário baseado na distribuição assintótica do estimador GPH do parâmetro d fracionário de modelos ARFIMA. Além destas metodologias bastante difundidas em literaturas econométricas, serão propostos dois testes baseados na técnica de reamostragem *bootstrap* e

que também utilizam a teoria fracionária a partir do estimador GPH, como será visto ao longo deste capítulo.

4.1 TESTES DE COINTEGRAÇÃO

Atualmente há diversos testes que verificam se duas ou mais séries são cointegradas. Existem testes cuja hipótese nula é a de inexistência de cointegração, entretanto há outros que definem como hipótese nula a existência de cointegração. No contexto clássico, têm-se três testes que mais se destacam na prática. O primeiro é o teste de Engle-Granger, proposto por Engle e Granger (1987) e que constitui um dos testes mais utilizados devido a sua facilidade de aplicação. O segundo, o teste de Phillips-Ouliaris, foi originalmente apresentado no artigo de Phillips e Ouliaris (1990). O terceiro e que surgiu mais recentemente é conhecido como procedimento de Johansen, desenvolvido por Johansen e Juselius (1990) e que passou a ser amplamente utilizado com o aperfeiçoamento de diversos *softwares*. Neste trabalho serão considerados os testes de Engle e Granger e o procedimento de Johansen. Além destes procedimentos clássicos, também serão utilizadas outras duas abordagens fracionárias. A primeira baseada na distribuição assintótica do estimador GPH e a segunda que também utiliza o estimador GPH, entretanto faz uso da técnica de reamostragem *bootstrap* para estimar a distribuição empírica de duas estatísticas de testes propostas. Todos os testes utilizados neste estudo consideram como hipótese nula a inexistência de cointegração entre as séries testadas.

4.1.1 Método de Engle Granger (EG)

Engle e Granger (1987) recomendam o uso do teste ADF do seguinte tipo:

$$\Delta \hat{e}_t = (\rho - 1)\hat{e}_{t-1} + \alpha + \beta t + \sum_{i=1}^{p-1} \theta_i \hat{e}_{t-i} + u_t \quad (4.2)$$

onde $u_t \sim IID(0, \sigma^2)$, ρ é o parâmetro de um processo AR(1) e os resíduos \hat{e}_t são obtidos através da equação de regressão dada em (4.1). A questão da inclusão dos termos de tendência e/ou constante na equação de regressão depende se estes termos aparecem em (4.1). Isto é, componentes determinísticos podem ser adicionados ou em (4.1) ou em (4.2), mas não em ambos. Entretanto, Hansen (1992) mostrou com base em experimentos de Monte Carlo, que independentemente de \hat{e}_t ter ou não uma tendência determinística, incluir uma tendência temporal em (4.2) resulta em uma perda de poder. Assim, é geralmente improvável que \hat{e}_t

obtido da estimação de (4.1) tenha média zero e, dado o resultado de Hansen, esta forma de testar a cointegração deveria ser baseada em (4.1) e (4.2) com B igual a zero.

Assim como no teste de raiz unitária, a hipótese nula de uma raiz unitária e, no caso, de não cointegração ($H_0: \rho - 1 = 0$), é baseada no teste- t com uma distribuição não normal. Entretanto, a menos que β já seja conhecido (e não estimado usando a equação (4.1)), não é possível utilizar as tabelas de valores críticos de Dickey-Fuller. Há duas razões para isto: primeiro, porque a maneira como o estimador de MQO é construído, fazendo com que os resíduos tenham a menor variância amostral, mesmo que as variáveis não sejam cointegradas, fazem e_t parecer o mais estacionário possível. Então, a distribuição Dickey-Fuller poderia tender a super-rejeitar a hipótese nula. Segundo, a distribuição da estatística de teste sob a hipótese nula é afetada pelo número de variáveis regressoras, incluídas em (4.1). Portanto, diferentes valores críticos são necessários, quando o número de variáveis regressoras muda. Como os valores críticos também mudam, dependendo da presença ou não de constante e/ou tendência na equação (4.2) e com o tamanho da amostra, há um grande número de permutações, cada uma com um conjunto de valores críticos diferentes, para testar a hipótese nula.

Mackinnon (1991) construiu uma tabela com os valores críticos apropriados para particulares testes, levando em consideração o número de variáveis regressoras e a presença ou ausência de constante e tendência.

4.1.2 Procedimento de Johansen

Quando se tem um modelo com mais de duas variáveis, pode ser que exista mais de um vetor cointegrante. Isto é, as k variáveis em um modelo podem se caracterizar como parte de diversas relações de equilíbrio que governam a evolução comum das variáveis. É possível existir até $k-1$ vetores cointegrantes linearmente independentes e somente quando $k=2$ é possível mostrar que o vetor cointegrante é único.

Assumir que há somente um vetor cointegrante, quando de fato existe mais de um, conduz à ineficiência no sentido que se pode somente obter uma combinação linear destes vetores ao estimar um modelo de equação única. Portanto, é útil estender o modelo com equação única para uma estrutura multivariada.

Johansen (1988) propôs um procedimento a partir do uso do método de máxima verossimilhança que é baseado na suposição de normalidade multivariada, embora os

argumentos assintóticos não dependam desta suposição. Johansen trabalha com o modelo de correção de erro (MCE) diretamente e adota uma estrutura que é baseada na suposição de que, introduzindo *lags* suficientes, isto irá permitir que o termo de perturbação seja bem comportado.

O modelo de correção de erro (MCE) se origina do vetor autorregressivo (VAR) que é dado por:

$$Y_t = \sum_{i=1}^p A_i Y_{t-i} + e_t \quad (4.3)$$

A partir do VAR chega-se ao MCE para um contexto multivariado:

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + e_t \quad (4.4)$$

sendo que:

Y_t = vetor de variáveis não estacionárias

A_i = matriz de parâmetros de ordem ($k \times k$), onde k é o número de variáveis no modelo.

e_t = vetor de resíduos

$$\Gamma_i = - \sum_{j=i+1}^p A_j \text{ (matriz dos coeficientes de curto-prazo)}$$

$$\Pi = \sum_{i=1}^p A_i - I \text{ (matriz dos coeficientes de longo-prazo)}$$

p = ordem do vetor autorregressivo (VAR).

A matriz $\Pi_{n \times n}$ contém as informações de longo prazo correspondentes a $\Pi = \alpha\beta'$ em que α representa o ajustamento do desequilíbrio, enquanto β constitui-se em uma matriz dos coeficientes de longo prazo.

No modelo descrito na equação do MCE para um contexto multivariado, o termo Π possui um papel fundamental, uma vez que contém as informações de longo prazo e de realimentação ou de ajuste de desequilíbrio do modelo. Dessa forma, esse método consiste em testar se os coeficientes da matriz Π contém as informações de longo prazo sobre as variáveis envolvidas. Existem três casos possíveis, considerando o *rank* ou posto R dessa matriz.

Primeiro, se esta matriz for de posto completo ou posto(Π) = k , ou se existem k colunas linearmente independentes, as variáveis em Y_t serão $I(0)$ ou estacionárias. Segundo, se o posto dessa matriz for igual a zero ou posto(Π) = 0, então não existe nenhum vetor de cointegração. Estes dois casos não são particularmente interessantes. Terceiro, se posto(Π) = $r \leq k-1$ existem $k-1$ vetores de cointegração, ou seja, o posto de Π indica o número de relações que cointegram.

Este número pode ser obtido utilizando dois testes de razão de verossimilhança: o teste Traço e o de Máximo Autovalor. A hipótese nula do primeiro teste é de que o número de vetores de cointegração é $r \leq p$ (em que $p = 1, 2, \dots, k-1$) e a hipótese alternativa é de que $r = k$, uma hipótese mais genérica. A idéia básica do segundo teste é de verificar a significância do maior autovalor, confrontando a hipótese nula de que r vetores de cointegração são significativos contra a alternativa de que o número de vetores significativos seja $r+1$, ou seja, $r=0$ contra $r=1$; $r=1$ contra $r=2$ e assim por diante. Estes testes são dados respectivamente por:

$$\lambda_{trace} = -n \sum_{i=r+1}^k \ln(1 - \hat{\lambda}_i) \quad r=0, 1, 2, \dots, k-1 \quad (4.5)$$

$$\lambda_{max} = -n \ln(1 - \hat{\lambda}_{r+1}) \quad (4.6)$$

onde $\hat{\lambda}$ é o autovalor associado a matriz Π .

Conforme Harris (1995), não é incomum os resultados desses dois testes divergirem, não indicando o mesmo número de vetores de cointegração, o que pode ser uma consequência de amostras pequenas. Além disso, quando estes testes divergirem, sugere-se usar o teste do Máximo Autovalor.

4.1.3 Teste de cointegração usando a distribuição assintótica do GPH

Seja Y_t o vetor de séries temporais. Inicialmente, deve-se verificar se cada série deste vetor Y_t possui raiz unitária. Em seguida, procede-se de maneira semelhante ao aplicado no teste de Engle e Granger, ou seja, ajusta-se um modelo de regressão entre as séries em estudo e extrai-se o resíduo estimado. De posse destes resíduos $\hat{\varepsilon}_t$ basta estimar o parâmetro d fracionário, por meio do estimador GPH e aplicar a mesma estatística baseada na distribuição assintótica de \hat{d}_{GPH} apresentada em (3.14). Neste caso as hipóteses do teste são: $H_0: \hat{\varepsilon}_t \sim I(1)$ (não são cointegradas) contra $H_1: \hat{\varepsilon}_t \sim I(1-b)$, $0 < b \leq 1$ (existe cointegração).

Percentis da ordem de 5% para a estatística do teste de cointegração baseado no estimador GPH podem ser encontrados em Santander et al. (2003).

4.1.4 Testes de cointegração baseados na técnica *bootstrap*

As estatísticas de teste apresentadas na Seção 3.1.4 propostas para detectar a presença de raiz unitária em uma série temporal também foram consideradas na avaliação de cointegração entre duas séries. Assim como no teste de Engle-Granger, aqui também é necessário aplicar o método de mínimos quadrados para propor um modelo de regressão linear entre as duas séries que se deseja testar cointegração. Em seguida é aplicado aos resíduos estimados da regressão o procedimento *bootstrap* descrito na Seção 3.1.4, com intuito de averiguar se estes apresentam raiz unitária. Abaixo são apresentados os passos para testar a existência de cointegração entre duas séries, fazendo uso da técnica de reamostragem:

2º) Estima-se o modelo de regressão linear simples para as duas séries em que se deseja testar a cointegração e extrai-se o vetor de resíduos estimados.

3º) A partir dos resíduos estimados são geradas B amostras *bootstraps* independentes.

4º) Para cada reamostragem é computado o valor da estatística de teste, que pode ser uma das estatísticas apresentadas na Seção 3.2.3.

5º) As B's estatísticas de teste são ordenadas e é calculado o percentil 5% que será o valor crítico do teste *bootstrap*.

6º) Compara-se o valor da estatística de teste da amostra com o valor crítico apontado pelo método *bootstrap*.

7º) Conclusão: se o valor da estatística de teste for menor do que o valor crítico a hipótese nula é rejeitada, ou seja, o resíduo é estacionário e, portanto, as séries são cointegradas.

5. RESULTADOS DE SIMULAÇÃO

Este capítulo traz os resultados de simulações para os testes de raiz unitária e cointegração apresentados neste estudo. Todos os testes serão avaliados segundo seu tamanho e poder, especificados da seguinte maneira:

Tamanho do teste: $\alpha = P(\text{rejeitar } H_0 / H_0 \text{ verdadeira})$, sendo assim no contexto de raiz unitária serão geradas séries com raiz unitária ($d=1$) e na cointegração serão geradas séries não cointegradas. O valor nominal de α será fixado em 5% para todos os testes.

Poder do teste: $1 - \beta = P(\text{rejeitar } H_0 / H_0 \text{ falsa})$, neste caso para avaliar a presença de raiz unitária serão geradas séries com $d < 1$ e para cointegração serão geradas séries fracionalmente cointegradas.

As *performances* dos testes de raiz unitária e cointegração foram investigadas através de simulações Monte Carlo (MC), considerando-se 1000 replicações Monte Carlo de tamanhos 100, 300 e 500 e 1000 reamostragens *bootstraps*. No caso dos testes de raiz unitária foram geradas séries a partir do modelo ARFIMA (0, d ,0) para $d \in (0.1, 1.0)$ e com $g(n)=n^{0.7}$. Para a cointegração também foram utilizados processos ARFIMA (0, d ,0) na geração das séries e_t e x_t e com $g(n)=n^{0.7}$. Na geração dos resíduos e_t considerou-se valores de $d \in (0.1, 1.0)$, enquanto para x_t o valor de d foi mantido fixo e igual a 1. A geração dos resíduos dos modelos ARFIMA (0, d ,0) foi feita utilizando a normal padrão. Os programas para as simulações foram desenvolvidos em R versão 2.6.2 e podem ser conferidos nos Anexos 1 e 2. O tempo de processamento gasto nas simulações depende do tamanho das séries, do número de replicações Monte Carlo, do número de reamostragens *bootstrap* e do computador utilizado. No caso de um Pentium 4 3GHz, 448 MB RAM e sistema operacional Windows XP Professional versão 2002 e considerando 1000 Monte Carlo e 1000 *bootstrap*, para séries de tamanho 100 o tempo gasto é de aproximadamente 5 horas, para 300 são necessárias 13 horas, já para 500, o tempo aumenta para 93 horas. Vale lembrar que o tempo gasto para simular os testes de raiz unitária é praticamente o mesmo gasto na cointegração.

5.1 PODER E TAMANHO DOS TESTES DE RAIZ UNITÁRIA

O objetivo desta seção é comparar o desempenho dos testes: Dickey-Fuller Aumentado (ADF), teste assintótico para o estimador GPH (GPH), teste baseado na técnica *bootstrap* e que utiliza o estimador do parâmetro d como estatística de teste (BOOTd) e teste

baseado na técnica *bootstrap* que também utiliza o estimador de d por meio da estatística de teste apresentada em 3.14 (**BOOTReg**). Esta comparação será feita sob dois aspectos: Tamanho e Poder dos testes.

A hipótese a ser testada é $H_0: y_t \sim I(1)$, ou seja, os componentes do vetor apresentam raiz unitária, como visto no capítulo 3.

Nas Tabelas 5.1, 5.2 e 5.3 são apresentados os resultados referentes ao poder e ao tamanho dos testes de raiz unitária estudados para séries de tamanho 100, 300 e 500, respectivamente. As Figuras 5.1, 5.2 e 5.3 trazem as curvas de poder dos quatro testes, para os três tamanhos de séries considerados anteriormente.

Na Tabela 5.1 e Figura 5.1 pode-se observar que para valores de d próximos de zero, todos os testes têm alto poder (próximo de 100%). Para $d \in (0,3;0,7)$ nota-se que o teste ADF apresenta poder consideravelmente menor que os demais, já os testes assintótico (GPH) e o *bootstrap* BOOTd são os que mais se destacaram com relação ao poder nesta faixa. Para d igual a 0,8 e 0,9 observa-se que o teste ADF melhorou em relação aos demais testes, chegando até a superá-los para $d = 0,9$. Quando $d = 1,0$ (tamanho do teste) verifica-se que as melhores *performances* ocorrem para os testes GPH e BOOTd que apresentaram tamanho mais próximo de 5% e o pior foi o teste BOOTReg. É interessante observar que os testes que apresentaram maior poder, também foram os que se destacaram com relação ao tamanho.

Tabela 5. 1 - Resultados de tamanho e poder para os testes de raiz unitária para séries de tamanho 100

d	ADF	GPH	BOOTDReg	BOOTd
0,10	99,80	100,00	100,00	100,00
0,20	97,80	100,00	99,80	100,00
0,30	92,30	99,70	99,60	99,70
0,40	81,60	98,30	96,70	98,40
0,50	65,10	92,30	87,80	92,50
0,60	46,00	76,00	66,40	76,30
0,70	33,00	52,80	43,90	52,90
0,80	22,80	29,40	19,40	30,00
0,90	12,20	11,80	7,70	11,60
1,00	6,10	4,90	3,10	5,10

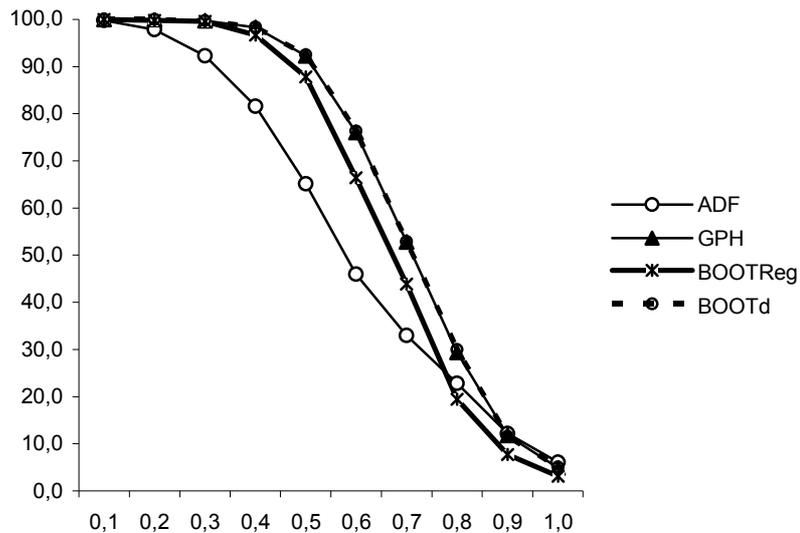


Figura 5. 1 Curva de Poder dos testes de raiz unitária para n=100

Na Tabela 5.2 e Figura 5.2 observa-se que para $d \in (0,1;0,6)$ todos os testes apresentaram poder igual ou muito próximo de 100%, exceto para o teste ADF quando $d = 0,5$ e $0,6$. É possível observar que o comportamento dos testes é relativamente parecido com o comportamento observado para séries de tamanho 100, entretanto, com o aumento da série houve um aumento considerável de poder quando consideramos $d \leq 0,8$. Aqui também se observou que os maiores poderes ocorreram para os testes GPH e BOOTd e os mais baixos foram para o ADF. Ainda é interessante observar que para $d = 0,9$ o teste ADF foi o que apresentou mais baixo poder, diferente do ocorrido quando consideramos séries de tamanho 100. Veja que o aumento da série contribuiu para aumentar de maneira significativa o poder dos testes que utilizam o estimador GPH, tanto o assintótico como os baseados na técnica *bootstrap*, quando $d = 0,9$. Para o tamanho do teste ($d = 1,0$) a melhor *performance* ocorreu para o teste GPH seguido do GPHd e a pior para o BOOTReg que tende a apresentar um tamanho bem menor do que o estipulado (5%).

Tabela 5. 2 - Resultados de tamanho e poder para os testes de raiz unitária para séries de tamanho 300

d	ADF	GPH	BOOTReg	BOOTd
0,1	100,00	100,00	100,00	100,00
0,2	100,00	100,00	100,00	100,00
0,3	100,00	100,00	100,00	100,00
0,4	99,50	100,00	100,00	100,00
0,5	91,60	99,90	99,80	99,90
0,6	69,20	99,00	98,60	99,00
0,7	48,50	87,40	83,70	88,20
0,8	27,70	56,60	49,40	57,10
0,9	12,8	19,8	15,2	21,3
1,0	5,90	5,30	3,60	5,40

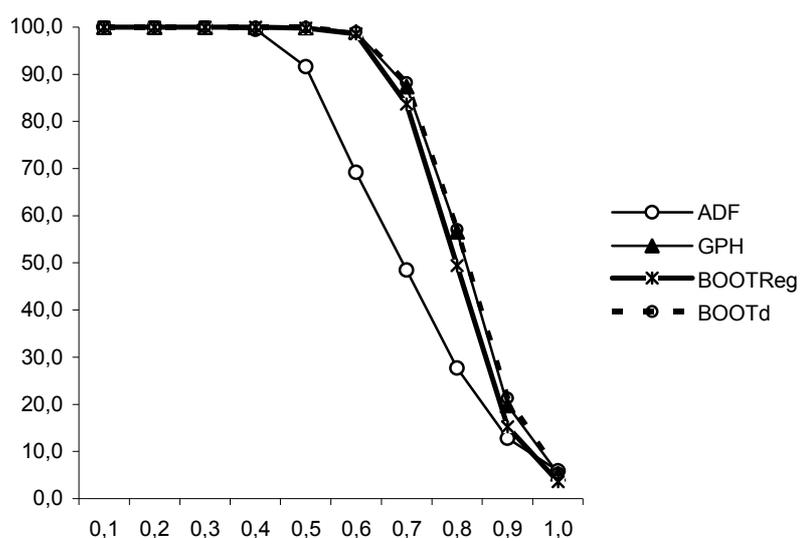


Figura 5. 2 Curva de Poder dos testes de raiz unitária para n=300

Os resultados de tamanho e poder para séries de tamanho 500 podem ser vistos na Tabela 5.3 e Figura 5.3. É possível notar que os resultados são bastante similares ao que foi observado para séries de tamanho 100 e 300. O teste clássico apresentou poder bastante inferior ao dos testes fracionários e com relação ao tamanho o teste GPH foi o que mais se destacou. Com o aumento do tamanho da série observa-se um melhoria considerável no poder e tamanho dos testes, exceto para o teste BOOTReg que apesar de ter um alto poder continua a apresentar um tamanho bem abaixo do valor nominal de 5%.

Tabela 5. 3 - Resultados de tamanho e poder para os testes de raiz unitária para séries de tamanho 500

d	ADF	GPH	BOOTReg	BOOTd
0,1	100,00	100,00	100,00	100,00
0,2	100,00	100,00	100,00	100,00
0,3	100,00	100,00	100,00	100,00
0,4	99,90	100,00	100,00	100,00
0,5	98,40	100,00	100,00	100,00
0,6	82,40	99,90	99,80	99,90
0,7	58,90	97,40	96,10	97,60
0,8	31,60	68,50	62,80	68,70
0,9	12,60	26,60	21,80	26,40
1,0	4,70	5,00	3,30	4,90

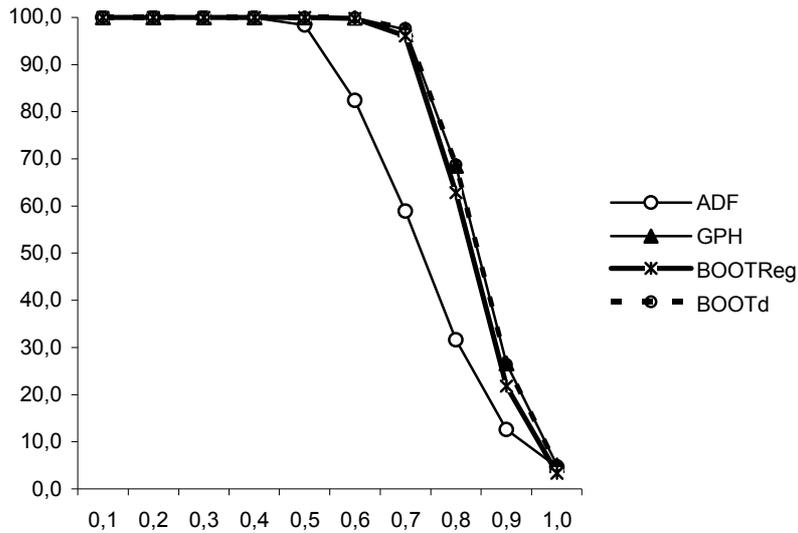


Figura 5.3 Curva de Poder dos testes de raiz unitária para n=500

5.2 PODER E TAMANHO DOS TESTES DE COINTEGRAÇÃO

Esta seção é dedicada à comparação dos testes de cointegração apresentados no Capítulo 4, a saber, teste de Engle e Granger (**EG**), Procedimento de Johansen (**Johansen**), teste assintótico para o estimador GPH (**GPH**), teste baseado na técnica *bootstrap* e que utiliza a distribuição empírica do d (**BOOTd**) e teste baseado na técnica *bootstrap* que utiliza a distribuição empírica mostrada em 3.14 (**BOOTReg**).

Como vimos anteriormente, para estudar o tamanho dos testes de cointegração foram simulados processos não cointegrados, já para o estudo de poder considerou-se processos fracionariamente cointegrados. Os passos para obtenção de poder e tamanho dos testes são descritos a seguir:

- Tamanho do teste – Processos não cointegrados
 1. Gerar $e_t \sim I(d)$, onde $d = 1$, utilizando uma $N(0,1)$;
 2. Gerar $x_t \sim I(d)$, onde $d = 1$, utilizando uma $N(0,1)$;
 3. Obter y_t considerando a relação $y_t = 2x_t + e_t$ (relação utilizada por Santander *et al.* (2003));
 4. No caso dos testes de Engle Granger, GPH e nos baseados na técnica *bootstrap*, deve-se ajustar um modelo de regressão linear entre x_t e y_t para obtenção dos resíduos estimados. No caso do procedimento de Johansen não é

necessário estimar o modelo de regressão, sendo o teste aplicado diretamente na matriz composta por x_t e y_t .

5. Aplicar os testes de raiz unitária aos resíduos do modelo de regressão ajustado com intuito de averiguar se $\hat{e}_t \sim I(d), d = 1$.

- Poder do teste – Processos cointegrados

O algoritmo para obtenção do Poder do teste é similar ao algoritmo do Tamanho do teste relacionado acima, exceto pelo primeiro passo. Para ele, este passo fica da seguinte forma, mantendo-se as demais etapas do procedimento:

1. Gerar $e_t \sim I(d)$, onde $d = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$, utilizando uma $N(0,1)$;

A Tabela 5.4 apresenta os resultados de tamanho e poder para os testes de cointegração considerando séries de tamanho 100. Os resultados para $n=300$ podem ser conferidos na Tabela 5.5 e para $n=500$ na Tabela 5.6. Já as curvas de poder para $n=100, 300$ e 500 , são apresentadas nas Figuras 5.4, 5.5 e 5.6, respectivamente.

Na Tabela 5.4 e Figura 5.4 observa-se que para valores de d entre 0,1 e 0,3, todos os testes apresentam alto poder (muito próximos de 100%), exceto para os testes EG (97,5) e o procedimento de Johansen (98,5) quando $d = 0.3$. Para $d \in (0,4;0,8)$, os testes baseados na técnica *bootstrap* e o assintótico apresentaram poderes mais elevados, já o EG foi o que apresentou menor poder nessa faixa. Ainda considerando $d \leq 0,8$, observa-se que o procedimento de Johansen teve melhor desempenho do que o EG, entretanto ficou com poder bem abaixo dos demais testes. Considerando $d = 0,9$, verifica-se que os testes BOOTd e EG foram os que tiveram maior poder, já o Johansen, foi o que apresentou pior performance (5,9) neste caso. Para o tamanho do teste, ou seja, quando $d=1,0$, verifica-se que as melhores *performances* ocorrem para os testes de Johansen e GPH que mais se aproximaram do nível nominal de 5%, enquanto a pior performance foi observada para o teste BOOTd com tamanho de 6,6. Na Figura 5.4 é possível notar que os testes *bootstraps* tiveram um comportamento semelhante ao do teste assintótico e de uma maneira geral, os testes fracionários apresentaram maior poder se comparado aos testes clássicos de Engle e Granger e Johansen.

Tabela 5. 4 - Resultados de tamanho e poder para os testes de cointegração para séries de tamanho 100

d	EG	GPH	Johansen	BOOTReg	BOOTd
0,1	100,00	100,00	100,00	100,00	100,00
0,2	99,90	100,00	99,60	100,00	100,00
0,3	97,50	99,80	98,50	99,80	99,90
0,4	85,80	97,10	89,20	96,00	98,10
0,5	64,80	90,10	70,20	86,90	92,30
0,6	40,70	68,50	41,50	64,20	74,40
0,7	24,90	41,70	25,00	37,00	48,60
0,8	14,40	21,30	14,00	18,20	27,10
0,9	9,50	8,80	5,90	7,70	11,30
1,0	5,60	4,70	5,10	4,10	6,60

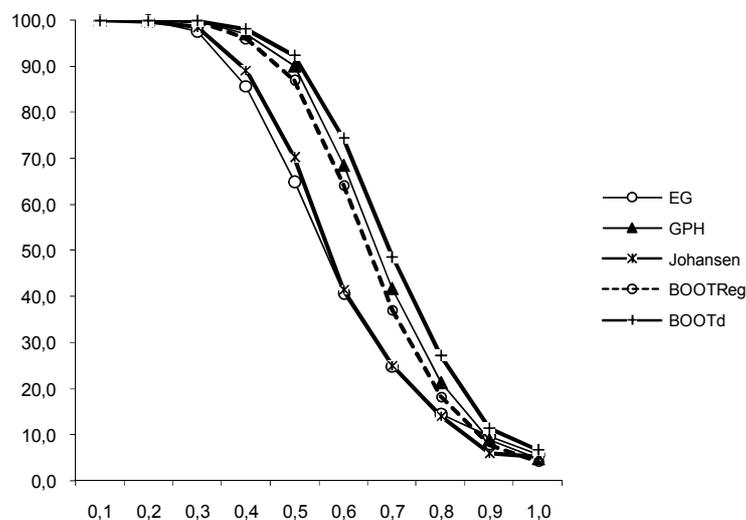


Figura 5. 4 Curva de Poder dos testes de cointegração para n=100

Quando $n=300$, pode ser visto na Tabela 5.5 e Figura 5.5 que, para valores de $d \leq 0,5$ há uma ótima *performance* (próximos de 100%) para todos os testes, exceto para o teste EG quando $d = 0,5$. Observe que para $d \leq 0,7$, os testes *bootstraps* e o assintótico apresentam poder muito elevado acima de 80%, sendo que o BOOTd é o que apresenta maior poder. Para $d=0,8$, o poder dos testes *bootstraps* e assintótica caem consideravelmente, mas ainda assim são maiores que os dos testes clássicos. O teste EG é o que apresenta menor poder se considerar $d \leq 0,8$. Para $d = 0,9$, o teste de Johansen foi o que apresentou pior desempenho, seguido do EG, um quadro um pouco diferente do que foi observado para tamanho 100, quando o teste EG teve um destaque com relação aos demais.

Na avaliação de tamanho do testes, ou seja, quando $d = 1,0$, o melhor desempenho ocorreu para o teste BOOTReg, seguido dos testes de Johansen e a pior *performance* ocorreu

para o teste BOOTd que passou a apresentar tamanho ainda maior do que o observado quando $n=100$.

Tabela 5.5 - Resultados de tamanho e poder para os testes de cointegração para séries de tamanho 300

d	EG	GPH	Johansen	BOOTReg	BOOTd
0,1	100,00	100,00	100,00	100,00	100,00
0,2	100,00	100,00	100,00	100,00	100,00
0,3	100,00	100,00	100,00	100,00	100,00
0,4	100,00	100,00	100,00	100,00	100,00
0,5	98,60	99,90	99,80	99,80	99,90
0,6	84,90	97,70	93,40	97,80	98,60
0,7	55,20	80,70	64,80	81,60	86,60
0,8	25,50	42,00	27,90	43,30	50,00
0,9	10,50	14,30	8,90	15,30	19,40
1,0	5,80	4,60	4,70	4,90	7,20

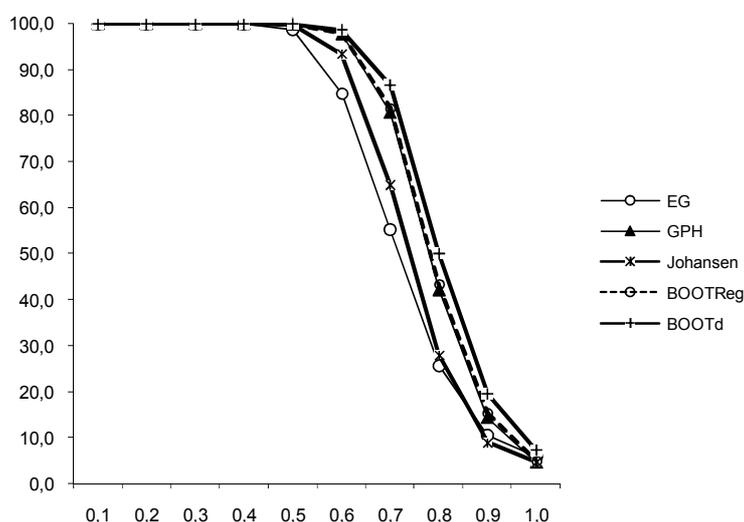


Figura 5.5 Curva de Poder dos testes de Cointegração para $n=300$

Os resultados para $n=500$ vistos na Tabela 5.6 e Figura 5.6 mostram que com o aumento da série houve um ganho de poder para todos os testes, sendo que os testes fracionários se mantiveram apresentando poder maior que os testes clássicos. Com relação ao tamanho, é possível observar que o teste EG foi o que mais se destacou, diferente do que vinha ocorrendo para n menores. O teste BOOTReg também apresentou tamanho bem próximo do nível nominal, já o outro bootstrap, o BOOTd, continua apresentando um tamanho muito alto.

Ainda é possível notar que assim como ocorreu nos testes de raiz unitária, o aumento do tamanho da amostra contribuiu de maneira significativa para aumentar o poder de todos os

testes, com destaque para os testes *bootstraps* e GPH que tiveram um poder acima de 90% quando $d \leq 0.7$.

Uma outra observação interessante é que o teste BOOTd havia se destacado bastante na avaliação de raiz unitária, não só em relação ao poder como também em relação ao tamanho, entretanto na avaliação de cointegração, ele apresentou um tamanho elevado e distante do nível nominal, o que pode ter contribuído para que este tenha apresentado poder mais elevado que os demais testes. Já o teste BOOTReg, apesar de ter apresentado um tamanho bem abaixo do nível nominal nos testes de raiz unitária, teve um ótimo desempenho na avaliação de cointegração.

Tabela 5.6 Resultados de tamanho e poder para os testes de cointegração para séries de tamanho 500

d	EG	GPH	Johansen	BOOTReg	BOOTd
0,1	100,00	100,00	100,00	100,00	100,00
0,2	100,00	100,00	100,00	100,00	100,00
0,3	100,00	100,00	100,00	100,00	100,00
0,4	100,00	100,00	100,00	100,00	100,00
0,5	99,80	100,00	100,00	100,00	100,00
0,6	94,80	99,70	99,10	99,70	99,80
0,7	70,10	93,40	82,10	93,00	95,10
0,8	37,20	59,10	38,50	60,50	66,40
0,9	15,70	20,00	12,60	22,00	27,60
1,0	5,00	4,50	5,30	5,10	6,70

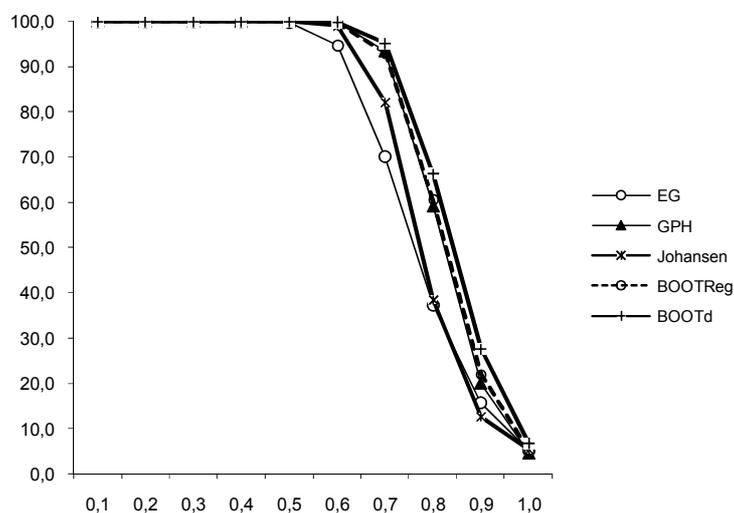


Figura 5. 6 Curva de Poder dos testes de Cointegração para n=500

Comparando as Figuras 5.4, 5.5 e 5.6, pode-se confirmar que com o aumento do tamanho amostral, houve uma melhoria na curva de poder dos testes aqui considerados.

6. APLICAÇÃO A DADOS REAIS

Neste capítulo é realizado um estudo de aplicação das técnicas de cointegração clássica e fracionária descritas neste trabalho. O objetivo é averiguar a existência de uma possível relação de longo prazo entre os mercados de ações do Brasil e Estados Unidos. Para isso, serão considerados os índices Ibovespa e Dow Jones que correspondem, respectivamente, às bolsas de valores de São Paulo e Nova York .

O Índice Bovespa¹ é o mais importante indicador do desempenho médio das cotações do mercado de ações brasileiro. Sua relevância advém do fato do Ibovespa retratar o comportamento dos principais papéis negociados na BOVESPA e também de sua tradição, pois o índice manteve a integridade de sua série histórica e não sofreu modificações metodológicas desde sua implementação em 1968.

O Índice Dow Jones² é um dos indicadores de maior confiabilidade mundial e tornou-se, desde a sua criação, uma referência importante para investidores e administradores de recursos estrangeiros, que se baseiam em sua *performance* para tomar decisões de investimentos.

A globalização e internacionalização do capital levam cada vez mais à integração dos mercados, com implicações importantes para a montagem de portfólios. Assim, conhecer o grau de relacionamento existente entre diversos mercados internacionais é de suma importância para os investidores. Sabe-se que, se dois mercados quaisquer forem integrados, o risco sistemático de investimento não pode ser eficazmente diminuídos por meio da diversificação (Gaió e Rolim, 2007).

O mercado financeiro tem sofrido várias transformações devido a essa tendência de globalização ocasionada principalmente pelo avanço tecnológico atual. Isso faz com que haja uma sensibilidade dos mercados às turbulências externas, ou seja, uma crise em um país economicamente ativo irá influenciar qualquer outra economia mundial.

Diante deste cenário, uma proposta interessante é averiguar se o mercado americano tem uma influência de longo prazo no mercado de capitais brasileiro. Para isso, serão utilizados os testes clássicos e fracionários mostrados neste trabalho, com o objetivo de verificar se as séries Ibovespa e Dow Jones são cointegradas.

¹ <http://www.bovespa.com.br>

² <http://www.cemig.com.br/choque/index.asp>

6.1 DESCRITIVA PARA AS SÉRIES IBOVESPA E DOW JONES

A Figura 6.1 apresenta os gráficos das séries históricas Ibovespa e Dow Jones no período de 11/03/1997 a 07/08/2008, totalizando quase três mil observações. Os dados foram retirados do site “<http://br.finance.yahoo.com>” e correspondem ao valor do fechamento diário das bolsas brasileira e americana.

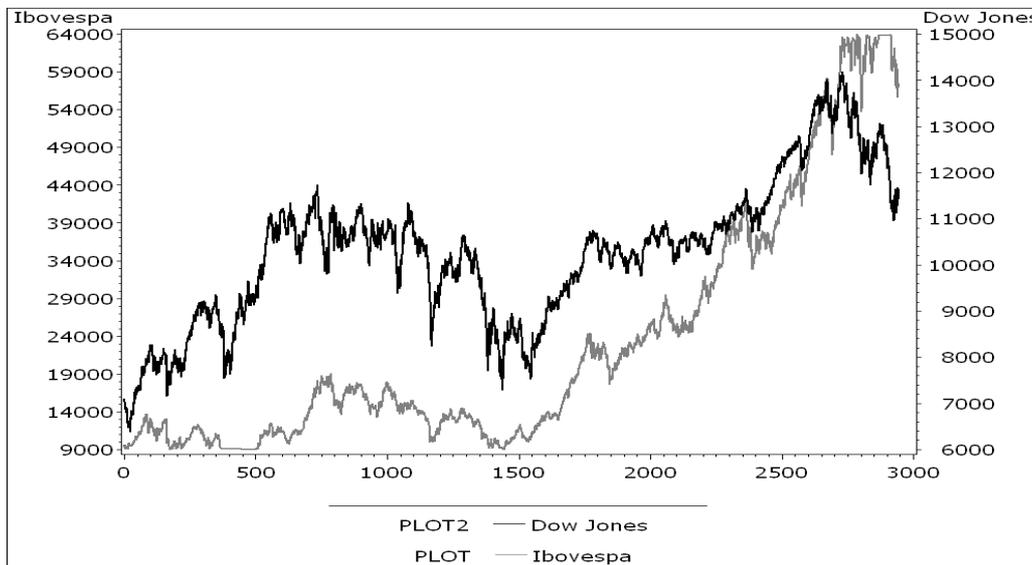


Figura 6. 1 Séries Ibovespa e Dow Jones

Observando a Figura 6.1, nota-se que o índice Ibovespa tem uma tendência crescente aparentemente não linear. Já o índice Dow Jones, apesar de apresentar uma tendência geral de crescimento, tem um comportamento bastante oscilatório. Nas primeiras 500 observações nota-se uma tendência de crescimento, nas próximas 500 há um comportamento oscilatório em torno de um mesmo intervalo. A partir da observação 1100 nota-se um movimento de queda seguido por outro de intensa alta, que não se mantém a partir da observação 2730, quando se começa a perceber uma tendência de decréscimo. O estimador GPH foi utilizado na estimação do parâmetro d das séries em estudo e obteve-se um $d=1,198$ para o índice Ibovespa e $d=1,036$ para o Dow Jones, o que indica uma provável existência de raiz unitária.

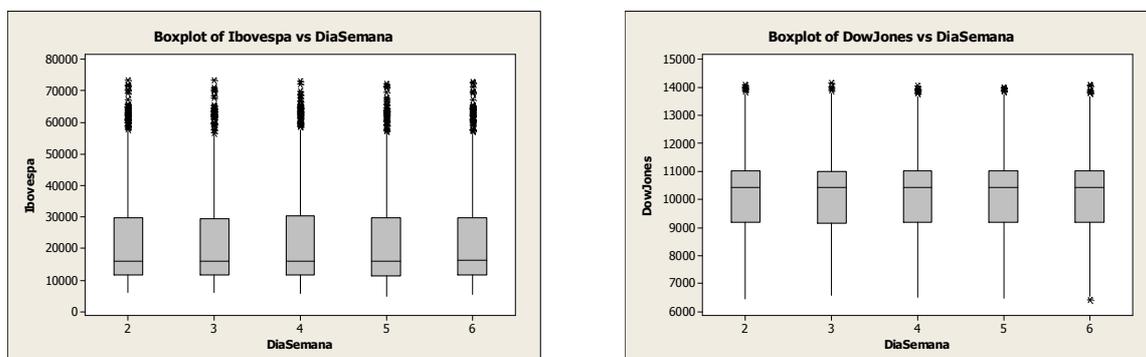


Figura 6. 2 Boxplot das séries Ibovespa e Dow Jones

Na Figura 6.2 pode ser visto que as séries em estudo não apresentam uma tendência de aumento ou queda na cotação em um dado dia da semana. Veja que o comportamento observado nos boxplots é relativamente parecido para qualquer dia da semana considerado.

A Figura 6.3 traz o gráfico de dispersão entre as duas séries em estudo. De uma maneira geral, é possível observar que há uma correlação positiva entre os índices, pois nota-se que quando um índice cresce o outro também tende a crescer.

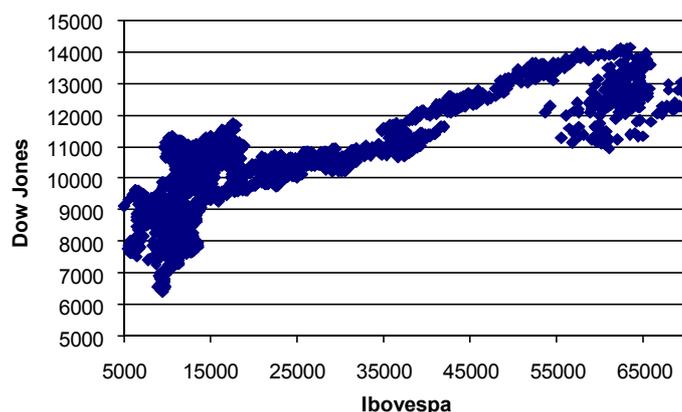


Figura 6. 3 Plot Ibovespa vs Dow Jones

6.1 VERIFICAÇÃO DA EXISTÊNCIA DE COINTEGRAÇÃO

O primeiro passo para avaliar se as séries são cointegradas é constatar a presença de raiz unitária em ambas. Para isto, serão aplicados os testes de raiz unitária vistos no Capítulo 3. Os testes foram realizados no R versão 2.6.2 e os resultados podem ser conferidos na Tabela 6.1.

Tabela 6. 1 Testes de Raiz Unitária para as séries

Séries	ADF	GPH	BOOTReg	BOOTd
Ibovespa	-1,606 (-3,411)	2,01 (-1,491)	2,01 (-1,789)	1,198 (0,847)
Dow Jones	-2,540 (-2,862)	0,366 (-1,491)	0,366 (-1,766)	1,036 (0,853)

Nota: Valores entre parênteses correspondem ao ponto crítico de 5%

Para a escolha do número de *lags* a serem considerados no teste ADF utilizou-se o procedimento automático baseado no critério de Informação de Akaike (AIC). Para a série Ibovespa foram considerados dois *lags*, intercepto e tendência; já para Dow Jones, o

intercepto e nove *lags* foram significativos. Todos os testes realizados mostram que não se pode rejeitar a hipótese nula de raiz unitária para nenhuma das séries consideradas no estudo.

Como as duas séries possuem a mesma ordem de integração, ou seja, ambas são $I(1)$, podemos verificar se as mesmas são cointegradas. Conforme mostrado na teoria, para que estas séries sejam cointegradas é necessário que o resíduo da regressão entre elas seja de ordem $I(d)$, com $d < 1$.

A verificação da suposição de cointegração também foi feita no R a partir dos testes considerados no Capítulo 4. A regressão linear ajustada foi feita no *software* Eviews 3.0, já os gráficos para averiguar as suposições de independência, normalidade e homogeneidade nos resíduos foram obtidos a partir do *software* Minitab 14.

Na Tabela 6.2 é apresentado o ajuste da regressão entre Ibovespa e Dow Jones. Verifica-se que tanto a constante quando o índice Dow Jones são significativos para o modelo de regressão ajustado. O R^2 mostra que o modelo sugerido está explicando cerca de 66% da variação ocorrida no índice Ibovespa. O teste de Fisher mostra que aparentemente o modelo é adequado. Entretanto, o valor da estatística de Durbin-Watson ($DW=0,0075$) é um indicativo de que há problemas de autocorrelação nos resíduos. Segundo Engle e Yoo (1987), a estatística DW não parece aconselhável para testar a existência de cointegração, mas valores próximos de zero são um bom indicativo de que não há cointegração. Maiores detalhes podem ser encontrados em Marques (1998).

Tabela 6. 2 Modelo de Regressão ajustado para as séries Ibovespa e Dow Jones

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-65167.54	1185.773	-54.95784	0.0000
DOWJONES	8.602982	0.114160	75.35924	0.0000
R-squared	0.658817	Mean dependent var		23205.93
Adjusted R-squared	0.658701	S.D. dependent var		16308.15
S.E. of regression	9527.345	Akaike info criterion		21.16240
Sum squared resid	2.67E+11	Schwarz criterion		21.16647
Log likelihood	-31138.47	F-statistic		5679.015
Durbin-Watson stat	0.007506	Prob(F-statistic)		0.000000

A Figura 6.3 apresenta os gráficos para os resíduos do modelo ajustado na Tabela 6.3. Pode-se constatar que as suposições de normalidade, independência e homocedasticidade não são verificadas, o que praticamente invalida o modelo ajustado anteriormente.

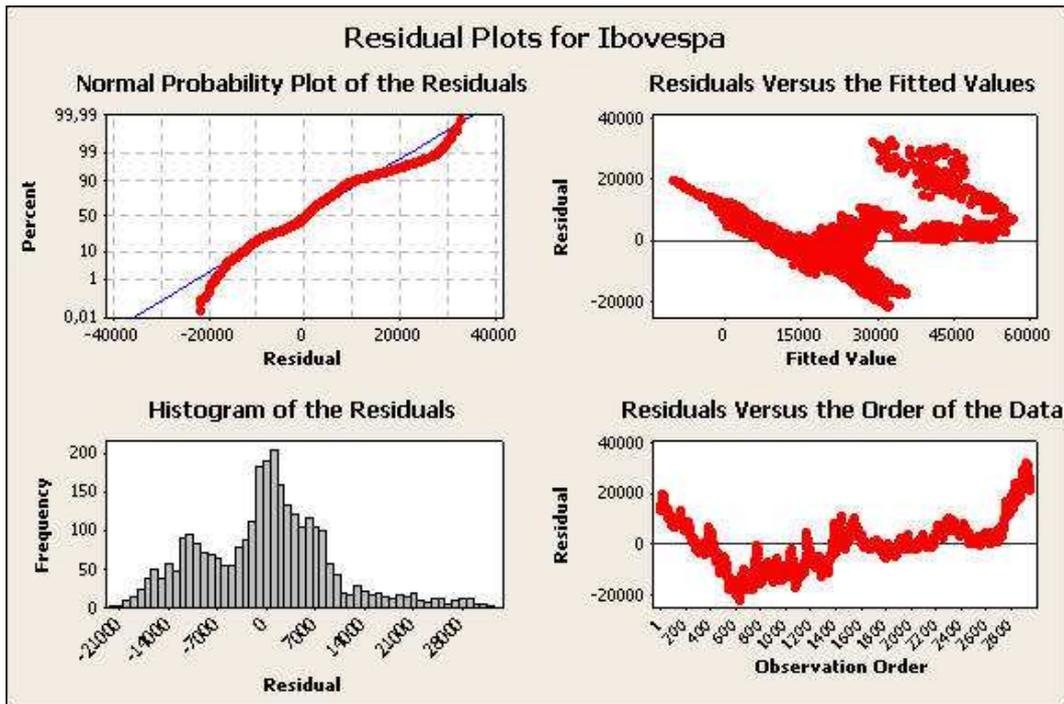


Figura 6. 4 Gráficos para os resíduos do modelo ajustado na Tabela 6.3

Foram aplicados os testes clássicos e fracionários mostrados no Capítulo 4 e os resultados podem ser conferidos na Tabela 6.3. Para o teste EG, o procedimento automático baseado no Critério de Informação de Akaike (AIC) apontou 3 *lags* significativos, sendo também significativos a tendência e o intercepto. Para o procedimento de Johansen, utilizou-se o Teste do Máximo Valor, considerando 2 *lags* e um termo de tendência linear. Comparando a estatística de teste com o ponto crítico em cada um dos testes, constata-se que a hipótese nula de não cointegração não foi rejeitada para nenhum dos testes considerados.

Tabela 6. 3 Testes de Cointegração estudados

Testes	EG	GPH	Johansen	BOOTReg	BOOTd
Estat. Teste	-3,1	-0,63	15,4	-0,63	0,938
P. Crítico (5%)	-3,56	-1,31	18,96	-1,676	0,866

De uma maneira geral, é possível observar que no modelo apresentado na Tabela 6.3 há problemas de regressão espúria, ou seja, aparentemente o modelo é adequado e está explicando boa parte do percentual da variação ocorrida na variável resposta. No entanto, a não verificação das suposições do modelo compromete de maneira significativa os parâmetros estimados pelo procedimento de mínimos quadrados. Assim, conclui-se que as séries Ibovespa e Dow Jones não são cointegradas, uma vez que os resíduos da regressão, além de não atender seus pressupostos, apresentam raiz unitária.

7. CONCLUSÕES

Neste trabalho são propostos alguns testes baseados na técnica *bootstrap* para testar a presença de raiz unitária e relações de cointegração em séries de longa dependência. Estes testes, aqui denominados BOOTReg e BOOTd, utilizam o estimador GPH proposto por Geweke e Porter-Hudak (1983). Além disso, os testes sugeridos tiveram suas *performances* comparadas a de outros testes amplamente divulgados na literatura. Para raiz unitária, foram considerados no estudo os testes ADF (Dickey e Fuller, 1981) e assintótico para o estimador GPH. Na avaliação de cointegração foram usados os testes de Engle e Granger (1987), Johansen (1990) e o baseado na distribuição assintótica do GPH. As comparações foram feitas via simulação Monte Carlo tomando por base dois aspectos principais: Poder e Tamanho dos testes.

Considerando os testes de raiz unitária, os resultados mostram que para valores de $d \leq 0,5$ que corresponde ao limite para processos de longa dependência estacionários, os testes apresentaram um alto poder, em geral acima de 90%, salvo para os testes ADF e BOOTReg quando $n=100$. Quando $d > 0,5$, o teste ADF sofre uma queda de poder considerável se comparado aos demais testes. O teste assintótico para o estimador GPH e o teste *bootstrap* BOOTd foram os que apresentaram melhor comportamento com relação ao poder. Com relação ao tamanho do teste ($d = 1,0$) a comparação foi feita considerando nível de significância de 5%. Assim, as melhores performances detectadas também ocorreram para os testes GPH (assintótico) e *bootstrap* BOOTd, já o pior desempenho foi do BOOTReg que ficou mais afastado do nível de significância estipulado e se mostrou um testes mais conservativo. De uma maneira geral, pode-se concluir que para detectar a presença de raiz unitária em modelos de longa dependência, os testes GPH e BOOTd são os mais indicados, uma vez que estes apresentaram melhores desempenhos que os demais, tanto com relação ao poder quanto ao tamanho.

Para cointegração, também se observou que para d na região de estacionaridade do modelo ARFIMA, os teste considerados no estudo apresentam poder acima de 90%, exceto os testes EG e BOOTReg quando $n=100$. De uma maneira geral, observou-se que o pior desempenho com relação ao poder ficou por conta do teste EG seguido do procedimento de Johansen. Os testes *bootstraps* tiveram comportamento semelhante ao observado no teste assintótico e estes foram os que apresentaram maior poder. Com relação ao tamanho, os testes que mais se destacaram foram o Johansen para $n=100$, o BOOTReg para $n=300$ e o EG para $n=500$. Verificou-se que o teste BOOTd foi o que apresentou pior desempenho de tamanho,

diferente do que foi observado na avaliação de raiz unitária, onde este apresentou tamanho muito próximo do valor nominal.

Uma outra observação importante é que à medida que o tamanho da amostra aumenta é possível notar que os testes de raiz unitária e cointegração apresentam uma melhora considerável em seus resultados.

Diante dos resultados obtidos via simulações Monte Carlo é possível constatar que a técnica *bootstrap* constitui uma ótima alternativa para a construção de testes de hipótese sobre o parâmetro fracionário d . Isso porque, de uma maneira geral, os testes aqui sugeridos apresentaram bons resultados frente aos testes comumente aplicados nos mais diversos estudos econométricos.

A título de ilustração, os testes de cointegração clássica e fracionária, considerados neste trabalho, foram aplicados a um conjunto de dados reais visando verificar a existência ou não de cointegração entre os mercados de ações do Brasil e Estados Unidos. Para o mercado brasileiro, foi considerado o índice Ibovespa, enquanto para o mercado americano considerou-se o índice Dow Jones. Os resultados mostraram que, tanto utilizando a abordagem clássica quanto a abordagem fracionária, ambas as séries apresentam raiz unitária. Na avaliação de cointegração, também não houve divergência entre os testes e todos apontaram que os índices Ibovespa e Dow Jones não são cointegrados, ou seja, não se pode afirmar que há uma relação de longo prazo entre os mercados de ações do Brasil e Estados Unidos.

Em trabalhos futuros, uma sugestão interessante seria estender os resultados deste trabalho para modelos ARFIMA(p,d,q) mais gerais, e também considerar outros tipos de estimadores para o parâmetro d . Uma outra possibilidade, seria considerar outros procedimentos *bootstraps*, tais como, o *bootstrap* aplicados aos resíduos da regressão, o *bootstrap* local (considerados nos estudos de Franco e Reisen, 2004) e o *Sieve bootstrap* (Bulman, 1997) para construir novos testes de raiz unitária e cointegração.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Alves, D.C.O., Cati, R.C. and Fava, V.L. (2001) Purchasing power parity in Brazil: a test for fractional cointegration, *Applied Economics*, 33, 1175-1185.
2. Andersson, M.K., Gredenhoff, M.P, (1998) Robust testing for fractional integration using the bootstrap. *Working Paper Series in Economics and Finance*, Stockholm School of Economics, No 218.
3. Andersson, M.K., Gredenhoff, M.P, (2000) Improving fractional integration tests with bootstrap distribution. *Working Paper*, 74, National Institute of Economic Research, Stockholm
4. Andrews, D.W.K., Lieberman, O., (2002) *Higher-order improvements of the parametric bootstrap for long-memory Gaussian processes*. Cowles Foundation, Discussion Paper 378, Yale University.
5. Arteche, J., Orbe, J., (2005) Bootstrapping the log periodogram regression. *Econometrics Letters*, 86, 79-85
6. Bisaglia, L. Guégan, D., (1998) A Comparison of Techniques of Estimation in Long Memory Processes, *Computational Statistics and Data Analysis*, 27, 61-81.
7. Box, G. E. P., Jenkins, G. M. (1976) *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
8. Andrews, D.W.K., Lieberman, O., (2002). *Higher-order improvements of the parametric bootstrap for long-memory Gaussian processes*. Cowles Foundation, Discussion Paper 1378, Yale University.
9. Chen, W. D. (2001) Evaluation of the long memory relationship between gross national product and saving in Taiwan. *Applied Economics Letters*, 8, 655-657.
10. Caporale, G.M, Gil-Alana, L.A. (2002) Unemployment and input prices: a fractional cointegration approach. *Applied Economics Letters*, 9, 347-351.
11. Davidson, J. (2002) *A Model of Fractional Cointegration, and Tests for Cointegration Using the Bootstrap*. *Jornal of Econometrics*.
12. Davidson, J. (2004) *Alternative Bootstrap Procedures for Testing Cointegration in Fractionally Integrated Processes*. *Jornal of Econometrics*.
13. Davison, A.C., Hinkley, D.V., (1997) *Bootstrap Methods and their Application*. Cambridge: Cambridge University Press.

14. Dickey, D.A., and Fuller, W.A. (1979) Distributions of The Estimators For Autoregressive Time Series With a Unit Root, *Journal of the American Statistical Association*, 75, 427-431.
15. Dickey, D.A., and Fuller, W.A. (1981) Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root, *Econometrica*, 49, 1057-1072.
16. Diebold, F.X. and Rudebusch, G.D. (1991) On the power of Dickey-Fuller tests against fractional alternatives, *Economics Letters*, 35, 155-160.
17. Dittmann, I., (2000). Residual-based tests for fractional cointegration: a Monte Carlo study. *Journal of Time Series Analysis*, v.6, n.21, pp.615-647
18. Efron, B., Tibshirani, R.J. (1993) *An Introduction to the Bootstrap*. New York: Chapman and Hall.
19. Efron, B. (1979) *Bootstrap* methods: another look at the Jackknife. *Annals of Statistics*, 7, 1-26.
20. Engle, R. F. and Granger, C. W. J. (1987) Co-integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55, 251-276.
21. Engle, R. F. and Yoo, B. S. (1987) Forecasting and testing in co-integrated systems, *Journal of Econometrics*, 35, 143-159.
22. Fox, R., Taqqu, M. S., (1986) Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series, *Annals of Statistics*, 14, 517-532.
22. Franco, G.C. and Reisen, V.A. (2004) *Bootstrap* Techniques in Semiparametric Estimation Methods for ARFIMA Models: A Comparison Study. *Computational Statistics*, 19,243-259.
24. Franco, G.C., Reisen, V.A., (2007) Bootstrap approaches and confidence intervals for stationary and non-stationary long-range dependence processes. *Physica. A*, **375(2)**, 546-562.
25. Fuller, W. (1976) *Introduction to Statistical Time Series*, John Wiley.
26. Gaio, L. E., Rolim, R. C.(2007) Interferência dos mercados externos sobre o Ibovespa. Universidade Federal de Lavras. X Semead – USP.
27. Granger, C. W. J., Joyeux, R., (1980). An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis*, **1**, 15-29.
28. Granger, C. W. J., Newbold, P., (1974). Spurious regression in econometrics. *Journal of Econometrics*, v. 2, pp. 111-120.
29. Granger, C. W. J., (1981). "Some Properties of Time Series Data and Their Use in Econometric Model Specification," *Journal of Econometrics*, 121-130.

30. Geweke, J., Porter-Hudak, S. (1983) The Estimation and Application of Long Memory Time Series Model. *Journal of Time Series Analysis*, 4(4):221-238.
31. Grau-Charles, P., (2004) Test for long memory processes: A bootstrap approach *Computational Economic Finance*, **111**
32. Hansen, B.E. (1992) Efficient estimation and testing of cointegration vectors in the presence of deterministic trends, *Journal of Econometrics*, 53, 87-121.
32. Harris, R.I.D. (1995) Using cointegration analysis in econometric modelling. London.
34. Hassler, U. and Wolters, J. (1994) On the power of unit root tests against fractional alternatives. *Economics Letters*, 45, 1-5.
35. Hosking, J. (1981), Fractional differencing. *Biometrika*, 68(1),165-175.
36. Hurvich, C. M. and Ray, B. K. (1995) Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes, *Journal of Time Series Analysis*. 19, 17-41.
37. Johansen, S. (1988) Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamic and Control*, 12, 231-254.
38. Johansen, S. and Juselius, K. (1990) Maximum Likelihood Estimation and Inferences on Cointegration - with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52, 169–210.
39. MacKinnon, J.G. (1991) Critical Values for Cointegration Tests. Chapter 13 in *Long-run Economic Relationships: Readings in Cointegration*, Oxford University Press.
40. Marques, C. R. (1998) Modelos dinâmicos Raízes Unitárias e Cointegração. Lisboa.
41. Murteira, B. J. F., Muller, D. A., Turkman, K. F. (1993) *Análise de Sucessões Cronológicas*. Editora McGraw-Hill de Portugal.
42. Peretti, C. (2002) Unilateral and bilateral bootstrap test for long memory, *Computational Economic Finance*, 334
43. Perron, P. (1997) Further evidence on breaking trend functions in macroeconomic variables. *Journal of Econometrics*, v. 80, p.355-385.
44. Phillips, P.C.B. and Durlauf, S.N. (1986) Multiple time series regression with integrated process, *Review of Economic Studies*, 53, 473-495.
45. Phillips, P.C.B. and Perron, P. (1988) Testing for a Unit Root in Time Series Regression, *Biometrika*, 75, 335–346.
46. Phillips, P.C.B., Ouliaris, S., (1990). Asymptotic properties of residual based tests for cointegration, *Econometrica*, v.58, n.1, pp. 165-193.

47. Rao, B. B., (1994) Trend, Unit Root and Structural Change in Macroeconomic Time Series, in *Cointegration for the Applied Economist*, Basingstoke: Macmillan Press, 113-146.
48. Reisen, V. A. (1994) Estimation of the fractional difference parameter in the ARIMA(p,d,q) model using the smoothed periodogram. *Journal of Time Series Analysis*, 15(3), 335-350.
49. Reisen, V. A. (1995) ARFIMA – O modelo ARIMA para o d fracionário. 6ª Escola de Séries Temporais e Econometria, UFES
50. Reisen, V. A., Abraham, B., Lopes, S.R.C. (2001) Estimation of parameters in ARFIMA process; Simulation study. *Communications in Statistics B*, v. 30, n. 4, p. 787-803.
51. Reisen, V.A., Abraham, B., Toscano, E.M.M. (2000) Parametric and Semi-parametric estimations of stationary univariate ARFIMA models. *Revista Brasileira de Probabilidade Estatística - REBRAPE*, v. 14, n. 2, p. 167-206.
52. Robinson, P.M. (1995), Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, 23(3), 1048-1072.
53. Robinson, P. M. (2003) Long memory time series, In: *Time series with long memory*. Oxford University Press, Oxford
54. Rubin, H., (1950). *Consistency of Maximum-Likelihood Estimates in the Explosive Case*. *Statistica Inference in Dynamic Economic Models*, ed by T. C Koopmans.
55. Said, S. E., Dickey, D.A. (1984) Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika*, 71, 599-607.
56. Santander, L.A.M., Reisen, V.A., Abraham, B. (2003) Non-cointegration tests and fractional ARFIMA process. *Statistical Methods*, 5(1): 1-22.
57. Schwert, G. W., (1989). Tests for unit roots: A Monte Carlo investigation. *Journal of Business and Economic Statistics*, v.7, 147-159.
58. Stock, J.H. (1987) Asymptotic properties of least-squares estimators fo co-integrating vectors, *Econometrica*, 55, 1035-1056.
59. Tse, Y.K., Anh, V.V., Tieng, Q., (1999), “No-cointegration test based on fractional differencing: Some Monte Carlo results”, *Jornal of Statistical Planning and Inference*, v. 80, pp. 227-267.

Anexo 1 – Programa para os Testes de Raiz Unitária

```
rm(list=ls(all=TRUE))
require(uroot)
require(fracdiff)

# ***** Função para Gerar ARFIMA(0,d,0) *****

GARFIMA <- function(Res, N, D3, NB){
  Phi <- rep(NA, NB)
  Y <- rep(0, NB)
  Phi[1] <- D3
  for(i in 2:NB){
    Phi[i] <- Phi[i-1] * ((i-1-D3)/i)
  }
  Y[1] <- Res[1]
  for(i in 2:NB){
    Y[i] <- Res[i] + sum(Phi[1:(i-1)]*Y[i-(1:(i-1))])
  }
  if((NB-N) > 0){
    Y <- Y[-c(1:(NB-N))]
  }
  Y
}

# ***** Função para extrair Resíduo de uma ARFIMA(0,d,0) *****

ExtraiResid <- function(Serie, D4, N){
  Phi <- rep(NA, N)
  SerieRes <- rep(0, N)

  Phi[1] <- D4

  for (i in 2:N){
    Phi[i] = Phi[i-1]*((i-1-D4)/i)
  }

  SerieRes[1] <- Serie[1]
  for (i in 2:N){
    SerieRes[i] <- Serie[i] - sum(Phi[1:(i-1)] * Serie[i-(1:(i-1))])
  }
  DP <- 1
  #DP <- sd(SerieRes)
  SerieRes2 <- (SerieRes - mean(SerieRes)) / DP
  SerieRes2
}

# ***** INICIALIZAÇÃO DE VARIÁVEIS *****

N <- 100           # Tamanho das séries
Burn <- 300        # Tamanho do Burn-In
NB <- N + Burn     # Tamanho da séries com Burn- In que será eliminado
Replic <- 2        # Número de Replicações Monte Carlo
NBoot <- 4         # Número de Bootstraps
D3 <- 1.0          # Valor simulado de d
#set.seed(99999)   # escolha da semente
Rep <- rep(NB, Replic) # Variável auxiliar para o lapply
MRes <- 0          # Média vetor normal
DPRes <- 1         # Variância vetor normal

Tempo <- system.time({

# ***** Geração das 1000 séries ARFIMA (0, d, 0) *****

Res <- lapply(Rep, mnorm, mean=MRes, sd=DPRes) # Geração Resíduos
Serie <- lapply(Res, GARFIMA, N, D3, NB) # Geração Séries
```

```

ListaD <- sapply (Serie, fdGPH, bandw.exp = 0.7) # estimação
ListaDgph <- as.numeric(ListaD[1,]) # d's estimados
ListaSDAssintGPH <- as.numeric(ListaD[2,]) # variância assintótica
ListaSDRegGPH <- as.numeric(ListaD[3,]) # variância regressão

#***** Testes de Raiz Unitária / GPH *****

SerieS <- lapply(Serie, ts) # transforma os vetores em objetos séries
SerieSpvalor<-rep(0, Replic)
ET <- (ListaDgph-1.0)/ListaSDAssintGPH # calcula ET para cada série
cont1 <- 0.0
PCritico <- -1.58 # p.critico para teste assintótico (1.58 p/100 e 300 e 1.56 p/500)
GPH <- mean(ET < PCritico)

#***** TESTE BOOTSTRAP (BASEADO GPH) / ADF *****

PCBootET <- rep(NA, Replic)
PCBootET1 <- rep(NA, Replic)
PCBootD <- rep(NA, Replic)

for(i in 1:Replic){
  SerieRes <- ExtraiResid(Serie[[i]], ListaDgph[i], N)
  SeriesRes2 <- rep(list(SerieRes), NBoot)
  ResBoot <- lapply(SeriesRes2, sample, size=N, replace=T)
  SerieBoot <- lapply(ResBoot, GARFIMA, N, 1, N) # Geração Séries
  MatrizDBoot <- sapply (SerieBoot, fdGPH, bandw.exp = 0.7)
  VetorDgphBoot <- as.numeric(MatrizDBoot[1,])
  VetorSDRegGPHBoot <- as.numeric(MatrizDBoot[3,])
  ETBoot <- (VetorDgphBoot - 1.0) / VetorSDRegGPHBoot
  PCBootET[i] <- quantile(ETBoot, probs = 0.05)
  PCBootD[i] <- quantile(VetorDgphBoot, probs = 0.05)
  TestAdf<-ADF.test(SerieS[[i]],c(0,0,0),regvar=0,selectlags=list(mode="aic", Pmax=5))
  SerieSpvalor[i]<-TestAdf@stat[4]
}
})

PercADF<-mean(SerieSpvalor<0.05)
GPHBootET <- mean(ET < PCBootET)
GPHBootD <- mean(ListaDgph < PCBootD)
Tempo
PercADF
GPH
GPHBootET
GPHBootD

#*****Impressão dos resultados*****

x <- list(PercADF=PercADF, GPH=GPH, GPHBootET=GPHBootET, GPHBootD=GPHBootD)
write.table(x, file = "D:\Fabiana\Resultados\Saida.txt", sep = "\t", row.names = F)

```

Anexo 2 – Programa para os testes de Cointegração

```

rm(list=ls(all=TRUE))
require(uroot)
require(fracdiff)
require(urca)

# *****Função para Regressão*****

GeraY <- function(X, E1, AlfaPar, BetaPar){
  Y = AlfaPar + (BetaPar*X) + E1
  Y
}

# ***** Função para Gerar ARFIMA(0,d,0) *****
GARFIMA <- function(Res, N, D3, NB){
  Phi <- rep(NA, NB)
  Y <- rep(0, NB)
  Phi[1] <- D3
  for(i in 2:NB){
    Phi[i] <- Phi[i-1] * ((i-1-D3)/i)
  }
  Y[1] <- Res[1]
  for(i in 2:NB){
    Y[i] <- Res[i] + sum(Phi[1:(i-1)]*Y[i-(1:(i-1))])
  }
  if((NB-N) > 0){
    Y <- Y[-c(1:(NB-N))]
  }
  Y
}

# ***** Função para extrair Resíduo de uma ARFIMA(0,d,0) *****

ExtraiResid <- function(Serie, D4, N){
  Phi <- rep(NA, N)
  SerieRes <- rep(0, N)

  Phi[1] <- D4

  for (i in 2:N){
    Phi[i] = Phi[i-1]*((i-1-D4)/i)
  }

  SerieRes[1] <- Serie[1]
  for (i in 2:N){
    SerieRes[i] <- Serie[i] - sum(Phi[1:(i-1)] * Serie[i-(1:(i-1))])
  }
  DP <- 1
  #DP <- sd(SerieRes)
  SerieRes2 <- (SerieRes - mean(SerieRes)) / DP
  SerieRes2
}

# *****Inicialização de Variáveis*****

Tam <- 300          # Tamanho das séries
Replic <- 1000     # Número de replicações
BurnIn <- 300      # Elimina as posições iniciais das séries
NBoot <- 1000     # Número de replicações bootstraps
DPRes <- 1        # Variância do vetor normal
MRes <- 0         # Média do vetor normal
AlfaY <- 0        # Intercepto para o modelo de regressão
BetaY <- 2        # Inclinação para o modelo de regressão
DResiduo <- 1.0   # Parâmetro d fracionário para os resíduos
DSerie <- 1.0     # Parâmetro d fracionário para as séries X
AlfaN <- 0.7      # Parâmetro que define o número de frequências a ser utilizada na estimação do d
NB <- Tam + BurnIn # Tamanho da Série mais o burn in que será eliminado
Rep <- rep(NB, Replic) # Variável auxiliar para o lapply

```

```

Tempo <- system.time({

# *****Geração da matriz de Resíduos a partir do ARFIMA*****

Res <- lapply(Rep, rnorm, mean=MRes, sd=DPRes) # Geração da Matriz de vetores normais
E <- sapply(Res, GARFIMA, Tam, DResiduo, NB) # Geração Séries Residuo

#***** Geração da matriz de Explicativas a partir do ARFIMA *****

ResX <- lapply(Rep, rnorm, mean=MRes, sd=DPRes) # Geração da Matriz de vetores normais
X <- sapply(ResX, GARFIMA, Tam, DSerie, NB) # Geração Séries Xs

#***** Geração da matriz de Respostas*****

Y <- GeraY(X, E, AlfaY, BetaY) #Geração Séries Ys

#***** TESTES PARA COINTEGRAÇÃO *****

ResTeste <-rep(0, (Replic))
ET <-rep(0, (Replic))
PJohanse <- rep(NA, 2)
PCBootET <- rep(NA, Replic)
PCBootET1 <- rep(NA, Replic)
PCBootD <- rep(NA, Replic)
ListaDgph <- rep(NA, Replic)
i <- 1
for (i in 1:(Replic)){
  YX<-data.frame(cbind(Y[,i], X[,i]))
  Regressao<-lm(YX[,1]~0+YX[,2])
  Residuo<-Regressao$residual

#***** ADF, GPH E JOHANSEN *****

  ResTestAdf<-ADF.test(ts(Residuo),c(0,0,0),regvar=0,selectlags=list(mode=c(0,1,2)))
  ResTeste[i] <- ResTestAdf@stat[3] # Salva apenas o valor da Estatística de Teste para o ADF
  EstimaD<- fdGPH (Residuo, bandw.exp = AlfaN)
  ET[i] <- (EstimaD$d - 1) / EstimaD$sd.as # Salva o valor da Estatística de Teste para o GPH
  ListaDgph[i] <- EstimaD$d # Necessário para o boot da distribuição empírica do D
  TesteJo <- ca.jo(YX, ecdet = "none", type="eigen", K=2,spec="transitory") # "transitory" "longrun"
  ResultJo <- c(TesteJo@teststat[2], TesteJo@cval[2,2]) # Extrai a Estatística de Teste e o ponto crítico a 5%
  PJohanse<- rbind(PJohanse, ResultJo)

#***** TESTES BOOTSTRAPS*****

  SerieRes <- ExtraiResid(Residuo, EstimaD$d, Tam)
  SeriesRes2 <- rep(list(SerieRes), NBoot)
  ResBoot <- lapply(SeriesRes2, sample, size=Tam, replace=T) #resíduos bootstraps
  SerieBoot <- lapply(ResBoot, GARFIMA, Tam, 1, Tam) # Geração das séries bootstraps
  MatrizDBoot <- sapply (SerieBoot, fdGPH, bandw.exp = 0.7) # estimação dos ds de todas as séries bootstraps
  VetorDgphBoot <- as.numeric(MatrizDBoot[1,]) #ds estimados
  VetorSDRegGPHBoot <- as.numeric(MatrizDBoot[3,]) # erro padrão da regressão
  ETBoot <- (VetorDgphBoot - 1.0) / VetorSDRegGPHBoot
  PCBootET[i] <- quantile(ETBoot, probs = 0.05)
  PCBootD[i] <- quantile(VetorDgphBoot, probs = 0.05)
}
})

PJohanse <- PJohanse[-1,] # Retira a primeira linha que é replicada
PercJo <- mean(PJohanse[,1] > PJohanse[,2])
PercADF <- mean(ResTeste < -2.76) # Ponto crítico Simulado
PercGPH<-mean(ET < -1.78) # Ponto crítico Simulado -1.73(100) e -1.78 (300 e 500)
GPHBootET <- mean(ET < PCBootET) # Utilizando o desvio da Regressão
GPHBootD <- mean(ListaDgph < PCBootD) # Utilizando a Distribuição Empírica do D

PercADF
PercGPH
PercJo

```

```
GPHBootET  
GPHBootD
```

```
Tempo # Tempo gasto na simulação em segundos
```

```
#*****Impressão dos resultados*****
```

```
x<-list(PercADF=PercADF, GPH=PercGPH, PercJo=PercJo, GPHBootET=GPHBootET, GPHBootD=GPHBootD,  
D=DResiduo)  
write.table(x, file = "c:\\Temp\\Dissertacao\\Resultados\\SaidaCoin.txt", sep = "\\t", row.names = F)
```