Estacionariedade em Séries Temporais com Quebras Estruturais

Paulo Evandro Dawid

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do grau
de
Mestre em Estatística

Área de concentração: **Séries temporais**
Orientadora: **Profa. Dra. Mariane Streibel**

Trabalho financiado pelo Programa de Pós-graduação do Banco Central do Brasil.

São Paulo - fevereiro - 2004

Estacionariedade em Séries Temporais com Quebras Estruturais

Este exemplar corresponde à redação final
da dissertação de mestrado devidamente
corrigida e defendida por
Paulo Evandro Dawid
e aprovada pela comissão julgadora.

São Paulo, fevereiro de 2004.

Banca examinadora:

- Profa. Dra. Mariane Streibel (Orientadora) - IME-USP
- Prof. Dr. Pedro Alberto Morettin - IME-USP
- Prof. Dr. Denisard Cnéio de Oliveira Alves - FEA-USP

Ao meu pai,
de quem herdei
o prazer pelo conhecimento,
e à minha querida mãe.

Agradecimentos

Concluída esta etapa, gostaria de agradecer

Ao Banco Central do Brasil que com seu programa de pós-graduação possibilitou e incentivou este trabalho.

À Profa. Mariane Streibel pela orientação e conselhos pertinentes.

Aos professores do IME-USP, sempre disponíveis e fundamentais à minha formação.

À minha esposa Sandra, minha incentivadora cotidiana.

Aos meus pais e à minha família, com os quais sempre posso contar.

Aos amigos do IME e do Banco Central, pelo companheirismo e apoio nesta jornada.

Resumo

Um dos principais problemas na análise de séries temporais é testar se uma determinada série é estacionária ou não. Séries estacionárias são, *grosso modo*, séries que mantêm o mesmo comportamento ao longo do tempo. A teoria e várias técnicas de análise já estão bem estabelecidas para essas séries. A análise de séries não-estacionárias, por sua vez, é mais complexa e não existe uma abordagem geral.

Os testes de raiz unitária são os mais utilizados para se testar um tipo específico de não-estacionariedade, denominada tendência estocástica ou passeio aleatório.

Este trabalho pretende explorar uma abordagem alternativa desenvolvida no contexto de modelos de componentes não-observáveis. São os chamados testes de estacionariedade, que diferem dos testes de raiz unitária pela inversão da hipótese nula.

Os testes de estacionariedade vêm apresentando generalizações recentes, possibilitando a aplicação a uma gama mais ampla de séries: com tendência, com sazonalidade e com quebras estruturais.

Nessa linha, pretende-se então revisar os principais conceitos relacionados e aplicar esses testes de estacionariedade a séries econômicas brasileiras.

Palavras-chave: Componentes não-observáveis, distribuição de Cramér-von Mises, estacionariedade, quebras estruturais, tendência estocástica, teste invariante localmente ótimo.

Abstract

Testing for stationarity is an important issue in time series analysis. Stationary series are, in a certain way, series that keep their behaviour as time goes by. The theory and many analysis tools are already available for stationary series. The analysis of nonstationary series is more difficult and there is not a general approach.

Unit root tests are the most popular ones to test the specific kind of nonstationarity called stochastic trend or random walk behaviour.

This work intends to explore an alternative approach developed for unobserved components models. The correspondent tests are called stationarity tests and differ from unit root tests by inverting the null hypothesis.

Stationary tests present some recent generalizations for applications in a broader class of series: with trend, sazonality or structural breaks.

Following this new generalizations, we intend to review the main related concepts and use the stationary tests in Brazilian economic data.

Keywords: Unobserved components, Cramér-von Mises distribution, stationarity, structural breaks, stochastic trend, locally best invariant test.

Sumário

1 Introdução **1**

1.1 Motivação econômica 1

1.2 Motivação estatística 3

1.2.1 Testes de raiz unitária 3

1.2.2 Testes de estacionariedade 4

1.3 Plano do trabalho 7

2 Definições e resultados básicos **9**

2.1 Processos estocásticos e estacionariedade 9

2.2 Movimento browniano e ponte browniana 10

2.3 Inversão numérica da função característica 11

2.4 Distribuições da família Cramér-von Mises 12

2.5 Teste invariante localmente ótimo - ILO 16

2.5.1 Teste localmente ótimo 16

2.5.2 Teste invariante 17

2.5.3 Teste ILO para o modelo linear normal 18

3 Variância de longo termo **20**

3.1 A variância de longo termo - σ_L^2 20

3.2 Metodologia proposta para a estimação de σ_L^2 26

3.3 Algoritmo para a estimação de σ_L^2 28

4 Testes de estacionariedade **30**

4.1 Série com nível constante 31

4.2 Série com tendência linear 32

4.3 Série com quebra no nível e na tendência 33

4.4 Série com quebra no nível apenas 35

4.5 Extensão para ruídos estacionários 36

4.6 Modelos com componente sazonal 37

5 Simulações **38**

5.1 Estimativa da variância de longo termo 38

5.2 Tamanho e poder para amostras finitas 40

5.3 Tamanho e poder do teste simplificado 43

6 Aplicações **47**

6.1 Análise das exportações brasileiras 47

6.2 Análise das importações brasileiras 53

6.3 Câmbio real brasileiro 61

Referências Bibliográficas **63**

Capítulo 1

Introdução

1.1 Motivação econômica

A não-estacionariedade é uma característica comum em muitas séries temporais financeiras e econômicas. Agregados macroeconômicos, como o PIB, a produção industrial, o investimento *etc.*, em geral apresentam tendência de crescimento. Isso reflete a própria dinâmica do desenvolvimento econômico, que por sua vez é explicado em diversas teorias de desenvolvimento por fatores como a acumulação de capital, o nível de poupança e investimento, o crescimento da população economicamente ativa, os ganhos de produtividade advindos da inovação tecnológica etc. Dada essa "tendência natural" ao crescimento, essas séries são passíveis de serem caracterizadas com um componente de tendência determinística.

Já as séries financeiras, como índices de preço de ações, taxas de câmbio em regime livre, preços de *commodities* etc., usualmente apresentam não-estacionariedade caracterizada por um componente que é um passeio aleatório, ou seja, uma tendência dita estocástica. Esse resultado empírico também pode ser explicado pela dinâmica de formação desses preços e está de acordo com modelos econômicos baseados em expectativas racionais: os agentes desses mercados, quando dotados de todas as informações relevantes dis-

poníveis, conseguem antecipar perdas ou ganhos futuros esperados. E isso reflete no valor atual do ativo, de forma que a série desses preços presentes corresponde a uma série acumulada de ruídos, ou seja, de choques aleatórios imprevisíveis por parte dos agentes econômicos. Esse comportamento é chamado de 'passeio aleatório'.

Quando um componente da série temporal é um passeio aleatório, costuma-se dizer que a série possui uma raiz unitária, ou que é integrada de ordem 1, $I(1)$. A relevância de se saber se uma série possui raiz unitária está no fato de que, em caso afirmativo, os choques externos causam um efeito permanente na série. Ao passo que, em uma série estacionária, há um retorno à média após um certo tempo. Por exemplo, se a série de um agregado macroeconômico possuir um componente que é um passeio aleatório, isto é, se sua tendência é estocástica, tem-se como principal consequência empírica que os choques externos significativos, reais ou monetários, terão um efeito permanente nessa série. Assim, as políticas monetárias e fiscais que afetam tal série passam a exigir maior atenção, pois seus efeitos não são transitórios.

Nessa linha, Nelson & Plosser (1982), num influente trabalho, analisaram 13 séries macroeconômicas dos Estados Unidos, aplicando a elas o teste de Dickey-Fuller (DF) para raiz unitária [v. abaixo]. Como resultado, obtiveram que, para a maioria das séries, não se pode rejeitar a hipótese de que possuem uma raiz unitária. Assim, puseram em questão a maneira tradicional de se modelar as séries macroeconômicas como uma tendência determinística mais um ruído estacionário. E lançaram uma agenda de pesquisa que de certa forma estende-se até hoje, haja vista a importância para a validação de modelos da teoria econômica. Do ponto de vista estatístico, o desafio é obter melhores testes, pois com base em amostras finitas não se pode afirmar com precisão absoluta se uma série temporal possui ou não um componente não observado que é um passeio aleatório, ou raiz unitária.

1.2 Motivação estatística

Para uma visão mais detalhada e abrangente, ver, por exemplo, Maddala & Kim (1998).

1.2.1 Testes de raiz unitária

O teste mais popular para a verificação da não-estacionariedade de uma série é o teste de Dickey-Fuller, ou sua versão aumentada (ADF). Ele consiste basicamente em testar se o modelo auto-regressivo (ou mesmo ARMA) que ajusta a série possui uma raiz unitária no polinômio auto-regressivo $\phi(L)$. Ajusta-se, por exemplo, a série y_t por um modelo AR(p):

$$\begin{aligned} \phi(L)y_t &= \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}\mathrm{id}\,(0,\sigma^2) \\ \phi(L) &= 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p \\ Ly_t &= y_{t-1} \quad : \quad L \text{ é o operador defasagem} \\ \mathcal{N}\mathrm{id} & \text{ indica distribuição normal e independente} \end{aligned} \tag{1.1}$$

E testa-se então se o polinômio $\phi(L)$ possui uma raiz no círculo unitário. A distribuição assintótica do estimador, sob a hipótese de raiz unitária, é diferente da t-Student obtida sob a alternativa da série ser estacionária, ou seja, quando as raízes de $\phi(L)$ caem fora do círculo unitário.

Na linha dos testes de raiz unitária, podem-se destacar sofisticações importantes, como o teste de Phillips-Perron[1] que permite que os erros ε_t possuam uma estrutura de dependência fraca. O que é mais realista no contexto de séries temporais e, em geral, é suficiente para abranger grande parte das séries econômicas reais. Perron (1989) obtém as distribuições assintóticas para séries contendo quebras estruturais, o que flexibiliza e aumenta o poder dos testes de raiz unitária, se tais quebras de fato existem.

[1] Phillips (1987) e Phillips & Perron (1988)

1.2.2 Testes de estacionariedade

Numa linha de pesquisa paralela, desenvolveram-se os chamados testes de estacionariedade. Como os testes de raiz unitária, eles também servem para verificar se uma série temporal é estacionária ou não, mas com uma modelagem distinta. Do ponto de vista estatístico, a principal diferença está na inversão da hipótese nula:

- Nos testes de raiz unitária, testa-se:

  $\mathbf{H_0}$: a série, condicionalmente ao modelo, possui raiz unitária e, portanto, não é estacionária.

  Contra a hipótese alternativa:

  $\mathbf{H_a}$: todas as raízes do modelo estão fora do círculo unitário e, portanto, a série é estacionária.

- Nos testes de estacionariedade, o modelo é a soma de um componente estacionário e um componente não-estacionário (passeio aleatório). Então, testa-se:

  $\mathbf{H_0}$: o passeio aleatório não está presente e, portanto, a série é estacionária.

  Contra a hipótese alternativa:

  $\mathbf{H_a}$: o passeio aleatório está presente e, portanto, a série não é estacionária.

Simplificadamente, o modelo para o teste de estacionariedade pode ser esquematizado como:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t \quad , \quad \varepsilon_t \sim \mathcal{EST}(0, \gamma_\varepsilon) \\ \mu_t &= \mu_{t-1} + \eta_t \quad , \quad \eta_t \sim \mathrm{RB}\,(0, \sigma_\eta^2) \\ \varepsilon_t &\ \text{e}\ \ \eta_t \text{ independentes entre si.} \end{aligned} \tag{1.2}$$

E o teste de hipóteses correspondente seria:

- H_0: $\sigma_\eta^2 = 0$
- H_a: $\sigma_\eta^2 > 0$

Um pouco de história

Em termos de aperfeiçoamento, os testes de estacionariedade seguiram mais ou menos a mesma trajetória do testes de raiz unitária. Com base no modelo (1.2), relata-se a seguir as principais evoluções dos testes de estacionariedade.

Nyblom & Makelainen (1983) desenvolvem um teste (NM) em que ε_t também é um ruído branco e ε_t e η_t possuem distribuições normais independentes. Nessas condições, NM é o teste invariante localmente ótimo (ILO ou LBI de *locally best invariant*) e, sob H_0, tem assintoticamente distribuição de Cramér-von Mises (CvM). A estatística de teste obtida ξ possui a seguinte região crítica:

$$\xi = \frac{\sum_{i=1}^{T}\left(\sum_{t=1}^{i} e_t\right)^2}{T^2\hat{\sigma}^2} > c(\alpha) \tag{1.3}$$

Em que T é o tamanho da série, e_t, $t = 1, \ldots, T$, são os resíduos do modelo ajustado, $c(\alpha)$ é o valor crítico dependente do nível de significância α e $\hat{\sigma}^2$ é a variância amostral:

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2 \tag{1.4}$$

A derivação da distribuição assintótica de ξ baseia-se em resultados da teoria de processos estocásticos [2]. Sob H_0, tem-se que a soma parcial dos resíduos, devidamente ajustada, converge fracamente para a chamada ponte browniana, $B(r)$, definida em termos do processo de Wiener padrão, $W(r)$:

$$\frac{1}{\sigma\sqrt{T}}\sum_{i=1}^{[Tr]} e_i \Rightarrow B(r) = W(r) - rW(1), \quad r \in [0,1] \tag{1.5}$$

'$\Rightarrow$' indica convergência fraca de medidas de probabilidade associadas. Tomando-se o limite em (1.3), tem-se:

[2] v., e.g., Tanaka (1996)

$$\xi \Rightarrow \int_0^1 B(r)^2 dr \tag{1.6}$$

que corresponde à distribuição de Cramér-von Mises.

Os demais resultados assintóticos obtidos nesses testes de estacionariedade, e mesmo nos testes de raiz unitária, derivam de raciocínios análogos.

Em 1986, Nyblom generaliza o modelo em (1.2) para permitir uma tendência linear:

$$y_t = \mu_t + \beta t + \varepsilon_t \tag{1.7}$$

A estatística utilizada ainda é a mesma ξ, mas a distribuição assintótica é uma generalização da CvM, denotada por CvM_2.

Kwiatkowski, Phillips, Schmidt & Shin (1992), analogamente a Phillips & Perron (1988) no contexto de raiz unitária, generalizam o teste NM de forma que ε_t possa ser um processo estacionário sob certas condições. De novo, uma hipótese mais realista para séries temporais empíricas. A estatística utilizada para esse teste, denominado KPSS, continua a ser a ξ acima, com as mesmas distribuições assintóticas, mas agora substituindo $\hat{\sigma}^2$ por um estimador consistente da chamada variância de longo termo da série, $\sigma_L^2(l)$, ou seja, requer-se que

$$\lim_{T\to\infty} \mathbb{P}(|\hat{\sigma}_L^2(l) - \sigma_L^2| > \epsilon) = 0 \quad , \quad \forall \epsilon > 0 \tag{1.8}$$

l indica o tamanho do núcleo $w(t,l)$ utilizada por Kwiatkowski *et al.* (1992) para obter um estimador consistente não-paramétrico, baseado na teoria espectral de séries temporais[3]. Utilizando uma janela de Bartlett[4], obtiveram:

$$\hat{\sigma}_L^2(l) = \frac{1}{T}\left[\sum_{i=1}^{T} e_i^2 + 2\sum_{t=1}^{l} w(t,l) \sum_{i=t+1}^{T} e_t e_{i-t}\right] \tag{1.9}$$

[3] v., *e.g.*, Priestley (1981)

[4] Outras janelas podem ser utilizadas, v., *e.g.*, Andrews (1991)

com

$$w(t,l) = 1 - \frac{t}{l+1} \quad , \quad t = 1, \ldots, l \tag{1.10}$$

Busetti & Harvey (2001) estenderam os testes NM e KPSS acima para séries contendo quebras estruturais (teste BH). Mais especificamente, séries com ou sem tendência linear que possuam descontinuidades (uma ou mais) no nível ou na inclinação. O modelo mais geral, como extensão de (1.2), seria:

$$y_t = \mu_t + \beta t + \delta_\mu w_t + \delta_\beta (w_t t) + \varepsilon_t \tag{1.11}$$

Supondo uma quebra, no nível e na inclinação, no instante τ, tem-se:

$$w_t \;=\; \mathbf{1}_{\{t>\tau\}} \quad , \quad t = 1, \ldots, T \quad , \quad 1 < \tau < T \tag{1.12}$$

Busetti & Harvey (2001) obtiveram as distribuições assintóticas, em termos do funcional $W(r)$, de quatro variações possíveis do modelo:

- Quebra no nível no modelo sem tendência.
- Quebra no nível e na inclinação no mesmo instante τ.
- Quebra apenas no nível no modelo com tendência.
- Quebra apenas na inclinação no modelo com tendência.

1.3 Plano do trabalho

No capítulo 2 apresentamos alguns conceitos básicos e resultados. Em particular, a metodologia para obtenção das distribuições da família Cramér-von Mises via inversão da função característica e os conceitos por trás dos testes invariantes e localmente ótimos utilizados. No capítulo 3 introduzimos o conceito de variância de longo termo, sua relação com a função de densidade espectral e propomos uma metodologia para sua estimativa, baseada em estimação espectral não-paramétrica. No capítulo 4 apresentamos

os modelos com tendência e quebras utilizados, seus respectivos testes de estacionariedade, uma forma simplificada para os testes e algumas extensões: para séries com ruído estacionário em vez de ruído branco e para séries com componentes sazonais não-aleatórias. No capítulo 5 realizamos algumas simulações para: (i) avaliar a qualidade da estimativa proposta para a variância de longo termo, (ii) avaliar o tamanho e poder empíricos do teste com quebra no nível e na tendência utilizando a estimativa proposta da variância de longo termo e (iii) avaliar em termos de tamanho e poder empíricos a qualidade do teste simplificado. E no capítulo 6, aplicamos alguns dos testes para dados econômicos brasilieiros: exportação, importação e taxa de câmbio real.

Capítulo 2

Definições e resultados básicos

Com o intuito de se fazer uma breve revisão e uma padronização da notação, definiremos neste capítulo alguns conceitos básicos relevantes. Também serão apresentados alguns resultados utilizados no decorrer deste trabalho.

2.1 Processos estocásticos e estacionariedade

Definição 2.1 (Processo estocástico). *Um processo estocástico é uma família de variáveis aleatórias* $\{X(t,\omega), t \in \mathbf{T}, \omega \in \Omega\}$ *definidas em um espaço de probabilidade* $(\Omega, \mathcal{F}, P)$. *Tais variáveis são indexadas pelo conjunto* $\mathbf{T}$, *que aqui possuirá uma relação de ordem definida: se* $t_1, t_2 \in \mathbf{T}$, *então* $t_1 \leq t_2$ *ou* $t_1 > t_2$.

Definição 2.2 (Processo normal). *Um processo estocástico* $\{X(t), t \in \mathbf{T}\}$ *é denominado normal se todas as suas distribuições finito-dimensionais são normais multivariadas, ou seja,*

$$(X(t_1), X(t_2), \ldots, X(t_n)) \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \forall t_1, t_2, \ldots, t_n \in \mathbf{T} \tag{2.1}$$

Como os processos a serem analisados são observados em intervalos de tempo regulares, discretos portanto, chamaremos de **série temporal** a uma realização de um processo estocástico para um conjunto enumerável $\mathbf{T}$. Ou seja, $\{y_t\}$ será uma série temporal se

$\{y_t\}_{t=1,2,\ldots,|\mathbf{T}|} = \{X(t,\omega) : t = 1,2,\ldots,|\mathbf{T}| \; ; \; \omega \in \Omega \; ; \; \omega$ fixo $\}$, para algum processo estocástico $X(t,\omega)$.

Definição 2.3 (Série estacionária). *Uma série temporal $\{y_t\}_{t\in\mathbf{T}}$ é denominada fracamente estacionária, ou estacionária de segunda ordem, ou, simplesmente, estacionária, se, para todo $t \in \mathbf{T}$,*

1. $\mathrm{E}\, y_t = \mu = cte$
2. $\mathrm{Var}\, y_t = \sigma^2 = cte < \infty$
3. $\mathrm{Cov}(y_t, y_{t+|h|}) = \gamma(h), \quad h \in \mathbb{Z}$, *ou seja, independe de* t

Notação : $y_t \sim \mathcal{EST}(\mu, \gamma(h))$

2.2 Movimento browniano e ponte browniana

Definição 2.4 (Movimento browniano). *O processo estocástico $\{W(t), t \in \mathbb{R}^+\}$ é um movimento browniano, ou processo de Wiener, se*

1. $W(0) = 0$
2. *As v.a.s* $(W(t_i) - W(t_{i-1}))_{i=2,\ldots,k}$ *são independentes para* $0 \leq t_1 \leq t_2 \leq \ldots \leq t_k$
3. $(W(t) - W(s)) \sim \mathcal{N}(0, t-s)$ *para* $0 \leq s \leq t$
4. *As trajetórias de* $W(t)$ *são contínuas com probabilidade* 1.

A prova da existência do movimento browniano pode ser vista, por exemplo, em Billinsgsley (1995). Como consequência da definição acima, tem-se que:

$$\mathrm{Cov}(W(s), W(t)) = \min(s,t) = s \wedge t \tag{2.2}$$

Definição 2.5 (Ponte browniana). *Um processo de ponte browniana $\{B(r), 0 \leq r \leq 1\}$ é um processo normal com funções amostrais contínuas e:*

1. $B(0) = B(1) = 0$

2. $\mathrm{E}\, B(r) = 0$, *para* $0 \leq r \leq 1$

3. $\mathrm{Cov}(B(r), B(s)) = r \wedge s - rs$, *para* $0 \leq r, s \leq 1$

A ponte browniana $B(r)$ pode ser escrita em termos do movimento browniano $W(t)$:

$$B(r) = W(r) - rW(1) \quad , \quad 0 \leq r \leq 1 \tag{2.3}$$

A ponte browniana de segundo nível $B_2(r)$ é uma generalização de $B(r)$, e em termos de $W(t)$ é definida como:

$$B_2(r) = W(r) - rW(1) + 6r(1-r)\left[\frac{1}{2}W(1) - \int_0^1 W(s)ds\right] \quad , \quad 0 \leq r \leq 1 \tag{2.4}$$

2.3 Inversão numérica da função característica

Os valores críticos e os níveis descritivos (p-valores) das distribuições utilizadas neste trabalho serão obtidos pela inversão numérica de suas funções características, apresentadas na seção seguinte.

Definição 2.6 (Função característica). *A função característica de uma variável aleatória Y, com função de distribuição F, é definida por*

$$\varphi_Y(t) = E(e^{itY}) = \int_{-\infty}^{\infty} e^{ity} dF \tag{2.5}$$

O valor de $F(y) = \mathbb{P}[Y \leq y]$ pode ser obtido pela fórmula de inversão de Lévy:

$$F(y) - F(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1 - e^{-ity}}{it}\varphi(t)dt \tag{2.6}$$

Fórmula de inversão alternativa:

$$F(y) = \frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty} \frac{e^{ity}\varphi(-t) - e^{-ity}\varphi(t)}{it}\, dt \tag{2.7}$$

Para o cálculo numérico da integral acima utilizamos o pacote $QUADPACK$[1] implementado na linguagem Ox [2].

[1] v. Piessens et al. (1983)

[2] v. Doornik (1999)

2.4 Distribuições da família Cramér-von Mises

O conceito relevante de convergência para as distribuições assintóticas apresentadas a seguir é o de convergência fraca de medidas de probabilidade associadas ou induzidas, que definimos abaixo. Intuitivamente, é uma generalização da convergência em distribuição para elementos aleatórios definidos em espaços mais gerais que o euclidiano.

Definição 2.7 (Convergência fraca de medidas de probabilidade associadas). *Sejam X_n, $n \geq 0$, elementos aleatórios no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ que recebem valores num espaço métrico separável e completo (M, d). Uma sequência de medidas de probabilidade P_n induzidas por X_n converge fracamente para P_0, induzida por X, se, para toda função $f : M \to \mathbb{R}$ uniformemente contínua (em relação a d) e limitada, tem-se que $\int f dP_n = \mathrm{E}\, f(X_n) \to \mathrm{E}\, f(X) = \int f dP_0$ quando $n \to \infty$.* **Notação:** $X_n \Rightarrow X$ *ou* $P_n \Rightarrow P_0$.

Distribuição de Cramér-von Mises

A distribuição de Cramér-von Mises, CvM, surge naturalmente no contexto de testes de qualidade de ajuste (*goodness of fit tests*). Mais especificamente, na avaliação assintótica da seguinte integral[3]:

$$C_n = \int_{-\infty}^{\infty} n[F_n(x) - F(x)]^2 dF(x) = \int_{-\infty}^{\infty} U_n(F)^2 dF = \int_0^1 U_n(r)^2 dr \tag{2.8}$$

F_n é a função de distribuição empírica de uma amostra aleatória $\{X_1, X_2, \ldots, X_n\}$ em que X_i tem função de distribuição $F(x)$ contínua:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(-\infty, x]}(x_i), \quad -\infty < x < \infty, \tag{2.9}$$

U_n é o processo empírico uniforme:

$$U_n(r) = \sqrt{n}[F_{U,n}(r) - r], \quad 0 \leq r \leq 1, \tag{2.10}$$

[3] *v., e.g.,* Shorack & Wellner (1986)

com $F_{U,n}(r)$ sendo a função de distribuição empírica de uma amostra aleatória com distribuição $Uniforme(0,1)$.

$U_n(r)$ converge fracamente para uma ponte browniana $B(r)$:

$$U_n(r) = \sqrt{n}[F_{U,n}(r) - r] \Rightarrow B(r) \tag{2.11}$$

E C_n converge então para a denominada distribuição de Cramér-von Mises (CvM):

$$C_n = \int_0^1 U_n(r)^2 dr \Rightarrow \int_0^1 B(r)^2 dr \sim \mathrm{CvM}$$

Função característica da distribuição CvM

Se $Y \sim \mathrm{CvM}$, então sua função característica $\varphi_Y(t)$ é dada por[4]:

$$\varphi_Y(t) = \sqrt{\frac{\sqrt{2it}}{\operatorname{sen}\sqrt{2it}}} \tag{2.12}$$

Os valores da distribuição serão obtidos numericamente pela fórmula de inversão (2.6). Para tanto, convém escrever $\varphi_Y(t)$ em termos da função $D_1(x) \equiv \operatorname{sen}(\sqrt{x})/\sqrt{x}$:

$$\varphi_Y(t) = [D_1(2it)]^{-\frac{1}{2}} \tag{2.13}$$

Dessa forma, os valores da distribuição podem ser obtidos pela seguinte expressão, apresentada em MacNeill (1978):

$$\mathbb{P}[Y \leq \xi] = 1 - \frac{1}{\pi}\sum_{j=1}^{\infty}(-1)^{j-1}\int_{\lambda_{1,2j-1}^{-1}}^{\lambda_{1,2j}^{-1}} \frac{e^{-\frac{x}{2}\xi}\, dx}{x\sqrt{-D_1(x)}} \tag{2.14}$$

com

$$\lambda_{1,k} = \frac{1}{\pi^2 k^2} \tag{2.15}$$

[4] *v., e.g.,* Anderson & Darling (1952) ou MacNeill (1978)

Distribuição de Cramér-von Mises de segundo nível

Uma possível generalização da distribuição de Cramér-von Mises pode se feita pela utilização da generalização do processo de ponte browniana subjacente.

Assim, em termos do processo de ponte browniana de segundo nível $B_2(r)$, dado por (2.4), a distribuição de Cramér-von Mises de segundo nível (CvM_2) é definida por

$$\int_0^1 B_2(r)^2 dr \sim \mathrm{CvM}_2 \tag{2.16}$$

Função característica da distribuição CvM_2

Se $Y_2 \sim \mathrm{CvM}_2$, então sua função característica $\varphi_{Y_2}(t)$ é dada por[5]:

$$\begin{aligned} \varphi_{Y_2}(t) &= [D_2(2it)]^{-\frac{1}{2}} \\ D_2(x) &\equiv \frac{12}{x^2}\left(2 - \sqrt{x}\sin\sqrt{x} - 2\cos\sqrt{x}\right) \end{aligned} \tag{2.17}$$

Para o cálculo numérico da inversão de φ_{Y_2}, pode-se utilizar, como em MacNeill (1978), a seguinte expressão:

$$\begin{aligned} \mathbb{P}[Y \leq \xi] &= 1 - \frac{1}{\pi}\sum_{j=1}^{\infty}(-1)^{j-1}\int_{\lambda_{2,2j-1}^{-1}}^{\lambda_{2,2j}^{-1}} \frac{e^{-\frac{x}{2}\xi}\,dx}{x\sqrt{-D_2(x)}} \\ \lambda_{2,2j-1} &= \frac{1}{4\pi^2 j^2} \\ \lambda_{2,2j} &= \frac{1}{4Z_j^2} \end{aligned} \tag{2.18}$$

Z_j é a j-ésima raiz positiva da função esférica de Bessel de primeira espécie e ordem 1:

$$J_1(x) = \frac{\operatorname{sen} x}{x^2} - \frac{\cos x}{x} \tag{2.19}$$

[5] *v., e.g.,* MacNeill(1978)

Distribuição de Cramér-von Mises com k graus de liberdade

Como em Anderson & Darling (1952), a distribuição de Cramér-von Mises pode ainda ser representada pela seguinte expansão:

$$\int_0^1 B(r)^2 dr \stackrel{d}{=} \sum_{j=1}^{\infty} \lambda_{1,j}\chi_j^2(1) \sim \mathrm{CvM} \tag{2.20}$$

em que $\chi_j^2(1)$, $j = 1, 2, \ldots$, são v.a.i.i.d.s com distribuição qui-quadrada com 1 grau de liberdade. Assim, a distribuição de Cramér-von Mises com k graus de liberdade $\mathrm{CvM}(k)$ será definida por

$$\sum_{j=1}^{\infty} \lambda_{1,j}\chi_j^2(k) \sim \mathrm{CvM}(k) \tag{2.21}$$

De forma análoga, obtém-se a distribuição de Cramér-von Mises de segundo nível com k graus de liberdade, $\mathrm{CvM}_2(k)$, $k = 1, 2, \ldots$:

$$\sum_{j=1}^{\infty} \lambda_{2,j}\chi_j^2(k) \sim \mathrm{CvM}_2(k) \tag{2.22}$$

Obs.: $\mathrm{CvM}(1) = \mathrm{CvM}$ e $\mathrm{CvM}_2(1) = \mathrm{CvM}_2$.

Funções características das distribuições $\mathrm{CvM}(k)$ e $\mathrm{CvM}_2(k)$

Se $X_1 \sim \chi^2(m)$, $X_2 \sim \chi^2(n)$ e X_1 é independente de X_2, então $X_1 + X_2 \sim \chi^2(m+n)$. Portanto, de (2.20), tem-se que uma v.a. com distribuição $\mathrm{CvM}(k)$ equivale à soma de k v.a.i.i.d.s com distribuição $\mathrm{CvM}(1)$ cada. A função característica da distribuição $\mathrm{CvM}(k)$ será dada então pelo produto de k funções características da distribuição $\mathrm{CvM}(1)$:

$$\varphi(t) = \left(\frac{\sqrt{2it}}{\operatorname{sen}\sqrt{2it}}\right)^{\frac{k}{2}} \tag{2.23}$$

Analogamente, a função característica da distribuição $\mathrm{CvM}_2(k)$ será:

$$\begin{aligned} \varphi(t) &= [D_2(2it)]^{-\frac{k}{2}} \\ D_2(x) &\equiv \frac{12}{x^2}\left(2 - \sqrt{x}\sin\sqrt{x} - 2\cos\sqrt{x}\right) \end{aligned} \tag{2.24}$$

2.5 Teste invariante localmente ótimo - ILO

2.5.1 Teste localmente ótimo

Seja Y um vetor aleatório com densidade $f(y|\theta)$, θ é um parâmetro sobre o qual deseja-se inferir a partir do valor observado de Y. Quando não há um teste uniformemente mais poderoso para um teste de hipóteses unilateral do tipo:

$$\begin{aligned} H_0 &: \theta \leq \theta_0 \\ H_a &: \theta > \theta_0 \end{aligned} \tag{2.25}$$

pode-se, seguindo Ferguson (1967), optar pelo teste localmente ótimo de Neyman e Pearson. Esse teste é localmente ótimo no sentido em que sua função poder possui a maior inclinação (derivada) em θ_0, entre todos os testes com mesmo tamanho em θ_0.

Seja $\phi(Y)$ um teste para (2.25). A função poder do teste $\phi(Y)$ é dada por

$$\beta_\phi(\theta) = E_\theta \phi(Y) \tag{2.26}$$

E sua derivada, supondo que se possa trocar a ordem entre derivação e integração,

$$\beta'_\phi(\theta) = \int \phi(y) \frac{\partial}{\partial \theta} f(y|\theta) dy \tag{2.27}$$

Assim, o teste ϕ_0 será localmente ótimo se, para todo teste ϕ de mesmo tamanho, ou seja, $\beta_\phi(\theta_0) = \beta_{\phi_0}(\theta_0)$, tem-se

$$\beta'_\phi(\theta_0) \leq \beta'_{\phi_0}(\theta_0) \tag{2.28}$$

Usando o lema de Neyman-Pearson, chega-se ao teste localmente ótimo ϕ_0 que rejeita H_0 se

$$\frac{\partial}{\partial \theta} \log f(y|\theta) \Big|_{\theta_0} > c_\alpha \tag{2.29}$$

c_α é uma constante que depende do nível de significância $\alpha = \beta_{\phi_0}(\theta_0)$.

2.5.2 Teste invariante

Seja $\mathcal{Y}$ o espaço amostral do vetor aleatório Y. Seja $\mathcal{G}$ um grupo de transformações (fechado, com inversa e identidade) biunívocas de $\mathcal{Y}$ em $\mathcal{Y}$. Um teste, ou uma regra de decisão, $\phi(Y)$ para testar hipóteses como (2.25) é invariante sob uma transformação $g \in \mathcal{G}$ se

$$\phi(gY) = \phi(Y), \ \forall g \in \mathcal{G} \tag{2.30}$$

Chamamos de órbita o conjunto $\{Y : Y = gX \quad g \in \mathcal{G}\}$, ϕ será então invariante se e somente se ϕ é constante em cada órbita de $\mathcal{G}$.

Um invariante maximal é uma função $M : \mathcal{Y} \to \mathbf{R}^k$, para algum $k \geq 1$, que assume valor constante em cada órbita de $\mathcal{G}$ e valores distintos para diferentes órbitas.

Teorema 2.1. *Seja M um invariante maximal em relação ao grupo de transformações $\mathcal{G}$. Então o teste ϕ é invariante se e somente se existe uma função h tal que $\phi(Y) = h(M(Y))$ para todo $Y \in \mathcal{Y}$.*

Prova:

$' \Leftarrow '$ Se $\phi(Y) = h(M(Y))$, então ϕ é invariante, pois

$$\phi(gY) = h(M(gY)) = h(M(Y)) = \phi(Y) \tag{2.31}$$

$' \Rightarrow ' \ M(Y_1) = M(Y_2)$ implica que Y_1 e Y_2 estão na mesma órbita e, portanto, $gY_1 = Y_2$ para algum $g \in \mathcal{G}$. Se ϕ é invariante, então $\phi(Y_2) = \phi(gY_1) = \phi(Y_1)$. Segue então que ϕ é função de M, pois $M(Y_1) = M(Y_2) \Rightarrow \phi(Y_2) = \phi(Y_1)$ e $M(Y_1) \neq M(Y_2) \Rightarrow \phi(Y_2) \neq \phi(Y_1)$ □

Portanto, se, em vez da distribuição de Y, utilizarmos a distribuição do invariante maximal $M(Y)$ na obtenção de um teste $\phi(Y)$, esse teste será invariante sob o respectivo grupo de transformações.

O teste invariante e localmente ótimo (ILO ou LBI de *locally best invariant*) será então o teste localmente ótimo de Neyman-Pearson (2.29), utilizando-se a f.d.p. de $M(Y)$ em vez da f.d.p. de Y.

2.5.3 Teste ILO para o modelo linear normal

Apresentamos aqui os resultados obtidos por King & Hillier (1985) para o modelo linear normal:

$$\begin{aligned} \mathbf{y} &= \boldsymbol{X}\boldsymbol{\beta} + \mathbf{u} \\ \mathbf{u} &\sim \mathcal{N}_n(\mathbf{0}, \sigma^2\boldsymbol{\Omega}(\theta)) \end{aligned} \tag{2.32}$$

$\mathbf{y}$ é um vetor aleatório $n \times 1$ e $\boldsymbol{X}$ é uma matriz $n \times k$ fixa de regressores. Os parâmetros do modelo são o vetor $\boldsymbol{\beta}_{k\times 1}$ e os escalares σ^2 e θ por meio da forma específica da matriz de covariâncias: $\sigma^2\boldsymbol{\Omega}(\theta)$, com $\boldsymbol{\Omega}(\mathbf{0}) = \boldsymbol{I}_n$, a matriz identidade de dimensão n.

O teste de hipóteses unilateral

$$\begin{aligned} H_0 &: \theta = 0 \\ H_a &: \theta > 0 \end{aligned} \tag{2.33}$$

é invariante sob o grupo de transformações

$$g(\mathbf{y}) = \gamma_0\mathbf{y} + \mathbf{X}\tilde{\gamma} \tag{2.34}$$

com γ_0 sendo um escalar positivo e $\tilde{\gamma}$ um vetor $k \times 1$.

O vetor $\mathbf{v} \in \mathbb{R}^{n-k}$:

$$\mathbf{v} = \frac{\mathbf{P}_1\mathbf{z}}{||\mathbf{P}_1\mathbf{z}||} \tag{2.35}$$

é um invariante maximal em relação a (2.34), $\mathbf{z}$ são os resíduos do ajuste de mínimos quadrados ordinário de (2.32) e $\mathbf{P}_1$ projeta $\mathbf{z}$ em $\mathbb{R}^{n-k}$, $\mathbf{P}_1$ é tal que $\mathbf{P}_1\mathbf{P}_1' = \mathbf{I}_{n-k}$ e

$\mathbf{P}_1'\mathbf{P}_1 = \mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

Pode-se então obter o teste ILO a partir de (2.29) e da densidade de $\mathbf{v}$:

$$\frac{\partial}{\partial\theta}\log f(\mathbf{v}|\theta)\Big|_{\theta=0} > c_1 \tag{2.36}$$

Assim, chega-se ao estimador correspondente ξ, com a seguinte região crítica:

$$\begin{aligned} \xi &= \frac{\mathbf{z}'\mathbf{A}_0\mathbf{z}}{\mathbf{z}'\mathbf{z}} < c \\ \mathbf{A}_0 &\equiv \frac{\partial}{\partial\theta}\mathbf{\Omega}(\theta)^{-1}\Big|_{\theta=0} \end{aligned} \tag{2.37}$$

Capítulo 3

Variância de longo termo

3.1 A variância de longo termo - σ_L^2

Os testes utilizados neste trabalho (cap. 4) foram obtidos primeiramente para séries com observações independentes (ruído branco). A fim de estendê-los para séries temporais com alguma estrutura de dependência, seguiremos o trabalho de Kwiatkowski *et al.* (1992). Nessa linha, tem-se que os testes obtidos para séries independentes continuam válidos assintoticamente para séries com certa estrutura de dependência se substituirmos a variância amostral por uma estimativa da variância de longo termo σ_L^2, definida na sequência.

Assim, este capítulo dedica-se à obtenção de uma estimativa da variância de longo termo.

Definição 3.1 (Variância de longo termo). *Seja* $\boldsymbol{\varepsilon} = \{\varepsilon_t\}_{t=1,\ldots,T}$ *uma série temporal estacionária* $\mathcal{EST}(0, \gamma(h))$, *ou seja,* $\forall t \in \{1, \ldots, T\}$,

1. $\mathrm{E}\,\varepsilon_t = 0$
2. $\mathrm{Var}\,\varepsilon_t = \sigma_\varepsilon^2 = \gamma(0) < \infty$
3. $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t+|h|}) = \gamma(h)$

3.1 A variância de longo termo - σ_L^2

A variância de longo termo do processo ε é definida por:

$$\sigma_L^2 = \lim_{T\to\infty} \frac{1}{T} \operatorname{Var}\left(\sum_{t=1}^{T} \varepsilon_t\right) = \lim_{T\to\infty} T \operatorname{Var} \bar{\varepsilon}\ , \quad \bar{\varepsilon} = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \tag{3.1}$$

Se $\varepsilon_t \sim \mathrm{RB}\,(0, \sigma^2)$, então

$$\begin{aligned} \sigma_L^2 &= \lim_{T\to\infty} \frac{1}{T} \operatorname{Var}\left(\sum_{t=1}^{T} \varepsilon_t\right) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \operatorname{Var} \varepsilon_t \\ &= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sigma^2 = \lim_{T\to\infty} \frac{1}{T}\, T\sigma^2 = \sigma^2 \end{aligned} \tag{3.2}$$

Portanto, no caso de ruído branco, a variância de longo termo de ε pode ser estimada consistentemente pela variância amostral

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} e_t^2 = \frac{\boldsymbol{e}'\boldsymbol{e}}{T}\ , \qquad e_t = \varepsilon_t - \bar{\varepsilon} \tag{3.3}$$

A relação (3.2) deixa de valer em geral quando ε possui alguma estrutura de autocorrelação. Para obter uma expressão mais geral para σ_L^2 utilizaremos o lema a seguir.

Lema de Kronecker. *Se uma sequência $\{a_j\}_{j=0,1,\ldots}$ é tal que $\sum_{j=0}^{\infty} |a_j| < \infty$, então*

$$\lim_{n\to\infty} \sum_{j=0}^{n} \frac{j}{n} |a_j| = 0 \tag{3.4}$$

Prova: Dado $\varepsilon > 0$, existe um $N < \infty$ tal que $\sum_{j=N+1}^{\infty} |a_j| < \varepsilon$. Portanto, para $n > N$,

$$\sum_{j=0}^{n} \frac{j}{n} |a_j| = \frac{1}{n}\sum_{j=0}^{N} j|a_j| + \sum_{j=N+1}^{n} \frac{j}{n}|a_j| < \frac{1}{n}\sum_{j=0}^{N} j|a_j| + \sum_{j=N+1}^{n} |a_j| < \frac{1}{n}\sum_{j=0}^{N} j|a_j| + \varepsilon$$

Como, para N fixo, $\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{N} j|a_j| = 0$, então

$$\lim_{n\to\infty} \sum_{j=0}^{n} \frac{j}{n} |a_j| < \varepsilon\ , \quad \forall \varepsilon > 0$$

□

3.1 A variância de longo termo - σ_L^2

Proposição 3.1. *Se* $\varepsilon_t \sim \mathcal{EST}(0, \gamma(h))$, *com* $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, *então*

$$\sigma_L^2 = \sum_{h=-\infty}^{\infty} \gamma(h)\,, \tag{3.5}$$

Prova:

$$\begin{aligned}
\sigma_L^2 &= \lim_{T\to\infty} \frac{1}{T}\left[\sum_{t=1}^{T} \operatorname{Var}\varepsilon_t + \sum_{t=1}^{T}\sum_{t\neq r=1}^{T} \operatorname{Cov}(\varepsilon_t, \varepsilon_r)\right] \\
&= \lim_{T\to\infty} \frac{1}{T}\left[\sum_{t=1}^{T} \gamma(0) + \sum_{t=1}^{T}\sum_{t\neq r=1}^{T} \gamma(t-r)\right] \\
&= \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\sum_{r=1}^{T} \gamma(t-r) \\
&= \lim_{T\to\infty} \frac{1}{T}\sum_{h=-T+1}^{T-1} (T-|h|)\,\gamma(h)\,, \quad \text{pois } \gamma(h) = \gamma(-h) \\
&= \lim_{T\to\infty} \sum_{h=-T+1}^{T-1} \left(1 - \frac{|h|}{T}\right)\gamma(h) \qquad (3.6)
\end{aligned}$$

Agora,

$$\begin{aligned}
\left|\lim_{T\to\infty} \sum_{h=-T+1}^{T-1} \frac{|h|}{T}\gamma(h)\right| &= \left|2\lim_{T\to\infty} \sum_{h=1}^{T-1} \frac{h}{T}\gamma(h)\right| \leq \\
&\leq 2\lim_{T\to\infty} \sum_{h=1}^{T-1} \frac{h}{T}|\gamma(h)| = 0\,, \text{ pelo lema de Kronecker}
\end{aligned}$$

E, de (3.6),

$$\sigma_L^2 = \sum_{h=-\infty}^{\infty} \gamma(h)$$

□

Se $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, então ε possui função densidade espectral $f(\lambda)$, definida pela tranformada de Fourier da função de autocovariância $\gamma(h)$:

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-ih\lambda} \tag{3.7}$$

e, de (3.5),

$$\sigma_L^2 = 2\pi f(0) \tag{3.8}$$

Sob a hipótese de que ε é estacionária com densidade espectral $f(\lambda)$, tem-se que σ_L^2 pode ser estimada consistentemente por um estimador consistente de $2\pi f(0)$. Será utilizado então um estimador suavizado de covariâncias para $f(\lambda)$:

$$\hat{f}(\lambda) = \frac{1}{2\pi} \sum_{|h|<T} w(h;l)\hat{\gamma}(h)e^{-i\lambda h} \tag{3.9}$$

- o núcleo $w(h;l)$ é uma função par em relação a h, com $|w(h;l)| \leq w(0,l) = 1$.
- l é um parâmetro positivo que indica a largura de faixa do núcleo $w(h;l)$.
- $\hat{\gamma}(h)$ é o estimador consistente usual de $\gamma(h)$, a covariância amostral:

$$\hat{\gamma}(h) = \frac{1}{T} \sum_{t=|h|}^{T-1} e_t e_{t-|h|} \tag{3.10}$$

e $\hat{\sigma}_L^2 = 2\pi\hat{f}(0)$ será o estimador para $\sigma_L^2 = 2\pi f(0)$:

$$\hat{\sigma}_L^2 = \sum_{|h|<T} w(h;l)\hat{\gamma}(h) \tag{3.11}$$

A consistência de $\hat{\sigma}_L^2$ será assegurada se $l \to \infty$ quando $T \to \infty$, mas com $l = o(T)$.

As escolhas do núcleo $w(h;l)$ e da largura de faixa l podem ser feitas de inúmeras formas seguindo-se apenas as restrições básicas acima. Pode-se chegar a escolhas ótimas, em algum sentido, com a imposição de outras propriedades desejáveis. Considerando núcleos da forma em que l é um parâmetro de escala, ou seja, $w(h;l) = w(h/l)$, chega-se, conforme Andrews (1991), ao núcleo espectral quadrático $w_{QS}(h/l)$ e à respectiva largura ótima l^* minimizando-se uma expressão assintótica envolvendo o erro quadrático médio

$$\mathrm{EQM}(\hat{\sigma}_L^2) = \mathrm{E}(\hat{\sigma}_L^2 - \sigma_L^2)^2 = [b(\hat{\sigma}_L^2)]^2 + \mathrm{Var}\,\hat{\sigma}_L^2 \tag{3.12}$$

3.1 A variância de longo termo - σ_L^2

Seguindo Parzen (1957), tem-se que o viés $b(.)$ de um estimador suavizado, como (3.11), depende da suavidade da função peso $w(x)$ em $x = 0$ e da suavidade da densidade espectral $f(\lambda)$ em $\lambda = 0$. Assim, definimos abaixo os números k_q e f_q, relacionados a essas suavidades, que serão utilizados nos resultados a seguir.

A suavidade da função peso $w(x)$ em $x = 0$ pode ser avaliada pelo valor de q na função

$$k_q = \lim_{x \to 0} \frac{1 - w(x)}{|x|^q}, \qquad q = 0, 1, \ldots \tag{3.13}$$

O maior q para o qual $k_q < \infty$, denominado expoente característico da função peso $w(x)$, é o valor de q que minimiza o EQM assintótico. Isso porque, conforme argumentado em Priestley (1981), a expressão que minimiza o EQM assintótico é $O(T^{-\frac{2q}{2q+1}})$, ver a proposição 3.2 abaixo. Tem-se também que quanto maior o valor do expoente característico, mais suave é a função peso na origem.

A suavidade da densidade espectral $f(\lambda)$ em $\lambda = 0$ pode ser indexada por

$$f_q = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} |h|^q \gamma(h), \qquad q = 0, 1, \ldots \tag{3.14}$$

$|f_q| < \infty$ se e somente se $f(\lambda)$ é diferenciável q vezes na frequência $\lambda = 0$.

Utilizaremos ainda as seguintes classes de funções peso:

$$\begin{aligned}
\mathcal{K}_1 &= \{\, w(.) : \mathbb{R} \to [-1, 1] \,|\, w(0) = 1, w(x) = w(-x)\ \forall x \in \mathbb{R}, \\
&\quad \int_{-\infty}^{\infty} w^2(x)dx < \infty,\ w(.) \text{ é contínua em zero e descontínua,} \\
&\quad \text{no máximo, num número finito de pontos } \} \\
\mathcal{K}_2 &= \{\, w(.) \in \mathcal{K}_1 \,|\, W(\lambda) \geq 0\ \forall \lambda \in \mathbb{R} \,\}, \qquad W(\lambda) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} w(x) e^{-ix\lambda} dx
\end{aligned}$$

Condição 3.1. *$\{\varepsilon_t\}$ é uma série estacionária de quarta ordem, com média zero, $\sum_{h=-\infty}^{\infty} \gamma(h) < \infty$ e $\sum_{j=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \kappa(0, j, m, n) < \infty$. Em que $\kappa(t, t+j, t+m, t+n)$ é o cumulante de quarta ordem de $(\varepsilon_t, \varepsilon_{t+j}, \varepsilon_{t+m}, \varepsilon_{t+n})$.*

Proposição 3.2 (Parzen (1957)). *Sob a condição 3.1, com $w(.) \in \mathcal{K}_1$, $l \to \infty$ e $l = o(T)$, tem-se*

(i) $\lim_{T\to\infty} \frac{T}{l} \operatorname{Var} \hat{\sigma}_L^2 = 2(\sigma_L^2)^2 \int_{-\infty}^{\infty} w^2(x)dx$

(ii) *Se* $\frac{l^q}{T} \to 0$ *para algum* $q < \infty$, *tal que* k_q, $f_q < \infty$, *então*

$$\lim_{T\to\infty} l^q(\mathrm{E}\,\hat{\sigma}_L^2 - \sigma_L^2) = -2\pi k_q f_q$$

(iii) *Se* $\frac{l^{2q+1}}{T} \to \alpha \in (0,\infty)$ *para algum* $q < \infty$, *tal que* k_q, $f_q < \infty$, *então*

$$\lim_{T\to\infty} \frac{T}{l} \mathrm{EQM}(\hat{\sigma}_L^2) = \frac{4\pi^2 k_q^2 f_q^2}{\alpha} + 2(\sigma_L^2)^2 \int_{-\infty}^{\infty} w^2(x)dx \tag{3.15}$$

Minimizando (3.15) em relação a l, obtém-se uma expressão para a largura ótima l^*:

$$l^* = \left[\frac{4\pi^2 q k_q^2 f_q^2 T}{(\sigma_L^2)^2 \int_{-\infty}^{\infty} w^2(x)dx}\right]^{\frac{1}{2q+1}} \tag{3.16}$$

De Epanechnikov (1969), ou Priestley (1981), tem-se que o núcleo em $\mathcal{K}_2$ que minimiza (3.15) é o núcleo espectral quadrático que, seguindo a normalização de Andrews(1991) é dado por

$$w_{QS}(h/l) = \frac{25l^2}{12\pi^2 h^2}\left[\frac{\operatorname{sen}(6\pi h/5l)}{6\pi h/5l} - \cos(6\pi h/5l)\right] \tag{3.17}$$

$q = 2$ para o núcleo w_{QS}, pois esse é o maior q tal que $k_q < \infty$. Assim,

$$k_2 = \lim_{x\to 0} \frac{1 - w_{QS}(x)}{x^2} = -\frac{1}{2!}\frac{d^2}{dx^2} w_{QS}(x)\Big|_{x=0} = \frac{18\pi^2}{125} \tag{3.18}$$

w_{QS} dado por (3.17) já se encontra normalizado, de forma que $\int_{-\infty}^{\infty} w^2(x)dx = 1$. Portanto, de (3.16) obtém-se a largura ótima para w_{QS}:

$$l^* = \left[\frac{8\pi^2 k_2^2 f_2^2 T}{(\sigma_L^2)^2}\right]^{\frac{1}{5}} \tag{3.19}$$

3.2 Metodologia proposta para a estimação de σ_L^2

Para a estimação de $\sigma_L^2 = 2\pi f(0)$ utilizaremos o método *plug-in* que, conforme Andrews (1991) argumenta, é o método mais adequado, até então proposto, para uma seleção automática da largura de faixa l baseada nos dados, para a estimação em uma frequência pontual, no caso em $\lambda = 0$.

O método *plug-in* consiste em obter um estimador $\widehat{l^*}$ para l^*, eq. (3.19), pela substituição das quantidades desconhecidas σ_L^2 e $f_2 = -f''(0)$ por seus respectivos estimadores:

$$\widehat{l^*}(\hat{\sigma}_L^2, \widehat{f}_2) = \left[\frac{8\pi^2 k_2^2 \widehat{f}_2^{\,2} T}{(\hat{\sigma}_L^2)^2}\right]^{\frac{1}{5}} \tag{3.20}$$

Substituindo (3.20) em (3.11), obtém-se

$$\hat{\sigma}_L^2 = \sum_{|h|<T} w_{QS}(h/\widehat{l^*}(\hat{\sigma}_L^2, \widehat{f}_2))\frac{1}{T}\sum_{t=|h|}^{T-1} e_t e_{t-|h|} \tag{3.21}$$

Nossa proposta é então obter o estimador $\hat{\sigma}_L^2$ resolvendo numericamente a equação acima.

Resta o problema de se encontrar um estimador para f_2. Algumas sugestões são:

1. Ajustar um modelo ARMA(p,q) para $\{e_t\}$: $\Phi(L)e_t = \Theta(L)a_t$, $a_t \sim \mathrm{RB}\,(0, \sigma_a^2)$ e utilizar os parâmetros estimados para obter uma estimativa de $f''(0)$, utilizando a relação

$$f(\lambda) = \frac{\sigma_a^2}{2\pi}\frac{|\Theta(e^{-i\lambda})|^2}{|\Phi(e^{-i\lambda})|^2} \tag{3.22}$$

2. Ajustar localmente um polinômio em torno de $\lambda = 0$ e utilizar o dobro do coeficiente quadrático como estimador de $f''(0)$.

3. Utilizar novamente o método *plug-in* na relação (3.14), substituindo $\gamma(h)$ por $\hat{\gamma}(h)$:

$$\widehat{f_2} = \frac{1}{2\pi} \sum_{|h|<T} h^2 \hat{\gamma}(h) \tag{3.23}$$

A sugestão 1, apresentada por Andrews (1991), não se mostrou muito adequada, pois as estimativas obtidas apresentam grandes variações para pequenas alterações dos parâmetros estimados do modelo ARMA(p,q).

Na sugestão 2, um bom ajuste, e uma consequente boa estimativa de f_2, ainda é muito dependente dos dados, pois o grau do melhor polinômio ajustado depende da forma de $\hat{f}(\lambda)$ em torno de $\lambda = 0$.

A sugestão 3 é apresentada em Newey & West (1994) para a estimativa de f_2, mas também para uma primeira estimativa de σ_L^2 num procedimento de duas etapas. Aqui utilizamos um procedimento distinto que apresenta estimativas finais razoavelmente estáveis.

Propomos uma variação da sugestão 3, utilizando novamente um estimador suavizado de covariâncias com o núcleo $w_{QS}(x)$:

$$\widehat{f_2} = \frac{1}{2\pi} \sum_{|h|<T} h^2 w_{QS}(h/l_2) \hat{\gamma}(h) \tag{3.24}$$

O problema agora volta a ser o original de se obter uma largura de faixa adequada, que denotaremos por l_2. Nesse estágio a estimação não é tão crítica, pois o resultado final é pouco afetado pela escolha do núcleo ou por pequenas variações da largura l_2, conforme argumentado, por exemplo, em Newey & West (1994) ou em Sheather & Jones (1991) no contexto de estimação de funções de densidade. Nesse sentido, utilizaremos um procedimento iterativo para obter, de forma aproximada, a menor estimativa $\widehat{l_2}$ para a qual $\widehat{l_2} \cong \widehat{l^*}$.

Uma idéia que surge é obter o $\widehat{l_2}$ que iguala $\widehat{l^*}$ resolvendo numericamente a equação (3.21) com $\widehat{f_2}$ dado por (3.24). No entanto, essa opção não se mostra viável, pois a solução numérica em geral não é única e por vezes não converge. Provavelmente, devido à alta volatilidade introduzida pelo termo h^2 em (3.24). Vale ressaltar que, nos dados reais analisados neste trabalho e nas simulações realizadas, o cálculo numérico de (3.21) sempre apresentou solução única e rápida convergência.

3.3 Algoritmo para a estimação de σ_L^2

O algoritmo proposto e utilizado para a estimação de σ_L^2 é então o seguinte:

1. Inicia-se com $\widehat{l_2} = 0$.

2. Calcula-se $\widehat{f_2}$ pela equação (3.24), com o núcleo w_{QS}:

$$\widehat{f_2} = \frac{1}{2\pi} \sum_{|h|<T} h^2 w_{QS}(h/\widehat{l_2})\hat{\gamma}(h)$$

3. Calcula-se $\hat{\sigma}_L^2$, equação (3.21), numericamente pelo algoritmo de Newton-Raphson:

$$\begin{aligned} g(\hat{\sigma}_L^2) &\equiv \hat{\sigma}_L^2 - \sum_{|h|<T} w_{QS}(h; \widehat{l^*}(\hat{\sigma}_L^2, \widehat{l_2})) \frac{1}{T} \sum_{t=|h|}^{T-1} e_t e_{t-|h|} \\ (\hat{\sigma}_L^2)_0 &= \frac{1}{T} \sum_{t=1}^{T} e_t^2 \\ (\hat{\sigma}_L^2)_{i+1} &= (\hat{\sigma}_L^2)_i - \frac{g((\hat{\sigma}_L^2)_i)}{g'((\hat{\sigma}_L^2)_i)} \end{aligned}$$

4. Estima-se l^* pelo seu estimador *plug-in* $\widehat{l^*}$, dado por (3.20):

$$\widehat{l^*} = \left[\frac{8\pi^2 k_2^2 \widehat{f_2}^2 T}{(\hat{\sigma}_L^2)^2} \right]^{\frac{1}{5}} \tag{3.25}$$

5. Incrementa-se $\widehat{l_2}$ e repete-se o procedimento a partir do item 2 até que $\widehat{l_2} \cong \widehat{l^*}$. O critério adotado, com bons resultados, foi $|\widehat{l^*} - \widehat{l_2}| < 0.1$

Capítulo 4

Testes de estacionariedade

Considere o modelo de componentes não-observados:

$$y_t = y_t^n + y_t^a, \qquad t = 1, \ldots, T \tag{4.1}$$

em que $\mathbf{y} = \{y_t\}$ é a série temporal de tamanho T a ser analisada, $\mathbf{y}^n = \{y_t^n\}$ é um componente não-aleatório e $\mathbf{y}^a = \{y_t^a\}$ é um componente aleatório (estocástico) a serem modelados.

Se $y^n = \mu_0 + \beta t$ é uma tendência linear e $y^a = \mu_t + \varepsilon_t$ é a soma de um passeio aleatório $\{\mu_t\}$ com um processo estacionário $\{\varepsilon_t\}$, tem-se, na forma estrutural, o modelo com tendência linear (MTL):

$$\begin{array}{rcll} y_t & = & \mu_t + \beta t + \varepsilon_t & , \ \varepsilon_t \sim \mathcal{EST}(0, \gamma_\varepsilon) \\ \mu_t & = & \mu_{t-1} + \eta_t & , \ \eta_t \sim \mathrm{RB}\,(0, \sigma_\eta^2) \\ \varepsilon_t & \text{e} & \eta_s \text{ independentes} & , \ \forall\, t, s = 1, \ldots, T \end{array} \tag{4.2}$$

Para testar se o componente aleatório $\mathbf{y}^a$ do modelo acima é estacionário, pode-se utilizar o seguinte teste de hipóteses:

$$\begin{array}{rcl} H_0 & : & \sigma_\eta^2 = 0 \\ H_1 & : & \sigma_\eta^2 > 0 \end{array} \tag{4.3}$$

Assim, sob a hipótese nula H_0, μ_t permanece constante e igual a μ_0 e $\mathbf{y}^a$ reduz-se ao processo estacionário $\{\varepsilon_t\}$. Sob a hipótese alternativa, μ_t segue um passeio aleatório e, portanto, $\mathbf{y}^a$ passa a ter um componente não-estacionário, no caso uma raiz unitária.

4.1 Série com nível constante

Consideremos o modelo (4.2) sem tendência ($\beta = 0$) e ε_t e η_t como sendo ruídos brancos gaussianos, ou seja,

$$\begin{array}{rcll} y_t & = & \mu_t + \varepsilon_t & , \ \varepsilon_t \sim \mathcal{N}\mathrm{id}\,(0, \sigma_\varepsilon^2) \\ \mu_t & = & \mu_{t-1} + \eta_t & , \ \eta_t \sim \mathcal{N}\mathrm{id}\,(0, \sigma_\eta^2) \\ \varepsilon_t & \text{e} & \eta_s \text{ independentes} & , \ \forall\, t, s = 1, \ldots, T \end{array} \tag{4.4}$$

Comparando (4.4) ao modelo linear normal (2.32) tem-se, com $\theta = \dfrac{\sigma_\eta^2}{\sigma_\varepsilon^2}$,

$$\mathbf{\Omega}(\theta) = \sigma_\varepsilon^2 \begin{bmatrix} 1+\theta & \theta & \theta & \cdots & \theta \\ \theta & 1+2\theta & 2\theta & \cdots & 2\theta \\ \theta & 2\theta & 1+3\theta & \cdots & 3\theta \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta & 2\theta & 3\theta & \cdots & 1+T\theta \end{bmatrix} \tag{4.5}$$

e

$$\mathbf{A}_0 = \left.\frac{\partial \mathbf{\Omega}(\theta)^{-1}}{\partial \theta}\right|_{\theta=0} = -\sigma_\varepsilon^2 \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & T \end{bmatrix} \tag{4.6}$$

Portanto, de (2.37), obtém-se o teste invariante localmente ótimo ILO para o modelo com nível constante (4.4):

$$\xi_\mu = \frac{\mathbf{e}'\mathbf{A}_0\mathbf{e}}{\mathbf{e}'\mathbf{e}} = \frac{\sum_{i=1}^{T}\left(\sum_{t=1}^{i} e_t\right)^2}{T^2\hat{\sigma}_\varepsilon^{\,2}} > c_\alpha \tag{4.7}$$

c_α é uma constante que depende do nível de significância α e e_t são os resíduos do ajuste de mínimos quadrados de y_t a um nível constante:

$$e_t = y_t - \bar{y}\ , \quad t = 1, \ldots, T \tag{4.8}$$

Isso porque, sob H_0, (4.4) se reduz a $y_t = \mu_0 + \varepsilon_t$.

Pela aplicação do teorema de Donsker (ver, por exemplo, Shorack & Wellner (1986)) tem-se que a soma parcial dos resíduos e_t, devidamente normalizada, converge fracamente a uma ponte browniana $B(r)$,

$$\frac{1}{\sigma_\varepsilon \sqrt{T}} \sum_{i=1}^{[Tr]} e_i \Rightarrow B(r) = W(r) - rW(1), \quad r \in [0,1] \tag{4.9}$$

e, assintoticamente,

$$\xi_\mu = \sum_{i=1}^{T} \left(\frac{\sum_{t=1}^{i} e_t}{T\hat{\sigma}_\varepsilon} \right)^2 \Rightarrow \int_0^1 B(r)^2 dr \tag{4.10}$$

que corresponde à distribuição de Cramér-von Mises.

4.2 Série com tendência linear

O acréscimo de uma tendência não-aleatória ao modelo (4.4) não altera a matriz $\Omega(\theta)$ em (4.5) e, consequentemente, o teste ILO continua a ser o dado por (4.7), porém a sua distribuição assintótica é alterada.

Considerando então o modelo com tendência linear (4.2), com ε_t e η_t sendo ruídos brancos gaussianos,

$$\begin{array}{lcll} y_t & = & \mu_t + \beta t + \varepsilon_t & , \ \varepsilon_t \sim \mathcal{N}\mathrm{id}\,(0, \sigma_\varepsilon^2) \\ \mu_t & = & \mu_{t-1} + \eta_t & , \ \eta_t \sim \mathcal{N}\mathrm{id}\,(0, \sigma_\eta^2) \\ \varepsilon_t & \text{e} & \eta_s \text{ independentes} & , \ \forall\, t, s = 1, \ldots, T \end{array} \tag{4.11}$$

o teste ILO será

$$\xi_\beta = \frac{\mathbf{e}'\mathbf{A}_0\mathbf{e}}{\mathbf{e}'\mathbf{e}} = \frac{\sum_{i=1}^{T}\left(\sum_{t=1}^{i} e_t\right)^2}{T^2\hat{\sigma}_\varepsilon^2} > c_\alpha \tag{4.12}$$

em que e_t são os resíduos do ajuste de mínimos quadrados de y_t a uma reta:

$$e_t = y_t - \hat{y}_t\ , \quad t = 1, \ldots, T \tag{4.13}$$

$$\hat{y}_t = \hat{\mu} + \hat{\beta}t \tag{4.14}$$

De MacNeill (1978), tem-se que a soma parcial desses resíduos converge para o funcional ponte browniana de segundo nível (B_2) e, por conseguinte,

$$\xi_\beta \Rightarrow \int_0^1 B_2(r)^2 dr = \mathrm{CvM}_2 \tag{4.15}$$

4.3 Série com quebra no nível e na tendência

Busetti & Harvey (2001) estenderam o modelo acima de forma a permitir quebras estruturais no nível e/ou na tendência. Nessa linha, consideremos o modelo com uma quebra simultânea no nível e na tendência:

$$\begin{aligned} y_t &= \mu_t + \beta t + \delta_\mu w_t + \delta_\beta(w_t t) + \varepsilon_t &&, \quad \varepsilon_t \sim \mathcal{N}\mathrm{id}\,(0, \sigma_\varepsilon^2) \\ \mu_t &= \mu_{t-1} + \eta_t &&, \quad \eta_t \sim \mathcal{N}\mathrm{id}\,(0, \sigma_\eta^2) \\ \varepsilon_t &\ \text{e}\ \eta_s \ \text{independentes} &&, \quad \forall\, t, s = 1, \ldots, T \end{aligned} \tag{4.16}$$

em que $w_t = \mathbf{1}_{\{t>\tau\}}$ e τ é o instante, conhecido, em que ocorre uma mudança estrutural.

O teste ILO continua a ser dado pela mesma estatística:

$${}_1\xi_\beta(\lambda) = \frac{\sum_{i=1}^{T}\left(\sum_{t=1}^{i} e_t\right)^2}{T^2\hat{\sigma}_\varepsilon^2} > c_\alpha \tag{4.17}$$

cuja distribuição assintótica agora depende de $\lambda = \tau/T$. O índice 1 em ${}_1\xi_\beta$ denota a presença de uma quebra.

Pela Proposição 3.2 de Busetti & Harvey (2001) tem-se que

$$
\begin{aligned}
{}_1\xi_\beta(\lambda) &= \lambda^2 \frac{\sum_{i=1}^{\tau}\left(\sum_{t=1}^{i} e_t\right)^2}{\tau^2\hat{\sigma}_\varepsilon^2} + (1-\lambda)^2 \frac{\sum_{i=\tau+1}^{T}\left(\sum_{t=\tau+1}^{i} e_t\right)^2}{(T-\tau)^2\hat{\sigma}_\varepsilon^2} && (4.18)\\
{}_1\xi_\beta(\lambda) &\Rightarrow \lambda^2 \int_0^1 B_2(r)^2 dr + (1-\lambda)^2 \int_0^1 B_2'(r)^2 dr && (4.19)
\end{aligned}
$$

em que $B_2(r)$ e $B_2'(r)$ são pontes brownianas de segundo nível independentes.

Portanto, a distribuição assintótica de ${}_1\xi_\beta(\lambda)$ é uma combinação linear de duas distribuições CvM_2. De (2.17), a função característica correspondente será dada por

$$
\begin{aligned}
\varphi_{{}_1\xi_\beta}(t) &= [D_2(2it\lambda^2)D_2(2it(1-\lambda)^2)]^{-\frac{1}{2}} && (4.20)\\
D_2(x) &\equiv \frac{12}{x^2}\left(2 - \sqrt{x}\,\mathrm{sen}\,\sqrt{x} - 2\cos\sqrt{x}\right)
\end{aligned}
$$

Teste simplificado para mais de uma quebra

O procedimento acima para o cálculo da região crítica torna-se bastante trabalhoso se pretende-se modelar uma série com mais de uma quebra. Busetti & Harvey (2001) propõem então um teste simplificado: eles constataram que a estatística (4.18) sem o parâmetro λ

$$
{}_1\xi_\beta^* = \frac{\sum_{i=1}^{\tau}\left(\sum_{t=1}^{i} e_t\right)^2}{\tau^2\hat{\sigma}_\varepsilon^2} + \frac{\sum_{i=\tau+1}^{T}\left(\sum_{t=\tau+1}^{i} e_t\right)^2}{(T-\tau)^2\hat{\sigma}_\varepsilon^2} \qquad (4.21)
$$

apresenta tamanho e poder semelhantes para $0.1 \leq \lambda \leq 0.9$. Além disso, ${}_1\xi_\beta^*$ sugere uma generalização para o modelo com k quebras no nível e na tendência:

$$
{}_k\xi_\beta^* = \sum_{j=1}^{k+1} \frac{\sum_{i=\tau_{j-1}+1}^{\tau_j}\left(\sum_{t=\tau_{j-1}+1}^{i} e_t\right)^2}{(\tau_j-\tau_{j-1})^2\hat{\sigma}_\varepsilon^2} \qquad (4.22)
$$

em que $0 = \tau_0 < \tau_1 < \ldots < \tau_k < T$ indicam os instantes em que se permite uma quebra simultânea no nível e na tendência linear.

Então, a distribuição assintótica de ${}_k\xi^*_\beta$ será a mesma que a da soma de $k+1$ variáveis aleatórias independentes com distribuição CvM_2 cada. Essa distribuição é denotada por $\mathrm{CvM}_2(k+1)$ em Busetti & Harvey (2001). E sua função característica será então

$$\begin{aligned} \varphi_{{}_k\xi^*_\beta}(t) &= [D_2(2it)]^{-\frac{k+1}{2}} \\ D_2(x) &\equiv \frac{12}{x^2}\left(2-\sqrt{x}\sin\sqrt{x}-2\cos\sqrt{x}\right) \end{aligned} \tag{4.23}$$

4.4 Série com quebra no nível apenas

Também de Busetti & Harvey (2001), tem-se o modelo mais simples, com quebra no nível apenas:

$$\begin{aligned} y_t &= \mu_t + \delta_\mu w_t + \varepsilon_t &&, \ \varepsilon_t \sim \mathcal{N}\mathrm{id}\,(0,\sigma^2_\varepsilon) \\ \mu_t &= \mu_{t-1} + \eta_t &&, \ \eta_t \sim \mathcal{N}\mathrm{id}\,(0,\sigma^2_\eta) \\ \varepsilon_t &\ \text{e}\ \eta_s \ \text{independentes} &&, \ \forall\, t,s = 1,\ldots,T \end{aligned} \tag{4.24}$$

com $w_t = \mathbf{1}_{\{t>\tau\}}$ e τ é o instante, conhecido, em que ocorre a quebra no nível.

O teste ILO continua a ser dado pela mesma estatística:

$${}_1\xi_\mu(\lambda) = \frac{\sum_{i=1}^T \left(\sum_{t=1}^i e_t\right)^2}{T^2\hat{\sigma}^2_\varepsilon} > c_\alpha \tag{4.25}$$

e sua distribuição assintótica será a combinação linear de duas v.a.s independentes com distribuição CvM cada. De (2.13), a função característica correspondente será dada por

$$\begin{aligned} \varphi_{{}_1\xi_\mu}(t) &= [D_1(2it\lambda^2)D_1(2it(1-\lambda)^2)]^{-\frac{1}{2}} \\ D_1(x) &\equiv \frac{\operatorname{sen}\sqrt{x}}{\sqrt{x}} \end{aligned} \tag{4.26}$$

Teste simplificado para mais de uma quebra

Analogamente, o teste simplificado para k quebras no nível no modelo (4.24) será

$$ {}_k\xi_\mu^* = \sum_{j=1}^{k+1} \frac{\sum_{i=\tau_{j-1}+1}^{\tau_j} \left(\sum_{t=\tau_{j-1}+1}^{i} e_t\right)^2}{(\tau_j - \tau_{j-1})^2 \hat{\sigma}_\varepsilon^2} \tag{4.27} $$

e a distribuição assintótica de ${}_k\xi_\mu^*$ tem a função característica

$$ \varphi_{{}_k\xi_\mu^*}(t) = [D_1(2it)]^{-\frac{k+1}{2}} \tag{4.28} $$

4.5 Extensão para ruídos estacionários

Condição 4.1 (Phillips & Perron (1988)). *Seja $\{\varepsilon_t\}_{t=1,\ldots,\infty}$ uma série temporal satisfazendo*

1. $\mathrm{E}\,\varepsilon_t = 0$, $\forall t$;
2. $\sup_t \mathrm{E}\,|\varepsilon_t|^\beta < \infty$, *para algum* $\beta > 2$;
3. $\sigma_L^2 = \lim_{T\to\infty} T^{-1}\,\mathrm{Var}\left(\sum_{t=1}^{T} \varepsilon_t\right)$ *existe e* $\sigma_L^2 > 0$;
4. $\{\varepsilon_t\}_{t=1,\ldots,\infty}$ *é um strong mixing, com coeficientes de mixing* α_m *satisfazendo* $\sum_{m=1}^{\infty} \alpha_m^{1-2/\beta} < \infty$.

Se os processos $\{\varepsilon_t\}$ nos modelos apresentados anteriormente respeitam a condição 4.1, então as estatísticas ξs correspondentes apresentam as mesmas distribuições assintóticas se substituirmos $\hat{\sigma}_\varepsilon^2$ por um estimador consistente da variância de longo termo σ_L^2. Isso porque

$$ \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} e_t \Rightarrow \sigma_L^2 B(r) \tag{4.29} $$

no caso em que $e_t = y_t - \bar{y}$, e

$$ \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} e_t \Rightarrow \sigma_L^2 B_2(r) \tag{4.30} $$

no caso em que $e_t = y_t - \hat{\mu} - \hat{\beta}t$.

Para mais detalhes, ver, por exemplo, Kwiatkovski *et al.* (1992).

4.6 Modelos com componente sazonal

Conforme demonstrado em Busetti (2002), a inclusão de um componente sazonal determinístico nos modelos anteriores não altera as distribuições assintóticas das respectivas estatísticas apresentadas.

O componente sazonal determinístico $\{z_t\}$ pode ser parametrizado, por exemplo, com a utilização de *dummies* sazonais:

$$\begin{aligned} z_t &= \sum_{i=1}^{s-1} \omega_i z_{it} \\ z_{it} &= \mathbf{1}_{\{t=i+ns\}} - \mathbf{1}_{\{t=ns\}}, \quad n = 0, 1, 2, \ldots, \quad i = 1, 2, \ldots, s-1 \end{aligned} \tag{4.31}$$

s é o período sazonal.

Capítulo 5

Simulações

5.1 Estimativa da variância de longo termo

A fim de avaliar a qualidade da estimativa da variância de longo termo σ_L^2, proposta no capítulo 3, realizamos um experimento de Monte Carlo em um processo AR(1), para o qual se tem a fórmula exata de σ_L^2.

Seja $\{\varepsilon_t\}_{t=1,\ldots,T}$ um processo AR(1) com parâmetro ϕ:

$$\varepsilon_t = \phi\varepsilon_{t-1} + a_t\,, \quad a_t \sim \mathcal{N}\mathrm{id}\,(0,1) \tag{5.1}$$

Utilizando as relações (3.8) e (3.22), σ_L^2 será dada por

$$\begin{aligned} \sigma_L^2 &= 2\pi f(0) \\ &= 2\pi\,\frac{1}{2\pi}\,\frac{1}{1-2\phi\cos 0+\phi^2} \\ &= \frac{1}{(1-\phi)^2}\,, \quad |\phi|<1 \end{aligned} \tag{5.2}$$

A tabela 5.1 apresenta os valores médios de $\hat{\sigma}_L^2$ obtidos de 1000 processos (5.1) simulados de tamanho T e parâmetro ϕ. A tabela 5.2 traz os valores médios de $\widehat{l^*}$.

5.1 Estimativa da variância de longo termo

Tabela 5.1: Estimativas médias da variância de longo termo σ_L^2 para 1000 amostras simuladas de tamanho T de um processo AR(1) com parâmetro ϕ.

	ϕ			
T	0	0.2	0.5	0.8
100	0.988	1.45	3.18	14.6
200	0.997	1.48	3.43	17.2
500	0.992	1.51	3.64	19.5
1000	0.996	1.53	3.77	20.9
σ_L^2	1	1.56	4	25

Comparando-se os valores médios de $\hat{\sigma}_L^2$ com o valor nominal de σ_L^2 (última linha da tabela 5.1), nota-se que em média as estimativas são menores que os valores nominais e os resultados são melhores para as séries com menor persistência (ϕ menor). Verifica-se também aproximações melhores para amostras maiores, o que é esperado, pois o valor nominal de σ_L^2 é um resultado assintótico.

Tabela 5.2: Estimativas médias da largura ótima l^* para 1000 amostras simuladas de tamanho T de um processo AR(1) com parâmetro ϕ.

	ϕ			
T	0	0.2	0.5	0.8
100	4.8	4.7	4.7	6.5
200	6.6	6.3	6.6	8.9
500	9.6	9.5	9.4	13.0
1000	13.4	12.8	13.0	16.4

Quanto aos valores médios obtidos para $\widehat{l^*}$ (tabela 5.2), nota-se um aumento para as séries mais persistentes ($\phi = 0.8$), o que também é esperado, pois quanto mais persistente a série, maior deveria ser a largura de faixa dos núcleos, para se dar maior peso às informações passadas.

5.2 Tamanho e poder para amostras finitas

Apresentamos aqui os resultados obtidos por experimento de Monte Carlo para avaliar empiricamente o tamanho e o poder do teste para séries com uma quebra simultânea no nível e na tendência (seção 4.3). Novamente considerando o ruído estacionário $\{\varepsilon_t\}$ como um processo AR(1), simulamos 1000 réplicas do processo

$$\begin{aligned} y_t &= \mu_t + 2t - [T/2]w_t - w_t t + \varepsilon_t \quad , \quad \varepsilon_t = \phi\varepsilon_{t-1} + a_t\,, \quad a_t \sim \mathcal{N}\mathrm{id}\,(0,1) \\ \mu_t &= \mu_{t-1} + \eta_t \quad , \quad \eta_t \sim \mathcal{N}\mathrm{id}\,(0,\theta) \\ w_t &= \mathbf{1}_{\{t>[\lambda T]\}} \quad , \quad t = 1,\ldots,T \end{aligned} \tag{5.3}$$

Ou seja, fixamos os parâmetros $\mu_0 = 0$, $\beta = 2$, $\delta_\mu = -[T/2]$, $\delta_\beta = -1$ e $\sigma_a^2 = 1$.

O teste ILO rejeita $H_0 : \theta = 0$ para

$$_1\xi_\beta(\lambda) = \frac{\sum_{i=1}^{T}\left(\sum_{t=1}^{i} e_t\right)^2}{T^2\hat{\sigma}_L^2} > c_\alpha \tag{5.4}$$

com $\hat{\sigma}_L^2$ obtido conforme o algoritmo da seção 3.3.

Os resultados obtidos para um tamanho nominal $\alpha = 5\%$ encontram-se nas tabelas 5.3 e 5.4, com $\lambda = 0.5$ e $\lambda = 0.1$, respectivamente.

Em linhas gerais, tem-se melhores resultados para um maior tamanho da amostra (eficiência assintótica) e para uma menor persistência do ruído (ϕ menor). Espera-se do poder que ele aumente monotonicamente ao distanciarmo-nos de $H_0 : \quad \theta = 0$ e tenda a 1 para $\theta \to \infty$. Essas características não se verificam aqui. Tipicamente, o poder apresenta um máximo para $\theta = 0.1$ ou 1 e volta a cair tendendo para um valor diferente de 1. Na figura 5.1 apresentamos esse comportamento mais detalhadamente para $\lambda = 0.1$, $\phi = 0$, $T = 100$, $10^{-4} \leq \theta \leq 10^3$ e 5000 réplicas. Uma explicação possível é que para valores mais altos de θ, com uma maior predominância relativa do componente de passeio aleatório, há uma maior probabilidade de ocorrer trechos de tendência

estocástica passíveis de serem espuriamente bem ajustados pelo método utilizado de mínimos quadrados ordinários. Essa explicação é corroborada se observarmos que os resultados obtidos para $\lambda = 0.1$ (tabela 5.4) são melhores do que os para $\lambda = 0.5$ (tabela 5.3). Isso porque, por questão de simetria, em um passeio aleatório é mais provável a existência de um bom ajuste espúrio com quebra no meio da amostra do que com quebra em uma extremidade, no caso em 10% da amostra.

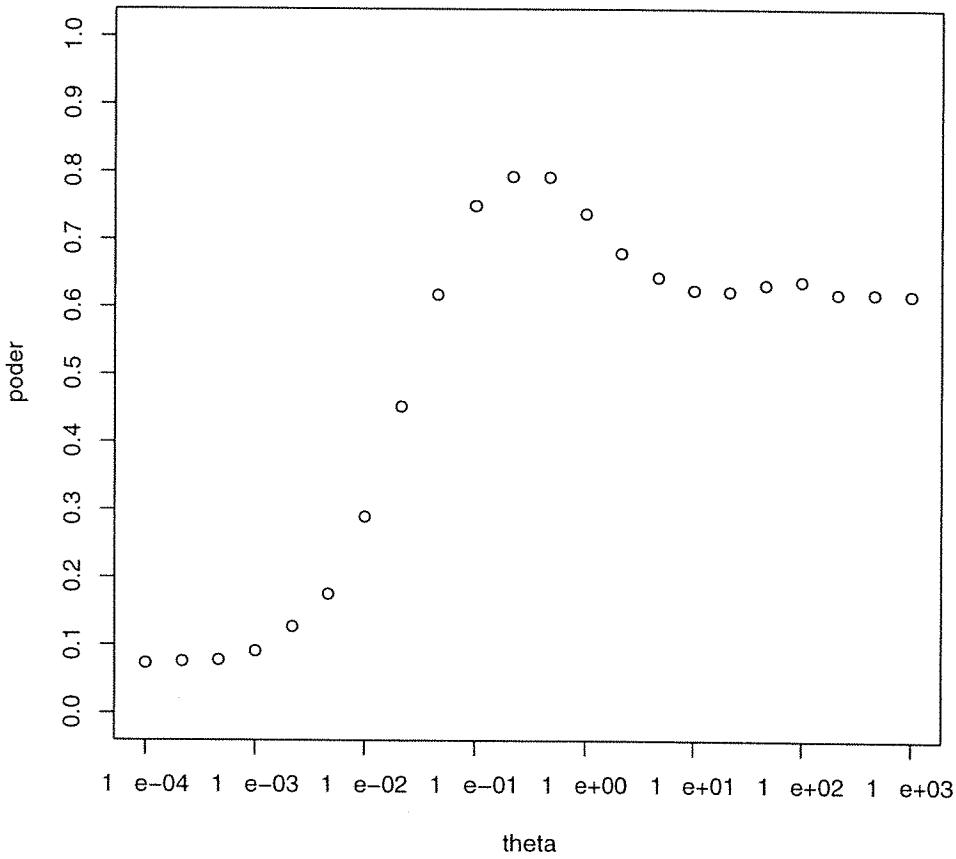

Figura 5.1: Poder empírico para o teste (5.4) com $\lambda = 0.1$, $\phi = 0$, $T = 100$, 5000 réplicas e $10^{-4} \leq \theta \leq 10^3$ (θ em escala logarítmica)

Tabela 5.3: Tamanho ($\theta = 0$) e poder ($\theta > 0$) empíricos para o teste (5.4), $\lambda = 0.5$ e 1000 réplicas.

		ϕ			
	θ	0	0.2	0.5	0.8
	0	0.087	0.137	0.289	0.419
$T = 100$	0.1	0.612	0.598	0.519	0.464
$\lambda = 0.5$	1	0.714	0.699	0.668	0.540
	10	0.604	0.594	0.606	0.585
	0	0.058	0.111	0.276	0.366
$T = 200$	0.1	0.878	0.846	0.765	0.512
$\lambda = 0.5$	1	0.712	0.693	0.719	0.655
	10	0.643	0.628	0.667	0.642
	0	0.049	0.084	0.221	0.315
$T = 500$	0.1	0.854	0.874	0.919	0.801
$\lambda = 0.5$	1	0.715	0.720	0.731	0.785
	10	0.707	0.701	0.702	0.720

Tabela 5.4: Tamanho ($\theta = 0$) e poder ($\theta > 0$) empíricos para o teste (5.4), $\lambda = 0.1$ e 1000 réplicas.

		ϕ			
	θ	0	0.2	0.5	0.8
	0	0.062	0.122	0.251	0.374
$T = 100$	0.1	0.745	0.717	0.613	0.463
$\lambda = 0.1$	1	0.718	0.704	0.714	0.603
	10	0.644	0.639	0.632	0.631
	0	0.044	0.078	0.190	0.326
$T = 200$	0.1	0.874	0.863	0.816	0.606
$\lambda = 0.1$	1	0.710	0.710	0.718	0.692
	10	0.682	0.683	0.692	0.682

5.3 Tamanho e poder do teste simplificado

Além dos testes apresentados no capítulo 4, Busetti & Harvey (2001) apresentam também os testes para quebra apenas na tendência linear ou apenas no nível no modelo com tendência. Ocorre no entanto que é mais trabalhosa a obtenção das distribuições assintóticas, via inversão da função característica, das respectivas estatísticas de teste. E também não é possível obter uma estatística simplificada como a apresentada nas seções 4.3 e 4.4, isso porque não é possível separar a distribuição em duas ou mais distribuições independentes [1].

A idéia então é utilizar o teste simplificado para quebra simultânea no nível e na tendência (estatística ${}_1\xi^*_\beta$) em séries geradas com quebra apenas na tendência, ou seja, a série é 'contínua' mas apresenta uma mudança de tendência em $t = \tau$.

Para comparar o desempenho dos dois testes, simulamos séries com quebra apenas na tendência e obtivemos o tamanho e o poder empírico dos testes para 10000 réplicas e um tamanho nominal $\alpha = 5\%$.

O processo gerador de dados é:

$$\begin{aligned} y_t &= \mu_t + t + z_t + \varepsilon_t \quad , \quad \varepsilon_t = \phi\varepsilon_{t-1} + a_t\,, \quad a_t \sim \mathcal{N}\mathrm{id}\,(0,1) \\ \mu_t &= \mu_{t-1} + \eta_t \quad , \quad \eta_t \sim \mathcal{N}\mathrm{id}\,(0,\theta) \\ z_t &= (t-\tau)\mathbf{1}_{\{t>\tau\}} \quad , \quad t = 1,\ldots,T \end{aligned} \tag{5.5}$$

Estatística do teste simplificado com quebra simultânea no nível e na tendência

$${}_1\xi^*_\beta = \frac{\sum_{i=1}^{\tau}\left(\sum_{t=1}^{i} e_t\right)^2}{\tau^2\hat{\sigma}^2_\varepsilon} + \frac{\sum_{i=\tau+1}^{T}\left(\sum_{t=\tau+1}^{i} e_t\right)^2}{(T-\tau)^2\hat{\sigma}^2_\varepsilon} \tag{5.6}$$

[1] ver também Kurozumi (2002)

Os resíduos $\{e_t\}$ são obtidos do ajuste de mínimos quadrados

$$\begin{aligned} y_t &= \hat{\mu}_0 + \hat{\beta}t + \hat{\delta}_\mu w_t + \hat{\delta}_\beta w_t t + e_t \\ w_t &= \mathbf{1}_{\{t>\tau\}} \end{aligned} \tag{5.7}$$

Estatística do teste com quebra apenas na tendência

$$_1\xi'_\beta = \frac{\sum_{i=1}^{T}\left(\sum_{t=1}^{i} e_t\right)^2}{T^2\hat{\sigma}_\varepsilon^2} \tag{5.8}$$

Os resíduos $\{e_t\}$ são obtidos do ajuste de mínimos quadrados

$$\begin{aligned} y_t &= \hat{\mu}_0 + \hat{\beta}t + \hat{\delta}_\beta z_t t + e_t \\ z_t &= (t-\tau)\mathbf{1}_{\{t>\tau\}} \end{aligned} \tag{5.9}$$

Os valores críticos da distribuição assintótica de $_1\xi'_\beta$, para um tamanho nominal de 5%, são 0.12716, 0.10956, 0.09563, 0.08643 e 0.08318 para $\lambda = 0.1, 0.2, 0.3, 0.4$ e 0.5, respectivamente (Kurozumi (2002)).

Como se depreende dos resultados obtidos (tabelas 5.5, 5.6 e 5.7), o teste simplificado apresenta uma leve vantagem em termos de tamanho (linhas com $\theta = 0$). E o teste "exato" apresenta uma vantagem, também não muito grande, em termos de poder (linhas com $\theta > 0$). Ambos os testes apresentam maior poder com o aumento da amostra, de $T = 100$ para $T = 200$, como esperado, pois a distribuição utilizada é assintótica.

Assim, pelos resultados dessas simulações, o teste simplificado para quebra no nível e na tendência aplicado a séries com quebra apenas na tendência apresenta desempenho semelhante do teste derivado para quebra apenas na tendência e que leva em conta a localização da quebra.

Tabela 5.5: Tamanho ($\theta = 0$) e poder ($\theta > 0$) empíricos, 10000 réplicas.

$T = 100$		$\lambda = \tau/T$				
	θ	0.1	0.2	0.3	0.4	0.5
teste	0	0.0532	0.0524	0.0537	0.0527	0.0520
simplificado	0.01	0.2365	0.2192	0.1979	0.1770	0.1709
com quebra	0.1	0.7840	0.7644	0.7627	0.7676	0.7608
no nível e	1	0.9752	0.9824	0.9898	0.9906	0.9915
na tendência	10	0.9908	0.9968	0.9987	0.9992	0.9991
teste	0	0.0567	0.0545	0.0510	0.0564	0.0503
com quebra	0.01	0.3281	0.2971	0.2572	0.2388	0.2153
na tendência	0.1	0.8692	0.8340	0.8028	0.8060	0.7952
apenas	1	0.9919	0.9890	0.9891	0.9913	0.9911
	10	0.9989	0.9985	0.9979	0.9984	0.9985

Tabela 5.6: Tamanho ($\theta = 0$) e poder ($\theta > 0$) empíricos, 10000 réplicas.

$T = 200$		$\lambda = \tau/T$				
	θ	0.1	0.2	0.3	0.4	0.5
teste	0	0.0515	0.0499	0.0515	0.0485	0.0515
simplificado	0.01	0.5980	0.5577	0.5249	0.5270	0.5084
com quebra	0.1	0.9708	0.9771	0.9794	0.9859	0.9840
no nível e	1	0.9997	0.9996	0.9999	1	0.9998
na tendência	10	0.9998	1	1	1	1
teste	0	0.0533	0.0585	0.0531	0.0520	0.0481
com quebra	0.01	0.6906	0.6528	0.6007	0.5878	0.5789
na tendência	0.1	0.9877	0.9827	0.9799	0.9855	0.9857
apenas	1	0.9998	0.9998	0.9999	1	1
	10	1	1	1	1	1

Tabela 5.7: Tamanho ($\theta = 0$) e poder ($\theta > 0$) empíricos, 10000 réplicas.

$T = 500$		$\lambda = \tau/T$				
	θ	0.1	0.2	0.3	0.4	0.5
teste	0	0.0509	0.0496	0.0461	0.0493	0.0518
simplificado	0.01	0.9611	0.9625	0.9645	0.9652	0.9681
com quebra	0.1	1	1	1	1	1
no nível e	1	1	1	1	1	1
na tendência	10	1	1	1	1	1
teste	0	0.0537	0.0491	0.0476	0.0506	0.0494
com quebra	0.01	0.9781	0.9720	0.9672	0.9682	0.9719
na tendência	0.1	1	1	1	1	1
apenas	1	1	1	1	1	1
	10	1	1	1	1	1

Capítulo 6

Aplicações

6.1 Análise das exportações brasileiras

O objetivo é analisar a série mensal de exportações brasileiras nas últimas décadas a fim de testar seu comportamento, utilizando-se da metodologia apresentada. Ou seja, deseja-se testar se o componente estocástico da série é ou não estacionário quando se modela a parte não-aleatória com componentes sazonais e de tendência linear. Permitindo-se ainda uma ou mais quebras estruturais no nível e na tendência.

Visto que a série claramente apresenta uma tendência crescente e que a variabilidade é proporcional ao nível (fig.6.1), trabalharemos com o logaritmo da série, que possui uma variabilidade mais regular (fig.6.2).

Assim, será utilizada a série de logaritmos (na base 10) das exportações mensais brasileiras, $x_t = \log_{10} X_t$, em dólar corrente, do período de janeiro/1969 a dezembro/2003.

6.1 Análise das exportações brasileiras

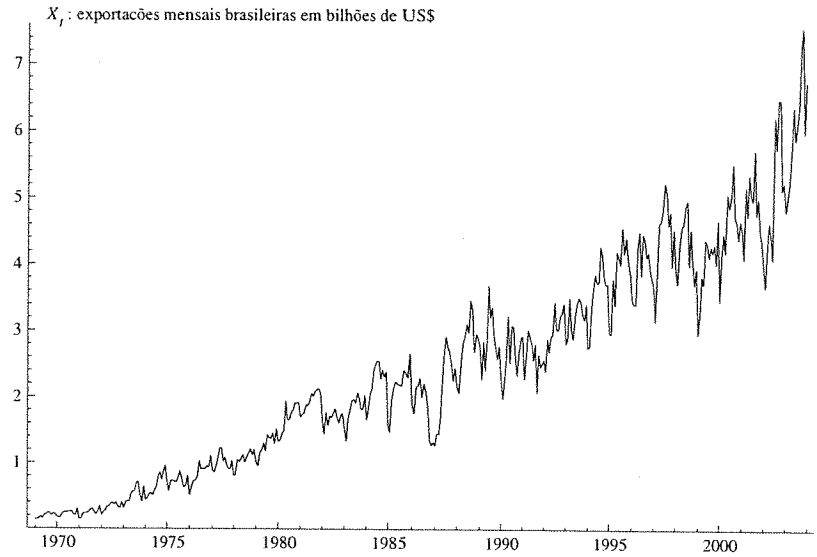

Figura 6.1: X_t = Exportações mensais brasileiras, em bilhões de US$, do período de janeiro/1969 a dezembro/2003, $t = 1, \ldots, T = 420$

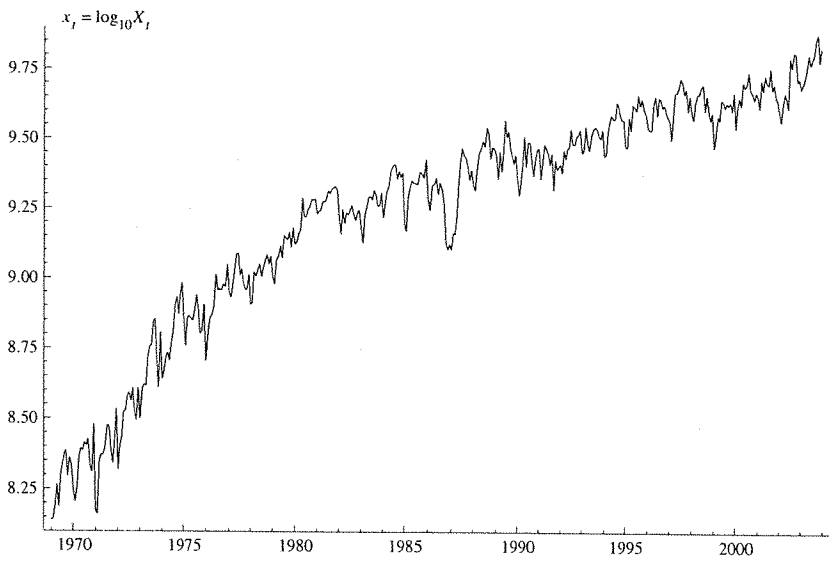

Figura 6.2: $x_t = \log_{10} X_t$

Modelo com tendência e sem quebras estruturais

Como primeira abordagem, testou-se o modelo com uma tendência linear, mas sem quebras estruturais:

$$\begin{aligned} x_t &= \mu_t + \beta t + \omega_t + \varepsilon_t \quad , \quad \varepsilon_t \sim \mathcal{EST}(0, \gamma_\varepsilon) \\ \mu_t &= \mu_{t-1} + \eta_t \quad , \quad \eta_t \sim \mathrm{RB}\,(0, \sigma_\eta^2) \end{aligned} \tag{6.1}$$

O componente sazonal determinístico ω_t é modelado com a parametrização trigonométrica, como em Harvey (1989):

$$\omega_t = \sum_{j=1}^{6} \left(\omega_{c,j} \cos \frac{2\pi jt}{12} + \omega_{s,j} \operatorname{sen} \frac{2\pi jt}{12} \right) \tag{6.2}$$

$\omega_{c,j}$ e $\omega_{s,j}$, $j = 1, \ldots, 6$ são os parâmetros associados aos componentes cosseno e seno, respectivamente. Como $\operatorname{sen} k\pi = 0$ para k inteiro, tem-se que $\omega_{s,6} \equiv 0$ e, portanto, o número total de parâmetros é 11, como na parametrização com *dummies* mensais. De fato, ambas parametrizações são equivalentes.

O teste de hipóteses, com $\theta = \sigma_\eta^2/\sigma_\varepsilon^2$, $\quad \sigma_\varepsilon^2 = \gamma_\varepsilon(0) > 0$, $\quad$ é:

$$\begin{aligned} H_0 &: \theta = 0 \\ H_1 &: \theta > 0 \end{aligned} \tag{6.3}$$

Os resíduos $\{e_t\}$ para o teste são obtidos pelo seguinte ajuste de mínimos quadrados:

$$x_t = \hat{\mu}_0 + \hat{\beta} t + \sum_{j=1}^{6} \left(\hat{\omega}_{c,j} \cos \frac{2\pi jt}{12} + \hat{\omega}_{s,j} \operatorname{sen} \frac{2\pi jt}{12} \right) + e_t \tag{6.4}$$

A figura 6.3 mostra a reta ajustada e a série $\{x_t\}$ subtraída do componente sazonal ajustado. As estimativas dos parâmetros são

$$\begin{aligned} \hat{\mu}_0 &= 8.57 \\ \hat{\beta} &= 0.00317 \\ \hat{\omega}_c &= (-0.0475,\ -0.0063,\ -0.0043,\ -0.0089,\ -0.0069,\ -0.0026) \\ \hat{\omega}_s &= (-0.0165,\ -0.0129,\ -0.0208,\ -0.0125,\ -0.0106,\ 0) \end{aligned}$$

6.1 Análise das exportações brasileiras

A estimativa de $\hat{\beta}$ corresponde a uma taxa de crescimento das exportações de $0.73\%\ a.m.$ ou $9.15\%\ a.a.$

Pela metodologia apresentada no capítulo 3, tem-se as seguintes estimativas para a largura de faixa ótima l^* e a variância de longo termo σ_L^2:

$$\begin{aligned} \widehat{l^*} &= 28.3 \\ \hat{\sigma}_L^2 &= 0.572 \end{aligned}$$

Chega-se então à estatística ξ_β do teste ILO e ao p-valor associado a sua distribuição assintótica ($\sim \mathrm{CvM}_2$):

$$\begin{aligned} \xi_\beta &= 0.262 \\ \text{p-valor} &= 0.00375 \end{aligned}$$

Ou seja, há forte evidência estatística para rejeitar a hipótese de estacionariedade em torno de uma tendência linear (*trend stationarity*) para a série de exportações analisada.

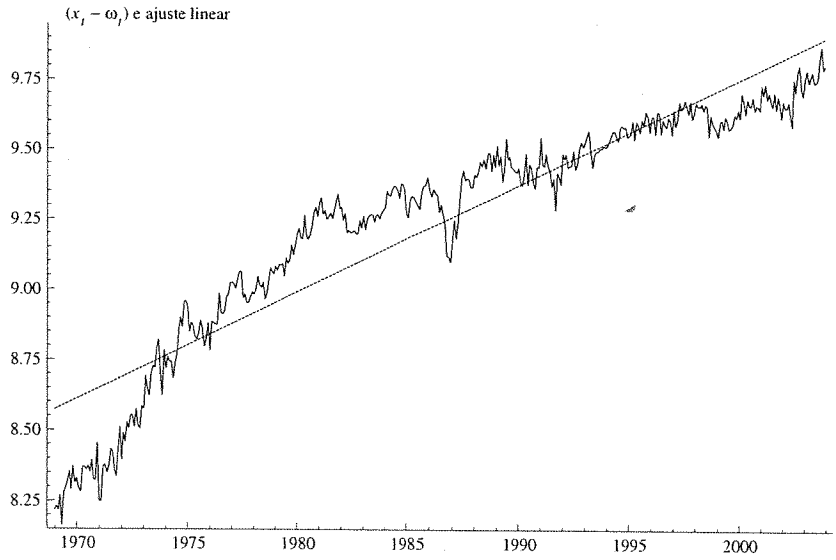

Figura 6.3: A série x_t dessazonalizada e a reta ajustada

Modelo com tendência e com uma quebra estrutural

Pela análise visual da série de exportações e pelo contexto histórico, conjecturamos uma quebra estrutural na série de exportações brasileiras em dezembro de 1981. Essa data pode ser identificada como o início de uma queda acentuada das reservas internacionais brasileiras devido às dificuldades de equilíbrio do balanço de pagamentos, decorrentes da crise da dívida externa no início da década de 1980. Essas dificuldades agravam-se ainda mais com a moratória mexicana em agosto de 1982 e, de certo modo, estendem-se até hoje devido à consequente redução da disponibilidade de crédito dos países desenvolvidos para os países em desenvolvimento.

Assim, introduzindo uma quebra estrutural no nível e na tendência em dezembro de 1981, tem-se o seguinte modelo

$$\begin{aligned}
x_t &= \mu_t + \beta t + \delta_\mu w_t + \delta_\beta (w_t t) + \omega_t + \varepsilon_t\,, \quad \varepsilon_t \sim \mathcal{EST}(0, \gamma_\varepsilon) \qquad (6.5)\\
\mu_t &= \mu_{t-1} + \eta_t\,, \quad \eta_t \sim \mathrm{RB}\,(0, \sigma_\eta^2)\\
\omega_t &= \sum_{j=1}^{6} \left(\omega_{c,j} \cos \frac{2\pi j t}{12} + \omega_{s,j} \operatorname{sen} \frac{2\pi j t}{12} \right)\\
w_t &= \mathbf{1}_{\{t>156\}}\,, \qquad \lambda = 156/420 = 0.37
\end{aligned}$$

O teste de hipóteses continua a ser como em (6.3) e os resíduos $\{e_t\}$ para o teste são obtidos pelo seguinte ajuste de mínimos quadrados:

$$x_t = \hat{\mu}_0 + \hat{\beta} t + \hat{\delta}_\mu w_t + \hat{\delta}_\beta (w_t t) + \hat{\omega}_t + e_t \qquad (6.6)$$

6.1 Análise das exportações brasileiras

As estimativas dos parâmetros são :

$$\begin{aligned}
\hat{\mu}_0 &= 8.24 \\
\hat{\delta}_\mu &= -0.125 \\
\hat{\beta} &= 0.00713 \\
\hat{\delta}_\beta &= -0.00513 \\
\hat{\omega}_c &= (-0.0468,\ -0.0056,\ -0.0035,\ -0.0081,\ -0.0062,\ -0.0023) \\
\hat{\omega}_s &= (-0.0138,\ -0.0116,\ -0.0201,\ -0.0121,\ -0.0104,\ 0)
\end{aligned}$$

As estimativas das inclinações correspondem a um crescimento das exportações de $21\%\, a.a.$ de janeiro/1969 a dezembro/1981 e de apenas $5.7\%\, a.a.$ de janeiro/1982 a dezembro/2004. A figura 6.4 mostra as retas ajustadas e a série sem o componente sazonal ajustado.

As estimativas para a largura de faixa ótima l^* e a variância de longo termo σ_L^2 são

$$\begin{aligned}
\widehat{l^*} &= 8.1 \\
\hat{\sigma}_L^2 &= 0.0211
\end{aligned}$$

A estatística ${}_1\xi_\beta$ do teste ILO e o p-valor associado a sua distribuição assintótica ($\sim \lambda^2\,\mathrm{CvM}_2 + (1-\lambda)^2\,\mathrm{CvM}_2$) são

$$\begin{aligned}
{}_1\xi_\beta &= 0.0544 \\
\text{p-valor} &= 0.12305
\end{aligned}$$

Ou seja, não se rejeita a hipótese de estacionariedade ao redor de uma tendência com quebra em dez/1981, até um nível de significância de 12%. Portanto, tendo como plausível a conjectura de que de fato houve uma mudança estrutural no comportamento das exportações a partir de 1982, conclui-se que o ingresso de divisas, em US$, advindas das exportações segue um processo estacionário ao redor de uma tendência determinística.

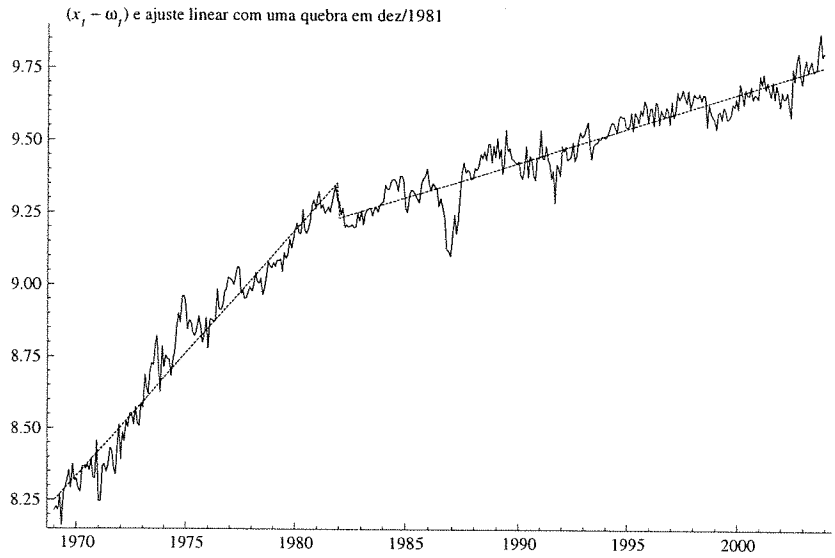

Figura 6.4: Série x_t dessazonalizada e retas ajustadas com uma quebra simultânea no nível e na tendência em dezembro de 1981.

6.2 Análise das importações brasileiras

Analisamos aqui a série mensal de importações brasileiras, em dólar corrente, no mesmo período anterior, janeiro de 1969 a dezembro de 2003, para fins de comparação com o comportamento das exportações. Em princípio, as exportações e as importações de um país deveriam apresentar comportamento semelhante. Isso porque o equilíbrio na balança comercial é uma condição de solvência do país no longo prazo. Déficits sistemáticos na balança comercial encontram cada vez menos possibilidade de serem financiados via empréstimos ou investimentos externos.

Assim, espera-se que a série da balança comercial (exportações menos importações) seja estacionária, ou, no contexto de cointegração, que as séries de exportação e de importação sejam cointegradas. Nesse sentido é fundamental saber o comportamento de cada série individualmente, se são estacionárias, se apresentam mesma tendência, quebras *etc.*

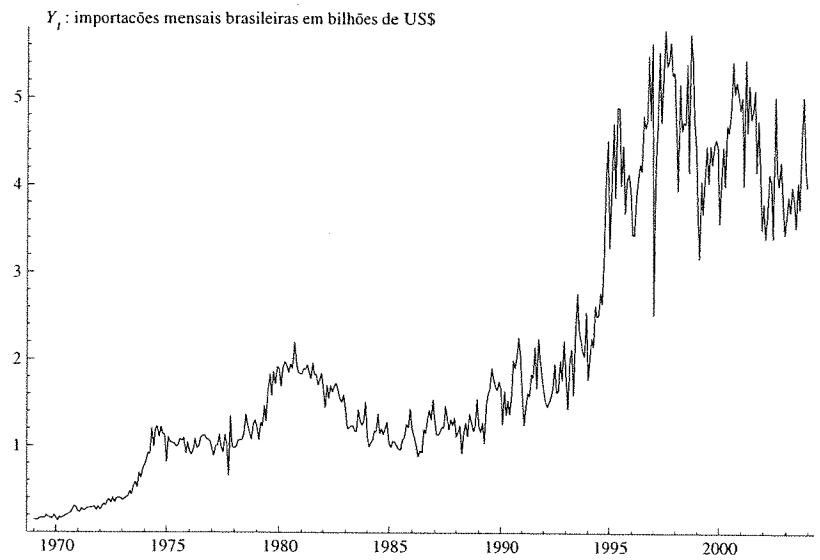

Figura 6.5: Y_t = Importações mensais brasileiras, em bilhões de US$, do período de janeiro/1969 a dezembro/2003, $t = 1, \ldots, T = 420$

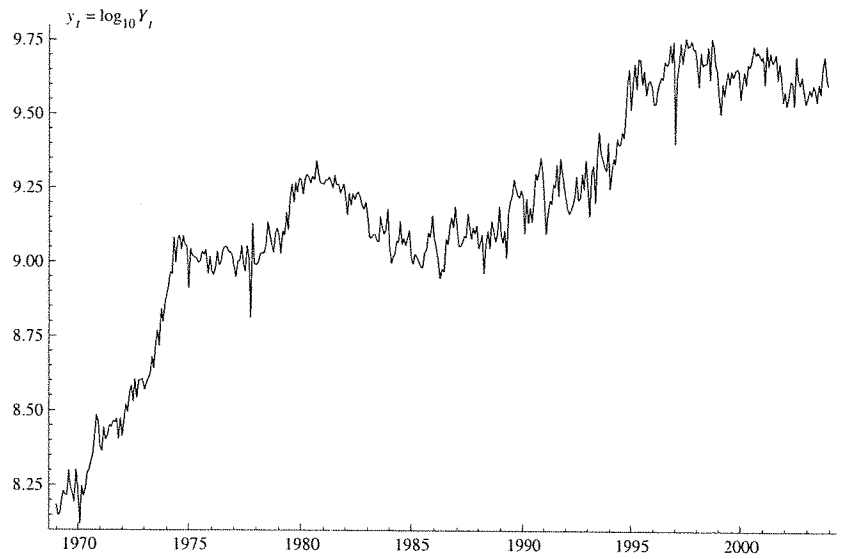

Figura 6.6: Logaritmo das importações mensais brasileiras, $y_t = \log_{10} Y_t$

Modelo com tendência e sem quebras estruturais

Novamente, o primeiro modelo testado é o com tendência, mas sem quebras estruturais:

$$\begin{aligned} y_t &= \mu_t + \beta t + \omega_t + \varepsilon_t\,, \quad \varepsilon_t \sim \mathcal{EST}(0, \gamma_\varepsilon) \\ \mu_t &= \mu_{t-1} + \eta_t\,, \quad \eta_t \sim \mathrm{RB}\,(0, \sigma_\eta^2) \\ \omega_t &= \sum_{j=1}^{6} \left(\omega_{c,j} \cos \frac{2\pi j t}{12} + \omega_{s,j} \operatorname{sen} \frac{2\pi j t}{12} \right) \end{aligned} \tag{6.7}$$

E o mesmo teste de hipóteses

$$\begin{aligned} H_0 &: \theta = 0\,, \qquad \theta \equiv \sigma_\eta^2/\sigma_\varepsilon^2 \\ H_1 &: \theta > 0 \end{aligned} \tag{6.8}$$

Os resíduos $\{e_t\}$ para o teste são obtidos pelo seguinte ajuste de mínimos quadrados:

$$y_t = \hat{\mu}_0 + \hat{\beta} t + \hat{\omega}_t + e_t \tag{6.9}$$

A figura 6.7 mostra a reta ajustada e a série y_t subtraída do componente sazonal ajustado. As estimativas dos parâmetros são

$$\begin{aligned} \hat{\mu}_0 &= 8.58 \\ \hat{\beta} &= 0.00283 \\ \hat{\omega}_c &= (-0.0166,\ -0.0041,\ -0.0094,\ -0.0044,\ -0.0037,\ 0.0000) \\ \hat{\omega}_s &= (-0.0278,\ -0.0100,\ -0.0128,\ -0.0052,\ -0.0132,\ 0) \end{aligned}$$

A estimativa de $\hat{\beta}$ corresponde a uma taxa de crescimento das importações de $0.65\%\ a.m.$ ou $8.14\%\ a.a.$

As estimativas para a largura de faixa ótima l^* e a variância de longo termo σ_L^2 são

$$\begin{aligned} \widehat{l^*} &= 23.9 \\ \hat{\sigma}_L^2 &= 0.734 \end{aligned}$$

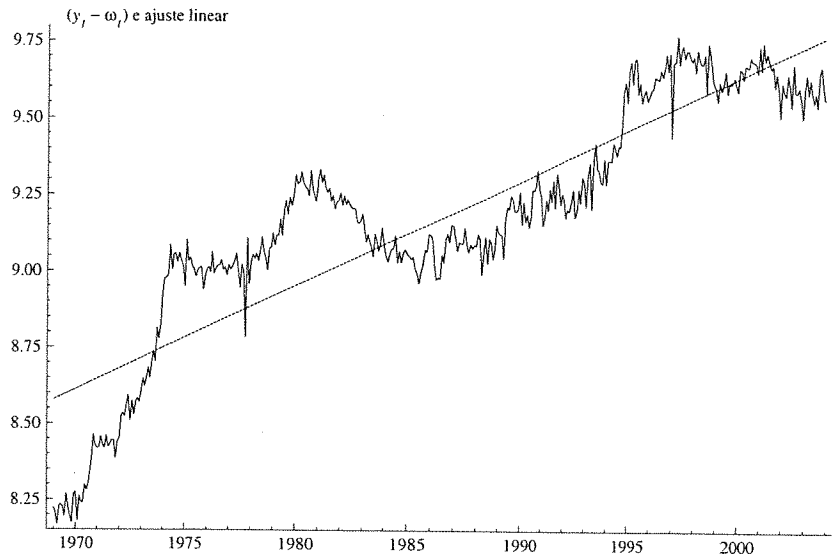

Figura 6.7: Série y_t dessazonalizada e a reta ajustada

A estatística ξ_β do teste ILO e o p-valor associado a sua distribuição assintótica ($\sim$ CvM_2) são

$$\begin{aligned} \xi_\beta &= 0.127 \\ \text{p-valor} &= 0.0821 \end{aligned}$$

Ou seja, a hipótese de estacionariedade ao redor de uma tendência linear seria rejeitada a um nível de significância de 10%, mas não seria a um um nível de significância de 5%, por exemplo. Portanto, diferentemente da série de exportações, o teste não apresenta forte evidência de não-estacionariedade, podendo até ser considerada estacionária em torno de uma tendência, a um nível de significância de 5%.

Modelo com tendência e com uma quebra estrutural

Pelo contexto histórico e pela análise visual da série de importações, identificamos preliminarmente quatro quebras estruturais plausíveis:

B1 : dez/1973 ($\tau = 60$), primeiro choque do petróleo.

B2 : dez/1981 ($\tau = 156$), crise da dívida externa.

B3 : mar/1986 ($\tau = 207$), Plano Cruzado.

B4 : jul/1994 ($\tau = 307$), Plano Real.

Modelo

$$
\begin{aligned}
y_t &= \mu_t + \beta t + \delta_\mu w_t + \delta_\beta (w_t t) + \omega_t + \varepsilon_t\,, \quad \varepsilon_t \sim \mathcal{EST}(0, \gamma_\varepsilon) \\
\mu_t &= \mu_{t-1} + \eta_t\,, \quad \eta_t \sim \mathrm{RB}\,(0, \sigma_\eta^2) \\
\omega_t &= \sum_{j=1}^{6} \left(\omega_{c,j} \cos \frac{2\pi j t}{12} + \omega_{s,j} \operatorname{sen} \frac{2\pi j t}{12} \right) \\
w_t &= \mathbf{1}_{\{t > \tau\}}\,, \qquad \lambda = \tau / T
\end{aligned}
\tag{6.10}
$$

Teste de hipóteses

$$
\begin{aligned}
H_0 &: \theta = 0\,, \qquad \theta \equiv \sigma_\eta^2 / \sigma_\varepsilon^2 \\
H_1 &: \theta > 0
\end{aligned}
\tag{6.11}
$$

Resultados

quebra	mês	τ	λ	$\widehat{l^*}$	$\hat{\sigma}_L^2$	${}_1\xi_\beta$	p-valor
B1	12/73	60	0.143	34.1	0.447	0.110	0.0507
B2	12/81	156	0.371	18.6	0.202	0.079	0.0278
B3	03/86	207	0.493	22.0	0.497	0.089	0.0057
B4	07/94	307	0.731	26.7	0.729	0.136	0.0056

Nota-se que a introdução de apenas uma quebra simultânea no nível e na tendência, nas datas factíveis acima, conduz o teste no sentido de rejeição de H_0.

Modelo com tendência e mais de uma quebra estrutural

Utilizando o teste simplificado da seção 4.3 e permitindo-se mais de uma quebra, nas mesmas datas acima, chega-se às seguintes combinações para as quais não se rejeita a hipótese de estacionariedade em torno de uma tendência com quebras simultâneas no nível e na tendência:

B12 : quebras em dez/1973 ($\tau_1 = 60$) e dez/1981 ($\tau_2 = 156$).

B14 : quebras em dez/1973 ($\tau_1 = 60$) e jul/1994 ($\tau_2 = 307$).

B134: quebras em dez/1973 ($\tau_1 = 60$), mar/1986 ($\tau_2 = 207$) e jul/1994 ($\tau_3 = 307$).

Resultados

quebras	$\widehat{l^*}$	$\hat{\sigma}_L^2$	${}_2\xi_\beta^*$	${}_3\xi_\beta^*$	p-valor
B12	18.9	0.149	0.240	–	0.2418
B14	24.8	0.109	0.184	–	0.5181
B134	25.4	0.081	–	0.344	0.1631

Das três combinações acima, destaca-se B14 que, com forte evidência estatística, não rejeita a hipótese de estacionariedade em torno de uma tendência com quebras em dez/1973 e jul/1994.

Uma possível interpretação sugerida é que os eventos que mais afetaram a série de importações brasileiras no período analisado foram a primeira crise do petróleo em 1973 e o Plano Real em 1994.

Importante destacar ainda que, pela metodologia utilizada, as séries de exportação e importação brasileiras apresentam comportamentos distintos, o que sugere uma maior dificuldade para se conseguir um equilíbrio na balança comercial. Aparentemente, as

exportações são mais influenciadas pela disponibilidade de crédito externo e abertura aos mercados dos países desenvolvidos, ou seja, predominam os fatores externos. Já as importações são sensíveis tanto a choques externos (crise do petróleo, crise da dívida externa) como a choques resultantes da política econômica interna (Plano Real).

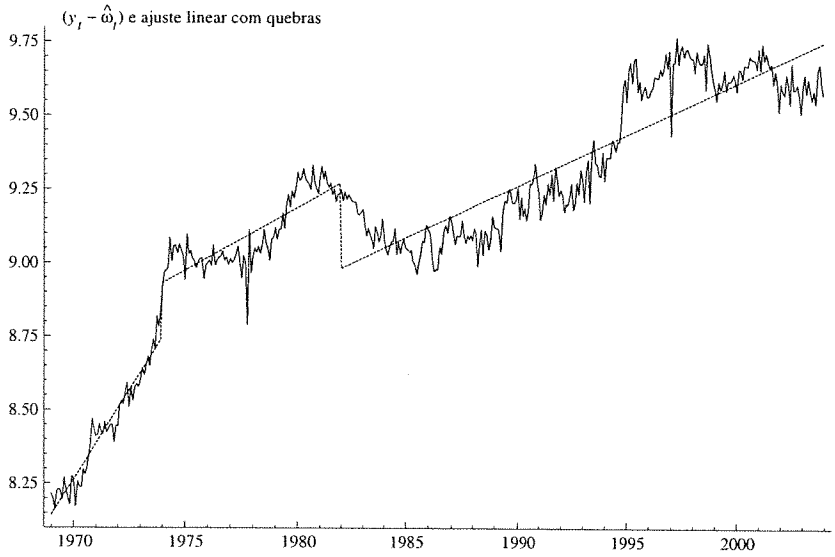

Figura 6.8: Série y_t dessazonalizada e retas ajustadas com quebras em dez/73 e dez/81.

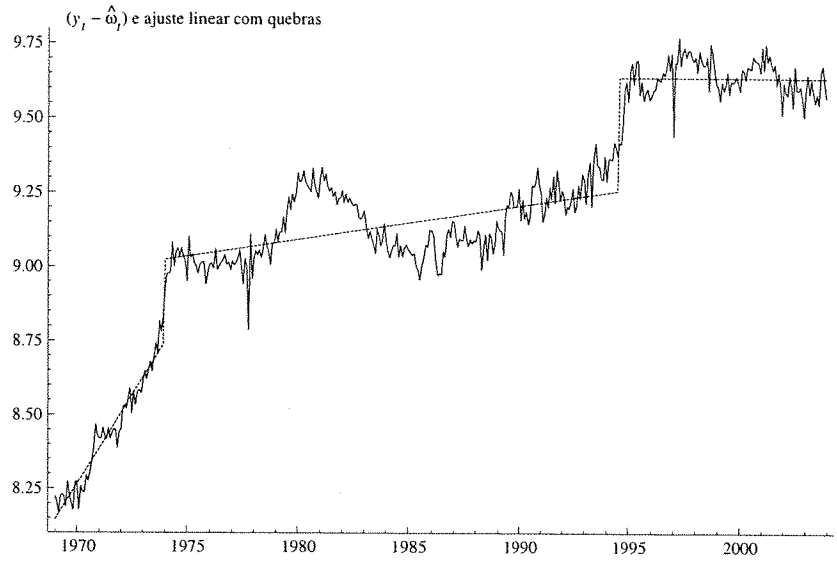

Figura 6.9: Série y_t dessazonalizada e retas ajustadas com quebras em dez/73 e jul/94.

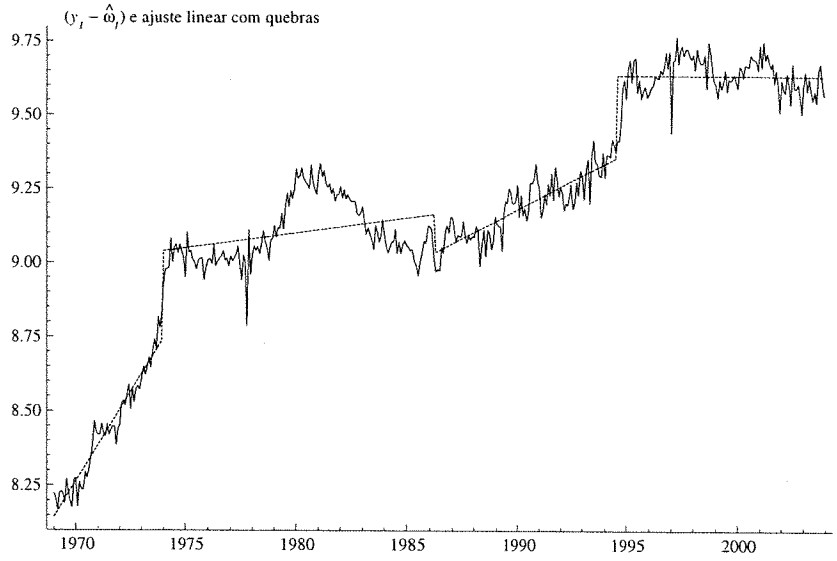

Figura 6.10: Série y_t dessazonalizada e retas ajustadas com quebras em dez/73 e jul/94.

6.3 Câmbio real brasileiro

A hipótese da paridade do poder de compra (PPC)[1], em sua forma relativa, estabelece que as variações da taxa de câmbio nominal C acompanham, principalmente no longo prazo, as variações dos índices de preço doméstico P_d e externo P_e. Posto de outra forma, espera-se que a taxa de câmbio real

$$c = C\,\frac{P_e}{P_d} \tag{6.12}$$

flutue em torno de um nível constante. Ou ainda, em termos estatísticos, que a série temporal de câmbio real c_t seja estacionária.

A fim de testar a hipótese da PPC, utilizamos então a metodologia apresentada no capítulo 4, no modelo sem tendência.

Para construir a série de câmbio real utilizou-se como índice de preço doméstico o IGP-DI geral mensal centrado de fim de período (fonte: IPEA), como índice de preço externo o IPC americano (fonte: FMI) e a taxa de câmbio nominal de venda de fim de mês (fonte: Banco Central do Brasil - Boletim de Balanço de Pagamentos). O período analisado é de janeiro de 1969 a dezembro de 2003.

A série construída encontra-se na figura 6.11.

Modelo sem tendência e sem quebras

$$\begin{aligned} c_t &= \mu_t + \varepsilon_t\,, \quad \varepsilon_t \sim \mathcal{EST}(0,\gamma_\varepsilon) \\ \mu_t &= \mu_{t-1} + \eta_t\,, \quad \eta_t \sim \mathrm{RB}\,(0,\sigma_\eta^2) \end{aligned} \tag{6.13}$$

[1] ver, por exemplo, Araújo & Silveira (2002)

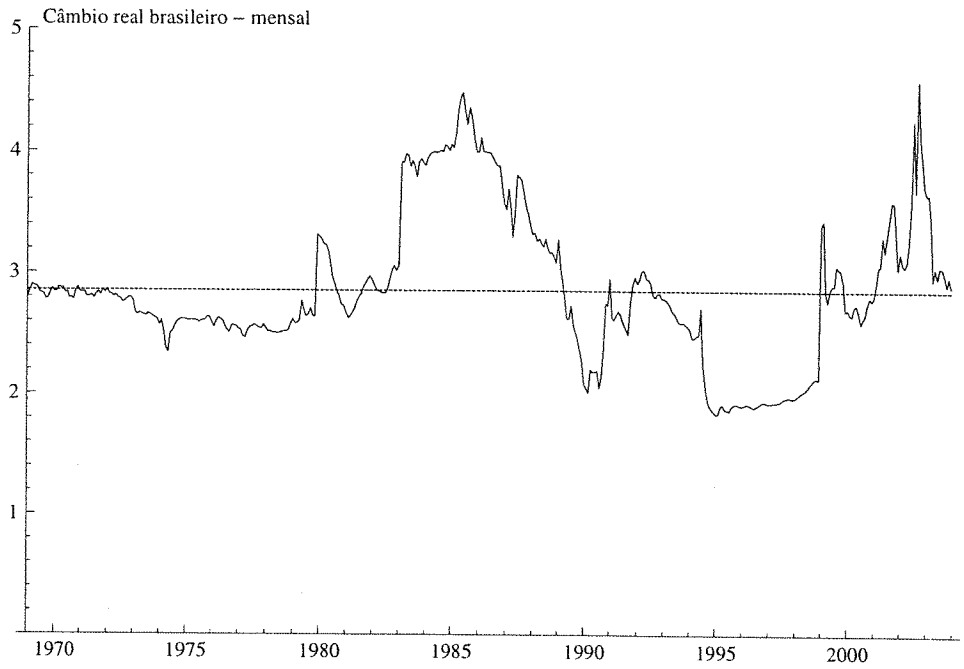

Figura 6.11: Taxa de câmbio real em R$/US$ - valor a cada final de mês de jan/1969 a dez/2003

Teste de hipóteses

$$\begin{aligned} H_0 &: \theta = 0\,, \qquad \theta \equiv \sigma_\eta^2/\sigma_\varepsilon^2 \\ H_1 &: \theta > 0 \end{aligned} \tag{6.14}$$

Resultados

A estimativa para $\hat{\mu}_0$ é de 2.8535 e as estimativas para a largura de faixa ótima l^* e a variância de longo termo σ_L^2 são

$$\begin{aligned} \widehat{l^*} &= 27 \\ \hat{\sigma}_L^2 &= 9.617 \end{aligned}$$

6.3 Câmbio real brasileiro

A estatística ξ_μ do teste ILO e o p-valor associado a sua distribuição assintótica ($\sim$ CvM) são

$$\begin{aligned} \xi_\mu &= 0.12473 \\ \text{p-valor} &= 0.4766 \end{aligned}$$

Portanto, não se rejeita a hipótese de estacionariedade para essa particular série de câmbio real brasileiro no período analisado. Assim, pela metodologia utilizada, há evidência estatística da validade da paridade do poder de compra.

Referências Bibliográficas

[1] Anderson, T. W. e Darling, D. A. (1952) Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Stat.*, **23**, 193-212.

[2] Andrews, D. J. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, **59**, 817-858.

[3] Araújo, C. H. V. e Silveira F°, G. B. (2002) Mudanças de regime no câmbio brasileiro *Working Paper Series - Banco Central do Brasil*, **WP41**

[4] Billingsley, P. (1995) *Probability and Measure* 3a. edição. New York: Wiley-Interscience.

[5] Busetti, F. e Harvey, A. C. (2001) Testing for the presence of a random walk in series with structural breaks. *J. Time Ser. Anal.*, **22**, 127-150.

[6] Busetti, F. (2002) Testing for (common) stochastic trends in the presence of structural breaks. *J. Forecast.*, **21**, 81-105.

[7] Doornik, J. A. (1999) *Object-Oriented Matrix Programming Using OX* 3a. edição. London: Timberlake Consultants Press.

[8] Epanechnikov, V. A. (1969) Non-parametric estimation of multivariate probability density. *Theory of Probability and Its Applications*, **14**, 153-158.

[9] Ferguson, T. S. (1967) *Mathematical Statistics: A Decision Theoretic Approach.* New York: Academic Press.

[10] Harvey, A. C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

[11] King, M. L. e Hillier, G. H. (1985) Locally best invariant tests of the error covariance matrix of the linear regression model. *J. R. Statist. Soc. B*, **47**, 98-102.

[12] Kurozumi, E. (2002) Testing for stationarity with a break. *J. Economet.*, **108**, 63-99.

[13] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. e Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. Econometrics*, **44**, 159-178.

[14] Macneill, I. B. (1978) Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *Ann. Stat.*, **6**, 422-433.

[15] Maddala, G. S. e Kim, I-M. (1998) *Unit Roots, Cointegration, and Structural Change*. Cambridge: Cambridge University Press.

[16] Nelson, C. R. e Plosser, C. I. (1982) Trends versus random walks in macroeconomic time series: some evidence and implications. *J. Monet. Econ.*, **10**, 139-162.

[17] Newey, W. K. e West, K. D. (1994) Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, **61**, 631-653.

[18] Nyblom, J. (1986) Testing for deterministic linear trend in time series. *J. Am. Stat. Assoc.*, **81**, 545-549.

[19] Nyblom, J. e Makelainen, T. (1983) Comparisons of tests for the presence of random walk coefficients in a simple linear model. *J. Am. Stat. Assoc.*, **78**, 856-864.

[20] Parzen, E. (1957) On consistent estimates of the spectrum of a stationary time series. *Ann. Math. Stat.*, **28**, 329-348.

[21] Perron, P. (1989) The Great Crash, the oil price shock, and the unit root hypothesis. *Econometrica*, **57**, 335-346.

[22] Phillips, P. C. B. (1987) Time series regression with a unit root. *Econometrica*, **55**, 277-301.

[23] Phillips, P. C. B. e Perron, P. (1988) Testing for a unit root in time series regression. *Biometrika*, **75**, 335-346.

[24] Piessens, R., Uberhuber, C. W. e Kahaner, D. K. (1983) *QUADPACK, A Subroutine Package for Automatic Integration.* Berlin: Springer-Verlag.

[25] Priestley, M. B. (1981) *Spectral Analysis and Time Series.* London: Academic Press.

[26] Sheather, S. J. e Jones, M. C. (1991) A reliable data-based bandwidth selection method for kernel density estimation. *J. R. Statist. Soc. B*, **53**, 683-690.

[27] Shorack, G. R. e Wellner, J. A. (1986) *Empirical Processes with Applications to Statistics.* New York: John Wiley & Sons.

[28] Tanaka, K. (1996) *Time Series Analysis: Nonstationary and Noninvertible Distribuiton Theory.* New York: Wiley-Interscience Publication.