



# Boletim



**SOCIEDADE PORTUGUESA  
DE ESTATÍSTICA**

*Publicação semestral*

*primavera de 2016*



## Séries Temporais e suas aplicações

### Modelação de séries temporais de contagem

Maria Eduarda Silva ..... 33

### Séries temporais em Economia

Paulo Teles e André Almeida ..... 41

### Agrupamento de séries temporais e sua aplicação na análise de processos geofísicos e ambientais

Manuel G. Scotto e Susana M. Barbosa ..... 75

### A dinâmica TGARCH de potência na evolução temporal da série das manchas solares de Wolfer

Esmeralda Gonçalves e Nazaré Mendes Lopes ..... 86

Editorial .....	1
Mensagem da Presidente .....	3
Notícias .....	4
<i>Enigmística</i> .....	10
SPE e a Comunidade .....	11
Ciência Estatística .....	94
Prémios “Estatístico Júnior” .....	98
Retrospectiva .....	100

### Informação Editorial

**Endereço:** Sociedade Portuguesa de Estatística.  
Campo Grande. Bloco C6. Piso 4.  
1749-016 Lisboa. Portugal.

**Telefone:** +351.217500120

**e-mail:** [spe@fc.ul.pt](mailto:spe@fc.ul.pt)

**URL:** <http://www.spestatistica.pt>

**ISSN:** 1646-5903

**Depósito Legal:** 249102/06

**Tiragem:** 500 exemplares

**Execução Gráfica e Impressão:** Gráfica Sobreireense

**Editor:** Fernando Rosado, [fernando.rosado@fc.ul.pt](mailto:fernando.rosado@fc.ul.pt)

**Sociedade Portuguesa de Estatística desde 1980**



SOCIEDADE PORTUGUESA  
DE ESTATÍSTICA

# PRÉMIO ESTATÍSTICO JÚNIOR 2016



Candidaturas até  
**01 DE JUNHO  
DE 2016**

## CONTACTOS

Sociedade Portuguesa de Estatística  
Bloco C6, Piso 4 – Campo Grande  
1749-016 Lisboa  
Telefone 21 750 04 27

[www.spestatistica.pt](http://www.spestatistica.pt)  
[spe@fc.ul.pt](mailto:spe@fc.ul.pt)

Com o apoio:

 Porto  
Editora®

# Editorial

## **... Boletim SPE, *decimus annum* em nova série, um projeto consolidado...**

Esta edição do Boletim SPE completa uma sequência de vinte nas quais, ao longo de dez anos, (re)visitámos as mais importantes áreas científicas e de investigação da Ciência Estatística.

Passado todo esse tempo e para atualização todos esses campos científicos devem e podem ser alvo de novas edições.

A volatilidade, o carácter efémero, da investigação científica assim exige!

Todos os Boletins, em cada edição, debruçam-se sobre um Tema Central. Na concretização deste, como Editor, tenho contado com o apoio de pelo menos um investigador sénior da área específica abordada. Assim, o Boletim tem (pelo menos) um co-editor que lhe dá corpo e substância. Ele, na prática, é “o criador” da correspondente edição.

Reconhecidamente, agradeço aos colegas que têm colaborado nessa tarefa ao longo de tantos anos e de tantas edições. Aqueles que têm alguma experiência editorial bem saberão avaliar o valor deste reconhecimento.

Tudo pode ser melhorado! Desiderato que, como editor, sempre desejei e que, com o maior empenho, quis cumprir em cada Boletim.

A evolução feita ao longo desta série editorial deu a experiência que, de algum modo, facilita e em todos os Boletins faz sobressair um líder científico. Assim, convidada pelo editor e renovada em cada edição, na informalidade que nos guia, foi criada uma “Comissão Editorial”, que sabemos muito bem vinda e desejada em todas as publicações científicas mesmo naquelas com índole de divulgação como se pretende com o Boletim SPE.

Aquela Comissão, na prática, é o máximo responsável pela congregação de autores convidados e o grande impulsionador do respetivo Tema Central. Dela depende o maior ou menor sucesso da missão editorial do Boletim no seu momento de apresentação do Estado da Arte numa determinada área do saber estatístico.

É um instrumento fundamental que consolida o Boletim e faz cumprir a missão Estatutária da SPE de divulgação da Estatística. O mesmo se pode dizer da relevante secção *SPE e a Comunidade* deste Boletim.

Desiderato cumprido?

A palavra deve dar-se à comunidade científica, aos leitores!

E, em momento “aniversário”, todos os autores, reconhecidamente, devem ser associados e também os leitores pelas muitas sugestões e comentários recebidos bem como o incentivo à continuidade!

No próximo Boletim Outono 2016 completam-se 10 anos nesta fase de edição iniciada em 2006, a convite da Direção.

O Boletim está consolidado!

Mas a consolidação também exige sempre melhoria...

No 10º aniversário desta nova série, está na hora de visitar os “Tema Central” e talvez o mais relevante na conjuntura seja A Investigação Estatística – um novo ponto da situação.

A comunidade tem a palavra!

## **... sobre a soma das partes que criam um Congresso...**

O Congresso SPE é o principal objectivo estatutário, como acontece e deve suceder em todas as associações congéneres.

É a grande oportunidade de iniciação para muitos estatísticos. É também a reunião magna!

O Congresso SPE é o resultado de uma soma de esforços: da Comissão Organizadora, da Comissão Científica e da Direção, que realizam uma iniciativa Presidencial mandatada pelos sócios e concretizam os grandes objetivos estatutários. Realizar congressos bienais é cumprir (apenas) metade do objetivo pois a máxima divulgação da Ciência Estatística e a reconhecida adesão de inscitos implicaria a periodicidade anual.

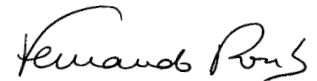
Do saber de experiência feito sabemos do esforço envolvido na concretização de um Congresso.

Mas a soma das partes que concorrem torna mais fácil cumprir a missão. Todos os sócios e, muito em especial, os Seniores que sustentam a SPE e o seu sucesso (também em congressos) devem tomar a iniciativa de oferecer os seus préstimos facilitando assim o trabalho de pesquisa da Direcção sobre “potenciais voluntários”. Invertamos assim o ónus da procura, em benefício da SPE. Decerto o trabalho da Direcção fica ainda mais estimulado.

De modo original e para que conste, este Boletim inclui a expressão da satisfação de missão cumprida apresentada pela Comissão Organizadora do XXII Congresso SPE.

É um grande estímulo!

O Tema Central do próximo Boletim SPE será *O Tema Central da Estatística*.



# Mensagem da Presidente

Caros Sócios da SPE

Passado 1 ano da tomada de posse da actual Direcção é tempo de reflectir sobre as actividades passadas e planear as actividades futuras. Este exercício foi feito na preparação da Assembleia Geral Ordinária de 2016 que teve lugar dia 24 de Fevereiro. Num resumo breve, as actividades nortearam-se segundo os dois grandes objectivos traçados por esta Direcção:

**Objectivo I - aumentar a sensibilização pública para a Estatística e aumentar a visibilidade da SPE na sociedade;**

**Objectivo II - aumentar a coesão interna da sociedade e apoiar o desenvolvimento da Estatística em Portugal.**

Assim, no âmbito do **Objectivo I** destaco o desenvolvimento das actividades *Prémio Estatístico Júnior, AEVAE e Explorística*. Relativamente ao **Objectivo II**, atribuiu-se o Prémio SPE 2015 a Ana Isabel Coelho Borges com o trabalho intitulado *Modelação Conjunta de Dados Longitudinais e de Sobrevivência de Cancro da Mama*; concederam-se 13 bolsas de participação no XXII Congresso da SPE a alunos de Mestrado e Doutoramento, apoiou-se a organização do XXII Congresso, iniciaram-se os trabalhos de organização do II Encontro Luso-Galaico de Biometria com aplicações à saúde, ecologia e ambiente, estabeleceu-se uma colaboração com a DStatG, Sociedade Alemã de Estatística, assegurou-se a representação da SPE em diversas organizações nacionais e internacionais. Por último, mas igualmente senão mais importante, atribuiu-se o Prémio Carreira 2015 à Professora Antónia Turkman e naturalmente, assegurou-se a continuidade deste Boletim. Os sócios podem consultar os detalhes destas actividades e de outras não referidas no Relatório de Actividades de 2015 que se encontra disponível, pela primeira vez, na página da SPE. A Direcção quer deixar aqui um agradecimento aos sócios que, certamente à custa de sacrifício pessoal, possibilitaram a realização das actividades no ano de 2015. O Plano de Actividades para 2016 que se norteia pelos mesmos objectivos, também se encontra disponível para consulta. A discussão destes documentos quer no Conselho Geral quer na Assembleia Geral foi interessante e esclarecedora.

A Sociedade é dos sócios e para os sócios e é, essencialmente, o que os sócios fizerem dela. Assim, a Direcção está aberta a apoiar todas as iniciativas em prol da Estatística em Portugal.

Porto, 25 de Fevereiro de 2016

Cordiais saudações

Maria Eduarda Silva

# Notícias

## • II Encontro Galaico-Português de Biometria



De 30 de Junho a 2 de Julho de 2016 vai-se realizar em Santiago de Compostela, na Faculdade de Matemática, o II Encontro Galaico-Português de Biometria.

Mais informações: <http://biometria.sgapeio.es/>

FR

## • 25th International Workshop on Matrices and Statistics



De 6 a 9 de junho, vai decorrer na Madeira o "25th International Workshop on Matrices and Statistics".

Mais informação em:

<http://www.iwms.ipt.pt>

Ana Abreu

## • Prémios SPE 2016

Com o objectivo de fomentar o interesse pela estatística em jovens investigadores, está prevista a abertura de concursos para atribuição dos seguintes prémios

**Prémio SPE - Jovem Investigador 2016**

**Prémio SPE - Iniciação à Investigação 2016**

Informação detalhada pode ser consultada em <http://spestatistica.pt/>

ES

## • Dia Europeu da Estatística

A FENStatS com a aprovação da ESAC e Eurostat designou o dia 20 de Outubro como o Dia Europeu da Estatística, "European Statistics Day", a partir de 2016 e a celebrar nos anos em que não é celebrado o Dia Mundial da Estatística, "World Statistics Day".

ES

## • Entrevista da Presidente da SPE à revista "Pontos de Vista"



A revista «Pontos de Vista» distribuída juntamente com o jornal «Público» no dia 14 de Novembro de 2015, entrevistou a Prof. Eduarda Silva, Presidente da SPE, no âmbito das comemorações do «Dia Mundial da Estatística».

A entrevista está disponível em:

<http://www.spestatistica.pt/images/noticias/entrevista.pdf>.

FR

## • XXII Congresso SPE 2015 pela Comissão Organizadora



O XXII Congresso da Sociedade Portuguesa de Estatística (SPE), organizado em conjunto com a Universidade do Algarve, que se realizou de 07 a 10 de Outubro de 2015, no Centro de Congressos Ria Formosa sediado no Real Marina Hotel & SPA, em Olhão, decorreu de forma excelente, tendo reunido perto de 200 participantes, na grande maioria portugueses mas também com a presença de outras nacionalidades.

Desta forma, foi possível, mais uma vez, fomentar o desenvolvimento da investigação na área da Probabilidade e da Estatística, promover a sua implantação junto da sociedade civil e o intercâmbio científico através do diálogo e colaboração entre os seus participantes.

Outro marco extremamente importante alcançado neste evento, foi a vincada divulgação de Portugal, da região do Algarve, da Universidade do Algarve e, em particular, da cidade de Olhão.



O programa científico contemplou 4 sessões plenárias, 7 sessões temáticas com 21 comunicações orais, 82 comunicações orais livres e 30 *posters*. Para as sessões plenárias foram convidados os oradores Luzia Gonçalves, da Universidade Nova de Lisboa, James W. Taylor, da Universidade de Oxford, Manuel Scotto, da Universidade de Lisboa e Peter Müller, da Universidade do Texas, Austin.



No dia 7, o início dos trabalhos foi precedido pelo habitual minicurso do congresso da SPE, este ano, intitulado "Estatística Bayesiana Computacional - uma introdução", e lecionado por Maria Antónia Amaral Turkman (CEAUL & FCUL) e por Carlos Daniel Paulino (CEAUL & IST).





A sessão de abertura do congresso decorreu ao fim da tarde, com a presença de Conceição Ribeiro e Clara Cordeiro (comissão organizadora), António Miguel Pina, presidente da Câmara Municipal de Olhão, Maria de Lurdes Cristiano (em representação do Reitor da UAlg e diretora da Faculdade de Ciências e Tecnologia), Maria Eduarda Silva (presidente da SPE), Carlos Daniel Paulino (presidente cessante da SPE) e Carlos Marcelo (representante do Instituto Nacional de Estatística).



Após a sessão de abertura seguiu-se a primeira plenária, intitulada "A Bayesian Feature Allocation Model for Tumor Heterogeneity", proferida por Peter Müller, da Universidade do Texas.



O primeiro dia terminou com a Receção de Boas-Vindas, no Chalé João Lúcio, cortesia da Câmara Municipal de Olhão. O ambiente foi sempre animado, tendo contribuído para isso, o local, a boa disposição dos participantes e as diversas iguarias da região, tudo muito bem confeccionado e apresentado pelos alunos da Escola Básica 2,3 Ciclos João da Rosa, de Olhão.

A importância dos patrocinadores foi diretamente proporcional ao sucesso do congresso, tendo alguns deles marcado presença através dos seus expositores.



No dia 8, a manhã foi preenchida por diferentes sessões de comunicações orais temáticas e livres, uma sessão de posters, e pela segunda sessão plenária, proferida por James Taylor, da Universidade de Oxford, e intitulada "Predicting the expected shortfall corresponding to value at risk forecasts produced by quantile models". À tarde decorreu a atribuição do Prémio SPE 2015 à aluna do

Doutoramento em Ciências - especialidade em Matemática da Escola de Ciências da Universidade do Minho, Ana Isabel Borges, com o trabalho “Modelação Conjunta de Dados Longitudinais e de Sobrevivência de Cancro da Mama”. O Prémio SPE é atribuído anualmente e tem por objetivo estimular a atividade de estudo e investigação científica em Probabilidades e Estatística entre os jovens que trabalham nestas áreas.



O 2º dia do congresso terminou com o passeio. Tendo o congresso sido organizado em Olhão, e uma vez que o tempo ajudou, realizou-se um passeio de barco pela ria Formosa. As características ambientais e paisagísticas inerentes à ria conferem-lhe um elevado valor científico, cultural, social e económico que se tem preservado ao longo do tempo. Para além uma visita geral pela ria, efetuaram-se duas paragens; a primeira à ilha da Armona, seguindo-se a ilha do Farol. Apesar das 3 horas de viagem os participantes estiveram sempre muito animados, inspirados pela beleza do local.

No dia 9, decorreu mais uma manhã preenchida por diferentes sessões de comunicações orais temáticas e livres, uma sessão de posters, e Luzia Gonçalves, do Instituto de Higiene e Medicina Tropical da Universidade Nova de Lisboa foi a oradora convidada na terceira sessão plenária intitulada "Dados, Ética e Modelos na Saúde Tropical: Constrangimentos e Desafios".



Um dos momentos mais marcantes deste dia do XXII Congresso da SPE foi a atribuição do prémio Carreira SPE a Maria Antónia Amaral Turkman, pela sua obra científica e dedicação ao desenvolvimento e divulgação da Estatística em Portugal.



À noite decorreu o jantar do congresso que juntou todos os participantes e acompanhantes num serão de boa disposição e excelente convívio, animado pelo Rancho Folclórico de Santo Estevão, Tavira.

No dia 10, dia de encerramento do XXII Congresso da Sociedade Portuguesa de Estatística, Manuel Scotto, do Instituto Superior Técnico da Universidade de Lisboa, falou sobre "O Operador Thinning na Modelação de Séries Temporais de Valores Inteiros", na última sessão plenária, durante mais uma manhã preenchida por diferentes sessões de comunicações orais temáticas e livres.



A sessão de encerramento do evento decorreu às 13h30, tendo a presidente da SPE tecido grandes elogios à comissão organizadora do congresso pelo excelente trabalho desenvolvido para concretizar o XXII Congresso da SPE. No entanto, estes resultados formidáveis só foram possíveis devido à colaboração de todos.

Um grande agradecimento a todos os participantes e a todos os patrocinadores,  
Novembro de 2015

A Comissão Organizadora

Nelson Antunes, Maria Helena Gonçalves, Clara Cordeiro, Conceição Ribeiro e Carlos Sousa



## Enigmística de mefqa

D A D U S

(A, N, Á, L, I, S, E)

No Boletim SPE outono de 2015 (p.13):

*distração  
distribuição  
distribuição  
distribuição*

mistura de distribuições

*índice*

índice de cauda

## SONDAGENS E ESTATÍSTICA

Rui Oliveira Costa, *geral@eurosondagem.pt*

*Eurosondagem e Universidade Lusófona*

Os Estudos de Opinião, mais conhecidos via os média por Sondagens, também são Estatística, mas não são só Estatística.

Situam-se na área das Ciências Sociais e não nas Ciências Puras. São nalguns casos indicadores de tendências, noutros medição “da Temperatura”, noutros ainda tentativas de perceção de futuras escolhas.

Têm evoluído a nível mundial e também em Portugal. Com altos (mais) e alguns baixos, têm vindo a cimentar-se e a credibilizar-se. Os atores políticos valorizam-nas mais ou menos conforme a “simpatia” para com eles dos resultados, mas de um modo geral já não as atacam. Digamos que se habituaram. É este o atual “Estado da Arte”.

Durante o ano de 2015, tiveram lugar eleições em Espanha, França, Grécia, Reino Unido, Dinamarca, Polónia, Turquia e noutros Países com menos atenção mediática.

Contando com Eleições Regionais (pelo menos Espanha e França), foram dezenas. Só os resultados do Partido Trabalhista no Reino Unido e do Partido Cidadãos em Espanha, ficaram a cerca de 5% do expectável. Em todos os outros (repito dezenas) não ocorreram surpresas. Vamos ver como correm as Presidenciais Norte Americanas.

Esta performance traz mais responsabilidade. Em Portugal correu bastante bem, tanto no que concerne às Eleições Legislativas, como nas Presidenciais.

Mas os Estudos de Opinião mantêm um “inimigo” – a abstenção.

Vejam: 80% dos inquiridos respondem ir votar e como. Daí os estudos transmitirem a existência de 20% de “indecisos”. Alguns julgam estar aí o problema. Nada mais errado. O problema não são os 20% que dizem ir votar, mas que ainda não decidiram. Destes, provavelmente 9 em cada 10 ficam indecisos até ao fim e irão engrossar a abstenção. O problema está entre os 80% que dizem ir votar e em quem, e dos quais 20% ou 30% não votam. Estes são os inimigos dos Estudos de Opinião pré-eleitorais. Por alguma razão (e é esta) as Sondagens pré-eleitorais no Brasil são tão apreciadas. Devem-no ao voto obrigatório, que coloca a abstenção à volta dos 3%.

A maior parte dos meus colegas Brasileiros aprenderam nos Estados Unidos da América. São melhores que os seus ex-professores? É que estes (os Norte-Americanos) deparam-se com taxas de abstenção algumas vezes perto dos 50%. Mas se a abstenção é um fator determinante nos Estudos pré-eleitorais, já no que concerne às previsões em dia de eleições (Sondagem Boca de Urna), só contam os que votam e aqui estamos todos (a nível mundial) em igualdade de circunstâncias.

Sujeitos a uma avaliação tremenda, cerca de 2 a 3 horas após as previsões (e estas sim são mesmo previsões, com intervalos e sem desculpas).

A Boca de Urna em Portugal é das melhores do Mundo (obviamente colocando aqui os meus leais concorrentes). O País acredita, e por vezes até de mais, pois trata as previsões como resultados.

Sem falsa modéstia, permitam que refira as últimas Eleições (as Presidenciais de 24 de Janeiro), apresentando a projeção que a Eurosondagem fez na SIC:

Às 19 Hs

Abstenção  
De: 48,8  
A: 53,2  
Ponto médio: 51,0%  
Resultado real: 51,2%

Às 20 Hs

Candidatos	Previsão	Resultado
<b>Henrique José de Sousa Neto</b>	<b>De: 0,7 A: 1,4</b>	<b>0,8%</b>
<b>António Manuel Seixas Sampaio da Nóvoa</b>	<b>De: 21,1 A: 24,8</b>	<b>22,9%</b>
<b>Cândido Manuel Pereira Monteiro Ferreira</b>	<b>De: 0,2 A: 0,6</b>	<b>0,2%</b>
<b>Edgar Freitas Gomes da Silva</b>	<b>De: 3,0 A: 4,1</b>	<b>4,0%</b>
<b>Jorge Manuel Pais Seara Rodrigues Sequeira</b>	<b>De: 0,2 A: 0,6</b>	<b>0,3%</b>
<b>Vitorino Francisco da Rocha e Silva</b>	<b>De: 2,6 A: 3,7</b>	<b>3,3%</b>
<b>Marisa Isabel das Santos Matias</b>	<b>De: 8,8 A: 11,3</b>	<b>10,1%</b>
<b>Maria de Belém Roseira Martins Coelho Henriques de Pina</b>	<b>De: 3,3 A: 4,8</b>	<b>4,2%</b>
<b>Marcelo Nuno Duarte Rebelo de Sousa</b>	<b>De: 50,1 A: 54,5</b>	<b>52,0%</b>
<b>Paulo Alexandre Baptista Teixeira de Morais</b>	<b>De: 1,4 A: 2,5</b>	<b>2,2%</b>

Como se constata, todos os candidatos (e foram 10) situam-se dentro dos intervalos, e com diferenças para o ponto médio de décimas.

Mérito da Estatística? Também. Esta ajuda e muito a estudar o Universo.

Há todo um histórico eleitoral a analisar. Mas nenhum candidato era repetente. Existem pois vários históricos (Legislativas, Presidenciais e não só).

Há segredo, não.

Há trabalho e muito. Cruza a estatística pura, dura e essencial, com a análise política e social.

E daqui resulta o essencial – a escolha da AMOSTRA.

É a relação que se pretende a mais perfeita que for possível, entre Estatística e Ciências Sociais.



# **Estatística, inquéritos de opinião e cidadania**

António Salvador, *a.salvador@intercampus.pt*

*Intercampus*

## **Introdução.**

Gostaria de começar esta minha (necessariamente) despretensiosa participação por um agradecimento ao amável e, por que não dizê-lo, honroso convite que me foi feito pelos responsáveis pelo Boletim da Sociedade Portuguesa de Estatística para escrever algumas palavras relacionadas com as relações entre a Estatística e os estudos de opinião e as sondagens eleitorais.

Foi quando reflecti um pouco sobre o tópico que me dei conta da responsabilidade que isto significava.

Afinal a Estatística tem uma grande importância na vida de todos nós, no nosso quotidiano, e essa importância reflecte-se em variadíssimos aspectos sociais. Então a Estatística não é a arte de narrar uma história com recurso a números? No entanto, Estatística é também um nome pomposo, que traduz a por vezes enorme dificuldade que encerra dentro de si. Nesse sentido, e apesar de trabalhar em investigação de mercado e em estudos de opinião há muitos anos, não me posso considerar um especialista em Estatística – o que explica o termo “despretensioso” de há pouco.

Começaria por sublinhar o quanto a ligação dos inquéritos de opinião à Estatística e às estatísticas é inquestionável. Podemos dizer, de forma simplista, que os inquéritos de opinião produzem estatísticas (ou seja, dados estatísticos) utilizando os conhecimentos da Estatística, sobretudo no que diz respeito à sua dimensão inferencial, pois todo o conhecimento que estes inquéritos produzem é baseado na utilização de amostras que desejamos representativas dos universos que pretendemos conhecer.

Deste modo, toda a reflexão e investigação que a Estatística desenvolveu ao longo de anos e anos no domínio da amostragem teve uma aplicação imediata nos inquéritos por sondagem e permitiu a produção de uma quantidade impressionante de informação sobre as sociedades em que vivemos, pois a consulta através de recenseamentos depressa se limitou a aplicações muito delimitadas e altamente dilatadas no tempo.

Começamos então pelo que os inquéritos de opinião devem à ciência Estatística.

## **1. A produção de informação.**

Perante as limitações do uso dos recenseamentos – devido à sua logística, duração de aplicação e custo – muito cedo os inquéritos por sondagem adquiriram um estatuto ímpar na produção de informação estatística sobre as características das populações dos países ocidentais. Esta prática foi desenvolvida, em simultâneo, pelas instituições públicas produtoras de «estatísticas oficiais» e pelas empresas privadas. Os inquéritos por sondagem conseguiram assim tornar-se instrumentos fundamentais de gestão pública e de gestão privada.

No primeiro caso, continuou a utilizar-se o conceito de inquérito de opinião (ou inquérito sociológico). E essa metodologia serviu para produzir informação sobre o país, mas também para conhecer as expectativas e os anseios da população para melhorar a gestão pública.

No segundo caso, passou a utilizar-se o conceito de «estudos de mercado». E essa metodologia serviu para produzir informação sobre os consumidores, os seus hábitos, as suas preferências e escolhas, até mesmo as suas intenções futuras.

Em todos estes estudos e projectos, a Estatística tem sido omnipresente. Não só na parte da constituição de boas amostras, que permitam inferir para o Universo com segurança, como na parte da Estatística descritiva (ou mesmo interpretativa), que permite resumir e analisar os resultados obtidos.

Resumindo, a estatística encontrou nestes inquéritos por sondagem um veículo fundamental para ajudar ao desenvolvimento das sociedades modernas e à emancipação dos cidadãos, através de um conhecimento acrescido da realidade que os circunda, o que tem contribuído para a sua capacidade de formar uma opinião e decidir por si. Este aspecto concreto tornou-se especialmente importante no âmbito dos estudos políticos e eleitorais.

## **2. Os inquéritos eleitorais.**

Ninguém duvida da contribuição que os estudos políticos e eleitorais, realizados de forma científica, têm tido para o fortalecimento da democracia.

Desde meados do século passado – quando Gallup impõe, nos EUA, o prestígio deste tipo de estudos – que as democracias ocidentais não conseguiram prescindir mais deste tipo de informação. Que podemos dividir em três grupos.

Em primeiro lugar, temos os estudos políticos em geral, realizados independentemente dos actos eleitorais. Têm frequentemente objectivos científicos, por exemplo, a compreensão do comportamento eleitoral das populações, mas têm também objectivos de gestão política, tal como os restantes inquéritos sociológicos. Assim, governos e câmaras municipais realizam este tipo de investigação junto de amostras representativas dos seus eleitores, de modo a melhor gerir as necessidades e as expectativas da população. Assim fazendo, contribuem para que a opinião dos cidadãos seja tida em consideração de um modo muito mais intenso do que acontecia no passado. O que faz aumentar de forma objectiva a sua influência, mesmo quando esses eleitores se sentem relativamente afastados da «política» e dos «políticos».

Em segundo lugar, tomaram grande relevo as chamadas «sondagens pré-eleitorais». Este conceito de «sondagem» é um pouco desajustado, mas tornou-se habitual, passando os «inquéritos» a designar as recolhas de opinião mais extensas e aprofundadas e as «sondagens» as recolhas rápidas sobre temas delimitados, como, neste caso, a intenção de voto, ou seja, em quem os eleitores pensam votar. No entanto, como se sabe, qualquer que seja o âmbito ou a extensão, estamos sempre em presença de inquéritos de opinião, realizados «por sondagem», que, em concreto, significa que a informação é extraída através de uma amostra representativa e não junto do Universo inteiro.

Estes inquéritos pré-eleitorais têm servido para os agentes políticos, por um lado, terem conhecimento do resultado das suas sucessivas decisões e actuações, e para os eleitores, por outro lado, terem informação em relação ao que pensam os seus concidadãos, de modo a poderem decidir com mais conhecimento de causa e, eventualmente, a ter um comportamento de «voto útil» nos actos eleitorais.

A sua importância pode ser bem ilustrada pelo facto de terem sido elas o veículo de promoção dos inquéritos por amostragem nas sociedades modernas. Tal aconteceu nos EUA, como aconteceu em Portugal. Mais uma razão para elas ocuparem, neste texto, um merecido lugar de destaque. De facto, até ao ano de 1971, para além dos Censos, o número dessas sondagens em Portugal foi zero. E, até



1974, foram efectuadas, quase em clandestinidade, três sondagens eleitorais (algumas das quais, da autoria da Serte), que chegaram a ser publicadas pelo novo semanário Expresso (em 1973). Em 1974, uma organização gerida a nível privado, o CEAD, promove uma grande sondagem que incluiu intenções de voto, entre outros temas. Falamos, portanto, de zero sondagens até 1971 e de quatro sondagens até 1974/1975.

E existe uma razão para estas sondagens eleitorais terem assumido essa função de prestigiar os inquéritos por sondagem: a possibilidade de comparação dos resultados dos inquéritos com os das próprias eleições, vieram mostrar, de forma empírica, as potencialidades deste tipo de metodologia e desta ciência.

Finalmente, existem as «sondagens de boca-de-urna». E podemos dizer que nenhum outro tipo de inquérito contribuiu tanto para o prestígio de que falava há pouco. Com efeito, enquanto as sondagens pré-eleitorais podem «errar» (ou seja, os seus resultados não coincidirem – muito naturalmente – com os resultados da eleição, até porque estamos a lidar com momentos diferentes no tempo e com universos diversos, o dos eleitores e o dos votantes), as sondagens «de boca-de-urna» constituem um desafio quase único no domínio da investigação científica. Isto, porque é muito raro um qualquer “cientista” ser confrontado com um julgamento implacável e inquestionável de objectividade como este. O “cientista” dá os resultados da sua previsão às 19 horas (previsão acerca dos resultados de uma eleição que acabou de terminar) e, passadas apenas 2 ou 3 horas, sabe, de forma inequívoca se acertou ou se falhou. Se foi competente ou se não foi. É muito difícil encontrar, como disse, uma outra qualquer profissão de índole científica que se defronte com um exame tão cruel.

Apesar deste risco, os profissionais têm sido muito corajosos. E, após quase 30 anos de sondagens de «boca-de-urna» (a primeira foi realizada em 1987), é extremamente interessante observar que tanto os agentes políticos como, sobretudo, a população, acreditam de forma total nas projecções das televisões quando fecham as urnas. Por exemplo, quando é prevista uma derrota de um partido, são muito poucos os apoiantes que se deslocam às sedes de campanha ou aos hotéis onde estão reunidos os dirigentes desse partido.

Ora, como dizia há pouco, esta credibilidade arrastou a confiança das populações para os outros estudos realizados através da mesma metodologia, o que faz com que essas populações acreditem nos estudos de opinião em geral e comecem a usar os dados estatísticos para formular as suas opiniões, não apenas em relação às questões políticas, mas no que diz respeito às questões globais das sociedades em que vivem.

Numa palavra, estes estudos contribuíram para o prestígio do método de inquérito por sondagem em específico, mas também para divulgar a importância da estatística em geral e, de uma forma ainda mais ampla, para o tão necessário respeito pela importância dos «números», outrora tão desprezada após décadas e décadas de ignorância matemática e de quase desprezo generalizado, por tal disciplina.

### **3. A experimentação empírica.**

Apesar de ser imperioso assumir o que os inquéritos de opinião devem à ciência Estatística, não podemos deixar de anotar a relevância da influência inversa. Assim, inquéritos por sondagem em geral, inquéritos eleitorais em particular e sondagens de «boca-de-urna» de forma ainda mais expressiva, também permitiram aprofundar, e por vezes relativizar, os conhecimentos estatísticos, nomeadamente os relacionados com a amostragem. Tanto no que diz respeito às dimensões das amostras, como em relação às técnicas de amostragem propriamente ditas.

Tem sido assim possível, ao longo de todos estes anos, testar empiricamente técnicas de amostragem inicialmente afastadas ou desprezadas pela ciência estatística, nomeadamente as técnicas não-probabilísticas. E foi assim que a inicialmente desconsiderada «amostra por quotas» passou a ser

progressivamente utilizada em todo o mundo, com estatuto de igual para igual em relação aos processos de tipo aleatório. E se criaram outros métodos alternativos, como o *random route*.

Por outro lado, tem sido possível observar as reduções de qualidade provocadas pelo uso de amostras menores, tendo-se concluído que o reforço dos métodos de estratificação pode compensar de forma satisfatória alguma redução significativa da dimensão das amostras e, conseqüentemente, dos seus custos.

Sobretudo, a evidência empírica tem permitido relativizar fortemente a importância do «erro de amostragem» em comparação com o «erro de medida». É algo que os estatísticos não gostam de ouvir, pois a medição deste último foi sempre – uma quimera.

É assim indiscutível alguma complementaridade, que talvez tenha sido, de facto, mais expressiva na estimativa da importância do «erro de medida». E é com esse aspecto que termino esta pequena reflexão.

Sabemos hoje, pelo confronto de diversas metodologias semelhantes (amostras experimentais), ou pela comparação de resultados de inquéritos com resultados de eleições, que o «erro de medida» pode atingir valores extremamente inesperados e incómodos. E que o investigador ganha muito mais em preocupar-se com o método de recolha (entrevista directa ou telefónica, por exemplo) ou com a formulação e ordem das perguntas nos questionários, do que com as margens de «erro de amostragem» (dentro de certos limites, claro está).

Em suma, a metodologia das Ciências Sociais, através da utilização dos inquéritos por sondagem, tem permitido atravessar fronteiras que a Estatística sempre evitou numa base teórica (o que é natural). E podemos hoje estimar o que antes não podíamos, como por exemplo, e pegando no exemplo das sondagens eleitorais, qual a diferença expectável dos «indecisos» quando a pergunta sobre a intenção de voto é obtida através da resposta a um entrevistador ou através da colocação de um boletim de voto fictício numa urna. Ou qual a diferença de resultados entre uma pergunta de intenção de voto colocada no início ou no fim de um questionário.

Infelizmente, as empresas que realizam sondagens eleitorais ainda são obrigadas por lei, a apresentar nas suas fichas técnicas, cálculos do «erro de amostragem» – o que sugere ao leitor desprevenido que essa é uma margem de erro global, ou seja, uma diferença entre os valores apresentados e a «realidade» –, quando, por um lado, o «erro de medida» é normalmente muito superior, e, por outro lado, quando tal cálculo se baseia (segundo as próprias leis da Estatística) num pressuposto de existência de amostras aleatórias, o que, de uma maneira geral, não existe. O que significa que temos todos aprendido muito, mas uns aprenderam mais do que outros.



## **Sondagens - Falham e acertam no que mostram, no modo como são lidas, e nos objetivos a que se propõem**

João H. C. António, [jantonio@cesop.lisboa.ucp.pt](mailto:jantonio@cesop.lisboa.ucp.pt)

*Centro de Estudos e Sondagens de Opinião (CESOP)  
Universidade Católica Portuguesa*

### **“As sondagens falham muito” é argumento que convém**

O discurso público sobre sondagens no período decorrente entre maio e setembro de 2015 ficou muito marcado por um fenómeno já conhecido como o *falhanço das sondagens no Reino Unido*. Até quase à véspera das eleições legislativas portuguesas de 4 de outubro de 2015, a desconfiança acerca das sondagens esteve sempre presente na discussão política em Portugal. Os motivos para este *falhanço* britânico poderão ter sido vários (e.g., limitação a telefones fixos, baixa percentagem de resposta, aumento de sondagens feitas por internet com base amostral não aleatória ou mesmo ajuste ou omissão de resultados de uma sondagem em função dos resultados de outras<sup>1</sup>) e os vários agentes manifestaram no próprio dia das eleições a necessidade de analisar aprofundadamente o que se teria passado. No dia seguinte às eleições foi anunciada a realização de um inquérito independente mandatado pelo Conselho de Sondagens Britânico. As conclusões deste inquérito, divulgadas na imprensa no dia 19 de janeiro de 2016, apontam a falta de representatividade das amostras utilizadas na maioria das sondagens como o principal fator para justificar o desvio entre os resultados das sondagens e os resultados eleitorais, embora não se possa excluir, por impossibilidade de verificação objetiva, a possibilidade de ter havido mudança de sentido de voto tardia ou em cima da hora<sup>2</sup>. Outros hipotéticos fatores explicativos, como problemas relacionados com a formulação e ordem das perguntas nos questionários, ou a deficiente cobertura do universo (devido aos votantes com residência fora do Reino Unido) foram excluídos pelo grupo de peritos.

Vários fatores terão contribuído para este *falhanço*. Para além dos destacados no parágrafo anterior, parece-nos relevante, para o caso em análise e para a sua interpretação a partir de Portugal, observar com mais atenção o sistema eleitoral do Reino Unido. Os 650 membros do parlamento (House of Commons) são eleitos através de 650 círculos eleitorais uninominais. Ou seja, embora as sondagens sejam feitas a pensar numa eleição nacional, estamos na realidade perante 650 eleições locais. Isso por si só poderia ainda assim não ser fonte de dificuldade para a realização de sondagens<sup>3</sup>. Muito mais relevante do que o número de círculos é o facto de as sondagens indicarem a distribuição percentual das intenções de voto, i.e., a distribuição proporcional das intenções de voto, enquanto a

---

<sup>1</sup> “(...) market research agency Survation said it had "chickened out" of publishing a telephone poll on Wednesday evening, which showed the Tories on 37% and Labour on 31%. Its chief executive Damian Lyons said he would "always regret" the decision, but the results seemed "so out of line" with previous polls.” <http://www.bbc.com/news/uk-32652104>.

<sup>2</sup> <http://www.bbc.com/news/uk-politics-35347948>.

<sup>3</sup> Por exemplo, os círculos eleitorais em Portugal não têm introduzido especial ruído na qualidade das sondagens.

distribuição de mandatos nas eleições para o parlamento britânico resulta de uma regra de não proporcionalidade, segundo a qual apenas o mais votado em cada círculo é eleito. Ou seja, num sistema eleitoral como o britânico, deve-se esperar que as sondagens sejam capazes de estimar corretamente a distribuição de votos presente na população mas não será de esperar que a distribuição de mandatos seja igualmente acertada. Para um bom cálculo da distribuição do número de lugares no parlamento teríamos de fazer sondagens em cada um dos 650 locais ou, pelo menos, em todos aqueles que não pudéssemos assumir como de vitória assegurada para uma das listas concorrentes.

As sondagens não erraram estrondosamente no Reino Unido, mas, pelo menos em Portugal, deixaram a ideia de o terem feito. Isso deve-se, em nosso entender, ao facto de elas estarem a servir para estimar aspetos – a distribuição de lugares, e consequente constituição de maiorias/minorias – para as quais, no caso em análise, têm pouquíssima capacidade preditiva. Até poderão acertar com precisão no número de lugares, por exemplo se grande parte dos círculos se comportar de acordo com a média nacional. Mas poderão ficar aquém das expectativas sempre que isso não aconteça (e parece ter sido esse o caso).

Um agregador de sondagens britânico, o *Poll of polls* da BBC, indicava no dia 6 de maio (véspera das eleições) as seguintes estimativas:

Tabela 1. Resultados do *Poll of polls* da BBC (na véspera das eleições) e Resultados eleitorais

Partidos	BBC Poll of Polls	Resultados
Conservative	34%	36,9%
Labour	33%	30,4%
UKIP	13%	12,6%
Liberal Democrats	8%	7,9%
Green	5%	3,8%
Others	7%	8,4%

Fonte: <http://www.bbc.com/news/politics/poll-tracker>

Na semana que antecedeu as eleições foram publicadas sondagens para todos os gostos, variando a diferença entre Conservadores e Trabalhistas entre +6 pontos percentuais (p.p.) e -3 p.p.. Nessa semana foram publicadas duas sondagens que indicavam 6 pontos percentuais de avanço aos Conservadores – a diferença que se veio a verificar nas eleições. Mas a maioria das sondagens indicavam diferenças mais pequenas entre os principais partidos.

As sondagens falharam no Reino Unido (ou houve uma viragem de última hora que favoreceu o partido no poder). Mas mais do que um efetivo falhanço, ficou no discurso e na memória das pessoas a ideia de um grande falhanço: afinal, de umas eleições empatadas, cenário retratado pelas sondagens, passou-se para uma maioria absoluta no Parlamento<sup>4</sup>. É essa memória do grande falhanço que estaria e esteve disponível e imediatamente acessível para responder a, e justificar, resultados de sondagens que contrariassem o esperado nas semanas que antecederam as eleições legislativas portuguesas.

### O que era esperado nas eleições legislativas portuguesas?

Também em Portugal há uma ferramenta muito interessante para o acompanhamento de sondagens políticas. Trata-se do site *popstar*<sup>5</sup>, desenvolvido por uma equipa multidisciplinar liderada por Pedro Magalhães e envolvendo investigadores do ICS-ULisboa, do INESC-ID, da FEUP e do NIPE-UM. O *popstar* recolhe e analisa vários aspetos relacionados com a opinião política e

<sup>4</sup> Esta maioria absoluta só foi possível devido à não proporcionalidade do sistema eleitoral do Reino Unido. Em Portugal, mantendo o atual sistema eleitoral, seria impossível um partido obter maioria absoluta com apenas 36,9% dos votos.

<sup>5</sup> <http://www.popstar.pt/index.php>

económica, entre os quais aquilo que neste artigo nos interessa: as sondagens políticas. A ferramenta construída por esta equipa para analisar as sondagens políticas trata-se, segundo os seus autores<sup>6</sup>, de um agregador de sondagens “tecnicamente sofisticado”, que estima as médias para cada partido com recurso ao filtro de Kalman<sup>7</sup>. Este agregador de sondagens tem como dados-fonte as sondagens políticas publicadas na comunicação social e consequentemente depositadas na ERC<sup>8</sup>. Por esse motivo, a qualidade dos seus resultados deriva da qualidade das sondagens produzidas. Pelo que nos foi possível compreender, o agregador é sensível à dimensão da amostra de cada sondagem mas não é sensível à entidade responsável pela mesma nem ao método de inquirição utilizado. Assim sendo, sondagens com amostras maiores têm maior peso nas estimativas produzidas e o peso dos dados de cada empresa nas estimativas é reflexo do número de sondagens realizadas e da dimensão da amostra de cada uma delas. O agregador contabiliza 158 sondagens políticas realizadas entre as Eleições Legislativas de 2011 e as Legislativas de 2015. A maior parte destas sondagens foram realizadas pela Eurosondagem (publicadas no Expresso e na SIC) e pela Aximage (Correio da Manhã e Jornal de Negócios) (respetivamente, 34% e 32% do total). Sendo a média de inquiridos das sondagens da primeira empresa (M=1050) muito superior à da segunda (M=638), o peso da Eurosondagem no agregador final acaba por ser superior ao de todas as outras empresas.

As expectativas geradas até setembro de 2015 incidiam numa vitória do PS. A generalidade das sondagens realizadas até meados de setembro indicavam mais intenções de voto para o PS do que para a coligação PàF (PSD.CDS)<sup>9</sup> (ver Tabela 2).

Tabela 2. Sondagens realizadas desde a formação da coligação PàF (PSD.CDS) até ao início da campanha eleitoral para as eleições legislativas de 2015

Instituto	N	Dia <sup>10</sup>	PàF	PS	BE	CDU	PDR	L/TDA	PàF-PS
Aximage	603	10/5	37,8%	<b>37,9%</b>	4,2%	7,8%	2,6%	2,0%	-0,1%
Eurosondagem	1021	12/5	33,6%	<b>38,1%</b>	4,8%		2,5%	1,8%	-4,5%
Aximage	598	4/6	38,2%	<b>39,0%</b>	4,3%	7,7%	2,1%	1,2%	-0,8%
Eurosondagem	1030	9/6	33,3%	<b>36,9%</b>	4,5%	10,5%	2,7%	2,0%	-3,6%
Católica-CESOP	1048	16/6	<b>38%</b>	37%	8%	10%			1%
Intercampus	1014	4/7	32,7%	<b>37,6%</b>	6,0%	11,0%	0,7%	0,6%	-4,9%
Eurosondagem	1025	7/7	34,6%	<b>36,7%</b>	4,8%	10,2%	2,5%	1,9%	-2,1%
Aximage	607	16/7	38,8%	<b>39,0%</b>	4,1%	7,2%	1,4%	1,3%	-0,2%
Eurosondagem	1030	4/8	34,8%	<b>36,3%</b>	5,0%	10,0%	2,3%	1,7%	-1,5%
Eurosondagem	1040	2/9	35,0%	<b>36,0%</b>	4,6%	10,4%	2,3%	1,7%	-1,0%
Aximage	602	7/9	<b>40,1%</b>	34,3%	4,7%	8,8%	1,8%	1,9%	5,8%
Eurosondagem	1510	16/9	34,0%	<b>35,5%</b>	5,2%	10,3%	2,2%	1,8%	-1,5%
Aximage	704	17/9	<b>38,5%</b>	37,8%	6,3%	9,1%			0,7%
Católica-CESOP	647	17/9	<b>41%</b>	34%	8%	7%			7%

<sup>6</sup> <http://www.pedro-magalhaes.org/filtrar-o-ruído-das-sondagens/>

<sup>7</sup> <http://www.cs.unc.edu/~welch/kalman/media/pdf/Kalman1960.pdf>

<sup>8</sup> <http://www.erc.pt/pt/sondagens/publicitacao-de-sondagens/depositos-de-2015>

<sup>9</sup> Desde que a coligação PàF foi formalizada até 18 de setembro foram realizadas 14 sondagens (Eurosondagem: 6; Aximage: 5, Católica-CESOP: 2; Intercampus: 1). Dez sondagens atribuíam ao PS mais intenções de voto do que à PàF e 4 sondagens atribuíam mais intenções de voto à PàF (a maior parte das diferenças entre estas duas listas concorrentes não era grande e estava dentro das margens de erro). A análise por empresa da diferença entre PàF e PS permite constatar que: (1) a Eurosondagem apresentou sempre o PS com mais intenções de voto (diferença PS-PàF entre 4,5 p.p. e 1 p.p.); (2) a Aximage apresenta PS à frente nas 3 primeiras sondagens (entre 0,1 p.p. e 0,8 p.p.) e PàF à frente nas duas últimas (5,8 p.p., a 7 de setembro e 0,7 p.p. a 17 de setembro); (3) a Católica-CESOP indicava a PàF com mais intenções de voto em junho (1 p.p.) e em setembro (7 p.p.); (4) a Intercampus indicava em julho uma vantagem de 4,9 p.p. para o PS.

<sup>10</sup> A data indicada nos quadros corresponde ao último dia de trabalho de campo (inquirição) para cada sondagem. No caso das sondagens diárias (tracking poll) com amostra parcialmente renovada dia a dia, são apenas consideradas as sondagens com toda a amostra renovada. Os dados podem ser obtidos em: <http://www.popstar.pt/dados.php>

Foram por isso recebidos com enorme surpresa os resultados da primeira sondagem diária realizada pela Católica-CESOP para a RTP. Esta sondagem, com trabalho de campo (inquirição) realizado nos últimos dias antes do início oficial da campanha indicava a coligação PàF à frente do PS com 7 pontos percentuais de diferença. Foi a primeira de uma série de sondagens realizadas pela Católica-CESOP para a RTP. Os seus resultados, aliados ao baixo número de inquiridos e ao elevado número de indecisos (quando comparados com as sondagens mais amplamente divulgadas até ao momento – as realizadas pela Eurosondagem para a SIC/Expresso) criaram um ambiente propício para que se gerasse alguma polémica na praça pública. Perante um resultado diferente do esperado, o *falhanço* britânico passou a estar ativado na memória de muitos dos que se interessam por eleições e sondagens políticas: “as sondagens falham muitas vezes”<sup>11</sup>.

Mas vejamos o que mostram as sondagens realizadas a partir do início da campanha eleitoral. Aquele que, como vimos, era um resultado surpreendente até à primeira divulgação da sondagem diária da Católica-CESOP, passou a ser um resultado normal. Em todas as sondagens que se seguiram, com a exceção de uma realizada pela Eurosondagem, a coligação PàF surgiu com mais intenções de voto do que o PS (ver Tabela 3). A diferença entre as duas listas concorrentes variou entre 3 pontos percentuais (na primeira sondagem da Intercampus) e 12 p.p. (na única realizada pela Markttest). Como seria de esperar, sondagens mais próximas das eleições tendem a apresentar valores mais próximos dos resultados eleitorais.

Tabela 3. Sondagens realizadas desde o início da campanha eleitoral para as eleições legislativas de 2015

Instituto	N	Dia	PàF	PS	BE	CDU	PDR	L/TDA	PàF-PS
Intercampus	753	20/9	<b>40,1%</b>	37,1%	4,0%	6,4%			3,0%
Católica-CESOP	828	22/9	<b>40%</b>	34%	8%	10%			6%
Eurosondagem	1467	22/9	35,5%	<b>36,0%</b>	5,0%	10,1%	1,9%	1,5%	-0,5%
Aximage	700	23/9	<b>40,1%</b>	36,4%	7,0%	8,2%			3,7%
Intercampus	1024	24/9	<b>37,0%</b>	32,3%	6,0%	9,1%			4,7%
Católica-CESOP	1075	26/9	<b>41%</b>	34%	7%	9%			7%
Católica-CESOP	3302	26/9	<b>38%</b>	32%	9%	9%	1%	1%	6%
Markttest	1607	27/9	<b>41,0%</b>	29,0%	8,7%	9,3%	0,7%	0,8%	12,0%
Intercampus	1008	28/9	<b>38,8%</b>	31,6%	7,8%	8,0%			7,2%
Católica-CESOP	1070	29/9	<b>39%</b>	34%	8%	10%			5%
Eurosondagem	2067	29/9	<b>37,7%</b>	32,7%	6,7%	9,4%			5,0%
Intercampus	1013	30/9	<b>37,2%</b>	32,9%	7,9%	8,8%			4,3%
Aximage	1387	1/10	<b>39,1%</b>	32,5%	9,0%	9,2%			6,6%

### A história das intenções de voto, contada a partir das sondagens

Ao contrário do que muitas vezes se pensa, prever com precisão quem ganhará as eleições não é a principal função das sondagens eleitorais realizadas antes das eleições. As sondagens permitem contar a *história das intenções de voto*. Por exemplo, quando olhamos as sondagens ao longo de um

<sup>11</sup> Note-se que uma sondagem realizada por uma outra empresa uns dias antes tinha já retratado uma distribuição do eleitorado em proporções muito semelhantes às desta sondagem. A diferença entre PàF e PS indicada pela sondagem da Aximage de 7 de setembro era de 6 pontos percentuais.

período de uma legislatura, elas permitem compreender como é que determinadas mudanças (sejam políticas, económicas ou outras) influenciam as intenções de voto nos vários partidos: a mudança de líder num partido da oposição; a descida do desemprego; um caso político/criminal/jornalístico; uma mudança significativa na Comissão Europeia... Os dados disponibilizados pelo agregador de sondagens são a melhor forma de retratar essa história. No caso português, a consulta do *popstar*<sup>12</sup> permite verificar a subida do PS desde as eleições de 2011 até 2014, seguindo-se um período de queda que se altera com a mudança de líder no último trimestre de 2014. Segundo este agregador, também se assiste a uma queda acentuada de cerca de 4 pontos percentuais nas intenções de voto no PS durante o mês de setembro de 2015 (o mês anterior às eleições).

Sendo de pouca validade, devido à deslocação de populações no período de férias, as sondagens realizadas em agosto, a questão que se coloca é se teremos motivos para pensar que a opinião dos portugueses mudou de forma considerável entre julho e setembro (ou durante o mês de setembro). É o que é que aconteceu aos eleitores que, segundo a Aximage e (principalmente) a Eurosondagem, iam votar no Partido Democrático Republicano (PDR) e no Livre/Tempo de Avançar (L/TDA). Referimo-nos a 3 a 4,5% do eleitorado que, segundo estas duas empresas, de maio até meados de setembro iam votar nestes partidos e que acabaram por se evaporar a 15 dias das eleições (o PDR teve 1,14% dos votos e o Livre teve 0,73%). Alguma mudança de voto fruto da campanha eleitoral, poderá explicar parte desta diferença, mas fica por explicar porque motivos as restantes empresas nunca registaram mais do que 2% dos votos para o total destes partidos. Poderá ter sido fruto do acaso ou das diferentes metodologias usadas pelas empresas. Com os dados de que dispomos não nos é possível avaliar o que poderá ter acontecido, as eventuais causas por trás dessas diferenças.

Sabemos que é normal e expectável que várias sondagens, independentes, com boas bases amostrais, e cumpridoras das melhores práticas disponíveis em função das limitações temporais e orçamentais a que todos estes trabalhos estão sujeitos, mostrem resultados diferentes entre si. Em situações em que, na população, dois partidos tenham proporções de intenções de voto muito semelhantes, é de esperar que umas sondagens apontem para a vantagem de um partido e outras para a vantagem do outro. É o que acontece com os partidos Amarelo e Verde na tabela seguinte. Na Tabela 4 simula-se a distribuição real de votantes na população (coluna População) e as distribuições de votantes encontradas por 5 sondagens independentes (colunas S1, S2, S3, S4 e S5). Na coluna “Média” mostram-se os valores médios das sondagens e na coluna seguinte, a diferença entre a proporção atribuída a cada partido pela média das sondagens e a proporção real de cada partido.

Tabela 4. Simulação da distribuição de eleitores na população e em 5 sondagens

Partidos	População (P)	Média					S1-5 (M)	Diferença (M-P)
		S 1	S 2	S 3	S 4	S 5		
Amarelo	<b>41%</b>	<b>43%</b>	38%	<b>45%</b>	39%	<b>42%</b>	<b>41,3%</b>	0,3%
Verde	38%	35%	<b>39%</b>	34%	<b>42%</b>	39%	37,8%	-0,2%
Azul	10%	9%	8%	12%	10%	9%	9,7%	-0,3%
Roxo	8%	10%	12%	7%	6%	7%	8,3%	0,3%
Cinzento	1%	2%	1%	1%	2%	3%	1,7%	0,7%
B/N	2%	1%	2%	1%	1%	3%	1,7%	-0,3%

<sup>12</sup> <http://www.popstar.pt/sondagens.php>

A mais-valia de um agregador de sondagens, seja o *popstar*, seja o *Poll of polls* da BBC, seja outro qualquer, é a possibilidade de, a partir da média das proporções referentes a cada partido indicadas por cada sondagem independente, estimar uma proporção para cada partido muito próxima da realmente existente na população. Deste modo, sendo os resultados de um agregador dependentes das sondagens que o alimentam, pode-se assumir a discrepância entre os resultados do agregador e os resultados eleitorais como um critério de avaliação da qualidade dessas sondagens.

Que proporções estimava o *popstar* antes das eleições legislativas? A duas semanas das eleições, 37% para a coligação PàF e 36,4% para o PS (ver Tabela 5). A uma semana, 37,9% e 35,2%, respetivamente. Na véspera, 38,3% para a coligação e 34,1% para o PS. Atendendo aos resultados das eleições (38,6% para a coligação PàF<sup>13</sup> e 32,3% para o PS), verifica-se que o *popstar* estimou muito bem o resultado da coligação mas sobrestimou ligeiramente o resultado do PS.

Na estimação dos partidos com menores votações este agregador teve um bom desempenho na estimação da CDU mas não estimou de forma tão correta os restantes partidos. Nos casos do PDR e do L/TDA, o agregador estimava na véspera o dobro da votação que estes partidos tiveram no dia da eleição. Este desvio deve-se a dois fatores, facilmente observáveis nas tabelas 2 e 3: (1) desequilíbrio do número de sondagens realizadas por cada instituto; (2) fraca variabilidade entre proporções atribuídas a cada partido nas várias sondagens realizadas por uma mesma empresa. Estes dois fatores somados contribuíram para enfraquecer as estimações do agregador.

Tabela 5. Resultados do agregador *popstar* em 3 momentos e resultados eleitorais

	PàF	PS	BE	CDU	PDR	L/TDA
19/09/2015	37%	36,4%	5,5%	9,3%	2,5%	1,9%
26/09/2015	37,9%	35,2%	6,1%	8,9%	2,4%	1,8%
3/10/2015	38,3%	34,1%	6,8%	9%	2,3%	1,7%
Resultados eleições	38,6%	32,3%	10,2%	8,3%	1,1%	0,7%

### O que poderá ser feito para diminuir incongruências entre realidade, sondagens e interpretação pública das sondagens?

É sempre importante lembrar que as sondagens não procuram prever resultados eleitorais mas sim indicar a distribuição das intenções de voto em determinado momento. Sabemos como ao longo do tempo os eleitores vão prestando atenção aos diversos programas, aos debates, ao comentário político, vão trocando opiniões com vizinhos, colegas, amigos, e vão formando a sua opção de voto, que para alguns só mesmo no dia das eleições fica fechada. Acontecimentos extraordinários entre o momento de inquirição das últimas sondagens pré-eleitorais e o dia das eleições podem mesmo alterar radicalmente o cenário político<sup>14</sup>. No entanto, salvaguardando a possibilidade de eventos extraordinários, a capacidade preditiva das últimas sondagens pré-eleitorais é genericamente boa. Em Portugal, as incongruências entre as *histórias* contadas pelas sondagens e os resultados eleitorais raramente são grandes ou mesmo relevantes, principalmente em eleições legislativas e presidenciais. Em eleições

<sup>13</sup> Esta percentagem inclui a votação de PSD e CDS nas Regiões Autónomas, onde estes partidos não concorreram coligados.

<sup>14</sup> Por exemplo, nas eleições legislativas de 2004 em Espanha, previa-se a vitória do PP. No entanto, o PSOE ganhou essas eleições. Entre as últimas sondagens pré-eleitorais e o ato eleitoral ocorreu o maior atentado terrorista em solo espanhol: o atentado de Atocha, a 3 dias das eleições. O atentado e as reações políticas a este terão sido decisivos na mudança de sentido de voto de muitos eleitores espanhóis.



autárquicas, regionais ou europeias, as alterações de sentido de voto nos últimos dias de campanha parecem ser mais frequentes e já aconteceram casos de diferenças significativas entre os resultados das últimas sondagens pré-eleitorais e os resultados das eleições.

Ainda assim, as empresas de sondagens encontram dificuldades crescentes para a boa condução das sondagens políticas publicadas na comunicação social. A primeira das quais, as dificuldades orçamentais dos seus clientes (jornais, rádios e televisões). As restrições orçamentais implicam redução do número de sondagens e redução do número de inquiridos, com correspondente aumento das margens de erro. Têm também levado a uma progressiva substituição das sondagens realizadas porta-a-porta por sondagens telefónicas, a maior parte das quais realizadas apenas com recurso a telefones fixos. Estes constrangimentos são ameaças claras à qualidade das sondagens pois podem influenciar o fator identificado pela comissão de inquérito nomeada pelo Conselho de Sondagens Britânico referido no início deste artigo: a falta de representatividade das amostras utilizadas. A segunda dificuldade está na leitura pública das sondagens. Como são interpretadas e analisadas no espaço mediático. Existe uma clara falta de literacia neste domínio de que são prova evidente as leituras de subida e descida de forças partidárias quando se comparam sondagens, quando na maioria dos casos os intervalos de confiança associados às proporções não permitem essas leituras, ou a dificuldade de compreensão de conceitos como *taxa de resposta*. A terceira dificuldade poderá estar a decorrer de algumas opções metodológicas que, por qualquer motivo que desconhecemos, estão a implicar baixa variabilidade de resultados entre sondagens. Maior variabilidade de empresas a fazer sondagens, por um lado, e maior variabilidade de resultados de sondagens (mesmo os realizados por uma só empresa), por outro, são fatores decisivos para uma maior precisão destes instrumentos de medida da opinião pública.



# Avaliação temporal do conhecimento matemático dos alunos da FCUP à entrada do ensino superior

Milton Severo<sup>1</sup>, *milton@med.up.pt*  
Paulo Trindade<sup>2</sup>, *trindade.paulo@gmail.com*  
A. Rita Gaio<sup>2,3</sup>, *argaio@fc.up.pt*

<sup>1</sup> *Faculdade de Medicina do Porto e  
EPI-Unit, Instituto de Saúde Pública da Universidade do Porto*  
<sup>2</sup> *Faculdade de Ciências da Universidade do Porto*  
<sup>3</sup> *Centro de Matemática da Universidade do Porto*

## 1. Introdução

Nos últimos anos, a Faculdade de Ciências da Universidade do Porto (FCUP) tem vindo a realizar testes diagnósticos de Matemática aos alunos do primeiro ano de várias das suas licenciaturas. Nesta publicação, vai-se analisar a qualidade psicométrica dos testes aplicados e avaliar a tendência temporal dos resultados obtidos. Serão usados conceitos da teoria clássica de testes, modelos da teoria de resposta ao item e técnicas de calibração.

Todos os anos surgem notícias sobre as classificações obtidas no exame nacional de Matemática do 12º ano, ora afirmando que os conhecimentos dos alunos pioraram, ora afirmando que melhoraram (Viana, 2013). Implicitamente, estas comparações assumem que os exames dos vários anos têm o mesmo nível de dificuldade, o que pode não ser verdadeiro. Quando um professor avalia o mesmo conhecimento em duas amostras de estudantes independentes através de dois testes de avaliação diferentes, pode acontecer que um dos grupos obtenha classificações melhores do que o outro por vários motivos. A explicação desejável é que um dos grupos consista, de facto, de alunos melhor preparados do que o outro; outra possibilidade é que um dos testes tenha sido mais fácil. Para percebermos qual das explicações é a correta podemos, sob algumas condições, utilizar técnicas de calibração de testes (Livingston, 2004).

Os testes diagnósticos de Matemática realizados na FCUP aos alunos do primeiro ano de diferentes licenciaturas têm estado em funcionamento desde 2009. São dois os principais objetivos deste procedimento: disponibilizar uma ferramenta que permita aos alunos identificarem os tópicos matemáticos onde falham e que portanto precisam de ser mais trabalhados; ajudar os professores a tomarem conhecimento das dificuldades dos alunos para que possam eventualmente adaptar as suas aulas e metodologias de ensino. O teste não tem um carácter obrigatório e consiste exclusivamente de perguntas de escolha múltipla.

O uso de testes constituídos por perguntas de escolha múltipla (PEMs) tem sido defendido por alguns autores (Schuwirth & Van Der Vleuten, 2004). Baseiam os seus argumentos nos seguintes factos: baixa ou nula variabilidade na obtenção da classificação final na situação de correção efetuada por vários professores; rapidez no cálculo e divulgação da classificação final; possibilidade de inclusão de um maior número de perguntas. Mas nem todos os testes com PEMs são bons testes; há que avaliar a qualidade intrínseca dos testes. Em primeiro lugar, as perguntas devem estar bem escritas e editadas (Norman, Van Der Vleuten, & Newble, 2002). Depois têm de ser *boas* perguntas (Haladyna, 2012).

No presente artigo, o trabalho é apresentado pela seguinte ordem: resumo das principais teorias sobre avaliação da qualidade psicométrica de testes de escolha múltipla, exposição da teoria usada sobre calibração de testes, aplicação aos testes diagnósticos realizados na FCUP e apresentação dos resultados, e finalmente enumeração das principais conclusões.

## 2. Avaliação da qualidade psicométrica de testes de escolha múltipla

Existem duas metodologias principais para a avaliação da qualidade psicométrica de testes de escolha múltipla: Teoria Clássica de Testes (TCT) e Teoria de Resposta ao Item (TRI). Enquanto a primeira tem como principais vantagens a simplicidade de cálculo das estatísticas e a facilidade de interpretação dos resultados, a segunda tem como ponto forte o facto de ser baseada em modelos probabilísticos. Qualquer uma das teorias inclui duas componentes: a avaliação da qualidade individual dos itens (as perguntas) e a avaliação da qualidade global do teste.

No que se segue, e até ao final deste artigo, designaremos apenas por “teste” um teste que consista unicamente de PEMs, e por “itens” as suas perguntas.

### 2.1. Teoria clássica de testes

No âmbito da TCT, existem dois indicadores principais para avaliar a qualidade dos itens: índice de dificuldade e índice de discriminação. Para avaliar a qualidade global do teste, é comum usarem-se o coeficiente alfa de Cronbach e o erro padrão de medição.

#### Índice de dificuldade

O índice de dificuldade de um item  $i$ , designado por  $P_i$ , corresponde à percentagem de estudantes que respondem corretamente ao item. Varia portanto entre 0 e 100, sendo que 0 significa dificuldade máxima e 100 significa dificuldade mínima (facilidade máxima). Para que se tenha uma boa pergunta, a recomendação é de que o índice de dificuldade varie entre 30% e 80% (Tavakol & Dennick, 2011).

#### Índice de discriminação de um item

O índice de discriminação de um item  $i$ , designado por  $D_i$ , identifica a capacidade desse item de distinguir os estudantes que obtiveram uma classificação final elevada dos que obtiveram uma classificação final baixa. Varia entre -1 e 1, sendo que 1 significa discriminação positiva máxima, 0 significa nenhuma discriminação e -1 significa discriminação negativa máxima. Os itens com índices de discriminação negativos e *próximos* de -1 devem ser revistos ou mesmo eliminados do teste, dado que isso significa que os estudantes com melhores classificações erraram mais esses itens do que os estudantes com piores classificações.

O método mais simples para calcular o índice de discriminação de um item consiste em considerar a diferença entre a proporção de respostas corretas ao item de entre as 27% melhores classificações finais e a proporção de respostas corretas ao item de entre as 27% piores classificações finais:

$$D_i = \frac{U_i}{NU} - \frac{L_i}{NL}$$

onde  $U_i$  (resp.  $L_i$ ) representa o número de respostas corretas ao item  $i$  no grupo de alunos que obtiveram 27% das classificações finais mais elevadas (resp. mais baixas) e  $NU$  (resp.  $NL$ ) representa o número de estudantes que obtiveram as 27% classificações finais mais elevadas (resp. mais baixas). Há autores que consideram 33% em vez dos anteriores 27% (Ebel, 1954).

Um outro parâmetro de avaliação da capacidade de discriminação de um item é o coeficiente de correlação bisserial pontual. Matematicamente, é equivalente ao coeficiente de correlação de Pearson (momento conjunto) entre a classificação obtida no item (variável dicotómica nominal) e a classificação final excluindo o item. Neste trabalho, seguir-se-á uma outra tendência (Lewis-Beck, Bryman, & Liao, 2003), que usa a correlação bisserial em lugar da correlação bisserial pontual. O coeficiente de correlação bisserial considera a classificação dicotómica do item como uma estratificação de uma variável artificial contínua seguindo uma distribuição normal. Para o item  $i$  é calculado de acordo com a seguinte fórmula:

$$D_{bis,i} = \frac{\bar{x}_1 - \bar{x}_0}{s_T} \left( \frac{n_1 n_0}{u(n_1 + n_0)^2} \right)$$

onde  $\bar{x}_1$  (resp.  $\bar{x}_0$ ) é a média das classificações finais dos estudantes que responderam corretamente (resp. incorretamente) ao item,  $n_1$  (resp.  $n_0$ ) é o número de estudantes que responderam corretamente (incorretamente) ao item e  $u$  é o valor da função densidade de probabilidade da

distribuição  $N(0,1)$  calculado em  $n_1/(n_1 + n_0)$ . Os valores da correlação bisserial são sempre mais elevados em termos absolutos do que os da correlação bisserial pontual.

Um índice de discriminação (aqui, coeficiente de correlação bisserial) superior a 0.4 tem sido associado a uma discriminação muito boa do item, entre 0.3 e 0.4 a uma discriminação boa, entre 0.2 e 0.3 a uma discriminação suficiente, e entre -1 e 0.2 a uma discriminação fraca (Ebel, 1954).

### Fiabilidade (Alfa de Cronbach)

A ideia central da teoria clássica de testes considera que a classificação obtida por um estudante num determinado teste corresponde à soma da classificação verdadeira do estudante com um erro aleatório. Aqui, denomina-se por classificação verdadeira do estudante o valor esperado para a classificação caso esse estudante respondesse a um número infinito de testes (ou de itens) equivalentes ou caso o teste não contivesse erros de medição.

O quadrado do coeficiente de correlação entre a classificação obtida e a classificação verdadeira é designado por fiabilidade do teste, e corresponde à percentagem da variação das classificações obtidas pelos estudantes que é explicada pela variação das classificações verdadeiras dos estudantes.

O coeficiente alfa de Cronbach ( $\alpha$ ) corresponde à quantidade

$$\alpha = \frac{(s_t^2 - \sum_{i=1}^I s_i^2)/(I - 1)}{s_t^2/I}$$

onde  $s_t$  é o desvio padrão (amostral) das classificações totais obtidas,  $s_i$  é o desvio padrão (amostral) das classificações obtidas no item  $i$  e  $I$  é o número total de itens. Na situação em que a covariância entre a classificação obtida num item e a correspondente classificação verdadeira é a mesma em todos os itens (tau-equivalência), o coeficiente  $\alpha$  constitui uma estimativa da fiabilidade do teste.

É importante realçar aqui que a fiabilidade de um teste aumenta com o número de itens do teste. Em particular, um teste com muitos itens possuindo baixa discriminação pode atingir níveis de fiabilidade altos (Cortina, 1993). Kehoe (Kehoe, 1995) sugeriu que, para testes com 10 a 15 itens, a fiabilidade deva ser superior a 0.5, e para testes com mais de 50 itens a fiabilidade deva ser superior a 0.8.

## 2.2 Teoria de resposta ao item

Os modelos da TRI estimam a probabilidade  $\pi_{ij}$  de um estudante  $j$  responder corretamente ao item  $i$ . Assume-se que esta probabilidade depende do conhecimento do estudante e das características do item (dificuldade e discriminação). A relação entre as características do estudante e do item e a probabilidade de responder corretamente ao item é modelada pela seguinte equação (modelo 2-PL):

$$\ln\left(\frac{\pi_i(z_j)}{1 - \pi_i(z_j)}\right) = a_i(z_j - b_i) \Leftrightarrow \pi_i(z_j) = \frac{1}{1 + e^{-a_i(z_j - b_i)}}$$

Aqui,  $z_j$  corresponde ao conhecimento (não observado) do estudante,  $a_i$  representa o declive da curva no ponto correspondente a uma probabilidade de acertar o item igual a 0.5, e  $b_i$  é o valor do conhecimento para o qual se tem uma probabilidade de 0.5 de responder corretamente ao item  $i$ . Os coeficientes  $a_i$  e  $b_i$  são designados por parâmetro de discriminação e parâmetro de dificuldade, respetivamente, enquanto que o gráfico da função  $\pi_i = \pi_i(z)$  é usualmente designado por curva característica do item  $i$ . Por regra, a variável  $Z$  encontra-se padronizada (média 0 e desvio padrão 1) e está diretamente associada ao conhecimento do aluno (quanto maior o valor de  $Z$ , maior o conhecimento).

Para interpretar  $a_i$  como o índice de discriminação do item  $i$ , padroniza-se o coeficiente para que este corresponda à correlação entre o item  $i$  e o conhecimento subjacente:

$$a_i^* = \frac{a_i}{(a_i^2 + 1)^{1/2}}$$

Assim, quanto mais afastado de 0 estiver o coeficiente  $a_i$ , maior é a capacidade de discriminação do item  $i$ .

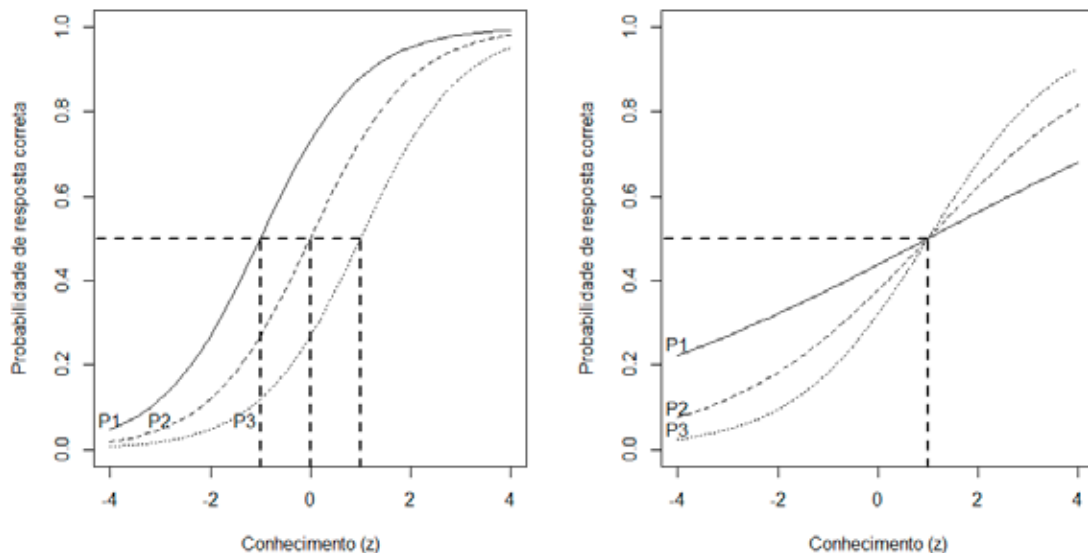


Figura 1. Curvas características de 3 itens (P1, P2 e P3) de um teste. No gráfico da esquerda, o valor do conhecimento correspondente a uma probabilidade de 0.5 de responder corretamente ao item 1 é de -1 enquanto que para o item 3 é de 1; o item 3 é portanto mais difícil do que o item 1. No gráfico da direita, o item 3 é o que apresenta maior declive no ponto correspondente a uma probabilidade de acertar o item de 0.5, portanto é o item mais discriminativo. O item 3 é também aquele que apresenta uma maior diferença entre os valores da probabilidade no grupo dos alunos com melhores conhecimentos e os valores da probabilidade no grupo dos alunos com piores conhecimentos. Já o item 1 é o menos discriminativo.

## Métodos de estimação

Os parâmetros de um modelo de TRI podem ser estimados por máxima verossimilhança conjunta. No entanto, esta estratégia apresenta uma tendência para gerar valores implausíveis na situação em que um estudante responde corretamente a todos os itens ou uma pergunta é respondida corretamente por todos os estudantes. Uma forma de contornar este problema consiste em fazer a estimação usando máxima verossimilhança marginal (utilizando o algoritmo EM, *Expectation-Maximization*). Uma descrição detalhada das metodologias de estimação pode ser encontrada em Baker e Kim (Baker & Kim, 2004).

### 2.3. Calibração de testes

Um processo de calibração de testes tem como objetivo primeiro a colocação de classificações obtidas em testes diferentes numa mesma escala, para que possam ser comparáveis. Os seus principais pressupostos são a equivalência entre testes e a simetria. Enquanto que a primeira propriedade refere que os testes medem o mesmo conhecimento, a segunda afirma que a calibração do teste X para o teste Y produz os mesmos resultados que a calibração do teste Y para o teste X. Se a estes pressupostos acrescentarmos a condição de que os testes apresentam uma fiabilidade semelhante, estamos na presença de um processo de equiparação. Estas noções são específicas da Psicometria.

Existem essencialmente dois tipos de estudos em calibração: usando grupos equivalentes ou usando grupos não equivalentes. Diz-se que os grupos são equivalentes quando as diferenças entre as classificações dos testes se devem exclusivamente a níveis de dificuldade e discriminação diferentes entre os testes; isto pode acontecer, por exemplo, porque ambos os testes foram aplicados aos mesmos estudantes ou porque os estudantes foram aleatoriamente distribuídos pelos dois testes. Nos grupos não equivalentes, assume-se que podem existir outras diferenças entre os conhecimentos dos estudantes, para além das anteriores. A calibração com grupos não equivalentes obriga à presença de perguntas comuns aos vários testes (designadas por perguntas âncora) (Livingston, 2004).

Dado que os testes diagnósticos da FCUP correspondentes a anos diferentes não foram sempre os mesmos, introduziram-se sistematicamente algumas perguntas âncora. Este procedimento permitiu a comparação entre resultados obtidos em anos diferentes e, em particular, a avaliação da tendência temporal das classificações.

Existem várias metodologias de calibração; como exemplos, temos a calibração da média, a calibração linear ou a calibração por percentil (não-linear). A primeira uniformiza apenas a média, a segunda uniformiza a média e o desvio padrão, e a terceira uniformiza a média, o desvio padrão, a simetria e a

curtose. Por regra, a escolha da metodologia depende dos tamanhos amostrais. As metodologias enumeradas anteriormente são recomendadas para tamanhos amostrais pequeno, intermédio e grande, respetivamente (Livingston, 2004).

Na situação concreta dos testes diagnósticos da FCUP, o tamanho amostral variou entre 164 e 363 estudantes, respetivamente em 2013 e 2009, pelo que a metodologia utilizada foi a calibração linear, de acordo com a TCT e de acordo com a TRI.

### Calibração linear de acordo com a teoria clássica de testes

Dados dois testes, digamos teste 1 e teste 2, diz-se que as classificações observadas  $x$  e  $y$ , respetivamente, estão calibradas de forma linear se:

$$\frac{y - \bar{y}}{s_y} = \frac{x - \bar{x}}{s_x} \Leftrightarrow y = \frac{s_y}{s_x} x - \frac{s_y}{s_x} \bar{x} + \bar{y}$$

Se estivermos na presença de grupos equivalentes, a média e o desvio padrão a utilizar devem ser os calculados nos grupos de indivíduos que se submeteram a cada um dos testes. No caso de grupos não equivalentes, é necessário que a informação dos itens âncora seja tida em consideração na estimação da média e do desvio padrão de cada teste (Kolen & Brennan, 2004).

### Calibração linear de acordo com a teoria de resposta ao item

Designemos por  $\mathbf{z}_o$  e  $\mathbf{z}_f$  o conhecimento estimado pelos modelos da TRI para os alunos dos testes 1 e 2, respetivamente. A equação de calibração linear que transforma  $\mathbf{z}_o$  (escala original – teste 1) em  $\mathbf{z}_f$  (escala final – teste 2) é:

$$\mathbf{z}_f = A\mathbf{z}_o + B.$$

Aqui, as constantes  $A$  e  $B$  são usadas para ajustar a média e o desvio padrão da escala original, respetivamente, para a escala final. Existem várias formas de estimação destas constantes. O método utilizado nos testes da FCUP foi o método média/média (Weeks, 2010) :

$$A = \frac{\mu(a_{oi})}{\mu(a_{fi})}, \quad B = \mu(b_{fi}) - A\mu(b_{oi})$$

que usa a média dos parâmetros de discriminação ( $\mu(a_{oi}), \mu(a_{fi})$ ) e de dificuldade ( $\mu(b_{oi}), \mu(b_{fi})$ ) dos itens âncora, do teste 1 e do teste 2 respetivamente.

## 3. Resultados

A tabela 1 mostra que o número de itens por teste variou entre 24 e 30, sendo que o número de perguntas âncora variou entre 4 (anos 2009 e 2010 com anos 2011, 2012 e 2103) e 28 (ano 2009 com ano 2010).

Tabela 1. Total de perguntas por teste diagnóstico (diagonal) e número de perguntas âncora.

	2009	2010	2011	2012	2013
2009	29 <sup>a</sup>	28	4	4	4
2010	---	30	4	4	4
2011	---	---	24 <sup>b</sup>	24	24
2012	---	---	---	24 <sup>b</sup>	24
2013	---	---	---	---	24 <sup>b</sup>

<sup>a</sup> Um item foi eliminado pelo facto de ter 2 opções iguais.

<sup>b</sup> As 30 perguntas foram organizados em 24 grupos de questões.

A mediana do índice de dificuldade, de acordo com a TCT, variou entre 46% (2010) e 61% (2012), encontrando-se portanto dentro dos níveis recomendados. Os itens usados apresentaram uma boa discriminação uma vez que o índice de discriminação variou entre 0.42 e 0.50. Quanto à fiabilidade, todos os testes exibiram valores superiores a 0.7, o valor recomendado (tabela 2).

Tabela 2. Descrição da qualidade dos itens de acordo com a teoria clássica de testes.

Ano	N	Índice de dificuldade			Índice de discriminação			Qualidade do teste Alfa de Cronbach
		Mediana	P25	P75	Mediana	P25	P75	
2009	363	0.47	0.39	0.57	0.43	0.35	0.47	<b>0.791</b>
2010	218	<b>0.46</b>	0.37	0.55	<b>0.42</b>	0.37	0.45	0.802
2011	328	0.50	0.46	0.60	0.48	0.44	0.51	0.825
2012	206	<b>0.61</b>	0.49	0.70	0.45	0.36	0.51	0.797
2013	164	0.57	0.48	0.70	<b>0.50</b>	0.46	0.55	<b>0.835</b>

A figura 2 mostra as CCIs obtidas dos modelos da TRI, para os testes efetuados nos vários anos. Todas as perguntas apresentaram discriminação positiva e bastante semelhante entre si, com exceção da pergunta 26 do ano 2009 que mostrou uma discriminação negativa. Esta questão dizia respeito ao número de interseções do gráfico de uma função trigonométrica com uma função do tipo exponencial e foi eliminada do teste, deixando também de constar nos testes dos anos seguintes.

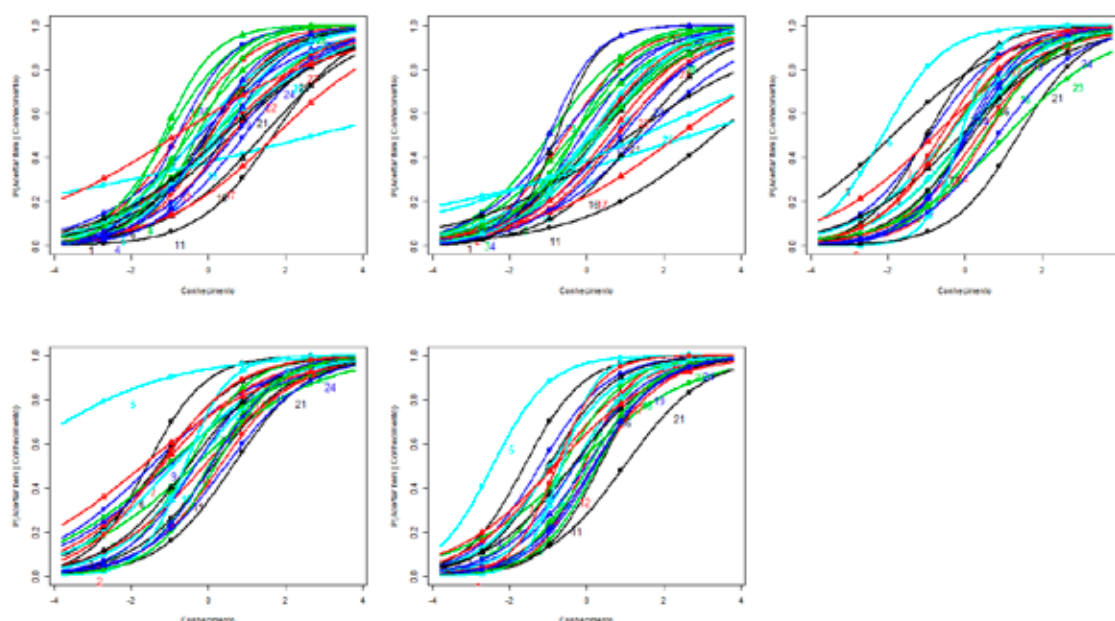


Figura 2. Curvas características dos itens dos testes, de 2009 a 2013 (da esquerda para a direita, e de cima para baixo).

O parâmetro de dificuldade variou entre -0.54 e 0.22, correspondendo a probabilidades de resposta correta para um estudante com conhecimento médio (zero) de 0.64 e 0.45, respectivamente (tabela 3). O parâmetro de discriminação variou entre 0.87 e 1.19, correspondendo a uma carga fatorial (correlação entre o conhecimento estimado e a variável latente que o item representa) de 0.66 e 0.76 (tabela 3). A semelhança entre os resultados obtidos pela TRI e aqueles obtidos pela TCT corroborou as boas propriedades psicométricas dos testes diagnósticos usados.

Tabela 3. Análise da qualidade dos itens de acordo com o modelo 2-PL da TRI.

	Parâmetro de Dificuldade Mediana (P25-P75)	$P(X=1 Z=0)$ <sup>1</sup> Mediana (P25-P75)	Parâmetro de Discriminação Mediana (P25-P75)	Carga Fatorial Mediana (P25-P75)	$p^2$
2009	0.08 (-0.46; 0.53)	0.47 (0.38; 0.59)	0.90 (0.70; 1.12)	0.67 (0.57; 0.74)	>0.05
2010	<b>0.22 (-0.27; 0.88)</b>	<b>0.45 (0.36; 0.57)</b>	<b>0.87 (0.76; 0.96)</b>	<b>0.66 (0.61; 0.69)</b>	>0.05
2011	-0.02 (-0.50; 0.21)	0.51 (0.44; 0.62)	1.03 (0.89; 1.16)	0.71 (0.66; 0.76)	>0.05
2012	<b>-0.54 (-1.22; 0.02)</b>	<b>0.64 (0.49; 0.73)</b>	0.94 (0.74; 1.10)	0.68 (0.60; 0.74)	>0.05
2013	-0.33 (-0.83; 0.08)	0.58 (0.48; 0.74)	<b>1.19 (1.04; 1.32)</b>	<b>0.76 (0.72; 0.80)</b>	>0.05

<sup>1</sup>: probabilidade de resposta correta para alunos com conhecimento matemático médio

<sup>2</sup>: valor-p do teste para a qualidade do ajustamento do modelo 2-PL, usando bootstrap.

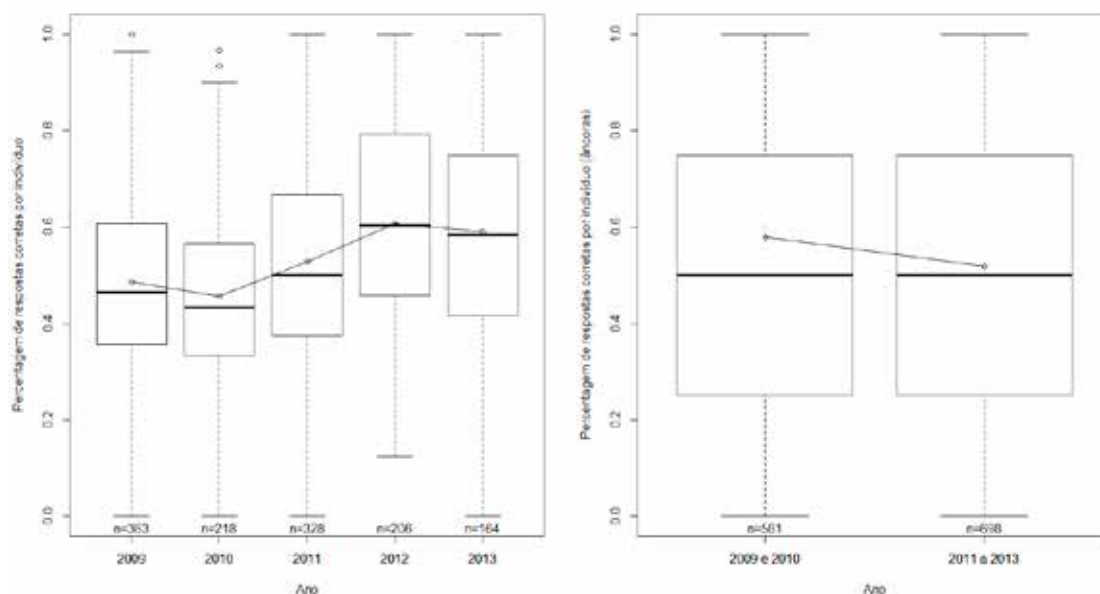


Figura 3. Diagrama em caixa de bigodes ilustrando a evolução temporal da distribuição da percentagem de respostas corretas, por indivíduo: (a) por ano e considerando todos os itens do teste; (b) por período temporal e considerando apenas os itens âncora. Os segmentos de reta unem as médias obtidas em cada ano.

A percentagem de respostas corretas, por indivíduo, apresentou uma tendência linear crescente estatisticamente significativa ao longo do tempo (figura 3(a);  $B1=0.03$ ,  $p<0.001$ ; estimação por regressão linear). No entanto, quando a mesma média se restringiu às perguntas âncora, a tendência foi a oposta (figura 3(b)).

Tabela 4. Constantes de calibração de acordo com a TCT e com a TRI.

Conversão	TCT		Conversão	TRI	
	A	B		A	B
2010->2009	0.924	0.240	2010->2009	0.917	-0.261
2011->2009	1.109	-1.555	2011->2010	1.081	-0.392
2012->2009	1.180	-3.489	2012->2011	0.887	0.500
2013->2009	1.094	-2.340	2013->2012	1.273	-0.203

Efetuuou-se então um processo de calibração. Segundo a TCT, e por comparação com o ano de 2009, as classificações dos anos 2011, 2012 e 2013 deviam ser menores entre 1.5 e 3.5 itens certos (constante B). Um efeito semelhante foi confirmado por aplicação da TRI: para qualquer dos anos analisados, o nível de conhecimento deveria diminuir em relação ao avaliado no ano anterior, com exceção de 2012. Para além disso, verificou-se a inexistência de uma tendência linear significativa ( $B1=-0.01$ ,  $p=0.138$ ) por aplicação da TCT, e a existência de uma tendência linear significativamente decrescente ( $B1=-0.07$ ,  $p<0.001$ ) por aplicação da TRI, como mostra a figura 4.



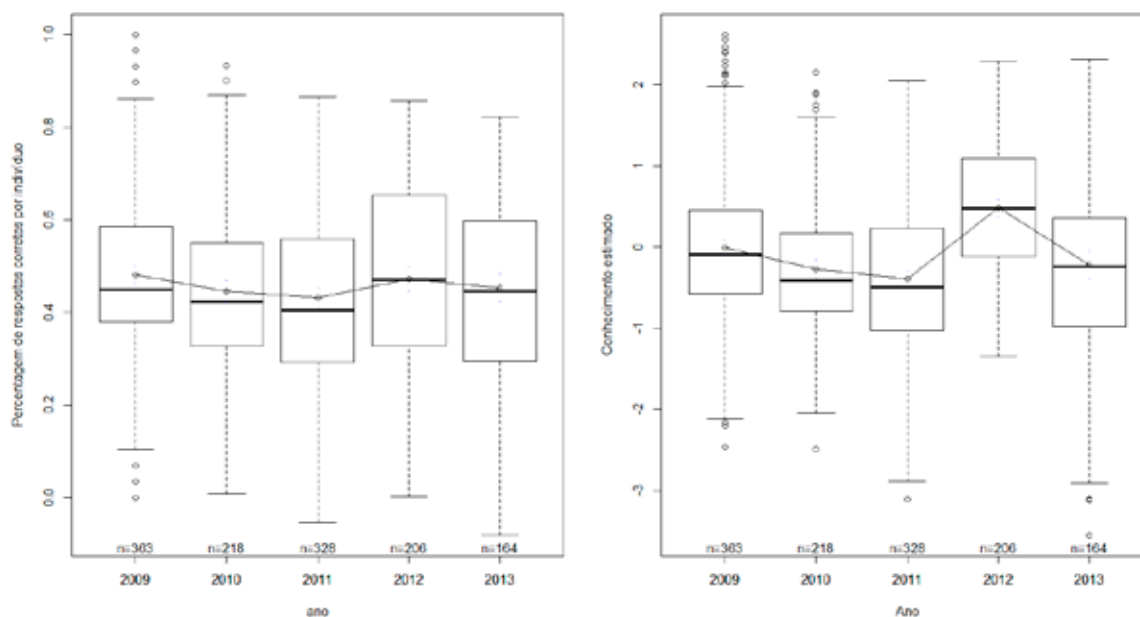


Figura 4. Diagrama em caixa de bigodes ilustrando a evolução temporal da distribuição da probabilidade média de resposta correta, por indivíduo: após calibração de acordo com a TCT (a); após calibração de acordo com a TRI (b). Os segmentos de reta unem as médias obtidas em cada ano.

Observa-se que, nos anos 2009 e 2010, e nos anos 2011, 2012 e 2013 poder-se-ia ter prescindido do processo de calibração, dado que nesses anos os testes foram (essencialmente) iguais. A calibração foi fundamental para conseguir estabelecer uma ponte entre estes dois períodos temporais.

#### 4. Conclusões

Os testes utilizados nos vários anos revelaram uma boa qualidade psicométrica. Inferiram-se diferenças estatisticamente significativas entre os conhecimentos matemáticos médios dos anos analisados, sendo que as classificações melhoraram linearmente de ano para ano. Contudo, após calibração, essa tendência crescente desapareceu: segundo a TCT, não existiu evidência estatística de qualquer tendência linear entre as várias classificações; a TRI identificou uma tendência linear decrescente entre os conhecimentos médios. Os anos com melhor conhecimento matemático médio foram 2009 e 2012.

A principal limitação do estudo aqui apresentado consistiu do número de itens âncora (4 itens apenas) entre os períodos temporais 2009-2010 e 2011-2013. Este facto pode justificar as diferentes conclusões obtidas da calibração, por TCT e TRI. De qualquer forma, a conclusão anterior ao processo de calibração afirmando que os estudantes estavam a melhorar de ano para ano será sempre falsa. A calibração veio mostrar que parte desse efeito se deveu ao facto de a dificuldade dos testes no segundo período temporal ser inferior.

#### Agradecimentos

A autora A. Rita Gaio foi parcialmente financiada pelo CMUP (UID/MAT/00144/2013), através da FCT e por fundos estruturais nacionais (MEC) e europeus (FEDER), sob o acordo de parceria PT2020.

#### REFERÊNCIAS BIBLIOGRÁFICAS

Baker, F. B., & Kim, S.-H. (2004). *Item response theory: Parameter estimation techniques*: CRC Press.

- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, 78(1), 98.
- Ebel, R. L. (1954). Procedures for the analysis of classroom tests. *Educational and Psychological Measurement*, 14(2), 352-364.
- Haladyna, T. M. (2012). *Developing and validating multiple-choice test items*: Routledge.
- Kehoe, J. (1995). *Basic item analysis for multiple-choice tests*: ERIC Clearinghouse on Assessment and Evaluation.
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking*: Springer.
- Lewis-Beck, M., Bryman, A. E., & Liao, T. F. (2003). *The Sage encyclopedia of social science research methods*: Sage Publications.
- Livingston, S. A. (2004). *Equating test scores*: Educational Testing Service Princeton, NJ.
- Norman, G. R., Van Der Vleuten, C. P., & Newble, D. I. (2002). *International handbook of research in medical education* (Vol. 7): Springer Science & Business Media.
- Schuwirth, L. W., & Van Der Vleuten, C. P. (2004). Different written assessment methods: what can be said about their strengths and weaknesses? *Medical Education*, 38(9), 974-979.
- Tavakol, M., & Dennick, R. (2011). Post-examination analysis of objective tests. *Medical teacher*, 33(6), 447-458.
- Viana, C. (2013). Matemática com pior resultado dos últimos sete anos. *Público*.
- Weeks, J. P. (2010). plink: An R package for linking mixed-format tests using IRT-based methods. *Journal of Statistical Software*, 35(12), 1-33.



## Modelação de séries temporais de contagem

Maria Eduarda Silva, *mesilva@fep.up.pt*

*Faculdade de Economia, Universidade do Porto & CIDMA*

### 1 Introdução

Séries temporais de valor inteiro e, em particular, séries temporais de contagem observam-se em diversas áreas e contextos, geralmente como contagens de acontecimentos ou objectos em intervalos de tempo específicos. Estas séries caracterizam-se frequentemente por contagens de valores baixos, distribuições assimétricas, excesso de zeros, sobre-dispersão, pelo que os modelos lineares de suporte contínuo, tradicionais na análise de séries temporais não são adequados. A título ilustrativo considere-se a série representada na figura 1(a) relativa ao número de endereços IP diferentes a aceder à página web do Departamento de Estatística da Universidade de Würzburg, observados em períodos de 2 minutos entre as 10h e as 18h do dia 29 de novembro de 2005. Esta série será, doravante, designada por IP.

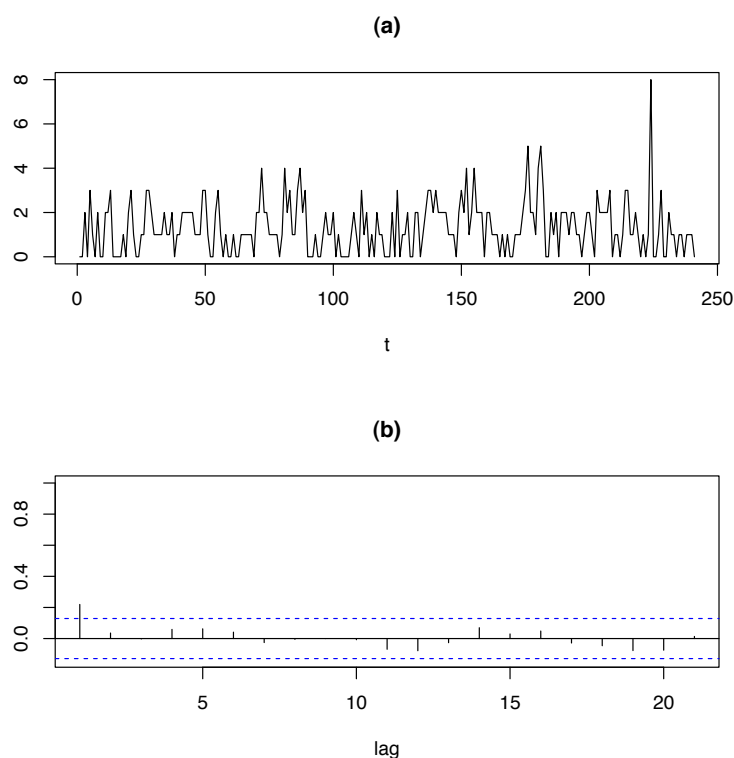


Figura 1: Número de endereços IP diferentes a aceder à página web do Departamento de Estatística da Universidade de Würzburg, observados em períodos de 2 minutos entre as 10h e as 18h do dia 29 de Novembro de 2005 (a) e correspondente função de autocorrelação amostral (b).

Encontram-se exemplos de séries temporais de contagem em muitas áreas de entre as quais destacamos: ambiente, economia e finanças, epidemiologia, ver Silva (2015) para uma lista mais extensa de exemplos. Em muitas situações, as variáveis discretas podem ser aproximadas por variáveis contínuas mas nem sempre essa aproximação é possível ou desejável. Consequentemente, têm sido propostas na literatura diversas estratégias para a análise estatística de séries de contagem. Uma das abordagens baseia-se em modelos do tipo GLM (Modelos Lineares Generalizados) e está descrita pormenorizadamente em Fokianos (2011). Outra abordagem muito popular na literatura, propõe modelos baseados em operações *thinning* (filtragem), operações aleatórias que preservam a característica inteira das variáveis. Estes modelos designados por modelos INARMA, acrónimo de *INteger AutoRegressive Moving Average*, resultam de substituir a operação de multiplicação usual nos modelos ARMA por uma operação *thinning*. Neste artigo ilustramos a modelação de séries temporais de contagem com modelos autorregressivos de valor inteiro de ordem 1, INAR(1).

## 2 Modelo INAR(1)

O modelo INAR(1), está definido no suporte discreto  $\mathbb{N}_0$  pela equação recursiva

$$X_t = \alpha \diamond X_{t-1} + \varepsilon_t \quad (1)$$

onde  $\{\varepsilon_t\}$  é uma sequência de variáveis aleatórias (v.a.) de suporte inteiro, independentes e identicamente distribuídas, independentes de  $X_{t-1}$  e de  $\alpha \diamond X_{t-1}$ , com valor esperado  $\mu_\varepsilon$  e variância  $\sigma_\varepsilon^2$ . ' $\diamond$ ' denota um operador aleatório, denominado *thinning* que operando sobre v.a. inteiras, produz v.a. inteiras e introduz autocorrelação na sequência  $X_t$ .

A operação *thinning* mais popular é a *thinning* binomial proposta por Steutel & VanHarn (1979) para generalizar as noções de v.a. auto-decomponível e estável a v.a. discretas. Se  $X$  é uma v.a. não negativa e  $\alpha \in [0, 1]$  então define-se  $\alpha \circ X$ ,  $\alpha$  *thinning* (binomial)  $X$  da seguinte forma:

$$\alpha \circ X := \sum_{i=1}^X Y_i, \quad (2)$$

onde  $\{Y_i\}_i$  é uma sequência de v.a. i.i.d. de Bernoulli com  $P(Y_i = 1) = \alpha$ , dita série de contagem de  $\alpha \circ X$ , e que é independente de  $X$ . Note-se que dado  $X$ ,  $\alpha \circ X$  tem uma distribuição binomial de parâmetros  $(X, \alpha)$ . Intuitivamente esta operação pode ser interpretada da seguinte forma: considere-se uma população com  $X$  elementos e que a probabilidade de sobrevivência de um qualquer desses elementos é  $\alpha$ ; se os indivíduos sobrevivem forma independente uns dos outros então  $\alpha \circ X$  é o número sobreviventes. A *thinning* binomial também pode ser interpretada como um esquema de reprodução em que cada fêmea dá origem a apenas uma fêmea com probabilidade  $\alpha$ . As propriedades principais da operação *thinning* binomial estão detalhadas em Silva & Oliveira (2004).

Consideremos o modelo (1) com a operação *thinning* binomial, proposto originalmente por Al-Osh & Alzaid (1987) e McKenzie (1988). Para cada  $t$ ,  $X_t$  é constituído por duas componentes:  $\alpha \circ X_{t-1}$  que pode ser interpretada como os sobreviventes dos elementos de  $X_{t-1}$ , cada um com probabilidade de sobrevivência  $\alpha$  e  $\varepsilon_t$  interpretada como os novos elementos que entram no sistema no intervalo  $]t - 1, t]$ . Algumas propriedades importantes do modelo INAR(1) são (Al-Osh & Alzaid, 1987)

1.  $E[X_t] = \frac{\mu_\varepsilon}{(1-\alpha)}$ .
2.  $V[X_t] = \frac{\alpha\mu_\varepsilon + \sigma_\varepsilon^2}{1-\alpha^2}$ .
3.  $\rho_k = \text{corr}(X_t, X_{t-k}) = \alpha^{|k|}$ ,  $k \in \mathbb{Z}$ .

4.  $E[X_t|X_{t-1}] = \alpha X_{t-1} + \mu_\varepsilon$ .
5.  $V[X_t|X_{t-1}] = \alpha(1 - \alpha)X_{t-1} + \sigma_\varepsilon^2$ .

Note-se que a função de autocorrelação de um processo INAR(1) é idêntica à de um AR(1), com a restrição de que o processo apresenta apenas correlações positivas. A média e a variância condicionais são ambas lineares em  $X_{t-1}$ . O INAR(1) difere do AR(1) relativamente à última propriedade. Adicionalmente,  $\{X_t\}$  é uma cadeia de Markov homogênea e portanto estritamente estacionária com probabilidades de transição dadas por

$$p(X_t|X_{t-1}) = P(X_t = k|X_{t-1} = l) = \sum_{j=0}^{\min\{k,l\}} \binom{l}{j} \alpha^j (1 - \alpha)^{l-j} P(\varepsilon_t = k - j), \quad (3)$$

onde  $\binom{\cdot}{\cdot}$  é o símbolo combinatório usual. Por outro lado, se  $G_Z$  denota a função geradora de probabilidades da v.a.  $Z$  então

$$G_{X_t}(s) = G_{X_{t-1}}(1 - \alpha + \alpha s) G_{\varepsilon_t}(s). \quad (4)$$

Isto significa que a distribuição marginal do processo  $X_t$  fica especificada pela escolha do processo de chegadas  $\varepsilon_t$ . Mais, (4) indica que qualquer distribuição auto-decomponível para inteiros pode ser usada como distribuição marginal. Distribuições como a de Poisson ou a Binomial-Negativa são exemplos de distribuições auto-decomponíveis. No entanto, se se pretender que a distribuição marginal de  $X_t$  e a distribuição das chegadas  $\varepsilon_t$  coincidam, à semelhança do que acontece nos processos AR(1) Gaussianos, a escolha está limitada à distribuição de Poisson. Neste caso, as distribuições marginal e de chegadas são ambas de Poisson mas com valores esperados diferentes. Este modelo designa-se por PoINAR(1). Assim, a distribuição de Poisson tem nos INAR(1) um papel semelhante à distribuição Gaussiana nos modelos AR usuais. Para mais propriedades do modelo PoINAR(1) consultar Al-Osh & Alzaid (1987) e McKenzie (1988).

Uma das características empíricas em muitas séries de contagem é a sobredispersão que não é passível de ser captada pelo modelo PoINAR(1). Na tentativa de obter modelos adequados para a sobredispersão e outras características empíricas das séries, têm sido propostas na literatura diversas generalizações da operação *thinning* binomial. Uma revisão de literatura actual sobre as operações *thinning* e modelos nelas baseados encontra-se em Scotto *et al* (2015).

Neste trabalho consideramos apenas o modelo PoINAR(1).

### 3 Modelação com PoINAR(1)

Seja  $\mathbf{x} = (x_1, \dots, x_n)$  a série temporal observada que pretendemos modelar com um modelo PoINAR(1) de parâmetros  $\boldsymbol{\theta} = (\alpha, \lambda)$ , onde  $\lambda = E(\varepsilon_t)$ .

#### 3.1 Estimação do modelo

A estimação do modelo PoINAR(1) é usualmente obtida recorrendo ao método da máxima verosimilhança, m.v.. Alternativamente podem ser considerados os métodos dos momentos, mínimos quadrados e abordagem Bayesiana, Silva (2015).

A função de verosimilhança condicional é dada por

$$L_n(\boldsymbol{\theta}|\mathbf{x}) = \prod_{t=2}^n \sum_{j=0}^{\min\{x_t, x_{t-1}\}} \binom{x_{t-1}}{j} \alpha^j (1 - \alpha)^{x_{t-1}-j} \frac{e^{-\lambda} \lambda^{x_t-j}}{(x_t - j)!} \quad (5)$$

Recorrendo a resultados para processos de Markov prova-se que sob certas condições os estimadores m.v. são consistentes, assintoticamente normais e eficientes, (Joe, 1997, pp. 318).

Geralmente, a maximização da verosimilhança é obtida numericamente. A inversa da matriz Hessiana calculada no máximo pode ser usada como uma estimativa da matriz de variância-covariância do estimador m.v.. As estimativas iniciais requeridas pelos métodos numéricos de otimização são baseadas no método dos momentos.

### 3.2 Ferramentas de diagnóstico

Um fase crucial na modelação de dados é a do diagnóstico, durante a qual se afere a adequação e qualidade estatística do modelo estimado. No contexto de séries temporais de contagem as principais ferramentas de diagnóstico são: métodos paramétricos de re-amostragem; análise dos resíduos; métodos baseados nas distribuições preditivas e *scores* e critérios de informação.

#### Reamostragem paramétrica

Tsay (1992) propôs um procedimento baseado em bootstrap paramétrico e funcionais específicos para avaliar características de interesse no ajuste do modelo. Aqui considera-se como funcional de interesse a função de autocorrelação da série. Assim, o modelo estimado é usado para gerar  $M$  séries temporais sintéticas com a mesma dimensão da série a modelar. A partir destas obtêm-se  $M$  funções de autocorrelação amostrais e, conseqüentemente uma distribuição empírica da função de autocorrelação. Então, para cada desfasamento (*lag*), calculam-se os quantis  $(1 - \alpha/2)$  e  $\alpha/2$  da respectiva distribuição empírica e que constituem um intervalo de aceitação. Forma-se deste modo um envelope de aceitação para a autocorrelação amostral. Diz-se que o modelo reproduz adequadamente as características de autocorrelação da série se a autocorrelação amostral se situa no envelope de aceitação. Notar que uma vez que a autocorrelação amostral é correlacionada em diferentes desfasamentos, o envelope de aceitação não é um intervalo de confiança conjunto para a autocorrelação amostral.

#### Resíduos

A adequação do modelo para representar a dinâmica da série e a sua dispersão pode ser verificada usando os resíduos de Pearson definidos como

$$r_t = \frac{X_t - E(X_t|X_{t-1})}{\text{Var}(X_t|X_{t-1})^{1/2}}, \quad (6)$$

onde as quantidades populacionais são substituídas por estimativas. Se o modelo estiver correctamente especificado, estes resíduos devem apresentar média zero, variância 1 e não devem ter autocorrelação.

Podemos ver a estrutura do modelo PoINAR(1) como a soma de duas componentes não observadas: uma componente,  $\alpha \circ X_{t-1}$ , especifica o número de sobreviventes (ou partidas) dos sistema entre  $t - 1$  e  $t$ , enquanto que  $\varepsilon_t$  representa o número de chegadas ao sistema em  $t$ . Esta interpretação sugere uma decomposição dos resíduos que permite a verificação da adequação de cada componente. Detalhes do procedimento em Freeland & McCabe (2004).

#### Distribuições preditivas

Uma ferramenta útil para verificar a adequação da assumções distribucionais é a transformada integral de probabilidade, PIT. Esta ferramenta tem sido usada na aferição de distribuições preditivas contínuas, Gneiting *et al* (2007). Czado *et al* (2009) propôs um ajustamento ao PIT para o caso de distribuições discretas. A avaliação do modelo baseada no desempenho das distribuições preditivas é ainda conseguida através de regras de scoring sugeridas por Czado *et al* (2009) e Jung & Tremayne (2011).

## Critérios de Informação

O critério de informação de Akaike, AIC, e suas variantes é uma das ferramentas mais populares para selecionar modelos. Alguns autores têm usado o AIC para escolher entre modelos não encaixados para séries de contagem apesar de não existirem estudos relativos ao desempenho deste critério neste contexto. Psaradakis *et al* (2009) estudam critérios de informação tais como AIC, BIC e o critério de Hannan-Quinn (HQ) para distinguir entre alguns dos modelos não lineares mais populares para séries temporais e, com base em simulações concluem que estes critérios podem ser úteis.

## 4 Ilustração

Esta secção ilustra o procedimento de modelação da série temporal IP descrita na secção 1. Esta série, estudada originalmente por Weiß (2007), exhibe autocorrelação significativa, como se pode ver na figura 1(b). A média e variância amostrais,  $\bar{x} = 1.31$  e  $\hat{\sigma}^2 = 1.39$  não indicam sobredispersão, pelo que se considera um modelo PoINAR(1).

As estimativas m.v. para os parâmetros do modelo são  $\hat{\alpha} = 0.24(0.00)$  e  $\hat{\lambda} = 1.01(0.01)$ . A reamostragem paramétrica e a análise da autocorrelação dos resíduos, figura 2 indicam que o modelo captura adequadamente a dinâmica da série. No entanto, o cronograma dos resíduos indica que a observação em  $t = 224$  apresenta um resíduo muito maior do que os restantes, originado pela componente das chegadas. Este resultado pode indicar a presença de um outlier aditivo cujo efeito não contamina as observações subsequentes.

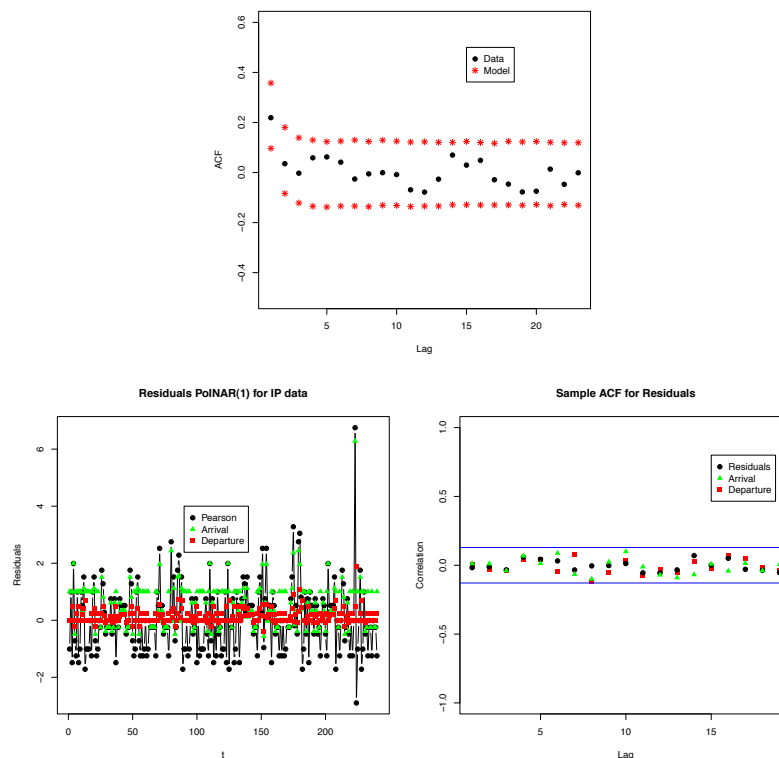


Figura 2: Resultados do bootstrap paramétrico com  $M = 500$  no painel superior, componentes dos resíduos no painel inferior à esquerda e respectivas autocorrelações à direita, para a série IP.

Silva & Pereira (2015) propuseram uma abordagem Bayesiana para a modelação de tais outliers assumindo que o processo observado  $Y_t$  é obtido de um processo não observado PoINAR(1),  $X_t$ , contaminando  $X_t$  com um outlier de dimensão  $\eta_t$  e com probabilidade de contaminação  $\delta_t$ .

$$Y_t = X_t + \eta_t \delta_t,$$

$$\text{com } X_t = \alpha \circ X_{t-1} + e_t \text{ e } \delta_t \sim Be(p_t) \quad (7)$$

onde  $\delta_1, \eta_1, \dots, \delta_n, \eta_n$  são independentes do processo latente  $X_t$  e  $\eta_t$ , a dimensão do outlier no tempo  $t$  é uma v.a. com o mesmo suporte que  $X_t$  e média  $\beta$ ,  $\eta_t \sim Po(\beta)$ . Aplicando o procedimento de detecção de outliers descrito em Silva & Pereira (2015), obtém-se, para cada  $t$  a probabilidade de  $Y_t$  estar contaminado, ver figura 3. O gráfico indica claramente que a observação em  $t = 224$  é, com grande probabilidade, um outlier.

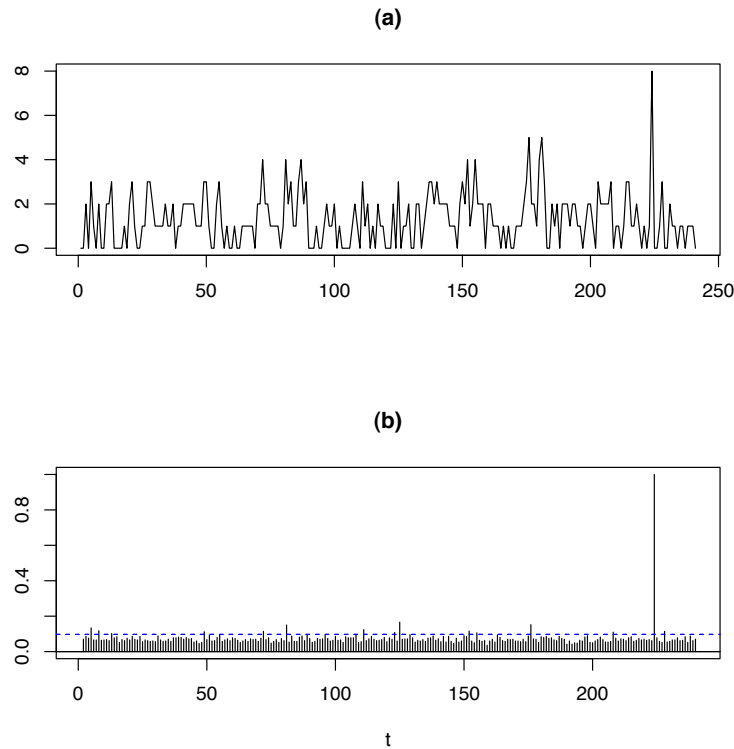


Figura 3: Série Temporal IP (a) e a probabilidade *a posteriori* de ocorrência de outlier para cada tempo  $t$ (b).

O modelo (7) estimado é o seguinte:

$$Y_t = X_t + 7I_{224},$$

$$X_t = 0.27 \circ X_{t-1} + e_t, \quad e_t \sim Po(0.89) \quad (8)$$

Os resíduos estão representados na figura 4. Note-se que o maior resíduo reduziu de 6.8 para 3.3, indicando uma melhor adequação do modelo aos dados. Adicionalmente, verifica-se que a soma dos quadrados do erro de previsão  $\sum_{t=2}^n (y_t - \hat{y}_t)^2$ , onde  $\hat{y}_t = E(y_t | y_{t-1} = y_{t-1}; \text{estimativas dos parâmetros})$  cai de 317.9 para 264.0 quando o outlier é incluído no modelo.

O processo PoINAR(1) pode ser interpretado como uma fila de espera com serviço infinito. O tempo de serviço é geométrico com parâmetro  $1 - 0.27$  e o processo de chegada é Poisson com média 0.89. Um resultado fundamental da teoria de filas de espera, a equação de Little Flow, permite então dizer que um novo IP acede ao servidor do Departamento em média durante 2 minutos e 44 segundos.



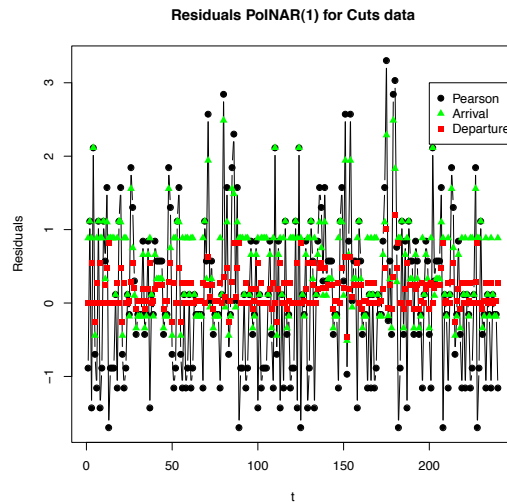


Figura 4: Série IP: resíduos do modelo com outlier (8).

## 5 Comentários Finais

Séries temporais de contagem ocorrem em muitas áreas de conhecimento. A literatura apresenta várias abordagens e diversos modelos para analisar adequadamente este tipo de dados, ver por exemplo Fokianos (2011), Davis & Wu (2009), Cui & Lund (2009). Este trabalho focou-se nos modelos INAR(1) que são uma classe de modelos *observation-driven*, apropriados para dados tipo *stock*. A generalização dos modelos INAR(1) a modelos de ordem superior não é única em virtude do seu carácter não linear. Os modelos INAR(p) propostos por Du & Li (1991) constituem a generalização mais popular na literatura. McKenzie (2003) propôs modelos com componentes MA, designados por INARMA, enquanto Monteiro *et al* (2010) estudou modelos INAR periódicos. A extensão ao caso bivariado foi considerada por Pedelli & Karlis (2011) para os modelos INAR e por Silva *et al* (2016) para modelo INMA.

**Agradecimentos** Este trabalho foi financiado parcialmente pelo CIDMA - Centro para a Investigação e Desenvolvimento em Matemática e Aplicações, e pela FCT- “FCT Fundação para a Ciência e a Tecnologia”), através do projecto PEst-OE/MAT/UI4106/2014.

## Referências

- [1] Al-Osh, M., Alzaid, A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis* **8**, 261-275.
- [2] Cui, Y., Lund, R. (2009). A new look at time series of counts. *Biometrika* **96**, 781–792.
- [3] Czado, C., Gneiting, T., Held, L. (2009). Predictive model assessment for count data. *Biometrics* **65**, 1254-1261.
- [4] Davis, R., Wu, R. (2009). A negative binomial model for time series of counts. *Biometrika* **96**, 735-749.
- [5] Fokianos, K. (2011). Some Recent Progress in Count Time Series, *Statistics* **45**, 49-58.
- [6] Freeland, R.K., McCabe, B.P.M. (2004). Analysis of low count time series data by Poisson autoregression. *Journal of Time Series Analysis* **25**, 701-722.

- [7] Gneiting, T., Balabdaoui, F., Raftery, A.E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**, 243-268.
- [8] Jin-Guan, D., Yuan, L. (1991). The integer-valued autoregressive INAR(p) model. *Journal of Time Series Analysis* **12**, 129-142.
- [9] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
- [10] Jung, R.C., Tremayne, A.R. (2011). Convolution-closed models for count time series with applications. *Journal of Time Series Analysis* **32**, 268-280.
- [11] McKenzie, E.(1988).Some ARMA models for dependent sequences of Poisson counts. *Advances in Applied Probability* **20**, 822-835.
- [12] McKenzie, E. (2003). Discrete variate time series, in: *Handbook of Statistics*, Rao, C., Shanbhag, D. (Eds.), Elsevier Science, Amsterdam, 573-606.
- [13] Monteiro, M., Pereira, I., Scotto, M.G. (2010). Integer-valued autoregressive processes with periodic structure. *Journal Statistical Planning and Inference* **140**, 1529-1541.
- [14] Pedeli, X., Karlis, D. (2011). A bivariate INAR(1) process with application. *Statistical Modelling* **11**, 325-349.
- [15] Psaradakis, Z., Sola, M., Spagnolo, F., Spagnolo, N.(2009). Selecting nonlinear time series models using information criteria. *Journal of Time Series Analysis* **30**, 369-394.
- [16] Scotto, M.G. Weiss, C.H., Gouveia, S. (2015). *Thinning*-based models in the analysis of integer-valued time series: a review. *Statistical Modelling* **15**, 590-618.
- [17] Silva, M.E. (2015). Modelling Time Series of Counts: An INAR Approach. *Textos de Matemática, Departamento de Matemática da Universidade de Coimbra* **47**, 107-122.
- [18] Silva, M.E., Oliveira, V. (2004). Difference equations for the higher-order moments and cumulants of the INAR(1) model. *Journal of Time Series Analysis* **25**, 317-333.
- [19] Silva, M. E., Pereira, I. (2015). Detection of Additive Outliers in Poison INAR(1) Time Series, *Mathematics of Energy and Climate Change, International Conference and Advanced School Planet Earth, Portugal, March 21-28, 2013*, Bourguignon, J.-P., Pinto, A.A., Viana, M., eds, CIM Series in Mathematical Sciences, **2**, 377-388.
- [20] Silva, I., Silva,M.E., Torres, C. (2016). Inference for bivariate integer-valued moving average models. Submetido.
- [21] Steutel, F., Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability* **7**, 893-899.
- [22] Tsay, R.S.(1992). Model checking via parametric bootstraps in time series analysis. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **41**, 1-15.
- [23] Weiß, C.H.(2008). Thinning operations for modelling time series of counts: a survey. *AStA - Advances in Statistical Analysis* **92**, 319-341.



# Séries temporais em Economia

Paulo Teles<sup>†</sup> e André Almeida<sup>‡</sup>, *pteles@fep.up.pt; 120411039@fep.up.pt*

<sup>†</sup>*Faculdade de Economia e LIAAD-INESC Porto LA, Universidade do Porto*

<sup>‡</sup>*Faculdade de Economia da Universidade do Porto*

## 1 Introdução

A análise empírica em Economia recorre frequentemente a dados temporais, tanto no domínio da macroeconomia como da microeconomia e das Finanças. Em resultado, tornou-se necessária em Economia Aplicada a intervenção conjunta da Análise de Séries Temporais e da Econometria, tendo assim surgido a Econometria de Séries Temporais que tem registado um desenvolvimento extremamente rápido. O seu principal objetivo consiste em desenvolver modelos razoavelmente simples capazes de prever, interpretar e testar hipóteses relativas a dados económicos. Alguns dos métodos mais frequentemente utilizados para dar resposta a estes desafios irão ser abordados em seguida.

Para este efeito, representa-se uma série temporal como  $X_t$ , com  $t = 1, \dots, n$ . Na análise de séries económicas, recorre-se frequentemente à transformação logarítmica, ou seja,  $X_t = \log(Y_t)$  onde  $\log$  designa o logaritmo natural e  $Y_t$  é a série temporal original.

## 2 Modelos com tendência

Admite-se frequentemente que as séries económicas são formadas por uma componente de curto/médio prazo, designada como componente cíclica, e por uma componente de longo prazo, designada por tendência. A identificação e separação dessas componentes é muito importante, uma vez que se encontra na origem da produção de dados económicos que, por um lado, serão utilizados pelos economistas em estudos empíricos ou na validação empírica dos seus modelos teóricos e que, por outro, irão também influenciar as decisões dos diferentes agentes, sejam eles consumidores, investidores ou decisores de política. No entanto, esta é uma tarefa que apresenta alguns obstáculos, como por exemplo a dificuldade em identificar, empírica e teoricamente, a magnitude e duração do impacto dos diferentes tipos de choques que afetem as variáveis de interesse. Note-se que uma identificação imprecisa destes choques resulta necessariamente numa incorreta identificação das duas componentes. Refira-se ainda que uma série com tendência é não estacionária na média, a forma de não estacionariedade a ser considerada aqui.

### 2.1 Tendência determinística ou estocástica

A média  $\mu_t$  de uma série não estacionária pode ser representada por uma tendência determinística que é uma função puramente determinística do tempo e que se supõe geralmente ser um polinómio de grau  $k$  ( $k = 1, 2, \dots$ ), ou seja,  $\mu_t = \beta_0 + \beta_1 t + \dots + \beta_k t^k$  onde  $\beta_i$  ( $i = 0, \dots, k$ ) são os coeficientes do polinómio. Logo,  $X_t$  pode ser escrita como

$$X_t = \mu_t + U_t = \beta_0 + \beta_1 t + \dots + \beta_k t^k + U_t \quad (1)$$

onde  $U_t$  é estacionária, pelo que  $X_t$  se desvia da tendência apenas temporariamente. Este modelo é designado por modelo de tendência determinística ou estacionário em tendência pois, com efeito, a diferença  $X_t - \mu_t = U_t$  já é estacionária e note-se ainda que  $U_t$  pode ser um modelo Autoregressivo e de Médias Móveis ARMA( $p, q$ ) tal que  $\phi(B)U_t = \theta(B)a_t$ , onde  $B$  designa o operador atraso, ou seja,  $B^i Z_t = Z_{t-i}$ ,  $\phi(B)$  e  $\theta(B)$  representam os operadores autoregressivo e de médias móveis respetivamente e  $a_t$  é um ruído branco. Outras funções são frequentemente propostas para  $\mu_t$ , destacando-se a curva seno-cosseno.

Por sua vez, nos modelos com tendência estocástica, o nível médio da série varia estocasticamente no tempo. Para compreender este fenómeno, considere-se o exemplo do passeio aleatório

$$X_t = X_{t-1} + a_t \Leftrightarrow (1 - B)X_t = a_t \quad (2)$$

onde  $a_t$  é um ruído branco de média zero. Dada a informação passada  $X_t, X_{t-1}, \dots$ , o nível da série no período  $t$  é  $\mu_t = X_{t-1}$ , que está sujeito a uma perturbação estocástica em  $(t - 1)$ , pelo que se diz que a série tem uma tendência estocástica e é portanto não estacionária. O modelo (2) pode ainda incluir uma constante ou *drift*, isto é,  $X_t = \delta + X_{t-1} + a_t$ . De uma forma geral, um modelo  $\varphi(B)X_t = \theta(B)a_t$  é não estacionário se existirem raízes do polinómio autoregressivo que não estejam fora do círculo unitário, ou seja, se  $\varphi(B) = \phi(B)(1 - B)^d$  para um qualquer inteiro  $d \geq 1$  e em que todas as raízes de  $\phi(B)$  estão fora do círculo unitário, pelo que este operador é estacionário. Verifica-se assim a existência de  $d$  raízes unitárias em  $\varphi(B)$ . Logo, uma série não estacionária pode ser tornada estacionária através do cálculo das suas diferenças  $(1 - B)^d X_t = \nabla^d X_t$ , sendo  $\nabla = (1 - B)$  o operador diferença. Consequentemente,  $(1 - B)^d X_t$  segue o modelo ARMA( $p, q$ )

$$\phi(B)(1 - B)^d X_t = \theta(B)a_t, \quad (3)$$

pelo que  $X_t$  segue o modelo ARIMA( $p, d, q$ ). Diz-se então que  $X_t$  é integrada de ordem  $d$ , designando-se por  $I(d)$ .

No exemplo do passeio aleatório acima, note-se que  $(1 - B)X_t = a_t$ , que é um ruído branco e portanto estacionário, ou seja, a aplicação de uma diferença ( $d = 1$ ) à série elimina a tendência estocástica e consequentemente a não estacionaridade. Em conclusão, a aplicação de um número adequado de diferenças a uma série não estacionária elimina uma tendência estocástica, tornando a série estacionária. Por isso, uma série deste tipo é frequentemente designada por estacionária em diferenças.

É muito habitual admitir-se que, no longo prazo, as variáveis macroeconómicas crescem seguindo uma tendência com taxa constante e que quaisquer desvios em relação a essa tendência acabam por desaparecer. A suposição de que a tendência não muda com o tempo conduz à prática comum de eliminar a tendência dos dados macroeconómicos recorrendo a uma equação de regressão determinística linear ou polinomial. No entanto, Nelson e Plosser (1982) desafiaram o método tradicional, demonstrando com recurso a séries de dados que muitas variáveis macroeconómicas de relevo (como o PIB real e nominal, a produção industrial, o emprego, a taxa de desemprego e os preços no consumidor, entre outras) tendem a ser estacionárias em diferenças e não em tendência. Esta conclusão significa que as variáveis macroeconómicas não crescem a uma taxa suave de longo prazo. Alguns choques macroeconómicos têm uma natureza permanente, pelo que os seus efeitos nunca são eliminados.

## 2.2 Separação Ciclo/Tendência

Tendo sido definidos os tipos de tendência que é habitual considerar, serão abordados em seguida alguns dos métodos mais utilizados na identificação e separação da tendência e da componente cíclica. De uma forma geral, estes podem ser divididos em duas classes distintas: Métodos estatísticos de remoção de tendência e Filtros. A principal diferença entre estes métodos reside no facto de os primeiros removerem a tendência transformando a série temporal e tornando-a estacionária (na média). Por outro lado, a aplicação de filtros não altera a série temporal, sendo apenas utilizados métodos estatísticos para eliminar

do periodograma da série as frequências que não correspondam à componente cíclica, ou seja o ruído de curto prazo e a tendência (Canova, 2007). Para que tal seja possível, considera-se que  $X_t = C_t + T_t$ , sendo  $C_t$  e  $T_t$  respetivamente a componente cíclica e a tendência, ambas não observáveis.

## 2.2.1 Métodos estatísticos de remoção da tendência

De entre os métodos mais utilizados de extração da tendência, o método do “growth cycle” é o mais simples. Reescrevendo  $X_t = X_t - X_{t-1} + X_{t-1} = \nabla X_t + X_{t-1}$ , este método consiste simplesmente em interpretar  $\nabla X_t$  como a informação relativa a  $C_t$ . Assim, basta aplicar uma diferença à série original de forma a obter  $C_t$  e, por conseguinte, a tendência resulta diretamente como  $T_t = X_t - C_t$ . Note-se que  $C_t$  não é necessariamente um processo com média nula. Deste modo, a identificação das diferentes fases de um ciclo requer a comparação de  $C_t$  com uma determinada taxa de crescimento de longo prazo,  $\mu$ . Isto é, uma recessão ocorre quando  $C_t < \mu$  e uma expansão quando  $C_t > \mu$ . No entanto, de uma forma geral, este método tem fraco desempenho, resultando em ciclos demasiado curtos e com elevada variância.

Alternativamente, é possível considerar que a tendência é uma função determinística do tempo, representada pelo polinómio (1), isto é,  $\mu_t = T_t = \beta_0 + \sum_{i=1}^k \beta_i t^i$ , em que  $\text{corr}(T_t, C_t) = 0$ . Desta forma, estimando  $T_t$  pelo método dos mínimos quadrados, obtém-se  $C_t$  como resíduo de estimação. A escolha da ordem do polinómio para  $T_t$  deverá ser baseada nalgum critério de seleção de modelos como o AIC ou o  $R^2$  ajustado. Contudo, este método entra em conflito com as conclusões de Nelson e Plosser (1982) referidas anteriormente.

Um método alternativo que apresenta algumas vantagens relativamente aos anteriores é a denominada transformação de Beveridge e Nelson (1981). Estes autores admitem que  $X_t$  é integrada de primeira ordem. Por conseguinte, a aplicação de uma diferença elimina a tendência estocástica, resultando na série estacionária  $\nabla X_t$  que, por definição de estacionaridade, possui a representação em médias móveis

$$\nabla X_t = \bar{X} + \sum_{i=0}^{\infty} \psi_i a_{t-i} = \bar{X} + \sum_{i=0}^{\infty} \psi_i B^i a_t = \bar{X} + \psi(B) a_t,$$

onde  $a_t$  é um ruído branco e  $\psi_i \rightarrow 0$  exponencialmente quando  $i \rightarrow \infty$ . Beveridge e Nelson (1981) mostraram que esta equação pode ser escrita como

$$\nabla X_t = \bar{X} + \psi(1) a_t + (1 - B) \Lambda(B) a_t = \bar{X} + \psi(1) a_t + \nabla \varepsilon_t$$

onde  $\Lambda(B) = [\psi(B) - \psi(1)] / (1 - B) = - \sum_{i=0}^{\infty} \left( \sum_{j=i+1}^{\infty} \right) B^j$  e  $\varepsilon_t = \Lambda(B) a_t$  é um processo estacionário de média nula. Consequentemente,

$$X_t = X_0 + \sum_{i=1}^t \nabla X_i = X_0 + \sum_{i=1}^t (\bar{X} + \psi(1) a_i + \nabla \varepsilon_i) = \psi(1) \sum_{i=1}^t a_i + \varepsilon_t + (t\bar{X} + X_0 - \varepsilon_0). \quad (4)$$

Portanto, ao admitirem que a série temporal tem exatamente uma raiz unitária, Beveridge e Nelson (1981) concluem que  $X_t$  corresponde à soma de uma tendência estocástica representada por um passeio aleatório

com *drift*,  $T_t = \psi(1) \sum_{i=1}^t a_i + (t\bar{X} + X_0 - \varepsilon_0) = T_{t-1} + \bar{X} + \psi(1) a_t$ , e de uma componente cíclica

representada por um processo estacionário de média nula,  $C_t = \varepsilon_t = \Lambda(B) a_t$ . Note-se que  $C_t$  e  $T_t$  se encontram perfeitamente correlacionados uma vez que ambos dependem de  $a_t$ . A principal vantagem deste método é o facto de impor apenas um pressuposto, referente ao número de raízes unitárias de  $X_t$ , não existindo nenhum pressuposto relativamente à correlação ente  $C_t$  e  $T_t$ . Contudo, não tem o melhor desempenho em termos práticos, gerando frequentemente padrões e *timings* para os ciclos diferentes dos que seriam teoricamente esperados.

## 2.2.2 Filtros

Pese embora a validade dos últimos métodos, a aplicação de filtros é indiscutivelmente a prática mais utilizada entre os economistas para extrair a componente cíclica. A utilização destes métodos permite também uma melhor previsão para a evolução da componente de longo prazo, sendo frequentemente utilizados para a estimação do denominado produto potencial, ou seja, da capacidade produtiva de uma economia. De entre estes destacam-se o filtro de Hodrick e Prescott (HP), o filtro de Baxter e King (BK) e, por último, o filtro de Kalman. A aplicação destes filtros tem por objetivo remover do periodograma da série as frequências que não correspondam a essa componente. Por exemplo, no que respeita ao PIB *per capita*, os dois principais centros de investigação especializados em análise de ciclos económicos (NBER e CEPR) consideram que um ciclo tem aproximadamente uma duração entre 6 e 32 trimestres, o que corresponde às frequências no intervalo  $[\pi/16; \pi/3]$ . Deste modo, o filtro ideal, designado por  $h(B)$ , eliminaria as restantes frequências, mantendo inalteradas as frequências neste intervalo tal que  $C_t = h(B) X_t$ . Ambos os filtros HP e BK não apresentam mudança de fase, o que significa que mantêm inalterados os *timings* dos picos e cavas de cada ciclo.

O filtro HP deve o seu nome aos dois economistas Hodrick e Prescott (1997) que tornaram este método célebre na análise de séries económicas. É o mais utilizado e resulta da seguinte otimização no domínio tempo:

$$\min_{\{T_t\}} \left\{ \sum_{t=1}^n (X_t - T_t)^2 + \lambda \sum_{t=2}^{n-1} [(T_{t+1} - T_t) - (T_t - T_{t-1})]^2 \right\},$$

em que  $\lambda$  representa o multiplicador de Lagrange e é escolhido de forma a aproximar o filtro do ideal, dependendo da periodicidade dos dados em questão. Na prática, e tomando como referência a análise do PIB *per capita*, os valores tipicamente utilizados são  $\lambda = 6400, 1600, 10$  para dados mensais, trimestrais e anuais respetivamente. Mais uma vez, esta escolha *a priori* do valor de  $\lambda$  tem por objetivo eliminar do periodograma da série as frequências associadas a ciclos de duração superior a 6 – 7 anos, por se considerar que correspondem à tendência (Canova, 2007). Dado  $C_t = X_t - T_t$ , esta otimização representa a minimização da variância da componente cíclica, isto é,  $(X_t - T_t)^2$ , penalizando ao mesmo tempo acelerações da tendência, uma vez que  $(T_{t+1} - T_t) - (T_t - T_{t-1}) = \nabla T_{t+1} - \nabla T_t$ . Note-se que esta penalização, e a conseqüente escolha do valor para  $\lambda$ , desempenham um papel fundamental na identificação de  $C_t$  e  $T_t$ . No caso de  $\lambda = 0$ , a solução da presente otimização implicaria  $C_t = 0$ . Por outro lado, o caso de  $\lambda \rightarrow \infty$  implicaria considerar uma função linear para  $T_t$ , ou seja, uma suavização excessiva da tendência. Assim, resolvendo este problema de otimização, obtém-se

$$HP(B) = \frac{\lambda(1-B)^2(1-B^{-1})^2}{1 + \lambda(1-B)^2(1-B^{-1})^2} \quad (5)$$

o que corresponde ao ganho de  $4\lambda [1 - \cos(\omega)]^2 / [1 + 4\lambda(1 - \cos(\omega))^2]$ . Da expressão de  $HP(B)$ , constata-se que valores passados, presentes e futuros de  $X_t$  serão importantes para a determinação de  $C_t$ . O peso relativo de cada um desses três conjuntos de valores é dado pelo valor de  $\lambda$ . A aplicação deste filtro remove completamente as frequências baixas correspondentes à tendência e mantém inalteradas as frequências elevadas, também denominadas de ruído, resultando num gráfico menos suave para  $C_t$ . Uma das desvantagens da aplicação dos filtros em geral, e que também é uma característica do filtro HP em particular, é a má qualidade de ajustamento da tendência nos períodos iniciais e finais. Deste modo, a sua utilização levanta algumas reservas relativamente às considerações a retirar destes períodos de tempo. Söderlind (1994) e Cogley e Nason (1995) chamam também a atenção para a aplicação do filtro HP a séries temporais com pouca variabilidade, uma vez que o mecanismo de penalização da suavização excessiva do ciclo pode resultar numa componente cíclica significativa (Canova, 2007).

Outro filtro resultante de uma otimização no domínio tempo, embora não tão frequentemente utilizado como o filtro HP, é o filtro *Exponential smoothing*. A diferença em relação ao filtro anterior consiste basicamente no mecanismo de penalização da suavização excessiva da componente cíclica. O problema

de otimização considerado é

$$\min_{\{T_t\}} \left\{ \sum_{t=1}^n (X_t - T_t)^2 + \lambda \sum_{t=2}^{n-1} (T_t - T_{t-1})^2 \right\}$$

o que significa que, contrariamente a penalizações na aceleração, são penalizadas variações na tendência.

Por outro lado, o filtro de Baxter-King resulta de uma otimização no domínio espectral:

$$\min_{\{\tilde{h}_j\}_{j=-K}^K} \frac{1}{2\pi} \left\{ \int_{-\pi}^{\pi} |h(e^{-i\omega}) - h_K(e^{-i\omega})|^2 d\omega \right\} \quad \text{s.r.} \quad h_K(B) = \tilde{h}_0 + \sum_{j=1}^K \tilde{h}_j (B^j + B^{-j})$$

onde  $h(e^{-i\omega})$  representa o filtro ideal e, por sua vez,  $h_K(e^{-i\omega})$  representa um filtro aproximado ao ideal, cujos coeficientes  $\tilde{h}_j$  se pretende otimizar de forma a que a diferença entre estes seja a menor possível. Estes coeficientes correspondem ao peso relativo que os valores passados e futuros de  $X_t$  terão na determinação do filtro aproximado. Note-se ainda que  $\tilde{h}_j = \tilde{h}_{-j}$  com  $j = 1, \dots, K$ , ou seja, valores de  $X_t$  com o mesmo desfasamento temporal possuem o mesmo peso relativo (Baxter e King, 1999).

Da resolução desta problema de otimização obtém-se  $\hat{h}_j = h_j - \left( h_0 + 2 \sum_{j=1}^K h_j \right) / (2K + 1)$ , sendo

$\sum_{j=-K}^K \hat{h}_j = 0$ . Este procedimento consiste numa aproximação do filtro ideal por uma representação em

médias móveis simétrica de ordem  $(2K + 1)$ , que elimina tanto a tendência como o ruído de curto prazo, resultando numa representação mais suavizada do ciclo. Refira-se que filtros com a capacidade de eliminar estes dois conjuntos de frequências são geralmente denominados de *band-pass filters*. A principal vantagem deste método é precisamente a capacidade para eliminar as frequências correspondentes ao ruído de curto prazo. Recorde-se que a aplicação de um filtro HP mantém inalterado este conjunto de frequências, resultando por isso numa representação para  $C_t$  menos suavizada em comparação com a representação resultante da aplicação do filtro BK. Em contrapartida, a principal desvantagem do filtro BK é o facto de se perderem  $2K$  observações, correspondentes aos  $K$  períodos iniciais e finais da série temporal. Considerando por exemplo uma série com dados trimestrais, é geralmente utilizado  $K = 12$ , o que implica a perda dos primeiros e últimos três anos de observações. Consequentemente, a aplicação deste filtro deve ter por base uma série temporal relativamente extensa. Por último, refira-se ainda que ambos os filtros HP e BK não apresentam mudança de fase, o que significa que mantêm inalterados os *timings* dos picos e cavas de cada ciclo (Baxter e King, 1999).

Por fim, um outro filtro geralmente utilizado pelos economistas em estudos macroeconómicos é o filtro de Kalman, originalmente desenvolvido por Kalman (1960) num contexto de sistemas lineares. A aplicação deste filtro requer a representação em espaço de estados de um sistema dinâmico da forma

$$\begin{aligned} \mathbf{Z}_t &= \mathbf{A}^T \mathbf{Z}_t + \mathbf{H}^T \zeta_t + \xi_t, \\ \zeta_{t+1} &= \mathbf{F} \zeta_t + \nu_{t+1} \end{aligned}$$

em que  $\mathbf{Z}_t$  é uma matriz com as variáveis pré-determinadas, os vetores  $\xi_t$  e  $\nu_t$  são ruídos brancos não correlacionados,  $\zeta_t$  é não observável e representa o vetor das variáveis de estado e, finalmente,  $\mathbf{A}$ ,  $\mathbf{H}$  e  $\mathbf{F}$  são matrizes de parâmetros. A primeira equação é geralmente denominada de equação de medida e a segunda de equação de transição. Este filtro consiste num algoritmo iterativo que, para uma dada condição inicial, procede à estimação de  $\zeta_t$  e sequencialmente recalcula essa estimação assim que novas observações se encontram disponíveis, gerando assim uma projeção linear do sistema (Wei, 2006). Em termos práticos, a aplicação do filtro de Kalman é bastante sensível à definição das condições iniciais e dos parâmetros do modelo. Esta característica, aliada à frequente não linearidade das funções a otimizar, impossibilita várias vezes a resolução do problema de otimização o que, naturalmente, representa uma séria desvantagem deste método.

### 3 Testes de raízes unitárias

As séries temporais de variáveis económicas são geralmente não estacionárias e, quando tal sucede, é necessário em primeiro lugar torná-las estacionárias (continuamos a referir-nos a não estacionaridade na média). Com efeito, a identificação de um modelo ARMA exige a determinação prévia de  $d$ , a ordem de integração da série, que é o número de raízes unitárias no polinómio autoregressivo e, portanto, o número de diferenças necessárias para tornar a série estacionária. Por este motivo, é fundamental testar a estacionaridade de uma série, tendo sido propostos muitos testes com este objetivo, pelo que serão apresentados os mais utilizados.

#### 3.1 Testes de Dickey-Fuller

O modelo AR(1)

$$X_t = \phi X_{t-1} + a_t \Leftrightarrow (1 - \phi B) X_t = a_t \quad (6)$$

é o mais simples que pode conter uma raiz unitária, onde  $a_t \sim N(0, \sigma_a^2)$  e  $t = 1, \dots, n$  com  $X_0 = 0$ . Se  $\phi = 1$  (raiz unitária),  $X_t$  é não estacionária, convertendo-se no passeio aleatório (2). Por isso, a hipótese nula em teste é a de que  $X_t$  é não estacionária, ou seja,  $H_0 : \phi = 1$ , contra a alternativa de estacionaridade

$H_1 : \phi < 1$ . O estimador de mínimos quadrados de  $\phi$  em (6) é  $\hat{\phi} = \frac{\sum_{t=1}^n X_{t-1} X_t}{\sum_{t=1}^n X_{t-1}^2}$ . Para testar as hipóteses, podia pensar-se em recorrer à distribuição normal ou  $t$ -Student, mas estas só são válidas no caso estacionário, ou seja, quando  $|\phi| < 1$ . Por este motivo, Dickey e Fuller (1979) propuseram as seguintes estatísticas-teste sob  $H_0$  e determinaram as suas distribuições assintóticas (ver também Chan e Wei, 1988):

$$n(\hat{\phi} - 1) = \frac{n^{-1} \sum_{t=1}^n X_{t-1} a_t}{n^{-2} \sum_{t=1}^n X_{t-1}^2} \xrightarrow{D} \frac{\frac{1}{2} \{ [W(1)]^2 - 1 \}}{\int_0^1 [W(z)]^2 dz} \quad (7)$$

$$T = \frac{\hat{\phi} - 1}{S_{\hat{\phi}_1}} = \frac{\hat{\phi} - 1}{\left[ \hat{\sigma}_a^2 \left( \sum_{t=1}^n X_{t-1}^2 \right)^{-1} \right]^{1/2}} \xrightarrow{D} \frac{\frac{1}{2} \{ [W(1)]^2 - 1 \}}{\left[ \int_0^1 [W(z)]^2 dz \right]^{1/2}} \quad (8)$$

onde  $\hat{\sigma}_a^2 = \frac{1}{n-1} \sum_{t=1}^n (X_t - \hat{\phi} X_{t-1})^2$ ,  $\xrightarrow{D}$  representa convergência em distribuição e  $W(z)$ , com  $z \in [0, 1]$ , designa o processo de Wiener ou movimento browniano *standard*. Note-se que, uma vez que  $W(z) \sim N(0, z)$ , então  $[W(1)]^2 \sim \chi_{(1)}^2$ . Conclui-se então que, sob  $H_0$ , as estatísticas-teste (7) e (8) não seguem distribuição normal nem  $t$ -Student respetivamente, pelo que seria errado recorrer às tabelas de percentis destas distribuições para obter os pontos críticos e respetivas regiões críticas para efetuar o teste. O modelo (6) pode ser escrito como  $\nabla X_t = \lambda X_{t-1} + a_t$  onde  $\lambda = (\phi - 1)$ , sendo  $H_0 : \lambda = 0$  e  $H_1 : \lambda < 0$ .

A média de  $X_t$  em (6) é zero. Se a série tiver uma média diferente de zero, é necessário considerar um modelo AR(1) com constante ou *drift*, ou seja,

$$X_t = \delta + \phi_\mu X_{t-1} + a_t \Leftrightarrow \nabla X_t = \delta + \lambda_\mu X_{t-1} + a_t. \quad (9)$$



As hipóteses em teste mantêm-se e as distribuições assintóticas das estatísticas-teste sob  $H_0$  a partir dos estimadores de mínimos quadrados foram também determinadas por Dickey e Fuller (1979):

$$n \left( \hat{\phi}_\mu - 1 \right) \xrightarrow{D} \frac{\frac{1}{2} \{ [W(1)]^2 - 1 \} - W(1) \int_0^1 W(z) dz}{\int_0^1 [W(z)]^2 dz - \left[ \int_0^1 W(z) dz \right]^2} \quad (10)$$

$$T_\mu = \frac{\hat{\phi}_\mu - 1}{S_{\hat{\phi}_\mu}} \xrightarrow{D} \frac{\frac{1}{2} \{ [W(1)]^2 - 1 \} - W(1) \int_0^1 W(z) dz}{\left\{ \int_0^1 [W(z)]^2 dz - \left[ \int_0^1 W(z) dz \right]^2 \right\}^{1/2}}. \quad (11)$$

Por vezes, uma série contém também uma tendência determinística linear:

$$X_t = \delta_0 + \delta_1 t + \phi_\tau X_{t-1} + a_t \Leftrightarrow \nabla X_t = \delta_0 + \delta_1 t + \lambda_\tau X_{t-1} + a_t. \quad (12)$$

As hipóteses em teste continuam as mesmas e as distribuições assintóticas das estatísticas-teste sob  $H_0$ ,  $n \left( \hat{\phi}_\tau - 1 \right)$  e  $T_\tau = \left( \hat{\phi}_\tau - 1 \right) / S_{\hat{\phi}_\tau}$ , foram também obtidas por Dickey e Fuller (1979). Estes autores obtiveram ainda as distribuições assintóticas dos estimadores de  $\delta$  no modelo (9) e de  $\delta_0$  e de  $\delta_1$  em (12), mas não são apresentadas aqui porque estes estimadores não são necessários para os testes.

Uma vez que nos três modelos (6), (9) e (12) as distribuições das estatísticas-teste não são habituais, Dickey (1976) e Fuller (1996) calcularam os percentis mais utilizados das respetivas distribuições empíricas para alguns valores de  $n$  recorrendo à simulação de Monte Carlo, obtendo assim os pontos críticos necessários para o teste. Por sua vez, recorrendo a modelos de regressão de superfície de resposta e também com base em simulação, MacKinnon (1991, 1996, 2010) desenvolveu formas funcionais que permitem aproximar os pontos críticos do teste para qualquer valor de  $n$  e para cada uma das estatísticas  $T$ ,  $T_\mu$  e  $T_\tau$ . Estes pontos críticos estão muito próximos dos obtidos por Dickey (1976) e Fuller (1996). Sejam quais forem os percentis utilizados, a região crítica encontra-se na cauda esquerda da distribuição, ou seja, rejeita-se  $H_0$  quando o valor observado da estatística-teste for inferior ao ponto crítico correspondente.

Se  $H_0$  não for rejeitada para  $X_t$ , conclui-se que o modelo contém (pelo menos) uma raiz unitária, ou seja, que a série é não estacionária. Consequentemente, é necessário calcular a primeira diferença da série  $\nabla X_t$  e utilizá-la para testar  $H_0$  novamente. Se  $H_0$  não for rejeitada para  $\nabla X_t$ , calcula-se a segunda diferença  $\nabla^2 X_t$  e testa-se novamente. Este procedimento é repetido incrementando a ordem das diferenças até que  $H_0$  seja rejeitada, determinando-se assim a ordem de integração de  $X_t$ , ou seja, o valor de  $d$ .

Estes testes baseiam-se na representação AR(1) de  $X_t$ , que pode não ser a mais adequada pois, com efeito, uma série pode ser gerada por um modelo AR de maior ordem. Além disso, qualquer processo estacionário pode ser aproximado com uma dada precisão por um modelo autoregressivo de ordem suficientemente elevada, pelo que Dickey e Fuller (1979) propõem o recurso a um modelo AR( $p$ ) com  $p \geq 1$  para o teste. Suponha-se então que a série é gerada por um modelo AR( $p$ )

$$\Phi(B) X_t = a_t \quad (13)$$

onde  $a_t \sim N(0, \sigma_a^2)$  é um ruído branco,

$$\Phi(B) = 1 - \Phi_1 B - \dots - \Phi_p B^p = (1 - \varphi_1 B - \dots - \varphi_{p-1} B^{p-1}) (1 - B) = \varphi(B) (1 - B)$$

e as raízes do polinómio  $\varphi(B)$  estão fora do círculo unitário, ou seja,  $\Phi(B)$  inclui uma raiz unitária. Logo, o modelo (13) pode ser escrito como

$$\begin{aligned} \Phi(B) X_t &= \varphi(B) (1 - B) X_t = (1 - \varphi_1 B - \dots - \varphi_{p-1} B^{p-1}) (1 - B) X_t \\ &= (X_t - X_{t-1}) - \sum_{i=1}^{p-1} \varphi_i (X_{t-i} - X_{t-i-1}) = a_t \Leftrightarrow X_t = X_{t-1} + \sum_{i=1}^{p-1} \varphi_i \nabla X_{t-i} + a_t. \end{aligned}$$

Então, o teste de uma raiz unitária equivale a testar  $H_0 : \phi = 1$  no modelo

$$X_t = \phi X_{t-1} + \sum_{i=1}^{p-1} \varphi_i \nabla X_{t-i} + a_t. \quad (14)$$

Este modelo é estimado pelo método dos mínimos quadrados e as estatísticas-teste propostas sob  $H_0$  são  $n(\hat{\phi} - 1)\hat{\psi}(1)$  e  $T = (\hat{\phi} - 1) / S_{\hat{\phi}}$ , sendo  $\psi(B) = 1/\varphi(B)$ . Dickey e Fuller (1979) mostraram que as distribuições assintóticas sob  $H_0$  destas estatísticas são as mesmas do modelo simples AR(1), pelo que as tabelas de pontos críticos também. Além disso, o estimador de  $\phi$  é independente dos estimadores de  $\varphi_1, \dots, \varphi_{p-1}$  e a distribuição assintótica destes últimos é a que é habitualmente obtida para os estimadores de mínimos quadrados, uma vez que é igual à que resulta da regressão de  $\nabla X_t$  sobre  $\nabla X_{t-1}, \dots, \nabla X_{t-p+1}$ . Consequentemente, a distribuição  $t$ -Student habitual pode ser utilizada para testar hipóteses sobre estes parâmetros. A seleção da ordem do modelo em (14), isto é, a seleção de  $p$ , pode ser efetuada com base em algum critério de seleção de modelos como o AIC, o BIC ou o critério de Schwartz. Alternativamente, pode começar por se estimar o modelo simples AR(1) e ir aumentando a ordem do modelo até que o último parâmetro estimado  $\hat{\varphi}_p$  seja não significativo, sendo então  $(p - 1)$  a ordem selecionada. A estratégia contrária também é possível, ou seja, começar por um modelo de ordem elevada e ir diminuindo essa ordem enquanto o parâmetro estimado de maior ordem for não significativa.

A expressão (14) pode ser escrita na forma equivalente  $\nabla X_t = \lambda X_{t-1} + \sum_{i=1}^{p-1} \varphi_i \nabla X_{t-i} + a_t$ , sendo  $H_0 : \lambda = 0$ .

O modelo AR( $p$ ) pode também incluir uma constante ou uma tendência determinística e Dickey e Fuller (1979) mostraram que as distribuições assintóticas das estatísticas-teste são também as mesmas do modelo simples AR(1). Na aplicação prática do teste, utiliza-se as  $(n - p)$  observações  $X_{p+1}, \dots, X_n$ , pelo que a primeira estatística-teste apresentada é normalizada por  $(n - p)$  em vez de  $n$ , o que não afeta a respetiva distribuição assintótica.

Por se basear num modelo AR( $p$ ) geral, este teste é vulgarmente conhecido como teste de Dickey-Fuller aumentado ou ADF (*augmented Dickey-Fuller*) e é também válido para um qualquer modelo ARMA (Said e Dickey, 1985).

### 3.2 Testes de Phillips-Perron

Em alternativa, Phillips (1987) propôs uma outra estatística para o teste num modelo ARMA geral. O modelo a considerar é  $X_t = \phi X_{t-1} + \varepsilon_t$  onde  $E(\varepsilon_t) = 0$  e  $\sup E|\varepsilon_t|^k < \infty$  para  $k > 2$ . Note-se que este modelo é muito mais geral do que pode parecer, não estando restrito a um simples AR(1), porque  $\varepsilon_t$  é um processo qualquer, não tendo que ser um ruído branco (se for, cai-se no modelo AR(1) dado em (6)). Com efeito, conforme a expressão (13), um modelo ARIMA com uma raiz unitária é  $\varphi(B)(1 - B)X_t = \theta(B)a_t \iff X_t - X_{t-1} = \varphi^{-1}(B)\theta(B)a_t$ , o que, sendo  $\varepsilon_t = \varphi^{-1}(B)\theta(B)a_t$ , pode ser escrito como  $X_t = X_{t-1} + \varepsilon_t$  que é o modelo a considerar sob  $H_0 : \phi = 1$ .

Neste contexto, começa por se estimar  $\phi$  pelo método dos mínimos quadrados como no modelo (6) e as estatísticas-teste são

$$Z_\phi = n(\hat{\phi} - 1) - \frac{1}{2} \frac{S_\ell^2 - S_\varepsilon^2}{\frac{1}{n^2} \sum_{t=1}^n X_{t-1}^2} \quad (15)$$

$$Z_T = \frac{\left(\sum_{t=1}^n X_{t-1}^2\right)^{1/2} (\hat{\phi} - 1)}{S_\ell} - \frac{1}{2} \frac{S_\ell^2 - S_\varepsilon^2}{S_\ell \left(\frac{1}{n^2} \sum_{t=1}^n X_{t-1}^2\right)^{1/2}} \quad (16)$$

onde  $S_\varepsilon^2 = \frac{1}{n} \sum_{t=1}^n \varepsilon_t^2 = \frac{1}{n} \sum_{t=1}^n (X_t - X_{t-1})^2$ ,  $S_\ell^2 = \frac{1}{n} \sum_{t=1}^n \varepsilon_t^2 + \frac{2}{n} \sum_{r=1}^{\ell} \left(1 - \frac{r}{\ell+1}\right) \sum_{t=r+1}^n \varepsilon_t \varepsilon_{t-r}$  e  $\ell$  é

o valor de truncagem do desfasamento para a correção das estatísticas-teste, sendo um número inteiro selecionado de modo a que a autocovariância de  $\varepsilon_t$  esteja próxima de zero para desfasamentos superiores a  $\ell$ , com  $\ell = o(n^{1/4})$ . Na prática, a observação da autocorrelação de  $\varepsilon_t = X_t - X_{t-1}$  (sob  $H_0$ ) pode ajudar nesta escolha e, uma vez que, sob  $H_0$ , a autocorrelação estimada da primeira diferença de  $X_t$  decresce exponencialmente, o valor de  $\ell$  é habitualmente bastante baixo. Normalmente, utiliza-se  $\ell = \omega(n/100)^{1/4}$  (a parte inteira) com  $\omega = 4$  ou  $\omega = 12$ . Os ponderadores  $[1 - r/(\ell+1)]$  (janela de Bartlett) foram sugeridos por Newey e West (1987) para garantir que  $S_\ell^2 > 0$  em amostras finitas. Em vez de usar  $\varepsilon_t = X_t - X_{t-1}$  no cálculo de  $S_\ell^2$ , pode usar-se os resíduos de mínimos quadrados do modelo  $\hat{\varepsilon}_t = (X_t - \hat{\phi}X_{t-1})$ . Phillips (1987) mostrou que, sob  $H_0$ , as estatísticas (15) e (16) têm as mesmas distribuições assintóticas das estatísticas de Dickey-Fuller para o modelo (6) dadas em (7) e (8) respetivamente, pelo que as tabelas de pontos críticos são iguais.

Phillips e Perron (1988) generalizaram estes resultados ao modelo com constante  $X_t = \delta + \phi_\mu X_{t-1} + \varepsilon_t$ , onde  $\varepsilon_t$  verifica condições bastante gerais semelhantes ao modelo anterior. Sendo  $\hat{\phi}_\mu$  o estimador de mínimos quadrados de  $\phi_\mu$ , as estatísticas-teste sob  $H_0$  são

$$Z_{\phi_\mu} = n \left( \hat{\phi}_\mu - 1 \right) - \frac{1}{2} \frac{S_\ell^2 - S_\varepsilon^2}{\frac{1}{n^2} \sum_{t=1}^n (X_t - \bar{X})^2} \quad (17)$$

$$Z_{T_\mu} = \frac{\left( \sum_{t=1}^n (X_{t-1} - \bar{X}_{-1})^2 \right)^{1/2} (\hat{\phi} - 1)}{S_\ell} - \frac{1}{2} \frac{S_\ell^2 - S_\varepsilon^2}{S_\ell \left( \frac{1}{n^2} \sum_{t=1}^n (X_t - \bar{X})^2 \right)^{1/2}} \quad (18)$$

onde  $\bar{X}_{-1} = \frac{1}{n} \sum_{t=1}^n X_{t-1}$ . As distribuições assintóticas destas estatísticas são também iguais às das estatísticas de Dickey-Fuller para o modelo (9) dadas em (10) e (11). Phillips e Perron (1988) propuseram ainda estatísticas-teste para o modelo com constante e tendência determinística  $X_t = \delta_0 + \delta_1 t + \phi_\tau X_{t-1} + \varepsilon_t$  e mostraram que as suas distribuições assintóticas são novamente as obtidas por Dickey-Fuller para o modelo (12).

### 3.3 Teste KPSS

Kwiatkowski, Phillips, Schmidt e Shin (1992) propuseram um teste que ficou conhecido pelas iniciais dos nomes dos seus autores, ou seja, KPSS. Ao contrário dos anteriores, neste teste a hipótese de estacionaridade é a hipótese nula e não a alternativa. Partindo da representação em componentes de uma série temporal, admite-se que esta é a soma de uma tendência determinística, de um passeio aleatório e de uma variável residual estacionária, ou seja,  $X_t = \delta t + Z_t + \varepsilon_t$ , onde  $Z_t = Z_{t-1} + a_t$ , com  $a_t$  i.i.d. de média 0 e variância  $\sigma_a^2$ , e se admite que  $\varepsilon_t$  satisfaz as condições definidas no teste de Phillips. O valor inicial  $Z_0$  é considerado fixo e representa o papel de uma constante. A hipótese nula de estacionaridade é  $H_0 : \sigma_a^2 = 0$ , o que significa que  $Z_t$  é constante (e naturalmente que  $H_1 : \sigma_a^2 > 0$ ). Com efeito, note-se que, sendo  $E(a_t) = 0$ ,  $\sigma_a^2 = 0$  implica que  $a_t = 0 \forall t$  e, portanto,  $Z_t = Z_{t-1} = Z_0$ . Uma vez que  $\varepsilon_t$  é estacionária, sob a hipótese nula  $X_t$  é estacionária em tendência, ou seja, estacionária em torno de uma tendência linear. Se  $\delta = 0$ ,  $X_t$  é estacionária em torno de um nível.

Assim, para estar a estacionaridade em tendência, ajusta-se a regressão de mínimos quadrados  $X_t = \delta_0 + \delta_1 t + \varepsilon_t$  e sejam  $\hat{\varepsilon}_t$  ( $t = 1, \dots, n$ ) os resíduos de estimação. O processo das suas somas parciais é

$\Omega_t = \sum_{i=1}^t \hat{\epsilon}_i$  e a variância de longo prazo de  $\epsilon_t$  define-se como  $\sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{n} E(\Omega_n^2)$ , cujo estimador é  $S_\ell^2$  definido no teste de Phillips e calculado com  $\hat{\epsilon}_t$ . No caso de se pretender testar a estacionaridade em nível ( $\delta = 0$ ), os resíduos devem ser obtidos da regressão  $X_t = \tau + \xi_t$ .

A estatística-teste, válida quer para  $\delta \neq 0$ , quer para  $\delta = 0$ , é

$$KPSS = \frac{1}{n^2} \sum_{t=1}^n \frac{\Omega_t^2}{S_\ell^2}. \quad (19)$$

Se  $X_t$  for não estacionária, o numerador pode crescer ilimitadamente, pelo que a estatística deverá assumir valores elevados para amostras grandes. Por isso, a região crítica é à direita. Os percentis da distribuição (cauda direita) da estatística *KPSS* a serem usados como pontos críticos no teste foram obtidos por Kwiatkowski *et al.* (1992) através de simulação de Monte Carlo.

### 3.4 Relevância dos testes

Nos dois primeiros testes, a hipótese nula é a existência de uma raiz unitária (não estacionaridade), enquanto no terceiro é a contrária (estacionaridade). Por isso, se aqueles conduzirem à decisão de não rejeitar  $H_0$ , este último deverá rejeitar a sua hipótese nula e vice-versa. No entanto, isso nem sempre sucede, sendo possível obter conclusões contraditórias.

Devido à importância deste problema, muitos testes de raízes unitárias foram propostos na literatura, destacando-se dois (além dos acima descritos) pela sua utilização, especialmente em Economia: os testes de Schmidt e Phillips (1992) e de Elliott, Rothenberg e Stock (1996). Com efeito, a conclusão destes testes assume uma relevância muito especial porque tem implicações na análise da evolução temporal das séries económicas e, em particular, dos ciclos económicos. Segundo a visão tradicional desses ciclos, o PIB e os níveis de produção são estacionários em tendência e não em diferenças. Esta perspetiva significa que  $\phi \neq 1$  no modelo (14) ou  $\lambda \neq 0$  na sua formulação alternativa. Caso contrário ( $\phi = 1$  ou  $\lambda = 0$ ), a série tem uma raiz unitária e é estacionária em diferenças. Os testes de raízes unitárias permitem esclarecer esta controvérsia, na linha de Nelson e Plosser (1982), com as consequências daí resultantes e já anteriormente referidas.

## 4 Cointegração de séries temporais

A literatura macroeconómica tem sido dominada nos anos mais recentes pela conclusão fundamental de que as teorias de equilíbrio envolvendo variáveis não estacionárias exigem a existência de uma combinação dessas variáveis que seja estacionária. Com efeito, as séries não estacionárias podem sofrer fortes variações mas, no entanto, as forças económicas e de mercado tendem a manter muitas dessas séries juntas, formando relações de equilíbrio de longo prazo em que os desvios em relação ao equilíbrio são temporários. Exemplos desta situação são a teoria da função consumo (hipótese do rendimento permanente), as taxas de juro de curto e de longo prazo, a procura de moeda (dependendo do nível de preços, do rendimento real e da taxa de juro) ou a arbitragem no mercado de bens e a teoria das paridades de poder de compra.

### 4.1 Definição

Todos estes exemplos ilustram o conceito de cointegração de Engle e Granger (1987). Para  $t = 1, 2, \dots$ , seja  $\mathbf{X}_t = (X_{1,t}, X_{2,t}, \dots, X_{m,t})^T$  uma série temporal vetorial, ou seja, um vetor ( $m \times 1$ ) de séries

temporais ( $T$  significa transposto). Diz-se que o vetor  $\mathbf{X}_t$  é cointegrado de ordem  $(d, b)$ , escrevendo-se  $CI(d, b)$ , com  $0 < b \leq d$ , se cada elemento de  $\mathbf{X}_t$  for  $I(d)$  mas existir algum vetor  $(m \times 1)$  de constantes  $\beta \neq \mathbf{0}$  tal que a combinação linear  $\beta^T \mathbf{X}_t$  é  $I(d - b)$ . O vetor  $\beta^T$  é designado por vetor cointegrante e não é único, pois, para qualquer constante não nula  $c$ , se  $\beta^T$  for um vetor cointegrante,  $c\beta^T$  também é. Habitualmente, utiliza-se uma das séries para normalizar o vetor de cointegração, fixando-se o respetivo parâmetro em 1. O caso mais comum é  $d = b = 1$  e será esse que vamos considerar na análise.

Para compreender melhor o conceito de cointegração, admita-se que um conjunto de variáveis económicas se encontra em equilíbrio de longo prazo quando  $\beta^T \mathbf{X}_t = 0$ . O desequilíbrio, ou seja, o desvio em relação ao equilíbrio, é designado por erro de equilíbrio e define-se como  $U_t = \beta^T \mathbf{X}_t$ . Se  $U_t \neq 0$ , o sistema encontra-se fora da situação de equilíbrio e note-se que, para que esta situação tenha significado, o erro de equilíbrio tem que ser estacionário, pelo que o desvio tem uma natureza temporária. O conceito de cointegração abriu a possibilidade de testar a existência de relações de equilíbrio de longo prazo tal como são sugeridas em Economia. No entanto, o termo “equilíbrio” é frequentemente utilizado com diferente significado pela teoria económica e pela econometria. Para a primeira, refere-se a uma igualdade entre transações desejadas e reais, enquanto para a segunda representa uma relação de equilíbrio de longo prazo entre variáveis não estacionárias.

De uma forma geral, se a série multivariada  $\mathbf{X}_t$  contiver mais do que duas componentes ( $m \geq 3$ ), todas  $I(1)$ , podem existir até  $(m - 1)$  vetores cointegrantes  $\beta_1^T, \dots, \beta_r^T$  linearmente independentes, com  $r = 0, 1, \dots, m - 1$ , isto é, vetores tais que  $\beta^T \mathbf{X}_t = \mathbf{U}_t = (U_{1,t}, \dots, U_{r,t})^T$  é um vetor  $(r \times 1)$  estacionário, onde a matriz  $\beta^T = (\beta_1 \dots \beta_r)^T$ , de dimensão  $(r \times m)$ , agrupa os vetores cointegrantes e se designa por matriz cointegrante, tendo característica  $r$ . Assim, se, para qualquer outro vetor  $\mathbf{b}^T$ , de dimensão  $(1 \times m)$ , linearmente independente dos vetores cointegrantes,  $\mathbf{b}^T \mathbf{X}_t$  for não estacionário, diz-se que  $\mathbf{X}_t$  é cointegrada de ordem  $r$ . Os vetores  $\beta_1^T, \dots, \beta_r^T$  não são únicos e formam uma base do espaço dos vetores cointegrantes designado por espaço de cointegração. Note-se no entanto que, no caso bivariado ( $m = 2$ ), o vetor cointegrante, se existir, é único (linearmente independente), ou seja,  $r = 1$ .

## 4.2 Teorema de representação de Granger

Para séries temporais não estacionárias cointegradas, existem diversas representações particularmente relevantes estabelecidas por Engle e Granger (1987) no que ficou conhecido como teorema de representação de Granger. Assim, mantendo a notação anterior,  $\mathbf{X}_t$  é um vetor  $(m \times 1)$  cointegrado com  $d = 1$  e  $b = 1$ , existindo  $r$  vetores cointegrantes agrupados na matriz  $\beta^T$  de dimensão  $(r \times m)$  e característica  $r$ . As representações mais importantes do vetor  $\mathbf{X}_t$  são apresentadas em seguida.

### 4.2.1 Representação em médias móveis

Uma vez que cada componente de  $\mathbf{X}_t$  é  $I(1)$ ,  $(1 - B) \mathbf{X}_t = \nabla \mathbf{X}_t$  é estacionária, pelo que pode ser escrita na representação em médias móveis

$$(1 - B) \mathbf{X}_t = \sum_{i=0}^{\infty} \Psi_i \mathbf{a}_{t-i} = \sum_{i=0}^{\infty} \Psi_i B^i \mathbf{a}_t = \Psi(B) \mathbf{a}_t, \quad (20)$$

onde  $\mathbf{a}_t$  é um ruído branco vetorial, de média zero e matriz de variâncias e covariâncias  $\Sigma_a$ ,  $\Psi_0 = \mathbf{I}_m$ , matriz identidade  $(m \times m)$ , e as matrizes de parâmetros  $\Psi_i$  ( $i = 1, 2, \dots$ ) são absolutamente somáveis. Além disso, tem-se  $\beta^T \Psi(1) = \mathbf{0}$  e  $\text{car}[\Psi(B)] = m - r$ . Para mostrar que  $\beta^T \Psi(1) = \mathbf{0}$ , começa por se escrever a matriz  $\Psi(B)$  como  $\Psi(B) = \Psi(1) + \Psi^*(B)(1 - B)$ , onde  $\Psi^*(B) = \sum_{i=0}^{\infty} \Psi_i^* B^i$  e  $\Psi_i^* =$

$\sum_{j=i+1}^{\infty} \Psi_j$  ( $i = 0, 1, 2, \dots$ ) que é uma sequência de matrizes cuja soma é absolutamente convergente.

Então, a representação (20) pode ser escrita como

$$\mathbf{X}_t = \Psi(1)(1-B)^{-1}\mathbf{a}_t + \Psi^*(B)\mathbf{a}_t \iff \beta^T \mathbf{X}_t = \beta^T \Psi(1)(1-B)^{-1}\mathbf{a}_t + \beta^T \Psi^*(B)\mathbf{a}_t. \quad (21)$$

Sendo  $\Psi^*(B)\mathbf{a}_t$  estacionário, e uma vez que  $(1-B)^{-1}\mathbf{a}_t$  não o é, (20) implica que  $\beta^T \mathbf{X}_t$  é estacionário se e só se  $\beta^T \Psi(1) = \mathbf{0}$ . Esta condição implica, por um lado, que qualquer matriz (ou vetor) que a satisfaça é uma matriz (vetor) cointegrante e, por outro, que  $|\Psi(1)| = 0$ , o que significa que  $\Psi(1)$  é singular. Logo, sendo esta uma matriz ( $m \times m$ ), a sua característica é inferior a  $m$  e mostra-se que  $\text{car}[\Psi(1)] = m - r$ , onde  $r$  é o número de vetores cointegrantes. Além disso,  $|\Psi(1)| = 0$  significa ainda que  $\Psi(B)$  é não invertível (existência de raízes unitárias). Então, a representação (20) implica que não é possível inverter o operador de médias móveis e representar uma série cointegrada na forma autoregressiva em termos de  $(1-B)\mathbf{X}_t$ . A representação AR de uma série cointegrada tem que ser diretamente em termos de  $\mathbf{X}_t$ .

#### 4.2.2 Representação autoregressiva

A representação autoregressiva da série não estacionária  $\mathbf{X}_t$  é

$$\mathbf{X}_t = \Phi_1 \mathbf{X}_{t-1} + \dots + \Phi_p \mathbf{X}_{t-p} + \mathbf{a}_t \iff \Phi(B)\mathbf{X}_t = \mathbf{a}_t, \quad (22)$$

onde  $\Phi(B) = (\mathbf{I}_m - \Phi_1 B - \dots - \Phi_p B^p)$  é tal que  $|\Phi(B)| = 0$  contém algumas raízes unitárias. Por outro lado, a representação (20) pode ser escrita como  $\mathbf{X}_t = (1-B)^{-1}\Psi(B)\mathbf{a}_t$ , pelo que, multiplicando por  $\Phi(B)$  nesta expressão, obtém-se

$$\Phi(B)\mathbf{X}_t = (1-B)^{-1}\Phi(B)\Psi(B)\mathbf{a}_t. \quad (23)$$

Então, a comparação de (22) e (23) mostra que  $\Phi(B)\Psi(B) = (1-B)\mathbf{I}_m$ , o que implica que

$$\Phi(1)\Psi(1) = \mathbf{0} \quad (24)$$

onde  $\mathbf{0}$  é a matriz nula ( $m \times m$ ). Além disso,  $\text{car}[\Phi(1)] = r < m$ , pelo que esta matriz é singular. Na representação em médias móveis acima, foi mostrado que  $\beta^T \Psi(1) = \mathbf{0}$  e que qualquer matriz (ou vetor) que a satisfaça é uma matriz (vetor) cointegrante. Logo, a expressão (24) implica que  $\Phi(1)$  pertence ao espaço gerado pelas linhas de  $\beta^T$ , ou seja,

$$\Phi(1) = \mathbf{M}\beta^T \quad (25)$$

onde  $\mathbf{M}$  é uma matriz ( $m \times r$ ) de característica  $r$  tal que  $\Psi(1)\mathbf{M} = \mathbf{0}$ .

#### 4.2.3 Representação em correção de erro

A partir da representação AR( $p$ ),  $\mathbf{X}_t = \Phi_1 \mathbf{X}_{t-1} + \dots + \Phi_p \mathbf{X}_{t-p} + \mathbf{a}_t$ , subtraindo  $\mathbf{X}_{t-1}$  em ambos os membros e rearranjando a expressão, obtém-se a representação em modelo corretor de erro (MCE) ou VMCE( $p-1$ ):

$$\nabla \mathbf{X}_t = \Pi \mathbf{X}_{t-1} + \Phi_1^* \nabla \mathbf{X}_{t-1} + \dots + \Phi_{p-1}^* \nabla \mathbf{X}_{t-p+1} + \mathbf{a}_t \iff \Phi^*(B)\nabla \mathbf{X}_t = \Pi \mathbf{X}_{t-1} + \mathbf{a}_t \quad (26)$$

onde  $\Pi = -(\mathbf{I}_m - \Phi_1 - \dots - \Phi_p) = -\Phi(1)$ ,  $\Phi_i^* = -\sum_{j=i+1}^p \Phi_j$  para  $i = 1, \dots, p-1$  e  $\Phi^*(B) = (\mathbf{I}_m - \Phi_1^* B - \dots - \Phi_{p-1}^* B^{p-1})$ . Sendo  $\mathbf{X}_t$  cointegrada com  $d = b = 1$ ,  $\nabla \mathbf{X}_t, \nabla \mathbf{X}_{t-1}, \dots, \nabla \mathbf{X}_{t-p+1}$  são estacionárias, mas  $\mathbf{X}_{t-1}$  não é. Logo, a representação (26) implica que  $\Pi \mathbf{X}_{t-1}$  seja também estacionário, o que significa que tem que conter as relações de cointegração. As matrizes  $\Phi_i^*$  ( $i = 1, \dots, p-1$ ) são frequentemente designadas por parâmetros de curto prazo e  $\Pi$  por matriz cointegrante ou

de longo prazo, sendo  $\Pi X_{t-1}$  designada por vezes por componente de longo prazo do modelo. Se o modelo VAR( $p$ ) dado em (22) contiver raízes unitárias, ou seja, se  $|\mathbf{I}_m - \Phi_1 B - \dots - \Phi_p B^p| = 0$  para  $|B| = 1$ , a matriz  $\Pi$  é singular e, com efeito, uma vez que  $\Pi = -\Phi(1)$ , a característica desta matriz é  $r < m$ , o que significa que existem  $(m - r)$  raízes unitárias. Além disso, a expressão (25) implica que  $\Pi = \alpha \beta^T$  (com  $\alpha = -M$ ), o que prova que  $\Pi X_{t-1}$  contém as relações de cointegração, pois  $\beta^T$  é a matriz de cointegração. Em resumo, se  $\text{car}(\Pi) = 0$ , esta matriz é nula e (26) converte-se num modelo VAR habitual nas primeiras diferenças. É a situação de inexistência de cointegração. Se  $\text{car}(\Pi) = m$ ,  $X_t$  é estacionária, mas esta situação tem pouco interesse, uma vez que uma combinação linear de séries estacionárias é também estacionária. Se  $\text{car}(\Pi) = r$  com  $0 < r < m$ ,  $X_t$  é cointegrada com  $r$  vetores cointegrantes e existe a representação em modelo corretor de erro (26).

Então, este modelo mostra que, se uma série  $X_t$  for cointegrada, a sua diferença  $\nabla X_t$  não pode ser representada apenas em termos das diferenças passadas  $\nabla X_{t-j}$  com  $j \geq 1$ . A representação tem que incluir um termo de “correção de erro”  $\Pi X_{t-1} = \alpha \beta^T X_{t-1} = \alpha U_{t-1}$ . Se se considerar que a relação de  $\nabla X_t$  com os seus valores desfasados  $\nabla X_{t-j}$  ( $j \geq 1$ ) representa o equilíbrio de longo prazo, o termo  $\beta^T X_{t-1} = U_{t-1}$  pode ser encarado como um erro relativamente ao equilíbrio, ou seja, como um erro de desequilíbrio. Por sua vez, a matriz  $\alpha$  mede a correção desse erro, isto é, o ajustamento necessário na direção do equilíbrio para corrigir o erro e os seus elementos são os pesos (“loadings”) das relações de cointegração nas equações individuais do modelo, pelo que medem a velocidade do ajustamento em direção ao equilíbrio. Esse ajustamento é precisamente a variação da série registada no período corrente, ou seja,  $\nabla X_t$ , o que significa que se admite que as variações da série ocorrem para corrigir o desequilíbrio do sistema e na direção do equilíbrio de longo prazo.

#### 4.2.4 Estimação

Note-se em primeiro lugar que a utilização de variáveis integradas em modelos de regressão pode produzir resultados muito satisfatórios, mas falsos. Com efeito, a regressão pode ser significativa, com um coeficiente de determinação elevado, mesmo que as variáveis não estejam relacionadas. Este problema ficou conhecido como regressão espúria e pode ocorrer quando o termo residual da regressão é não estacionário (Granger e Newbold, 1974, 1986; Phillips, 1986). No caso estacionário, os estimadores de mínimos quadrados possuem as propriedades habituais, nomeadamente a normalidade assintótica. No caso  $I(1)$ , estes estimadores convergem a uma taxa de  $n$ , superior à habitual  $n^{1/2}$ , propriedade que foi designada por “super-consistência”. Além disso, as respetivas estatísticas  $t$  habituais não têm distribuição  $t$ -Student, pelo que os testes da significância estatística dos parâmetros estimados baseados nesta distribuição não se aplicam, o mesmo sucedendo com os testes do qui-quadrado e  $F$ -Snedcor usuais. Por exemplo, no modelo AR(1) dado em (6), se  $X_t$  for  $I(1)$  e, portanto,  $\phi = 1$ , o estimador de  $\phi$  tem uma distribuição assintótica que não é normal. Com efeito,  $n^{1/2}(\hat{\phi} - 1)$  converge em probabilidade para zero, enquanto  $n(\hat{\phi} - 1)$  tem a distribuição assintótica (7). Em conclusão, no caso não estacionário, um modelo de regressão deve ser estimado com as variáveis em diferenças.

Assim, as equações de um modelo VAR( $p$ ) estacionário podem ser estimadas separadamente pelo método dos mínimos quadrados e, se o pressuposto da normalidade se verificar, estes estimadores são idênticos aos de máxima verosimilhança. Os estimadores têm as propriedades assintóticas habituais e, em particular, têm distribuição normal. No caso não estacionário (variáveis integradas ou cointegradas), a normalidade assintótica mantém-se, mas a respetiva matriz de variâncias e covariâncias é singular, contrariamente ao que sucede no caso estacionário, o que invalida as distribuições  $t$ -Student, qui-quadrado e  $F$ -Snedcor, conforme foi referido. Apesar deste resultado geral, mesmo em modelos VAR com variáveis  $I(1)$ , a inferência ainda é possível em muitos casos. Se a hipótese nula em teste não restringir parâmetros de todas as matrizes do modelo VAR( $p$ ), as estatísticas habituais mantêm as suas distribuições assintóticas *standard* (Lütkepohl, 2004, 2005). Por exemplo, se  $p \geq 2$ , as estatísticas  $t$  mantêm a sua distribuição assintótica  $t$ -Student (que pode ser aproximada à normal reduzida), uma vez

que são estatísticas adequadas para testar a nulidade de um único coeficiente. Com efeito, elas testam uma hipótese nula que restringe apenas um coeficiente numa das matrizes de parâmetros, deixando as outras matrizes sem restrições. De qualquer modo,  $(1 - B) \mathbf{X}_t = \nabla \mathbf{X}_t$  é estacionária, o que permite a estimação de mínimos quadrados (ou de máxima verosimilhança, sob o pressuposto de normalidade) com as variáveis em diferenças, obtendo-se estimadores com as propriedades desejadas.

Para estimar o modelo VMCE dado em (26), começa por se estimar a matriz  $\Pi$  recorrendo a um método conhecido como análise de correlação canónica ou, de forma equivalente, a uma regressão de característica reduzida, uma vez que  $\text{car}(\Pi) = r < m$  (Johansen, 1988, 1991, 1995). Para obter este estimador, forma-se uma matriz envolvendo  $\nabla \mathbf{X}_t, \nabla \mathbf{X}_{t-1}, \dots, \nabla \mathbf{X}_{t-p+1}$  e  $\mathbf{X}_{t-1}$  ( $t = 1, \dots, n$ ) e calcula-se os seus valores próprios  $\lambda_1 \geq \dots \geq \lambda_m$  e os correspondentes vetores próprios  $\mathbf{b}_1, \dots, \mathbf{b}_m$ . Os estimadores dos vetores cointegrantes são obtidos escolhendo  $\hat{\beta}^T = (\hat{\beta}_1 \dots \hat{\beta}_r)^T = (\mathbf{b}_1 \dots \mathbf{b}_r)^T$ , o que permite em seguida estimar  $\alpha$  pelo método dos mínimos quadrados a partir de uma regressão envolvendo esta matriz e  $\beta$ . O estimador de  $\Pi$  é então  $\hat{\Pi} = \hat{\alpha} \hat{\beta}$ . Por fim, reescrevendo o modelo VMCE com  $\nabla \mathbf{X}_t - \hat{\Pi} \mathbf{X}_{t-1}$  no primeiro membro em (26), os parâmetros das matrizes  $\Phi_i^*$  ( $i = 1, \dots, p - 1$ ) podem ser estimados pelos mínimos quadrados em cada equação separadamente. Supondo que o ruído branco  $\mathbf{a}_t$  tem distribuição normal, o método da máxima verosimilhança conduz aos mesmos estimadores (foi esta a abordagem adotada originalmente por Johansen, 1988).

Em condições muito gerais,  $\hat{\Phi}_1^*, \dots, \hat{\Phi}_{p-1}^*$  e  $\hat{\Pi}$  são consistentes e assintoticamente normais, convergindo à taxa habitual de  $n^{1/2}$ . A matriz de variâncias e covariâncias da distribuição conjunta dos primeiros é regular, pelo que é válida a inferência habitual sobre os parâmetros de curto prazo  $\Phi_i^*$  ( $i = 1, \dots, p - 1$ ). Pelo contrário, a matriz de variâncias e covariâncias de  $\hat{\Pi}$ , que é uma matriz  $(m^2 \times m^2)$ , tem característica  $mr$ , pelo que é singular se  $r < m$ . Para resolver este problema, a normalização dos vetores próprios consegue tornar único o estimador da matriz cointegrante  $\hat{\beta}^T$  e  $\hat{\alpha}$  é ajustado em conformidade. No entanto, estas não são restrições econométricas de identificação. Portanto, apenas o espaço de cointegração é estimado consistentemente, mas não os parâmetros de cointegração. Para estimar as matrizes  $\alpha$  e  $\beta$  consistentemente, são necessárias restrições de identificação, sem as quais apenas o produto  $\Pi = \alpha \beta^T$  o pode ser. Se não estiverem disponíveis restrições adequadas (por exemplo, da teoria económica), é sempre possível proceder à normalização da forma  $\beta^T = (\mathbf{I}_r; \beta_{m-r}^T)$ , onde  $\mathbf{I}_r$  é a matriz identidade  $(r \times r)$  e  $\beta_{m-r}^T$  é uma matriz  $(r \times (m - r))$ . Para  $r = 1$ , esta restrição equivale a normalizar o coeficiente da primeira variável de  $\mathbf{X}_t$  de modo a ser igual a 1. Esta normalização pode ser feita em relação a qualquer variável e por isso exige algum cuidado na escolha da ordem das variáveis. Se possível, tal ordem deve resultar da teoria económica, devendo conduzir a relações de cointegração que tenham significado e interpretação económica quando se impõem restrições de normalização. Além disso, a normalização assegura a identificação dos parâmetros em  $\beta_{m-r}^T$ , tornando a inferência possível. Em termos gerais, Johansen (1995) mostrou que o estimador de  $\beta_{m-r}^T$  converge em distribuição a uma taxa de  $n$  ("super consistência") para uma mistura de normais, o que significa que a inferência é válida como para estimadores assintoticamente normais. Por exemplo, é possível definir estatísticas  $t$ -Student (assintoticamente) utilizando os desvios padrão estimados de  $\hat{\beta}_{m-r}^T$  dados pela raiz quadrada dos elementos da diagonal principal da matriz de variâncias e covariâncias estimada. Por sua vez, em condições gerais, o estimador de  $\alpha$  converge à taxa habitual de  $n^{1/2}$  para uma distribuição normal, o que significa que se comporta como os estimadores habituais num modelo com variáveis estacionárias.

### 4.3 Testes de cointegração

Antes de estimar um modelo VAR ou VMCE, é necessário testar se as séries estão cointegradas. Para este efeito, existem diversas abordagens e irão ser apresentadas as principais.

O primeiro passo para testar a cointegração consiste em testar a hipótese nula de existência de uma raiz unitária em cada série univariada componente do vetor  $\mathbf{X}_t$  recorrendo aos testes apresentados acima. Se



essa hipótese não for rejeitada para todas as séries, o passo seguinte consiste então em testar se elas estão cointegradas, isto é, testar se  $\beta^T \mathbf{X}_t$  é estacionário para alguma matriz (ou vetor)  $\beta^T$ .

Por vezes, a escolha da matriz ou do vetor cointegrante é baseada nalguma consideração teórica. Por exemplo, se  $\mathbf{X}_t = (X_{1,t}, X_{2,t})^T$ , onde  $X_{1,t}$  é o rendimento e  $X_{2,t}$  a despesa, pode-se pretender testar se o rendimento e a despesa estão numa relação de equilíbrio de longo prazo e, logo, se  $X_{1,t} - X_{2,t} = U_t$  é estacionário. Neste caso, o vetor cointegrante é  $\beta^T = (1, -1)^T$ . Se a matriz (ou vetor) cointegrante for conhecida, testar a existência de cointegração resume-se a testar a hipótese nula de  $U_t = \beta^T \mathbf{X}_t$  conter uma raiz unitária, para o que se pode utilizar os testes de raízes unitárias descritos acima. Se a hipótese nula for rejeitada, conclui-se que  $\mathbf{X}_t$  é cointegrada. Se a matriz cointegrante for desconhecida, foram propostos vários testes abordados em seguida.

### 4.3.1 Teste de Engle-Granger

Engle e Granger (1987) propuseram um teste onde se considera a regressão de cointegração

$$X_{1,t} = \eta_2 X_{2,t} + \dots + \eta_m X_{m,t} + U_t, \quad (27)$$

normalizando em ordem ao coeficiente de  $X_{1,t}$ , que se supõe ser diferente de zero (mas note-se que se pode normalizar em ordem ao coeficiente de qualquer outra série). Se  $U_t$  for  $I(1)$  e contiver uma raiz unitária,  $\mathbf{X}_t$  não é cointegrada. Pelo contrário, se  $U_t$  for  $I(0)$  (estacionária), então  $\mathbf{X}_t$  é cointegrada com o vetor cointegrante  $\beta^T = (1, -\eta_2, \dots, -\eta_m)^T$ . A equação de regressão (27) pode incluir também um termo constante ou uma constante e uma tendência determinística.

Portanto, testar a cointegração entre as variáveis de  $\mathbf{X}_t$  consiste em testar se  $U_t$  é  $I(1)$ , o que significa testar a hipótese nula de não cointegração. Para este efeito, pode recorrer-se aos testes de raízes unitárias, sendo os mais comuns os testes de Dickey-Fuller a partir dos modelos (6) ou (14) onde se substitui  $X_{t-j}$  por  $U_{t-j}$  ( $j = 0, 1, \dots, p-1$ ). No entanto,  $U_t$  é desconhecido, pelo que é necessário estimar a regressão (27) pelo método dos mínimos quadrados, obter os respetivos resíduos  $\hat{U}_t$  e utilizá-los na realização do teste. Assim, o modelo a usar para o teste ADF é

$$\nabla \hat{U}_t = \lambda \hat{U}_{t-1} + \sum_{i=1}^{p-1} \varphi_i \nabla \hat{U}_{t-i} + a_t \quad (28)$$

e as hipóteses são  $H_0 : \lambda = 0$  contra  $H_1 : \lambda < 0$  (se  $p = 1$ , obtém-se o modelo AR(1) simples do teste de Dickey-Fuller). O recurso a  $\hat{U}_t$  em vez do verdadeiro termo de erro  $U_t$  torna desadequado o recurso às tabelas de percentis de Dickey-Fuller habituais, pelo que Engle e Granger (1987) obtiveram os pontos críticos para o teste com  $T = \hat{\lambda} / S_{\hat{\lambda}}$  através de simulação com o modelo (28), distinguindo os casos  $p = 1$  e  $p > 1$ . Além disso, Engle e Yoo (1987) obtiveram tabelas de pontos críticos mais completas e, por sua vez, a fórmula de MacKinnon (1991, 1996, 2010) também permite calcular pontos críticos para este teste. Os pontos críticos dependem do número de variáveis na equação de regressão (27) e de esta incluir uma constante ou uma constante e uma tendência determinística. A hipótese nula é rejeitada quando o valor observado de  $T$  é inferior ao ponto crítico, concluindo-se que existe cointegração entre as variáveis de  $\mathbf{X}_t$ . À semelhança dos testes de raízes unitárias, o modelo (28) pode incluir ainda uma constante ou uma constante e uma tendência determinística e, se algum destes termos for usado na regressão (27), não deve ser incluído em (28). Se a hipótese nula de não cointegração for rejeitada, pode-se em seguida definir e estimar o modelo VMCE, tendo este método ficado conhecido como o procedimento em dois passos de Engle e Granger.

### 4.3.2 Teste de Phillips-Ouliaris

Phillips e Ouliaris (1990) propuseram usar as estatísticas dos testes de raízes unitárias de Phillips-Perron dadas em (15) e (16) com os resíduos  $\hat{U}_t$ , ou seja, calcular estas estatísticas a partir do modelo  $\hat{U}_t =$

$\phi\hat{U}_{t-1} + \varepsilon_t$ , que pode também incluir uma constante ou uma constante e uma tendência determinística. As respectivas distribuições assintóticas dependem do número de variáveis na regressão de cointegração (27) e são iguais às do teste de cointegração de Dickey-Fuller para o modelo AR(1), ou seja, quando  $\varepsilon_t$  é um ruído branco independente. Phillips e Ouliaris (1990) e Haug (1992) obtiveram pontos críticos para este teste (e logo para o teste anterior) através de simulação. Portanto, estes dois testes são baseados em resíduos (de mínimos quadrados) e têm o grande inconveniente de apenas conseguirem estimar um vetor cointegrante, saído da regressão de cointegração, mesmo que exista mais do que um.

### 4.3.3 Teste de Johansen

Foi referido anteriormente que, num vetor  $\mathbf{X}_t$  não estacionário com  $m$  séries, poderão existir até  $(m - 1)$  vetores cointegrantes linearmente independentes. Uma vez que é possível estimar o modelo VMCE (26) e, em particular, os vetores cointegrantes pela regressão de característica reduzida ou, de forma equivalente, supondo normalidade, pela máxima verosimilhança conforme descrito acima, Johansen (1988, 1991) e Johansen e Juselius (1990) propuseram testes de razão de verosimilhanças que permitem identificar o número de vetores cointegrantes e assim testar a existência de cointegração. Para este efeito, designe-se por  $H_r$  a hipótese da existência de  $r$  vetores cointegrantes, sob a qual  $\mathbf{\Pi} = \alpha\beta^T$  ou, de forma equivalente,  $\text{car}(\mathbf{\Pi}) = r$ . Por um lado, maximiza-se a função de verosimilhança sob esta restrição, ou seja, sob  $H_r$  e, por outro, maximiza-se a mesma função sem restrições, construindo-se assim a razão de verosimilhanças.

O número de vetores cointegrantes distintos pode ser obtido através da verificação da significância dos valores próprios estimados de  $\mathbf{\Pi}$ , pois a característica de uma matriz é igual ao número dos seus valores próprios não nulos (em particular, se não existir cointegração, a característica de  $\mathbf{\Pi}$  é zero e todos os valores próprios são nulos). Assim, sendo  $\lambda_1 \geq \dots \geq \lambda_r \geq \lambda_{r+1} \geq \dots \geq \lambda_m$  os valores próprios ordenados de  $\mathbf{\Pi}$ , são os  $r$  maiores valores que têm associados os  $r$  vetores cointegrantes e  $\lambda_{r+1} \geq \dots \geq \lambda_m$  deverão ser zero para as combinações não cointegrantes. Na prática,  $\mathbf{\Pi}$  e os seus valores próprios têm que ser estimados e testa-se a hipótese nula  $H_r$  da existência de, no máximo,  $r$  vetores cointegrantes ( $0 \leq r < m$ ) e, conseqüentemente,  $(m - r)$  raízes unitárias, contra a hipótese alternativa  $H_m$  da existência de  $m$  vetores cointegrantes (estacionaridade). Este teste é designado por teste do traço (recorde-se que se trata de um teste de razão de verosimilhanças) e a respetiva estatística é

$$LR_{\text{tr}}(r|m) = -n \sum_{i=r+1}^m \log(1 - \hat{\lambda}_i) \quad (29)$$

onde  $\log$  representa o logaritmo natural e  $\hat{\lambda}_i$  ( $i = r + 1, \dots, m$ ) são os  $(m - r)$  menores valores próprios estimados. A distribuição assintótica desta estatística sob  $H_r$  depende apenas de  $(m - r)$  e foi obtida por Johansen (1988, 1991) e por Reinsel e Ahn (1992), que obtiveram também os respetivos pontos críticos através de simulação de Monte Carlo. Note-se que, se os valores próprios forem nulos, a estatística também é zero e, quanto mais afastados de zero eles estiverem, maior é  $\log(1 - \hat{\lambda}_i)$  em valor absoluto, sendo negativo, pelo que maior é  $LR_{\text{tr}}(r|m)$ . Logo, rejeita-se  $H_r$  quando (29) for superior ao ponto crítico adequado. A estatística  $LR_{\text{tr}}(r|m)$  mede o custo em termos de verosimilhança da omissão de  $(m - r)$  combinações lineares dos elementos de  $\mathbf{X}_{t-1}$ . A estratégia do teste deve seguir a sequência  $LR_{\text{tr}}(0|m)$ ,  $LR_{\text{tr}}(1|m)$ ,  $\dots$ ,  $LR_{\text{tr}}(m - 1|m)$ , isto é, começa por se testar  $H_0$  contra  $H_m$  e, se  $LR_{\text{tr}}(0|m)$  for não significativa, o teste pára e não se rejeita a hipótese de que  $\text{car}(\mathbf{\Pi}) = 0$ , ou seja, de que não existe cointegração; caso contrário, conclui-se que existe pelo menos um vetor cointegrante e prossegue-se com o teste de  $H_1$  contra  $H_m$ . Se  $LR_{\text{tr}}(1|m)$  for não significativa, o teste pára e não se rejeita a hipótese de existir um vetor cointegrante ( $r = 1$ ); caso contrário, conclui-se que existem pelo menos dois vetores cointegrantes e continua-se para testar  $H_2$  contra  $H_m$  e assim sucessivamente. Portanto, o teste pára quando  $LR_{\text{tr}}(r|m)$  for não significativa, concluindo-se que existem  $r$  vetores cointegrantes.

Johansen e Juselius também desenvolveram o teste da hipótese nula de existência de  $r$  contra a alternativa de  $(r + 1)$  vetores cointegrantes que se designa por teste do máximo valor próprio e cuja estatística é

$$LR_{\max}(r|r+1) = -n \log(1 - \hat{\lambda}_{r+1}). \quad (30)$$

Os pontos críticos da distribuição assintótica desta estatística dependem apenas de  $(m - r)$ , tal como no teste do traço, e foram obtidos por Johansen e Juselius (1990) através de simulação, rejeitando-se a hipótese nula quando  $LR_{\max}(r|r+1)$  for superior ao ponto crítico adequado. A estratégia consiste em começar por testar a hipótese nula  $H_0 : r = 0$  contra a alternativa  $H_1 : r = 1$ . Se  $LR_{\max}(0|1)$  for não significativa, não se rejeita  $H_0$  e o teste pára. Caso contrário, conclui-se que existe pelo menos um vetor cointegrante e testa-se  $H_1 : r = 1$  contra  $H_2 : r = 2$ . Se  $LR_{\max}(1|2)$  for não significativa, não se rejeita  $H_1$  e o teste pára, tomando-se  $r = 1$ . Caso contrário, passa-se ao teste de  $H_2 : r = 2$  contra  $H_3 : r = 3$  e assim sucessivamente.

O modelo (26) pode incluir também um vetor  $(m \times 1)$  de constantes ou uma tendência determinística (cujo parâmetro é um vetor) na componente de curto prazo (isto é, na equação) ou no termo corretor de erro. A distribuição assintótica destas estatísticas e, em consequência, os respetivos pontos críticos dependem da especificação adotada, pelo que Johansen e Juselius (1990), Osterwald-Lenum (1992) e MacKinnon, Haug e Michelis (1999) obtiveram tabelas para as diferentes formas do modelo (os dois últimos são os mais utilizados).

#### 4.4 Causalidade de Granger

É muitas vezes necessário em Economia definir um conceito de “causalidade” que permita identificar o tipo de relação entre duas ou mais variáveis. Por exemplo, pode ser útil saber se existe influência do rendimento sobre o consumo. Para este efeito, é fundamental dispor de uma definição de causalidade que seja operacional e possa ser testada. Assim, Granger (1969) propôs um conceito de causalidade que ficou conhecido como causalidade de Granger e cuja ideia base é a de que a causa não pode ocorrer depois do efeito, ou seja, o futuro não pode causar o presente. Consequentemente, diz-se que uma variável  $X$  causa outra variável  $Y$  no sentido de Granger se os valores passados de  $X$  contribuírem para melhorar as previsões do valor corrente de  $Y$ , tudo o resto constante. Se se incluir o valor corrente de  $X$ , além dos seus valores passados, então diz-se que  $X$  causa instantaneamente  $Y$  no sentido de Granger.

Mais rigorosamente, seja  $\mathcal{I}_t$  o conjunto de toda a informação passada e presente existente no momento  $t$ ,  $\mathcal{I}_{X,t}$  o conjunto de toda a informação passada e presente existente sobre a variável  $X$  no momento  $t$ , isto é,  $\mathcal{I}_{X,t} = \{X_t, X_{t-1}, \dots\}$ ,  $Y_t$  o valor corrente da variável  $Y$  ( $Y_t \in \mathcal{I}_t$ ) e  $\hat{Y}_t$  um previsor não enviesado de  $Y_t$ . Então, diz-se que  $X$  causa  $Y$  no sentido de Granger se  $EQM(\hat{Y}_t | \mathcal{I}_{t-1}) < EQM(\hat{Y}_t | \mathcal{I}_{t-1} \setminus \mathcal{I}_{X,t-1})$ , onde  $EQM(\hat{Y}_t) = E(Y_t - \hat{Y}_t)^2$  designa o erro quadrático médio da previsão. Diz-se também que  $X$  causa instantaneamente  $Y$  no sentido de Granger se  $EQM(\hat{Y}_t | \mathcal{I}_t \setminus Y_t) < EQM(\hat{Y}_t | \mathcal{I}_t \setminus \mathcal{I}_{X,t}, Y_t)$ .

Na prática, é necessário substituir o conjunto de toda a informação pelo conjunto de informação “relevante”. Este último pode ser constituído pela informação incluída num modelo VAR, admitindo estacionaridade, pelo que testar se  $X$  causa  $Y$  equivale a saber se  $X$  deve ou não ser incluída na equação de  $Y$ , o que pode ser feito através do teste da nulidade dos parâmetros de  $X$  nessa equação. Assim, para ser mais fácil a compreensão deste conceito, suponhamos uma série temporal bivariada estacionária

$\mathbf{X}_t = (X_{1,t}, X_{2,t})^T$  que segue um modelo VAR( $p$ ) cuja equação de  $X_{2,t}$  é  $X_{2,t} = \delta_2 + \sum_{i=1}^p \phi_{21,i} X_{1,t-i} + \sum_{i=1}^p \phi_{22,i} X_{2,t-i} + a_{2,t}$ . Então,  $X_{1,t}$  não causa  $X_{2,t}$  no sentido de Granger se e só se  $\phi_{21,1} = \dots = \phi_{21,p} = 0$ ,

pelo que testar a causalidade de Granger pode ser feito recorrendo ao teste  $F$  habitual da hipótese  $H_0 : \phi_{21,1} = \dots = \phi_{21,p} = 0$ , em que a estatística-teste é

$$F = \left[ (\hat{\mathbf{a}}_{2,R}^T \hat{\mathbf{a}}_{2,R} - \hat{\mathbf{a}}_{2,L}^T \hat{\mathbf{a}}_{2,L}) / p \right] / \left[ \hat{\mathbf{a}}_{2,L}^T \hat{\mathbf{a}}_{2,L} / (n - (2p + 1)) \right],$$

onde  $\hat{\mathbf{a}}_{2,L}$  e  $\hat{\mathbf{a}}_{2,R}$  representam respetivamente o vetor dos resíduos de estimação da equação de  $X_{2,t}$  acima sem qualquer restrição (modelo livre) e sob  $H_0$  (modelo restrito). Se  $X_{1,t}$  causar  $X_{2,t}$  e  $X_{2,t}$  causar  $X_{1,t}$ , então  $(X_{1,t}, X_{2,t})$  diz-se um sistema de “feedback”.

A definição de causalidade de Granger pode ser facilmente generalizada ao caso de várias variáveis. Considere-se uma série temporal multivariada estacionária  $\mathbf{X}_t = (X_{1,t}, \dots, X_{m,t})^T$  que segue um modelo VAR( $p$ ). Diz-se que  $X_{j,t}$  não causa  $X_{i,t}$  ( $i, j = 1, \dots, m; i \neq j$ ) no sentido de Granger se os parâmetros dos valores desfasados da primeira  $X_{j,t-k}$  ( $k = 1, \dots, p$ ) na equação da segunda forem todos nulos, ou seja, se  $\phi_{ij,1} = \dots = \phi_{ij,p} = 0$ .

Conforme foi referido, esta análise é válida para séries estacionárias. No caso não estacionário, supondo primeiro que as variáveis não estão cointegradas, a causalidade de Granger pode ser testada da mesma maneira com as variáveis em diferenças, ou seja, a partir do modelo VAR( $p$ ) do vetor  $\nabla^d \mathbf{X}_t$ . Se existir cointegração, a causalidade já não pode ser testada a partir de um modelo VAR, pois  $\nabla^d \mathbf{X}_t$  não admite tal representação, mas sim a de VMCE. Neste último, o termo corretor de erro será significativo em pelo menos uma das equações envolvidas, o que significa que existe causalidade de Granger, ou seja, a cointegração implica causalidade em pelo menos uma das direções. Para melhor compreender este aspeto, considere-se novamente o exemplo de duas séries  $X_{1,t}$  e  $X_{2,t}$ , mas agora suponha-se que são ambas  $I(1)$  e cointegradas, isto é, o vetor  $\mathbf{X}_t = (X_{1,t}, X_{2,t})^T$  é CI(1, 1). Então, a equação de  $X_{2,t}$  no

VMCE bivariado é  $\nabla X_{2,t} = \delta_2 + \alpha_2 \hat{U}_{t-1} + \sum_{i=1}^{p-1} \phi_{21,i}^* \nabla X_{1,t-i} + \sum_{i=1}^{p-1} \phi_{22,i}^* \nabla X_{2,t-i} + a_{2,t}$ , onde  $\hat{U}_{t-1} = \beta^T \mathbf{X}_{t-1}$  e  $\alpha_2 \hat{U}_{t-1}$  é o termo corretor de erro. O teste da hipótese de que  $X_{1,t}$  não causa  $X_{2,t}$  no sentido de Granger consiste em testar  $H_0 : \phi_{21,1}^* = \dots = \phi_{21,p-1}^* = \alpha_2 = 0$ . Uma vez que existe cointegração, pelo menos um dos parâmetros  $\alpha_1$  (o homólogo de  $\alpha_2$  na equação de  $\nabla X_{1,t}$ ) ou  $\alpha_2$  será diferente de zero, assegurando a causalidade em pelo menos uma das direções. Em conclusão, quando existe cointegração, a causalidade de Granger pode verificar-se através dos parâmetros de curto prazo ou através da relação de longo prazo, ou seja, do termo corretor de erro.

## 5 Modelos estruturais e análise impulso-resposta

A utilização de modelos VAR ou VMCE (sem restrições) é um instrumento importante em termos de previsão. No entanto, à exceção da definição das variáveis a incluir no modelo, estes estão desprovidos de qualquer conteúdo económico, sendo frequentemente aplicados de forma mecânica. Assim, de forma a auxiliar na definição e avaliação de políticas económicas, os economistas utilizam geralmente a teoria económica na construção de modelos VAR ou VMCE estruturais (Breitung *et al.*, 2004). O objetivo passa por melhorar a precisão da previsão e extrair informação sobre os efeitos de um determinado choque, o que não é possível num modelo sem restrições. Para tal, é necessário impor restrições, fundamentadas teoricamente, a um determinado número de parâmetros do modelo dependendo da dimensão do sistema que o descreve. Diferentes métodos de imposição de restrições irão ser apresentados em seguida, destacando-se o modelo AB e o modelo de restrições de longo prazo utilizado por exemplo no esquema de identificação de Blanchard e Quah (1989).

### 5.1 Modelos estruturais

Os modelos VAR e VMCE apresentados anteriormente não incluem explicitamente relações contemporâneas entre as variáveis endógenas que são os elementos de  $\mathbf{X}_t$  e por isso se diz que são modelos na

forma reduzida. Na prática, é muitas vezes necessário modelar também relações contemporâneas, pelo que é útil considerar a forma estrutural de um VAR( $p$ ),

$$\mathbf{A}\mathbf{X}_t = \mathbf{\Gamma}_1\mathbf{X}_{t-1} + \dots + \mathbf{\Gamma}_p\mathbf{X}_{t-p} + \mathbf{V}_t \quad (31)$$

em que  $\mathbf{A}$  é uma matriz ( $m \times m$ ) regular que contém as relações contemporâneas entre as variáveis em  $\mathbf{X}_t$ ,  $\mathbf{\Gamma}_i$  ( $i = 1, \dots, p$ ) são matrizes ( $m \times m$ ) de parâmetros e  $\mathbf{V}_t$  é um ruído branco vetorial ( $m \times 1$ ) com média zero e matriz de variâncias e covariâncias  $\Sigma_V$ . Sendo  $\mathbf{A}$  regular, a forma reduzida deste modelo é dada pela expressão (22), sendo obtida pré-multiplicando (31) por  $\mathbf{A}^{-1}$ , com  $\mathbf{\Phi}_i = \mathbf{A}^{-1}\mathbf{\Gamma}_i$  ( $i = 1, \dots, p$ ) e  $\mathbf{a}_t = \mathbf{A}^{-1}\mathbf{V}_t$ .

Por sua vez, a representação em modelo corretor de erro estrutural (SVMCE) é

$$\mathbf{A}\nabla\mathbf{X}_t = \mathbf{\Pi}^*\mathbf{X}_{t-1} + \mathbf{\Gamma}_1^*\nabla\mathbf{X}_{t-1} + \dots + \mathbf{\Gamma}_{p-1}^*\nabla\mathbf{X}_{t-p+1} + \mathbf{V}_t, \quad (32)$$

onde  $\mathbf{\Pi}^*$  e  $\mathbf{\Gamma}_i^*$  ( $i = 1, \dots, p-1$ ) são matrizes ( $m \times m$ ) de parâmetros. A forma reduzida deste modelo, dada por (26), é obtida pré-multiplicando (32) por  $\mathbf{A}^{-1}$ , sendo  $\mathbf{\Pi} = \mathbf{A}^{-1}\mathbf{\Pi}^*$  e  $\mathbf{\Phi}_i^* = \mathbf{A}^{-1}\mathbf{\Gamma}_i^*$  ( $i = 1, \dots, p-1$ ).

Admite-se geralmente que  $\mathbf{V}_t$  resulta da ocorrência de choques estruturais (exógenos) não observáveis que afetam a economia, como um choque petrolífero, cambial ou monetário. Esses choques, representados por  $\epsilon_t$ , são imprevisíveis a partir do passado do processo e são o input de um sistema dinâmico linear que gera a série temporal multivariada  $\mathbf{X}_t$ . Por isso, são um ruído branco e estão relacionados com  $\mathbf{V}_t$ , considerando-se habitualmente que este é formado pela combinação linear de  $\epsilon_t$ , isto é,  $\mathbf{V}_t = \mathbf{B}\epsilon_t$ , onde  $\mathbf{B}$  é uma matriz ( $m \times m$ ) de parâmetros e  $\epsilon_t$  tem média nula. Por conseguinte, os modelos (31) e (32) podem ser reescritos substituindo  $\mathbf{V}_t$  por  $\mathbf{B}\epsilon_t$  e o ruído branco da forma reduzida destes modelos pode também ser expresso em função dos choques estruturais, uma vez que  $\mathbf{a}_t = \mathbf{A}^{-1}\mathbf{V}_t = \mathbf{A}^{-1}\mathbf{B}\epsilon_t$ .

A definição dos modelos (31) e (32) exige a identificação dos parâmetros da forma estrutural. Refira-se em primeiro lugar que é consensual supor que os choques estruturais são não correlacionados (ortogonais). Esta suposição é necessária para se poder medir o impacto dinâmico de um choque isolado, uma vez que, se os choques fossem correlacionados, seria necessário ter em conta as relações entre eles. Em consequência, a matriz de variâncias e covariâncias de  $\epsilon_t$  é  $\mathbf{I}_m$ , a matriz identidade ( $m \times m$ ). Então, ficam por identificar os parâmetros das matrizes  $\mathbf{A}$  e  $\mathbf{B}$ , sendo necessário impor restrições sobre eles. É possível verificar a identificação de um modelo SVAR ou SVMCE recorrendo a uma condição de ordem semelhante à utilizada num sistema de equações simultâneas. Sendo  $\mathbf{a}_t = \mathbf{A}^{-1}\mathbf{B}\epsilon_t$ , a sua matriz de variâncias e covariâncias é  $\Sigma_a = \mathbf{A}^{-1}\mathbf{B}\mathbf{B}^T(\mathbf{A}^{-1})^T$ , uma vez que a matriz de variâncias e covariâncias de  $\epsilon_t$  é  $\mathbf{I}_m$ . A matriz  $\Sigma_a$  tem  $[m(m+1)/2]$  elementos não redundantes, pelo que a sua estimação fornece informação sobre  $[m(m+1)/2]$  parâmetros da forma estrutural, isto é, permite a identificação destes. Mas, como  $\mathbf{A}$  e  $\mathbf{B}$  são matrizes ( $m \times m$ ), contêm um total de  $2m^2$  parâmetros, o que implica que é ainda necessário impor (pelo menos)  $2m^2 - [m(m+1)/2] = m^2 - [m(m-1)/2]$  restrições adicionais, livremente distribuídas entre estas duas matrizes, e note-se que não é suficiente impor restrições em apenas uma delas, permanecendo em falta a imposição de  $[m(m-1)/2]$  restrições. Os diferentes modelos de imposição de restrições sobre os parâmetros distinguem-se pela forma como estas são distribuídas e podem ser classificados da forma seguinte.

- (i)  $\mathbf{B} = \mathbf{I}_m$ : uma opção pouco utilizada em termos práticos é considerar  $\mathbf{B} = \mathbf{I}_m$  e impor as restantes  $[m(m-1)/2]$  restrições em  $\mathbf{A}$ . Este modelo supõe que os diferentes choques estruturais que afetam a economia são linearmente independentes e portanto  $\mathbf{A}\mathbf{a}_t = \epsilon_t$ , o que geralmente não é considerado um bom ponto de partida teórico.
- (ii)  $\mathbf{A} = \mathbf{I}_m$ : é mais frequente considerar  $\mathbf{A} = \mathbf{I}_m$  e impor as restantes restrições em  $\mathbf{B}$  recorrendo à decomposição de Choleski. Esta decomposição consiste em fatorizar uma matriz simétrica definida positiva  $\mathbf{C}$  na forma  $\mathbf{C} = \mathbf{D}\mathbf{D}^T$ , onde  $\mathbf{D}$  é uma matriz triangular inferior com elementos

da diagonal principal estritamente positivos. Então, procede-se à decomposição de Choleski da matriz  $\Sigma_a$  (simétrica definida positiva) como  $\Sigma_a = \mathbf{B}\mathbf{B}^T$ , obtendo-se assim a matriz  $\mathbf{B}$  triangular inferior, o que assegura as  $[m(m-1)/2]$  restrições exigidas. Daqui resulta que os choques estruturais ortogonalizados são  $\epsilon_t = \mathbf{B}^{-1}\mathbf{a}_t$ , ou seja,  $\mathbf{a}_t = \mathbf{B}\epsilon_t$ , tornando possível a existência de dependência linear dos choques estruturais  $\epsilon_t$ . Este tipo de restrições também pode ser obtido a partir da definição de um “esquema temporal” para os choques em que se supõe que estes podem afetar um subconjunto de variáveis diretamente no período corrente, enquanto outro subconjunto de variáveis só é afetado com algum desfasamento. Um exemplo deste esquema é a identificação triangular (ou recursiva) de Sims (1980) em que os choques entram nas equações sucessivamente de modo a que o choque da segunda equação não afeta a variável explicada pela primeira equação no período corrente, o choque da terceira equação não afeta as variáveis explicadas pelas duas primeiras equações no mesmo período e assim sucessivamente. Este esquema triangular de Sims equivale a restringir  $\mathbf{B}$  a ser uma matriz triangular inferior, de forma semelhante ao resultado obtido a partir da decomposição de Choleski de  $\Sigma_a$ . Contudo, a alteração da ordem das equações que compõem o sistema dinâmico, ou seja, da ordem pela qual as variáveis estão dispostas no vetor  $\mathbf{X}_t$ , influencia a forma como estes choques se relacionam entre si e como afetam o sistema, gerando assim diferentes resultados. Esta elevada sensibilidade à ordem das equações do sistema dinâmico representa a principal desvantagem desta alternativa.

- (iii) Modelo AB: o denominado modelo AB combina restrições sobre  $\mathbf{A}$  e  $\mathbf{B}$ , de que resulta  $\mathbf{A}\mathbf{a}_t = \mathbf{B}\epsilon_t$ .
- (iv) Informação *a priori*: pode existir informação *a priori* a partir da teoria económica sobre os efeitos de longo prazo de alguns choques. Esses efeitos são medidos através da resposta das variáveis do sistema a tais choques. Um exemplo de um esquema de identificação deste tipo é o utilizado por Blanchard e Quah (1989). Estes autores admitem a existência de dois tipos de choques não correlacionados que afetam o produto nacional bruto (PNB) e que são choques sobre a oferta agregada e sobre a procura agregada. Estes últimos produzem efeitos de curto prazo devido à existência de rigidez nominal, ou seja, devido à existência de fatores como os contratos celebrados e os custos de obtenção da informação que impedem os preços de se ajustarem imediatamente para o respetivo valor de equilíbrio o que, por sua vez, produz efeitos no produto. Contudo, os mesmos autores admitem que estes efeitos se dissipam no longo prazo ou, de forma equivalente, consideram que no longo prazo estes efeitos serão consideravelmente menores do que os efeitos provocados por choques na oferta agregada, sendo portanto negligenciáveis. Desta forma, apenas os choques sobre a oferta agregada, tipicamente os choques de produtividade, produzem efeitos de longo prazo no produto. Assim, fazendo uso da teoria económica e admitindo o denominado princípio da neutralidade da procura agregada, Blanchard e Quah (1989) impõem restrições sobre os efeitos de longo prazo de forma a identificar o modelo. Por fim, note-se ainda que, teoricamente, é possível proceder à identificação de um modelo SVAR ou SVMCE através da combinação entre restrições de curto e longo prazo. Deste modo, a consideração de restrições sobre os efeitos de longo prazo permite assim reduzir o número de restrições necessárias a impor sobre  $\mathbf{A}$  e  $\mathbf{B}$ .

## 5.2 Análise impulso-resposta

Suponhamos primeiro que  $\mathbf{X}_t$  é  $I(0)$ , ou seja, estacionária, e que segue um modelo VAR( $p$ ),  $\Phi(B)\mathbf{X}_t = \mathbf{a}_t$ , com  $\Phi(B) = (\mathbf{I}_m - \Phi_1 B - \dots - \Phi_p B^p)$  em que as raízes de  $|\Phi(B)| = 0$  estão fora do círculo unitário. Os efeitos dos choques nas variáveis do sistema são mais facilmente visíveis na representação em médias móveis de ordem infinita

$$\mathbf{X}_t = \Phi^{-1}(B)\mathbf{a}_t = \Psi(B)\mathbf{a}_t = \sum_{i=0}^{\infty} \Psi_i \mathbf{a}_{t-i}, \quad (33)$$

onde  $\Psi_0 = \mathbf{I}_m$  e

$$\Psi_i = \sum_{j=1}^i \Psi_{i-j} \Phi_j \quad i = 1, 2, \dots \quad (34)$$

o que significa que as matrizes  $\Psi_i$  podem ser calculadas recursivamente a partir da forma reduzida do modelo VAR de  $\mathbf{X}_t$  em níveis. Os coeficientes desta representação podem ser interpretados como refletindo as respostas a impulsos que atingem o sistema. Com efeito, (34) mostra que os elementos de ordem  $(u, v)$  das matrizes  $\Psi_i$ , encarados como função de  $i$ , medem a resposta esperada de  $X_{ut}$  a uma variação unitária de  $a_{vt,t-i}$ , mantendo constantes todos os outros elementos do ruído branco. De forma semelhante, os mesmos elementos de  $\Psi_i$  também medem a resposta esperada de  $X_{u,t+i}$  a uma variação unitária de  $a_{vt}$  (mantendo constantes todos os outros elementos do ruído branco). Por isso, diz-se que os elementos de  $\Psi_i$  representam as respostas a impulsos das componentes de  $\mathbf{X}_t$  relativamente às variáveis residuais de  $\mathbf{a}_t$  e designam-se por funções impulso resposta, sendo também designadas por respostas a impulsos do erro de previsão porque  $\mathbf{a}_t$  é o erro de previsão a um passo de  $\mathbf{X}_t$ . Conforme já foi referido anteriormente, neste caso  $I(0)$ ,  $\Psi_i \rightarrow \mathbf{0}$  quando  $i \rightarrow \infty$ , pelo que o efeito de um impulso é transitório e acaba por desaparecer com o passar do tempo. Por vezes, interessa também analisar o efeito acumulado dos impulsos que é facilmente obtido somando as matrizes  $\Psi_i$ . O efeito acumulado em  $n$  períodos é dado por  $\sum_{i=0}^n \Psi_i$  e o efeito acumulado em todos os períodos, isto é, o efeito total de longo prazo, é obtido

fazendo  $n \rightarrow \infty$ , resultando em  $\Psi = \sum_{i=0}^{\infty} \Psi_i = (\mathbf{I}_m - \Phi_1 - \dots - \Phi_p)^{-1}$ . Recorde-se que esta matriz

inversa existe porque  $\mathbf{X}_t$  é  $I(0)$ . A estacionaridade de  $\mathbf{X}_t$  também implica que  $\Psi_i \rightarrow \mathbf{0}$  quando  $i \rightarrow \infty$ , o que permite calcular o efeito acumulado total e significa que não existem efeitos permanentes ou de longo prazo de um impulso (ou choque) numa série estacionária, isto é, só existem efeitos de curto prazo. A representação gráfica das funções impulso resposta em função de  $i$  mostra de uma forma sugestiva o comportamento das séries em resposta aos vários impulsos ou choques.

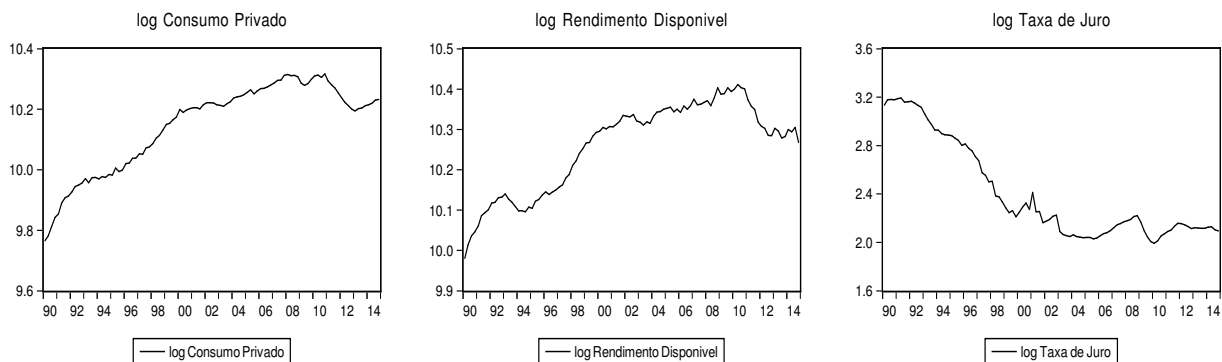
Se  $\mathbf{X}_t$  for não estacionária, a representação em médias móveis de ordem infinita dada em (33) não existe (quer haja cointegração, quer não). No entanto, as matrizes de resposta a impulsos podem ser calculadas da mesma forma que em (34) a partir de um VAR com variáveis integradas ou de um VMCE em níveis. A diferença é que as matrizes  $\Psi_i$  podem não convergir para zero quando  $i \rightarrow \infty$ , o que significa que alguns choques podem ter efeitos permanentes ou de longo prazo e que a resposta a impulsos acumulada total geralmente não existe. Relembre-se que, se  $\mathbf{X}_t$  for  $I(d)$  com  $d \geq 1$ , a matriz  $(\mathbf{I}_m - \Phi_1 - \dots - \Phi_p)$  não tem inversa (matriz singular). Portanto, a existência de efeitos permanentes ou de longo prazo dos choques e a impossibilidade de calcular o efeito acumulado total constituem a principal diferença em relação ao caso estacionário.

Por fim, refira-se que esta análise pode também ser efetuada com choques ortogonais do mesmo género dos indicados a propósito dos modelos estruturais.

## 6 Exemplo de aplicação

Recorrendo aos métodos descritos, procedeu-se em seguida à estimação de uma função consumo para Portugal, procurando-se também extrair diversas conclusões sobre o comportamento das séries envolvidas. Para este efeito, recolheu-se informação sobre as despesas de consumo privado em milhões de euros a preços constantes de 2011 e corrigidas da sazonalidade, sobre o rendimento disponível dos particulares em milhões de euros a preços correntes e corrigido da sazonalidade e sobre a taxa de juro de crédito a particulares para consumo e outros fins (taxas médias de novas operações das instituições financeiras monetárias). A série do rendimento disponível dos particulares foi convertida para preços constantes de 2011 com recurso ao deflator do consumo privado. Os dados são trimestrais abrangendo o período do primeiro trimestre de 1990 ao quarto trimestre de 2014 (última informação disponível), num total de

Figura 1: Cronogramas das séries logaritmizadas



$n = 100$  observações. A fonte é o Banco de Portugal: Boletim estatístico para a taxa de juro (Estatísticas monetárias e financeiras) e Séries trimestrais para a economia portuguesa (Boletim económico de Junho de 2015). As séries foram logaritmizadas com o logaritmo natural (designado por log) e os seus cronogramas encontram-se na figura 1.

## 6.1 Filtros

Em primeiro lugar, para exemplificar os métodos de separação ciclo/tendência, procedeu-se à aplicação dos diferentes filtros descritos anteriormente: Hodrick e Prescott, *Exponential Smoothing*, Baxter e King e, por último, o filtro de Kalman.

Da aplicação do filtro de Hodrick e Prescott, considerando  $\lambda = 1600$  por se tratarem de dados trimestrais, resultaram os gráficos apresentados na figura 2. No que respeita ao comportamento do logaritmo do consumo privado, a componente cíclica sugere um aumento da volatilidade, caracterizado por variações mais bruscas, desde o período de adesão à moeda única. Conforme seria de esperar, além do agravamento da tendência de queda no (log) consumo privado desde o final de 2009 com os primeiros sinais da crise da dívida soberana nacional, este período destaca-se como o ciclo de maior amplitude ilustrando a severidade da recente recessão. No entanto, recorde-se que, devido à má qualidade do ajustamento nos instantes iniciais e finais, se deve ter bastantes reservas nas considerações a retirar relativamente ao comportamento do ciclo e da tendência destas séries. Assim, será necessário aguardar pela disponibilização de mais dados estatísticos de forma a proceder a uma análise mais rigorosa. Porém, note-se que o ciclo e a tendência do logaritmo do rendimento disponível, embora com maior ruído de curto prazo, são em tudo semelhantes ao ciclo e tendência do (log) consumo privado, ilustrando a forte interligação entre estas duas variáveis. Refira-se ainda que, da aplicação deste filtro, resultou uma tendência bastante alisada, reflexo de constar no problema de otimização um termo de penalização de acelerações da tendência. Este tipo de comportamento da tendência representa uma vantagem na leitura económica dos resultados obtidos, como por exemplo no cálculo de taxas de crescimento. Uma vez que as séries foram previamente logaritmizadas, a taxa de crescimento pode ser lida diretamente pelo declive, pelo que se conclui que, no período entre o primeiro trimestre de 1990 e o último de 2009, a taxa de crescimento média anual situou-se aproximadamente em 5.5% para o (log) consumo privado e em 4% para o (log) rendimento disponível. Estes valores espelham a mudança nos hábitos de consumo dos portugueses ao longo destas duas décadas, sugerindo um certo grau de insustentabilidade, um comportamento que foi de certa forma encorajado pela acentuada quebra na taxa de juro, conforme sugere a figura 2. A evolução desta taxa de juro acompanhou a que foi observada nas taxas de juro do mercado interbancário como consequência de um movimento de liberalização financeira e de aumento de concorrência no setor bancário que se seguiu à adesão de Portugal à Comunidade Económica Europeia.



Figura 2: Filtro de Hodrick e Prescott

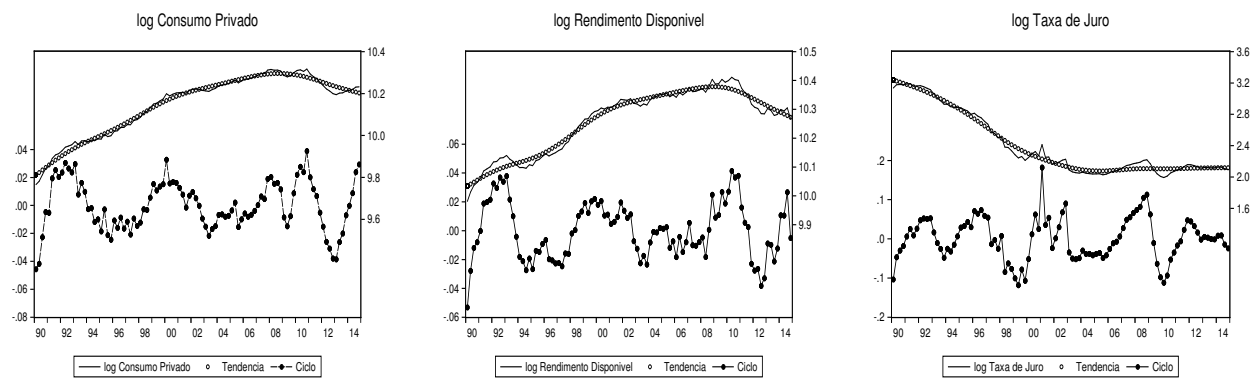
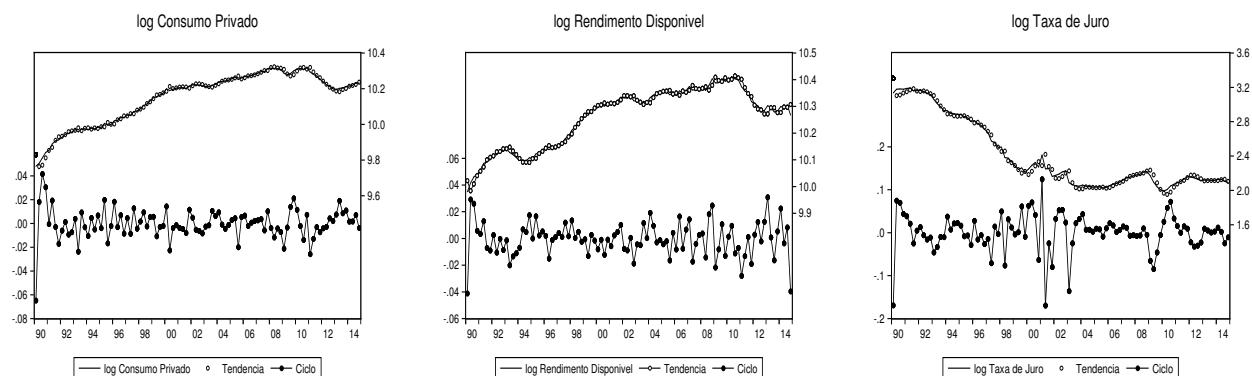


Figura 3: Filtro *Double Exponential Smoothing*

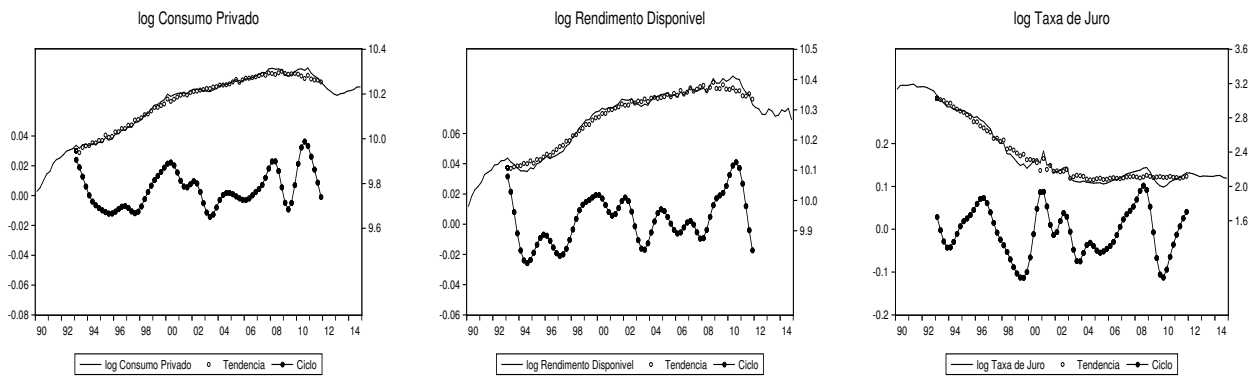


Passando à aplicação do filtro *Exponential Smoothing* (figura 3), a variante deste método mais aconselhável para acomodar a existência de uma tendência linear é o *Double Exponential Smoothing*, uma vez que o método simples não apresentou resultados satisfatórios. Tal como o filtro Hodrick e Prescott, este também não removeu as frequências elevadas do periodograma das séries, resultando numa componente cíclica com bastante ruído de curto prazo perceptível nos gráficos abaixo e obtendo-se uma tendência com um comportamento menos alisado. Por conseguinte, obteve-se um número superior de ciclos, geralmente de menor amplitude. Ficou também clara a vantagem da aplicação do filtro Hodrick e Prescott em detrimento do *Exponential Smoothing*, uma vez que a componente cíclica obtida por este último forneceu pouca informação além de evidenciar a volatilidade destas variáveis.

Por outro lado, por se tratar de um *band-pass filter*, o filtro de Baxter-King removeu completamente o ruído de curto prazo originando uma componente cíclica bastante mais alisada (figura 4). Recorde-se que uma das desvantagens deste filtro é a perda dos primeiros e últimos três anos de observações visível na figura 4. Esta particularidade, aliada à reduzida qualidade de ajustamento nos instantes iniciais e finais, limita a informação passível de ser extraída dos dados. No entanto, a identificação dos diferentes ciclos tornou-se mais evidente.

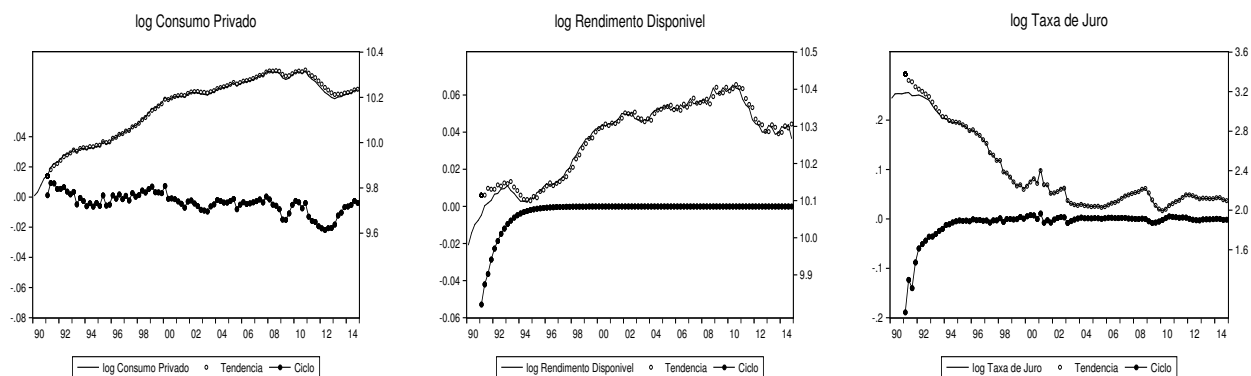
Por último, a aplicação do filtro de Kalman (figura 5) é por vezes incapaz de resolver o problema de otimização subjacente, resultando numa matriz de covariância singular. Para ultrapassar este obstáculo, o *software* proporciona várias soluções, como alterar o algoritmo de otimização utilizado, selecionando por exemplo o algoritmo de Marquardt ou de Berndt-Hall-Hall-Hausman e alterando simultaneamente a magnitude do critério de convergência, se pertinente. Outra solução consiste na atribuição manual das condições iniciais em alternativa à seleção automática. Por fim, uma solução geralmente utilizada consiste em considerar que a tendência possui um *drift* estocástico, o que implica uma definição alternativa para o espaço de estados. Refira-se ainda que as equações de medida e de transição devem ser estimadas várias vezes, selecionando o *output* com base nos habituais critérios de seleção de modelos como o AIC.

Figura 4: Filtro de Baxter-King



O problema de otimização associado à (log) taxa de juro foi o que apresentou maiores obstáculos que foram ultrapassados recorrendo à última solução sugerida, ou seja, considerando um *drift* estocástico, juntamente com a atribuição manual das condições iniciais. Por outro lado, os obstáculos encontrados na resolução do problema de otimização associado ao (log) rendimento disponível revelaram-se facilmente ultrapassáveis apenas com a atribuição destas condições. Na figura 5 fica evidente a má qualidade do ajustamento da tendência nos instantes iniciais. No entanto, o ajustamento tornou-se cada vez mais preciso, de tal forma que reduziu a amplitude dos ciclos e, em alguns casos, como no gráfico referente à (log) taxa de juro e ao (log) rendimento disponível, esvaziou esta componente de qualquer tipo de informação.

Figura 5: Filtro de Kalman



## 6.2 Raízes unitárias

O cronograma das séries (logaritmizadas) na figura 1 sugere que estas são não estacionárias. Esta conclusão foi confirmada pelas funções de autocorrelação (FAC) e de autocorrelação parcial (FACP) estimadas (não apresentadas) que mostram um decaimento para zero muito lento. Além disso, procedeu-se ao teste de raízes unitárias. Começando com o teste ADF, foi necessário decidir primeiro qual a ordem do modelo AR a utilizar. A FAC e a FACP estimadas após uma diferença (também não apresentadas) sugeriram as ordens 3, 3, 2 para o (log) consumo privado, o (log) rendimento disponível e a (log) taxa de juro respetivamente.

Para o (log) consumo privado, conforme sugerido pelo cronograma, optou-se inicialmente por ajustar um modelo com constante e tendência determinística, mas o parâmetro estimado desta última era inferior (em valor absoluto) ao seu desvio padrão estimado, mostrando claramente que era estatisticamente não significativo. Logo, optou-se pelo modelo  $AR(p - 1)$  com constante (sem tendência), tendo-se ensaiado

diversos valores de  $p$ , quer aumentando a partir de  $p = 1$  (teste DF), quer diminuindo a partir de  $p = 9$  (um valor suficientemente elevado) e verificando-se a significância estatística do último parâmetro estimado. Ambas as estratégias conduziram ao mesmo modelo AR(3), concordando também com a ordem sugerida pela FAC e pela FACP estimadas da primeira diferença da série. Este modelo foi estimado pelos mínimos quadrados e verificou-se que os respectivos resíduos são um ruído branco, confirmando que um AR(3) é adequado. Assim, designando o consumo privado no trimestre  $t$  por  $CP_t$ , o modelo estimado é (desvios padrão estimados entre parênteses)

$$\widehat{\log CP_t} = 0.162 + 0.984 \log CP_{t-1} - 0.002 \nabla \log CP_{t-1} + 0.252 \nabla \log CP_{t-2} + 0.230 \nabla \log CP_{t-3}$$

(0.084) (0.008) (0.099) (0.093) (0.097) (35)

ou, de forma equivalente,

$$\nabla \widehat{\log CP_t} = 0.162 - 0.016 \log CP_{t-1} - 0.002 \nabla \log CP_{t-1} + 0.252 \nabla \log CP_{t-2} + 0.230 \nabla \log CP_{t-3}$$

(0.084) (0.008) (0.099) (0.093) (0.097) (36)

perdendo-se por isso 4 observações, pelo que restaram  $(100 - 4) = 96$  observações. Assim, as estatísticas-teste são  $96 \left( \hat{\phi}_\mu - 1 \right) \hat{\psi}(1) = 96 (0.984 - 1) / (1 + 0.002 - 0.252 - 0.23) = 96 (-0.016) / (1 + 0.002 - 0.252 - 0.23) = -2.91$  e  $T_\mu = \left( \hat{\phi}_\mu - 1 \right) / s_{\hat{\phi}_\mu} = -0.016 / 0.008 = -1.92$ . Os pontos críticos tabelados por Dickey (1976) para um nível de significância de 5% e uma série de 100 observações (o mais próximo de 96 tabelado) são  $-13.7$  e  $-2.9$  respetivamente. De forma semelhante, o ponto crítico de MacKinnon (1991, 1996, 2010) para  $T_\mu$  é  $-2.89$  (quase coincidente com o obtido por Dickey) e o valor-p é 0.324. Como os valores das estatísticas são superiores aos respetivos pontos críticos e o valor-p é superior a 5% (e aos níveis de significância habituais), não se rejeitou a hipótese nula de não estacionaridade da série do consumo privado. Consequentemente, foi necessário calcular a primeira diferença e testar novamente a presença de uma raiz unitária. Como se tratava da série das diferenças, o teste baseou-se num modelo sem constante cuja estimação conduziu a um AR(2) sem constante (não apresentado) com  $\hat{\phi} = 0.671$ , sendo as estatísticas-teste  $96 \left( \hat{\phi} - 1 \right) \hat{\psi}(1) = -16.5$  e  $T = \left( \hat{\phi} - 1 \right) / s_{\hat{\phi}} = -3.31$ . Os pontos críticos de Dickey são  $-13.7$  e  $-1.95$  respetivamente e o de MacKinnon para  $T$  é  $-1.94$ , sendo o valor-p de 0.001. Logo, rejeitou-se claramente a hipótese de uma raiz unitária na série após uma diferença, o que significa que se rejeitou a hipótese de o (log) consumo privado (em níveis) ser  $I(2)$ , pelo que se concluiu que é  $I(1)$  (uma só raiz unitária).

Além do teste ADF, aplicou-se também os testes de Phillips-Perron e KPSS, recorrendo às estatísticas (16) e (19) respetivamente, tendo-se optado pelo modelo com constante e tendência determinística, de acordo com o cronograma. Uma vez que a autocorrelação das primeiras diferenças da série mostrou que um baixo número de desfasamentos era suficiente, fixou-se  $\omega = 4$  de que resulta  $\ell = 3$ , o menor valor habitualmente utilizado para o cálculo do desfasamento (outros valores de  $\omega$  foram ensaiados, incluindo  $\omega = 12$  ou  $\ell = 11$ , o outro valor mais utilizado, e os resultados quase não sofreram alteração). No teste de Phillips-Perron, o ponto crítico (a 5% de significância) é o mesmo de Dickey ou o de MacKinnon, ou seja,  $-3.45$  e  $-3.46$  respetivamente (quase iguais de novo). O valor da estatística-teste é  $Z_{T_r} = -1.54$ , superior ao ponto crítico e correspondendo a um valor-p (MacKinnon) de 0.810, pelo que, confirmando o resultado do teste ADF, não se rejeitou a hipótese de existência de raiz unitária. Assim, o teste com a série em diferenças num modelo apenas com constante (sem tendência) permitiu obter  $Z_{T_\mu} = -7.29$ . O ponto crítico é agora  $-2.9$  (Dickey) ou  $-2.89$  (MacKinnon) e o valor-p (MacKinnon) é aproximadamente 0, pelo que se rejeitou a hipótese de existência de raiz unitária. Concluiu-se portanto que a série é  $I(1)$ , tal como no teste ADF.

Para o teste KPSS, a estatística-teste é  $KPSS = 0.46$ , superior ao ponto crítico de 0.146 (a 5% de significância), pelo que se rejeitou a hipótese nula de estacionaridade (relembre-se que neste teste a hipótese nula é a de inexistência de raízes unitárias e que se rejeita esta hipótese para valores elevados da estatística-teste). Para a série das diferenças (modelo sem tendência) a estatística-teste é  $KPSS = 1.087$ ,

superior ao ponto crítico, que é 0.463, pelo que se rejeitou novamente a hipótese de estacionaridade (só a partir de um desfasamento igual a 7 e a 1% de significância é que não se rejeitava esta hipótese). Testando a série das segundas diferenças (novamente modelo sem tendência), obteve-se  $KPSS = 0.026$ , muito inferior ao ponto crítico, pelo que não se rejeitou a hipótese de estacionaridade, o que significa que se concluiu que a série é  $I(2)$ . Este resultado contraria os que foram obtidos pelos outros dois testes mas, conforme foi referido, os diferentes testes podem conduzir a conclusões contraditórias, como aconteceu neste caso. Apesar do resultado do teste KPSS, pareceu mais adequado concluir que a série do (log) consumo privado é  $I(1)$ .

Prosseguindo com o teste das séries (logaritmizadas) do rendimento disponível (designado por  $RD_t$ ) e da taxa de juro ( $TXJ_t$ ), as tabelas 1, 2 e 3 mostram os resultados dos testes ADF, Phillips-Perron e KPSS respetivamente (onde se incluem também os resultados relativos ao consumo privado de modo a dispôr dos resultados completos). No teste ADF, e tal como para o consumo privado, utilizou-se o modelo com constante para as séries em níveis e sem constante para as suas diferenças, estando a ordem dos modelos também indicada na tabela. Concluiu-se que ambas as séries são  $I(1)$ , tal como o consumo privado.

Tabela 1: Resultados do teste ADF

Série	Níveis				Primeiras diferenças			
	$p - 1$	$T_\mu$	ponto crítico	valor-p	$p - 1$	$T$	ponto crítico	valor-p
$\log CP_t$	3	-1.92	-2.90 <sup>(1)</sup> ; -2.89 <sup>(2)</sup>	0.324	2	-3.31	-1.95 <sup>(1)</sup> ; -1.94 <sup>(2)</sup>	0.001
$\log RD_t$	3	-1.75	-2.90 <sup>(1)</sup> ; -2.89 <sup>(2)</sup>	0.402	2	-3.31	-1.95 <sup>(1)</sup> ; -1.94 <sup>(1)</sup>	0.001
$\log TXJ_t$	2	-2.16	-2.90 <sup>(2)</sup> ; -2.89 <sup>(2)</sup>	0.222	1	-4.36	-1.95 <sup>(1)</sup> ; -1.94 <sup>(2)</sup>	0.000

(1) Dickey; (2) MacKinnon

No teste de Phillips-Perron, para as séries em níveis, utilizou-se o modelo com constante e com tendência determinística para o rendimento disponível e o modelo com constante para a taxa de juro, pelo que, na tabela 2,  $Z$  representa as estatísticas  $Z_{T_T}$  no primeiro caso e  $Z_{T_\mu}$  no segundo. Para as séries em diferenças, utilizou-se o modelo com constante (sem tendência) para a primeira e o modelo sem constante (nem tendência) para a segunda, pelo que  $Z$  na tabela 2 representa  $Z_{T_\mu}$  no primeiro caso e  $Z_T$  no terceiro. Utilizou-se  $\omega = 4$ , ou seja,  $\ell = 3$ . Da mesma forma que no teste ADF, concluiu-se que ambas as séries são  $I(1)$ , tal como o consumo privado.

Tabela 2: Resultados do teste Phillips-Perron

Série	Níveis				Primeiras diferenças			
	$\ell$	$Z$	ponto crítico	valor-p	$\ell$	$Z$	ponto crítico	valor-p
$\log CP_t$	3	-1.54	-3.45 <sup>(1)</sup> ; -3.46 <sup>(2)</sup>	0.810	3	-7.29	-2.90 <sup>(1)</sup> ; -2.89 <sup>(2)</sup>	0.000
$\log RD_t$	3	-0.58	-3.45 <sup>(1)</sup> ; -3.46 <sup>(2)</sup>	0.978	3	-7.92	-2.90 <sup>(1)</sup> ; -2.89 <sup>(2)</sup>	0.000
$\log TXJ_t$	3	-1.84	-2.90 <sup>(1)</sup> ; -2.89 <sup>(2)</sup>	0.351	3	-9.65	-1.95 <sup>(1)</sup> ; -1.94 <sup>(2)</sup>	0.000

(1) Dickey; (2) MacKinnon

No teste KPSS, para as séries em níveis, utilizou-se o modelo com tendência determinística para o (log) rendimento disponível e o modelo com constante para a (log) taxa de juro. Para as séries em diferenças, utilizou-se para todas o modelo com constante (sem tendência). Os resultados estão na tabela 3 (manteve-se  $\omega = 4$ ), concluindo-se que o rendimento disponível é  $I(2)$ , tal como para o consumo privado, sendo igualmente preferível considerar que esta série é  $I(1)$  em face das conclusões dos outros dois testes. Por sua vez, concluiu-se que a (log) taxa de juro é  $I(1)$ , concordando com os outros testes.

### 6.3 Cointegração

Uma vez que as três séries são integradas com a mesma ordem, pois são todas  $I(1)$ , é possível testar se estão cointegradas e, em caso afirmativo, estimar o modelo VMCE apropriado de modo a estimar uma

Tabela 3: Resultados do teste KPSS

Série	Níveis			Primeiras diferenças			Segundas diferenças		
	$\ell$	KPSS	ponto crítico	$\ell$	KPSS	ponto crítico	$\ell$	KPSS	ponto crítico
$\log CP_t$	4	0.460	0.146	3	1.087	0.463	3	0.026	0.463
$\log RD_t$	4	0.419	0.146	3	0.875	0.463	3	0.058	0.463
$\log TXJ_t$	4	1.722	0.463	3	0.427	0.463			

função consumo e a perceber como as séries se ajustam na direção do equilíbrio de longo prazo.

Começando com o teste de Engle-Granger, definiu-se a regressão de cointegração com termo constante  $\log CP_t = \delta_0 + \eta_1 \log RD_t + \eta_2 \log TXJ + U_t$  onde se normalizou em ordem ao coeficiente do consumo privado, uma vez que se pretendia estimar uma função consumo (mas recorde-se que poderia ter sido qualquer das outras variáveis). Esta equação foi estimada pelos mínimos quadrados e testou-se a presença de raízes unitárias nos respectivos resíduos  $\hat{U}_t$  com o teste ADF a partir de (28), tendo-se revelado mais adequado  $p = 1$ , isto é, o teste DF, pelo que se perdeu uma observação. A estatística-teste resultante é  $T = -3.09$ , sendo o ponto crítico de Engle e Yoo (1987) igual a  $-3.93$  (5% de significância), para 100 observações e 3 variáveis na regressão de cointegração acima. Por sua vez, o ponto crítico de MacKinnon é  $-3.83$  e o valor-p é 0.22. Uma vez que o valor de  $T$  é superior ao ponto crítico e o valor-p é maior que 5%, não se rejeitou a hipótese de existência de raízes unitárias em  $\hat{U}_t$ , o que significa que não se rejeitou a hipótese de não existir cointegração entre as séries (o que se afigura surpreendente, pois seria de esperar que existisse uma relação de equilíbrio de longo prazo entre estas variáveis).

Para o teste de Phillips e Ouliaris, fixou-se  $\omega = 4$ , o que corresponde a  $\ell = 3$ , sendo o valor da estatística-teste igual a  $Z_{T_\mu} = -2.92$ . O ponto crítico é igual ao do teste ADF e o valor-p (MacKinnon) é 0.288, pelo que não se rejeitou a hipótese de não cointegração, tal como no teste anterior.

Prosseguindo com o teste de Johansen, optou-se por um modelo VMCE com constante quer na relação de cointegração (componente de longo prazo), quer na componente de curto prazo, com  $p = 2$  nesta última componente, isto é, um desfasamento igual a 1 para o vetor das diferenças das séries. A tabela 4 mostra os resultados do teste do traço, incluindo o número de vetores cointegrantes  $r$  (relembre-se que, uma vez que são três séries, existem no máximo dois vetores cointegrantes), os respectivos valores próprios, a estatística-teste (29), os pontos críticos (a 5% de significância) de Osterwald-Lenum (1992) e os de MacKinnon, Haug e Michelis (1999) e o valor-p destes últimos, conforme assinalado. Testou-se em primeiro lugar  $H_0 : r = 0$  (0 vetores cointegrantes, ou seja, não cointegração) contra  $H_3 : r = 3$  e rejeitou-se  $H_0$ , uma vez que  $LR_{\text{tr}}(0|3) = 40$  é superior ao ponto crítico (32.56 ou 29.8) e o valor-p, igual a 0.002, é inferior a 5%. Continuou-se portanto com o teste de  $H_1 : r = 1$  contra  $H_3 : r = 3$  e concluiu-se que  $LR_{\text{tr}}(1|3) = 7.37$  é não significativa, uma vez que este valor é inferior ao respetivo ponto crítico e o valor-p é superior a 5%. Logo, não se rejeitou a hipótese de existir um único vetor cointegrante e o teste parou. Concluiu-se portanto que as séries são cointegradas, contrariando o resultado dos dois testes anteriores.

Tabela 4: Resultados do teste do traço (Johansen)

	$\lambda$	$LR_{\text{tr}}(r 3)$	ponto crítico	valor-p
$r = 0$	0.283	40.00	32.56 <sup>(1)</sup> ; 29.80 <sup>(2)</sup>	0.002
$r = 1$	0.052	7.37	17.52 <sup>(1)</sup> ; 15.49 <sup>(2)</sup>	0.535
$r = 2$	0.022	2.15	4.95 <sup>(1)</sup> ; 3.84 <sup>(2)</sup>	0.143

(1) Osterwald-Lenum; (2) MacKinnon, Haug e Michelis

Por sua vez, no teste do máximo valor próprio (tabela 5) começou por se testar  $H_0 : r = 0$  contra  $H_1 : r = 1$ , sendo a estatística-teste  $LR_{\text{max}}(0|1) = 32.63$  superior ao ponto crítico (23.09 ou 21.13) e o

valor-p inferior a 5%. Logo, rejeitou-se  $H_0$  e concluiu-se que existe pelo menos um vetor cointegrante, pelo que se continuou para testar  $H_1 : r = 1$  contra  $H_2 : r = 2$ , não se rejeitando  $H_1$ . Consequentemente, concluiu-se que  $r = 1$ , tal como no teste do traço.

Tabela 5: Resultados do teste do máximo valor próprio (Johansen)

	$\lambda$	$LR_{\max}(r r+1)$	ponto crítico	valor-p
$r = 0$	0.283	32.63	23.09 <sup>(1)</sup> ; 21.13 <sup>(2)</sup>	0.001
$r = 1$	0.052	5.23	16.05 <sup>(1)</sup> ; 14.26 <sup>(2)</sup>	0.713
$r = 2$	0.022	2.15	4.95 <sup>(1)</sup> ; 3.84 <sup>(2)</sup>	0.143

(1) Osterwald-Lenum; (2) MacKinnon, Haug e Michelis

Os testes de Engle-Granger e de Phillips-Ouliaris conduziram à conclusão contrária de ambos os testes de Johansen. Assim, de acordo com estes últimos, optou-se por considerar que as séries são cointegradas, não só porque estes testes são os mais utilizados, mas principalmente porque a existência de uma relação de equilíbrio de longo prazo entre estas variáveis faz todo o sentido.

## 6.4 Modelo corretor de erro

Uma vez que se concluiu que existe cointegração, a representação em modelo corretor de erro existe, sendo possível estimá-lo pela regressão de característica reduzida ou pela máxima verosimilhança (Johansen, 1988, 1991, 1995), de modo a extrair conclusões sobre a dinâmica quer de curto, quer de longo prazo e assim estimar uma função consumo (de longo prazo). Conforme referido acima nos testes de Johansen, uma vez que as séries do (log) consumo privado e do (log) rendimento disponível incluem uma tendência determinística, optou-se pela especificação de um VMCE com constante na relação de cointegração e na componente de curto prazo (componente autoregressiva das diferenças) com um desfasamento igual a 1. Refira-se que a existência de constante em ambas as componentes acarreta algumas dificuldades na estimação, pois a decomposição (nas duas componentes) não é identificada de forma única. Para resolver este problema, existem diversos métodos que partem da identificação da constante na relação de cointegração, como é o caso de Johansen (1995). O modelo estimado encontra-se em (37) (desvios padrão estimados entre parênteses).

$$\begin{aligned}
 & \begin{pmatrix} \widehat{\nabla \log CP}_t \\ \widehat{\nabla \log RD}_t \\ \widehat{\nabla \log TXJ}_t \end{pmatrix} = \begin{pmatrix} \hat{\delta}_{CP} \\ \hat{\delta}_{RD} \\ \hat{\delta}_{TXJ} \end{pmatrix} + \hat{\alpha} \hat{\beta}^T \begin{pmatrix} \log CP_{t-1} \\ \log RD_{t-1} \\ \log TXJ_{t-1} \\ 1 \end{pmatrix} + \hat{\Phi}_1^* \begin{pmatrix} \nabla \log CP_{t-1} \\ \nabla \log RD_{t-1} \\ \nabla \log TXJ_{t-1} \end{pmatrix} \\
 = & \begin{pmatrix} 0.004 \\ (0.001) \\ 0.002 \\ (0.001) \\ -0.012 \\ (0.005) \end{pmatrix} + \begin{pmatrix} -0.198 \\ (0.034) \\ -0.098 \\ (0.044) \\ 0.184 \\ (0.151) \end{pmatrix} \begin{pmatrix} 1 & -0.590 & 0.136 & -4.422 \\ & (0.129) & (0.035) & \end{pmatrix} \begin{pmatrix} \log CP_{t-1} \\ \log RD_{t-1} \\ \log TXJ_{t-1} \\ 1 \end{pmatrix} \\
 + & \begin{pmatrix} 0.003 & 0.131 & 0.012 \\ (0.103) & (0.097) & (0.023) \\ 0.096 & 0.042 & -0.009 \\ (0.135) & (0.128) & (0.031) \\ 0.176 & -0.039 & -0.059 \\ (0.459) & (0.435) & (0.105) \end{pmatrix} \begin{pmatrix} \nabla \log CP_{t-1} \\ \nabla \log RD_{t-1} \\ \nabla \log TXJ_{t-1} \end{pmatrix}. \tag{37}
 \end{aligned}$$

Começando a análise do modelo estimado pela componente de curto prazo, verifica-se que os elementos da matriz  $\hat{\Phi}_1^*$  são não significativos (relembe-se que as estatísticas  $t$ -Student habituais podem ser usadas

para estes estimadores), pelo que se poderia ter estimado um modelo sem esta componente, restando apenas a componente de longo prazo. No entanto, optou-se por mantê-la no modelo, a título exemplificativo e porque esta componente tem uma importância secundária. Não obstante, foi ainda ensaiado o modelo com a componente de longo prazo apenas e os resultados quase não sofreram alteração. Por sua vez, os elementos do vetor de constantes são significativos com uma exceção, que não está longe de o ser (as estatísticas *t*-Student também são válidas aqui). Em termos do longo prazo, a função consumo pretendida é a relação de cointegração estimada,

$$\widehat{\log CP}_t = 4.422 + 0.590 \log RD_t - 0.136 \log TXJ_t \iff \widehat{CP}_t = 83.258RD_t^{0.590}TXJ_t^{-0.136}. \quad (38)$$

Este modelo, expresso em termos das variáveis originais (e não em logaritmos), é vulgarmente conhecido como modelo de elasticidades constantes cujas estimativas podem agora ser interpretadas. Assim, quando o rendimento disponível varia em 1% e a taxa de juro permanece constante, o consumo privado varia em 0.59% no mesmo sentido, o que mostra que este último responde positivamente às variações do rendimento, conforme seria de esperar, e que é inelástico relativamente a essas variações (elasticidade inferior a 1). Esta segunda conclusão mostra que a resposta do consumo à variação do rendimento é apenas moderada. Por sua vez, quando a taxa de juro varia 1% e o rendimento disponível permanece constante, o consumo privado varia 0.136% em sentido inverso, o que seria de esperar, uma vez que aumentos (diminuições) da taxa de juro tornam o crédito ao consumo mais caro (barato) e tornam a poupança mais (menos) atrativa em detrimento (a favor) do consumo. Este comportamento reflete os efeitos de restrição de liquidez sobre as famílias e os efeitos de substituição inter-temporal associados à taxa de juro. Conclui-se também que o consumo é muito inelástico relativamente à taxa de juro (elasticidade substancialmente inferior a 1 em valor absoluto), o que significa que reage muito fracamente às variações desta. Conclui-se portanto que o consumo responde mais fortemente às variações do rendimento do que da taxa de juro.

Por fim, os dois primeiros pesos estimados (“loadings”) da relação de cointegração nas equações individuais do modelo são significativos, mas o terceiro não (recorde-se que as estatísticas *t* habituais são válidas aqui). Mesmo assim, optou-se por manter o modelo estimado nesta forma. Então, o peso estimado na equação do (log) consumo privado é  $-0.198$  o que significa que um desvio positivo (negativo) do consumo em relação ao equilíbrio de longo prazo provoca uma redução (aumento) do próprio consumo no curto prazo. Portanto, um “excesso” de consumo provoca uma redução do próprio consumo no curto prazo e vice-versa, isto é, os desvios em relação ao equilíbrio induzem uma variação do consumo de sinal contrário, o que é compreensível. O mesmo sucede relativamente ao (log) rendimento disponível, uma vez que o respetivo peso estimado também é negativo ( $-0.098$ ), concluindo-se que um desvio positivo (negativo) do consumo em relação ao equilíbrio de longo prazo provoca uma redução (aumento) do rendimento disponível no curto prazo. O sentido desta variação pode explicar-se pelo impacto do desvio no próprio consumo, que se repercute também no rendimento. Pelo contrário, o peso estimado da (log) taxa de juro é positivo ( $0.184$ ), o que significa que os desvios do consumo em relação ao equilíbrio de longo prazo provocam em efeito oposto ao das outras variáveis. Assim, um “excesso” (“défice”) de consumo provoca um aumento (redução) da taxa de juro no curto prazo, o que também é compreensível, devido à pressão que o aumento (diminuição) da procura de crédito pelos particulares exerce sobre as taxas de juro no sentido da alta (baixa) para responder às necessidades de financiamento do consumo. Note-se no entanto que este impacto será apenas ligeiro, pois este parâmetro estimado é não significativo, embora não esteja longe de o ser. Finalmente, lembrando que estes pesos medem a velocidade do ajustamento em direção ao equilíbrio de longo prazo, verifica-se que a velocidade mais elevada se regista para o próprio consumo, o que permite concluir que o impacto do desequilíbrio é mais forte sobre esta variável.

Para complementar estes resultados, é ainda muito interessante analisar a regressão de cointegração com as variáveis originais e não em logaritmos, de modo a estimar a propensão marginal ao consumo e o consumo autónomo. Uma análise com as séries originais semelhante à efetuada acima conduziu às mesmas conclusões, isto é, indicou que as séries são cointegradas da mesma ordem. Da estimação do modelo VMCE adequado, importa reter a regressão de cointegração,  $\widehat{CP}_t = 10422.820 + 0.611RD_t -$

$181.475TXJ_t$ , de que se destaca uma propensão marginal ao consumo de 0.611, o que significa que, quando o rendimento disponível varia em um milhão de euros e a taxa de juro permanece constante, o consumo privado varia em 611 mil euros no mesmo sentido. Logo, conclui-se que apenas pouco mais de 60% do aumento do rendimento é destinado ao consumo, uma propensão ao consumo moderada e inferior ao que poderia ser esperado, podendo explicar-se pelos efeitos sobre esta variável da crise económica dos anos mais recentes. Destaca-se também o termo constante desta regressão (10 422.8 milhões de euros) que é o consumo médio estimado quando o rendimento disponível e a taxa de juro são nulos e se designa por consumo autónomo. Considerando a dimensão da população portuguesa, este pode ser considerado um valor bastante baixo.

## 6.5 Causalidade de Granger

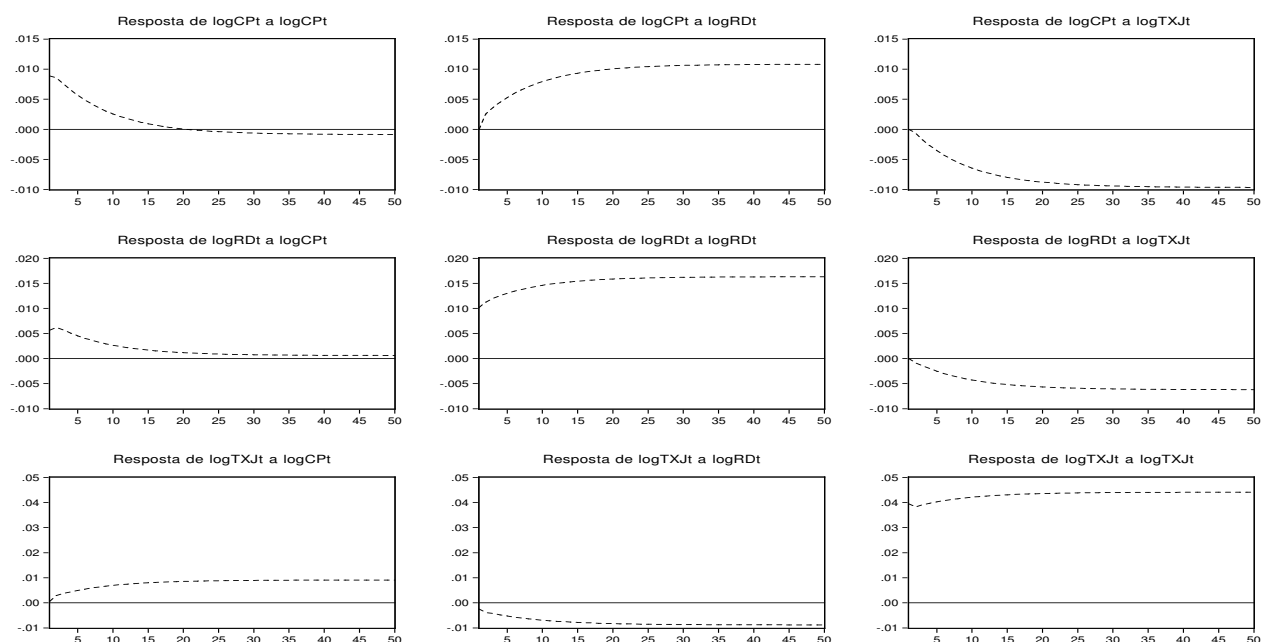
Uma vez que existe cointegração, a existência de causalidade em pelo menos uma das direções está assegurada, seja pela via dos parâmetros de curto prazo no modelo VMCE, seja pela via dos de longo prazo na relação de cointegração. Como o modelo estimado (37) é um VMCE(1), o teste de causalidade de Granger equivale a um teste da significância individual dos parâmetros da matriz  $\hat{\Phi}_1^*$  e dos parâmetros do termo corretor de erro dados no vetor  $\alpha$ . Assim, a observação da primeira equação de (37) permitiu rejeitar a hipótese nula de inexistência de causalidade, porque o parâmetro estimado do termo corretor de erro  $\hat{\alpha}_{11} = -0.198$  (longo prazo) é estatisticamente significativo ( $t = -5.863$ ). Pelo contrário, em termos do curto prazo, não se rejeitou as hipóteses de o (log) rendimento disponível não causar o (log) consumo privado ( $\hat{\Phi}_{1,12}^* = 0.131$  com  $t = 1.350$ ) e de a (log) taxa de juro não causar o (log) consumo privado no sentido de Granger ( $\hat{\Phi}_{1,13}^* = 0.012$  com  $t = 0.501$ ), pois os respetivos parâmetros estimados não são estatisticamente significativos. Relativamente ao (log) rendimento disponível, a segunda equação conduziu a conclusões semelhantes, pois existe causalidade no longo prazo, uma vez que  $\hat{\alpha}_{21} = -0.098$  é estatisticamente significativo ( $t = -2.205$ ) e nem o (log) consumo privado nem a (log) taxa de juro causam no sentido de Granger o (log) rendimento disponível no curto prazo, uma vez que os respetivos parâmetros estimados são não significativos ( $\hat{\Phi}_{1,21}^* = 0.96$  com  $t = 0.708$  e  $\hat{\Phi}_{1,23}^* = -0.009$  com  $t = -0.292$ ). Finalmente, a terceira equação conduziu à conclusão de que as outras duas variáveis não causam a (log) taxa de juro, nem no curto, nem no longo prazo.

## 6.6 Análise impulso-resposta

Optou-se por efetuar esta análise com choques ortogonais recorrendo à decomposição de Choleski da matriz de variâncias e covariâncias residual do modelo VMCE estimado. A figura 6 mostra as funções impulso-resposta para cada par de variáveis, permitindo observar o impacto dos impulsos numa série sobre outra. Verifica-se que quase todas as funções não tendem para zero, o que significa que o efeito desses impulsos não desaparece com o tempo e se deve à não estacionaridade das séries. Apenas nos casos da resposta do (log) consumo privado e do (log) rendimento disponível a impulsos no primeiro se observa que a função tende para zero. Além disso, observa-se respostas positivas do (log) consumo privado a impulsos no (log) rendimento disponível e vice-versa e da (log) taxa de juro a impulsos no (log) consumo privado. Mas, enquanto a primeira e a terceira são crescentes no tempo, a segunda vai desaparecendo gradualmente (tende para zero). Observa-se ainda respostas negativas e crescentes em valor absoluto do (log) consumo privado e do (log) rendimento disponível a impulsos na (log) taxa de juro e desta última a impulsos no (log) rendimento disponível. Estes comportamentos estão em geral de acordo com o esperado.



Figura 6: Funções impulso-resposta



## 7 Conclusão

A utilização de séries temporais em Economia originou a Econometria de séries temporais, cuja evolução tem sido extremamente rápida. A grande quantidade de investigação desenvolvida e de literatura produzida neste domínio prova-o à evidência. A proposta de novos métodos serviu para tratar problemas que se mantinham por resolver satisfatoriamente e para aplicar a novos temas, proporcionando progressos substanciais no conhecimento da Economia. A possibilidade de verificação empírica da teoria económica permitiu grandes avanços da própria teoria, com a descoberta de outras perspectivas sobre os sistemas económicos ou sobre mercados específicos.

Pretendeu-se apresentar resumidamente os temas mais importantes e os métodos mais utilizados, destacando-se as raízes unitárias e cointegração de séries temporais e os modelos econométricos relacionados. Mas, ficaram ainda por descrever muitos dos assuntos mais relevantes e utilizados em Econometria aplicada, como a consideração da sazonalidade, os modelos de heteroscedasticidade condicionada (GARCH), os modelos não lineares ou a deteção de quebras estruturais. Mesmo os temas aqui tratados ficaram muito incompletos devido às naturais limitações de espaço, mas o objetivo era precisamente apresentá-los de forma resumida, sem grande profundidade e detalhe, com caráter informativo e de modo a despertar o interesse por eles. Por isso, foi ainda feita a aplicação prática dos métodos descritos através de um exemplo ilustrativo que permitiu extrair diversas conclusões sobre as variáveis envolvidas.

## Referências

- [1] Baxter, M. e King, R. (1999). Measuring business cycles: approximate band-pass filters for economic time series. *R. of Economics and Statistics*, **81**, 575-593.
- [2] Beveridge, S. e Nelson, C.R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the “business cycle”. *J. of Monetary Economics*, **7**, 151-174.

- [3] Blanchard, O. e Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic R.*, **79**, 655-673.
- [4] Breitung, J., Brüggemann, R. e Lütkepohl, H. (2004). Structural vector autoregressive modeling and impulse responses. In Lütkepohl, H. e Krätzig, M. editors, *Applied Time Series Econometrics*, Cambridge University Press, Cambridge, 159-196.
- [5] Canova, F. (2007). *Methods for Applied Macroeconomic Research*. Princeton University Press, Princeton.
- [6] Chan, N.H. e Wei, C.Z. (1988). Limiting distribution of least squares estimates of unstable autoregressive processes. *Annals of Statistics*, **16**, 367-401.
- [7] Cogley, T. e Nason, J.M. (1995). Output dynamics in real-business-cycle models. *American Economic R.*, **85**, 492-511.
- [8] Dickey, D.A. (1976). *Estimation and Hypothesis Testing in Nonstationary Time Series*, Tese de Doutorado não publicada, Iowa State University.
- [9] Dickey, D.A. e Fuller, W.A. (1979). Distribution of the estimates for autoregressive time series with a unit root. *J. of the American Statistical Association*, **74**, 427-431.
- [10] Elliott, G., Rothenberg, T.J. e Stock, J.H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, **64**, 813-836.
- [11] Enders, W. (2015). *Applied Econometric Time Series*. 4<sup>a</sup> ed., John Wiley and Sons, Hoboken.
- [12] Engle, R.F. e Granger, C.W.J. (1987). Co-integration and error correction: representation, estimation and testing. *Econometrica*, **55**, 251-276.
- [13] Engle, R.F. e Yoo, B.S. (1987). Forecasting and testing in co-integrated systems. *J. of Econometrics*, **35**, 143-159.
- [14] Fuller, W.A. (1996). *Introduction to Statistical Time Series*. 2<sup>a</sup> ed., John Wiley and Sons, Nova Iorque.
- [15] Granger, C.W.J. (1969). Investigating causal relationships by econometric models and cross-spectral methods. *Econometrica*, **37**, 424-438.
- [16] Granger, C.W.J. e Newbold, P. (1974). Spurious regressions in Econometrics. *J. of Econometrics*, **2**, 11-120.
- [17] Granger, C.W.J. e Newbold, P. (1986). *Forecasting Economic Time Series*. 2<sup>a</sup> ed., Academic Press, Nova Iorque.
- [18] Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press, Princeton.
- [19] Haug, A. (1992). Critical values for the  $\hat{Z}_\alpha$  Phillips-Ouliaris test for cointegration. *Oxford Bulletin of Economics and Statistics*, **54**, 473-480.
- [20] Hodrick, R.J. e Prescott, E.C. (1997). Postwar U.S. business cycles: An empirical investigation. *J. of Money Credit and Banking*, **29**, 1-16.
- [21] Johansen, S. (1988). Statistical analysis of cointegration vectors. *J. of Economic Dynamics and Control*, **12**, 231-254.
- [22] Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, **59**, 1551-1581.

- [23] Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.
- [24] Johansen, S. e Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, **52**, 169-210.
- [25] Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *J. of Basic Engineering*, **82**, 35-45.
- [26] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. e Shin, Y. (1992). Testing the null of stationarity against the alternative of a unit root: How sure are we that the economic time series have a unit root?. *J. of Econometrics*, **54**, 159-178.
- [27] Lütkepohl, H. (2004). Vector autoregressive and vector error correction models. In Lütkepohl, H. e Kräzig, M. editors, *Applied Time Series Econometrics*, Cambridge University Press, Cambridge, 86-158.
- [28] Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin.
- [29] Lütkepohl, H. (2007). Econometric analysis with vector autoregressive models. *EUI Working Papers ECO 2007/11*, European University Institute, Department of Economics, San Domenico di Fiesole.
- [30] MacKinnon, J.G. (1991). Critical values for cointegration tests. In Engle, R.F. e Granger, C.W.J. editors, *Long-Run Economic Relationships: Readings in Cointegration*, Oxford University Press, Oxford.
- [31] MacKinnon, J.G. (1996). Numerical distribution functions for unit root and cointegration tests. *J. of Applied Econometrics*, **11**, 601-618.
- [32] MacKinnon, J.G. (2010). Critical values for cointegration tests. *Queen's Economics Department Working Paper N° 1227*, Department of Economics, Queen's University, Ontario, 267-276.
- [33] MacKinnon, J.G., Haug, A. e Michelis, L. (1999). Numerical distribution functions of likelihood ratio tests for cointegration. *J. of Applied Econometrics*, **14**, 563-577.
- [34] Maddala, G.S. e In-Moo, K. (1998). *Unit Roots, Cointegration and Structural Change*. Cambridge University Press, Cambridge.
- [35] Nelson, C. e Plosser, C. (1982). Trends and random walks in macroeconomic time series: some evidence and implications. *J. of Monetary Economics*, **10**, 130-162.
- [36] Newey West Newey, W.K. e West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55**, 703-708.
- [37] Osterwald-Lenum, M. (1992). A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics*, **54**, 461-471.
- [38] Pfaff, B. (2008). *Analysis of Integrated and Cointegrated Time Series with R*. 2<sup>a</sup> ed., Springer, Nova Iorque.
- [39] Phillips, P.C.B. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics*, **33**, 311-340.
- [40] Phillips, P.C.B. (1987). Time series regression with a unit root. *Econometrica*, **55**, 277-301.

- [41] Phillips, P.C.B. e Ouliaris, S. (1990). Asymptotic properties of Residual Based Tests for Cointegration. *Econometrica*, **58**, 165-193.
- [42] Phillips, P.C.B. e Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, **75**, 335-346.
- [43] Reinsel, G.C. e Ahn, S.K. (1992). Vector autoregressive models with unit roots and reduced rank structure: estimation, likelihood ratio test and forecasting. *J. of Time Series Analysis*, **13**, 353-375.
- [44] Robalo, C.R. (1998). *Modelos Dinâmicos, Raízes Unitárias e Cointegração*. Edinova, Lisboa.
- [45] Said, S. e Dickey, D.A. (1985). Hypothesis testing in ARIMA( $p, 1, q$ ) models. *J. of the American Statistical Association*, **80**, 369-374.
- [46] Schmidt, P. e Phillips, P.C.B. (1992). LM tests for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics*, **54**, 257-287.
- [47] Sims, C.A. (1980). Macroeconomics and reality. *Econometrica*, **48**, 1-48.
- [48] Söderlind, P. (1994). Cyclical properties of a real business cycle model. *J. of Applied Econometrics*, **9**, S113-S122.
- [49] Wei, W.W.S. (2006). *Time Series Analysis: Univariate and Multivariate Methods*. 2<sup>a</sup> ed., Pearson, Nova Iorque.



# Agrupamento de séries temporais e sua aplicação na análise de processos geofísicos e ambientais

Manuel G. Scotto, *manuel.scotto@tecnico.ulisboa.pt*

*CEMAT e Departamento de Matemática,  
Instituto Superior Técnico, Universidade de Lisboa*

Susana M. Barbosa, *susana.a.barbosa@inesctec.pt*

*INESC TEC-INESC Technology and Science*

## 1 Introdução

A análise de séries geofísicas e ambientais, em particular de séries temporais de variáveis climáticas, tem suscitado um interesse crescente pela sua relevância científica e socio-económica e pelo seu papel na identificação, compreensão e mitigação de alterações climáticas. Por outro lado, a disponibilidade cada vez mais generalizada, e de forma aberta e gratuita, de dados meteorológicos de estações, satélites, e saídas de modelos numéricos, torna cada vez mais evidente a necessidade de ferramentas para resumir a informação contida nesse enorme volume de dados. Neste contexto, as técnicas para agrupamento/agregação (*clustering*) de séries temporais assumem um papel de relevo.

O processo de agrupamento é uma técnica para criar grupos (classes) de dados/objetos, de acordo com o seu grau de semelhança. O agrupamento é feito de tal forma que os objetos pertencentes à mesma classe sejam o mais semelhantes (homogéneos) entre si do que objetos em classes diferentes, de acordo com algum critério definido *a priori*. Embora uma parte significativa dos métodos de agrupamento propostos na literatura sejam para classificar dados sem estrutura de dependência temporal, nos últimos anos têm vindo a ser propostos vários métodos para classificar séries temporais. O problema do agrupamento de séries temporais surge num vasto leque de áreas de investigação incluindo a economia, finanças, medicina, ecologia e estudos ambientais, entre outros.

Uma questão fundamental que surge sempre, em qualquer processo de agrupamento de séries temporais, ou de objetos em geral, é definir a noção de “semelhança”, ou seja, definir uma medida de (dis)similaridade adequada entre objetos. No contexto da classificação de séries temporais a definição de tal medida torna-se particularmente complexa, devido ao carácter dinâmico das séries, o que torna, em geral, as medidas clássicas de agrupamento pouco adequadas para dados dependentes do tempo, uma vez que esta interdependência é, normalmente, ignorada. Para ultrapassar esta dificuldade, têm sido propostas nos últimos anos um conjunto alargado de medidas de similaridade para o agrupamento de séries temporais que, grosso modo, podem ser agrupadas em 4 categorias; (a) baseadas em modelos (*model-based*), (b) baseadas em características das séries (*feature-based*), (c) baseadas em previsões (*future-information-based*) e, (d) baseadas em medidas de complexidade (*complexity-based*). Uma breve descrição sobre a literatura existente para cada uma destas categorias é apresentada a seguir.

As medidas de (dis)similaridade baseadas em modelos assumem que as séries temporais são bem modeladas por modelos paramétricos, por exemplo, do tipo ARIMA. O procedimento neste caso é ajustar,

primeiro, um modelo paramétrico a cada uma das séries e representar cada uma delas através do correspondente vetor de estimativas pontuais dos parâmetros do modelo. A seguir constrói-se uma medida de distância entre o conjunto de vetores dois a dois. Para processos ARIMA invertíveis, Piccolo (1990) introduziu a distância de Minkowski de segunda ordem (distância euclidiana) em que os vetores de parâmetros contêm as estimativas pontuais dos parâmetros do modelo na sua representação autorregressiva, simbolicamente  $AR(\infty)$ . Maharaj (2000, 1996) considerou, para processos ARMA estacionários e invertíveis, uma versão ponderada da distância de Minkowski de segunda ordem para construir a medida de distância. Galeano e Peña (2000) introduziram uma medida semelhante à de Maharaj (1996) embora para estimar a distância entre os resíduos dos modelos. Kalpakis et al. (2001) também recorreram à distância euclidiana para comparar vetores contendo estimativas dos coeficientes LPC (*Linear Predictive Coding*) cepstral.

Em relação às medidas baseadas em características das séries temporais, vários autores têm proposto medidas de similaridade relacionadas com a estrutura de autocorrelação das séries. Galeano e Peña (2000), por exemplo, propuseram uma versão ponderada da distância euclidiana para comparar as funções de autocorrelação estimadas a partir das séries. Na mesma linha D'Urso e Maharaj (2009) propuseram também uma versão ponderada (*fuzzy*) da distância euclidiana, em que os pesos representam o grau de pertença associada à atribuição das séries nas diferentes classes (*membership degree*). Recentemente, Lafuente-Rego e Vilar (2015) propuseram uma maneira alternativa de olhar para o agrupamento de séries temporais. Estes autores introduziram uma medida de dissimilaridade com base na comparação entre pares de funções de autocovariâncias quantílicas. Esta função é uma ferramenta muito útil na análise de, por exemplo, séries temporais não lineares e/ou não Gaussianas, para as quais, a função de autocovariância fornece, em geral, pouca informação sobre a estrutura de dependência do processo. Além desta medida obtida a partir da estrutura de autocovariância quantílica, outras medidas baseadas em quantis têm sido propostas na literatura. Monteiro et al. (2012) e Barbosa et al. (2011) introduziram a distância  $L_2$  de Wasserstein ponderada para estimar a distância entre distribuições de coeficientes de regressão quantílica associados a quantis definidos *a priori*. Por outro lado, medidas de distância baseadas no comportamento extremal conjunto de séries temporais também têm vindo a ser propostas nos últimos anos. Durante et al. (2015) e De Luca e Zuccolotto (2015, 2011) propuseram medidas de distância relacionadas com o valor dos coeficientes de dependência de cauda (*tail dependence coefficients*) entre pares de séries temporais.

Embora todas as medidas apresentadas até aqui tenham sido definidas no domínio do tempo, na literatura também têm sido propostas medidas no domínio da frequência. É importante referir que a análise no domínio da frequência é, em geral, complementar à tradicional análise no domínio do tempo. Através da análise espectral é possível avaliar a importância das várias frequências no comportamento individual de cada uma das séries, assim como aferir a relação entre séries em diferentes frequências e se essa relação evolui ao longo do tempo. Kakizawa et al. (1998) introduziram, por exemplo, uma medida de dissimilaridade entre séries temporais baseada no rácio das respetivas densidades espectrais estimadas. Maharaj (2002) usou um método baseado no espectro evolucionário para distinguir duas séries temporais não estacionárias em variância, de modo a tomar em consideração as alterações de estrutura ao longo do tempo. Por outro lado, Caiado et al. (2006) propuseram a distância euclidiana para o cálculo do grau de semelhança entre os periodogramas normalizados de duas séries temporais. D'Urso et al. (2014) e D'Urso e Maharaj (2012) introduziram uma outra medida de distância particularmente adequada na análise de séries temporais em que várias componentes periódicas contribuem significativamente para a variabilidade total das séries. A medida é composta por duas componentes: uma correspondente à soma das distâncias entre as variâncias de onduletas (*wavelet variances*) associadas a pares de séries temporais, para um conjunto de escalas definidas *a priori*; e uma outra equivalente para as correspondentes covariâncias de onduletas (*wavelet covariances*).

É importante salientar que todas as medidas de (dis)similaridade acima referidas apresentam a limitação de não incorporar qualquer informação sobre o, eventual, comportamento futuro das séries temporais.

Contudo, este conhecimento torna-se fundamental quando o objetivo é agrupar séries tendo em conta a sua evolução de curto e/ou de longo prazo (Scotto et al., 2013). Com este propósito Alonso et al. (2006) introduziram uma medida de dissimilaridade para comparar densidades de previsão estimadas para modelos  $AR(p)$ , considerando um horizonte de previsão definido *a priori*. Vilar et al. (2010) generalizaram a metodologia proposta por Alonso et al. (2006) na qual as séries temporais são modeladas através de modelos autorregressivos não paramétricos. Uma das limitações das abordagens de Vilar et al. (2010) e Alonso et al. (2006) é o facto de fornecerem unicamente estimativas pontuais da distância entre densidades de previsão. Liu et al. (2014) propuseram uma generalização das medidas propostas por esses autores na qual é possível avaliar a significância estatística das distâncias estimadas. É importante referir que as medidas de Alonso, Vilar, Liu e co-autores são particularmente adequadas para agrupar séries temporais em função de previsões de curto/médio prazo. Quando o objetivo é fazer o agrupamento tendo em conta previsões de longo prazo, estas metodologias tornam-se pouco adequadas. Neste sentido, Scotto et al. (2014, 2010) propuseram métodos de agrupamento baseados na distância  $L_2$  de Wasserstein ponderada, para o cálculo da distância entre distribuições de valores de retorno associadas a diferentes períodos de retorno definidos *a priori*.

Em relação às distâncias baseadas em medidas de complexidade, é importante salientar que, neste caso, a similaridade entre duas séries temporais não depende nem da estrutura de dependência temporal das séries nem do conhecimento específico dos modelos subjacentes, mas sim da quantificação do nível de informação partilhada por ambas as séries. Esta informação pode ser quantificada através do conceito de Complexidade de Kolmogorov (CK). Têm sido propostas várias metodologias para estimar a CK embora as mais utilizadas sejam, por um lado, algoritmos baseados na compressão de dados (Keogh et al. 2007, Cilibrasi e Vitányi, 2005), e por outro o agrupamento de permutações de distribuição (Brandmaier, 2015).

## 2 Medidas de similaridade/dissimilaridade para séries temporais

Nesta secção são apresentadas as definições de algumas das medidas de (dis)similaridade acima referidas, nomeadamente medidas baseadas em modelos, características das séries e previsões.

### 2.1 Medidas baseadas em modelos

Dispondo de  $N$  observações  $\mathbf{X}_N = (X_1, \dots, X_N)'$  e  $\mathbf{Y}_N = (Y_1, \dots, Y_N)'$  dos processos  $X = (X_t, t \in \mathbb{Z})$  e  $Y = (Y_t, t \in \mathbb{Z})$ , Piccolo (1990) introduziu a medida de distância

$$d^{Pic}(\mathbf{X}_N, \mathbf{Y}_N) := \sqrt{\sum_{i=1}^p (\pi_{i,X} - \pi_{i,Y})^2}, \quad (1)$$

onde  $p = \max(p_1, p_2)$ , sendo  $p_1, p_2 \in \mathbb{N}$  as ordens dos modelos ajustados às séries  $\mathbf{X}_N$  e  $\mathbf{Y}_N$ , na sua representação autorregressiva, e  $\pi_X = [\pi_{1,X} \ \pi_{2,X} \ \dots \ \pi_{p_1,X}]'$  e  $\pi_Y = [\pi_{1,Y} \ \pi_{2,Y} \ \dots \ \pi_{p_2,Y}]'$  os correspondentes parâmetros associados a cada um dos modelos. Note-se que  $\pi_{i,X} = 0$  se  $i > p_1$  e  $\pi_{i,Y} = 0$  se  $i > p_2$ .

Maharaj (1996) considerou uma extensão da medida em (1) em que as distâncias entre os parâmetros dos modelos são ponderadas pelas matrizes de autocovariância,  $W_X(p)$  e  $W_Y(p)$ , dos modelos autorregressivos ajustados às séries  $\mathbf{X}_N$  e  $\mathbf{Y}_N$ , e as correspondentes variâncias do ruído branco,  $\sigma_{\varepsilon_X}^2$  e  $\sigma_{\varepsilon_Y}^2$ , associado a cada um dos modelos. A medida de distância de Maharaj é dada pela expressão

$$d^{Mah}(\mathbf{X}_N, \mathbf{Y}_N) := \sqrt{N}(\pi_X - \pi_Y)'W^{-1}(\pi_X - \pi_Y),$$

sendo  $W = \sigma_{\varepsilon_X}^2 W_X^{-1}(p) + \sigma_{\varepsilon_Y}^2 W_Y^{-1}(p)$ .

Kalpakis et al. (2001) também adotaram a medida (1) embora substituindo  $\pi_{i,X}$  e  $\pi_{i,Y}$  pelos coeficientes LPC cepstral. Para cada uma das séries, estes coeficientes são calculados a partir de  $\pi_{i,X}$  (ou  $\pi_{i,Y}$ ) e da expressão recursiva

$$\phi_h = \begin{cases} \pi_1, & h = 1 \\ \pi_h + \sum_{j=1}^{h-1} (\pi_j - \phi_{h-j}), & h = 2, \dots, p, \\ \sum_{j=1}^p (1 - j/h) \pi_j \phi_{h-j}, & h > p \end{cases}$$

sendo  $\pi_i \equiv \pi_{i,X}$  (ou  $\pi_{i,Y}$ ). Uma das vantagens da utilização deste método é o facto de, em geral, serem necessários um número reduzido de coeficientes para obter uma estimativa bastante precisa da distância entre as séries, sem perder informação relevante. Uma outra vantagem desta abordagem é apresentar uma potência discriminante muito superior à das distâncias  $d^{Pic}$  e  $d^{Mah}$ . É importante salientar também que esta medida é invariante a deslocamentos de localização, de escala e de tempo.

É importante referir que na prática, no entanto, só é possível obter uma estimativa pontual destas distâncias, uma vez que os parâmetros dos modelos são desconhecidos pelo que têm de ser substituídos por estimativas pontuais.

## 2.2 Medidas baseadas em características das séries temporais

Quando o objetivo é classificar séries temporais em função da sua estrutura de autocorrelação, Galeano e Peña (2000) propuseram a seguinte versão ponderada da distância euclidiana para comparar funções de autocorrelação amostrais,  $\hat{\rho}_X := (\hat{\rho}_X(1), \dots, \hat{\rho}_X(h))'$  e  $\hat{\rho}_Y := (\hat{\rho}_Y(1), \dots, \hat{\rho}_Y(h))'$ , considerando um desfasamento de  $h$  unidades, obtidas a partir de séries  $\mathbf{X}_N$  e  $\mathbf{Y}_N$ . A medida proposta por estes autores é

$$d^{ACF}(\mathbf{X}_N, \mathbf{Y}_N) := \sqrt{(\hat{\rho}_X - \hat{\rho}_Y)' \Omega (\hat{\rho}_X - \hat{\rho}_Y)},$$

sendo  $\Omega$  uma qualquer matriz de pesos. Caiado et al. (2006) consideraram três definições para  $\Omega$ : a primeira é ser a matriz identidade. Neste caso a medida  $d^{ACF}$  torna-se na distância euclidiana. A segunda é considerar pesos com um decaimento geométrico sendo que, nesse caso, a medida  $d^{ACF}$  é dada pela expressão

$$\sqrt{\sum_{i=1}^h \alpha (1 - \alpha)^i (\hat{\rho}_X(i) - \hat{\rho}_Y(i))^2}, \quad 0 < \alpha < 1.$$

Finalmente, a terceira proposta dos autores é usar a distância de Mahalanobis entre as autocorrelações, sendo o  $\Omega$  a matriz de covariâncias amostrais dos coeficientes de autocorrelação com elementos dados pela fórmula de Bartlett truncada.

Uma das limitações destas medidas é o facto de ter que se assumir que a estrutura de dependência do processo permanece inalterada ao longo do tempo. No entanto, em muitos casos (por exemplo, na análise de séries ambientais e financeiras) a dinâmica das séries temporais muda ao longo do tempo, seja por causa de mudanças de regime, por mudanças de forma ou por uma combinação das duas. Nestas situações torna-se possível que em determinados intervalos ao longo do tempo, uma série seja mais “provável” pertencer à uma determinada classe enquanto, noutras janelas temporais, seja mais “provável” pertencer a uma outra classe. Para quantificar este grau de incerteza no processo de agrupamento das séries, D’Urso e Maharaj (2009) propuseram o método A-FCM (*Autocorrelation-based Fuzzy C-Mean*). Este método assenta no cálculo dos pesos associados à atribuição das séries em cada uma das  $C$  classes definidas *a priori*. Este cálculo é feito a partir da minimização da função objetivo

$$\sum_{k=1}^K \sum_{c=1}^C u_{k,c}^m \sum_{r=1}^h (\hat{\rho}_{X(k)}(r) - \hat{\rho}_c(r))^2,$$



sujeita às restrições

$$\sum_{c=1}^C u_{k,c}^m = 1, u_{k,c}^m > 0, \quad (2)$$

$$-1 \leq \hat{\rho}_c(r) \leq 1, \quad (3)$$

em que  $u_{k,c}$  representa o grau de pertença da  $k$ -ésima série temporal na  $c$ -ésima classe,  $m > 1$  controla o grau de imprecisão (*fuzziness*) da partição em  $C$  classes,  $\hat{\rho}_{X_{(k)}}(r)$  é o valor estimado da função de autocorrelação da  $k$ -ésima série temporal e  $\hat{\rho}_c(r)$  representa o valor estimado da função de autocorrelação da série temporal representativa (*centroid time series*) da  $c$ -ésima classe.

Uma maneira alternativa de agrupar séries temporais é comparar as suas funções de autocorrelação quantílicas. Esta foi a abordagem proposta por Lafuente-Rego e Vilar (2015). A função de autocorrelação quantílica proporciona informação adicional sobre a estrutura de dependência entre as observações associadas a pares de quantis, para cada uma das séries e também entre séries. Para um conjunto de quantis  $q_{\tau_1}, q_{\tau_2}, \dots, q_{\tau_r}$ , associados às ordens  $0 < \tau_1 < \dots < \tau_r < 1$ , definidas *a priori*, e um conjunto de defasamentos  $h_1 < h_2 < \dots < h_L$ , os autores propuseram a seguinte medida de distância entre as séries  $\mathbf{X}_N$  e  $\mathbf{Y}_N$

$$d^{QCF}(\mathbf{X}_N, \mathbf{Y}_N) := \sum_{i=1}^L \sum_{i=1}^r \sum_{j=1}^r \left( \hat{\gamma}_{h_i}^{\mathbf{X}_N}(\tau_i, \tau_j) - \hat{\gamma}_{h_i}^{\mathbf{Y}_N}(\tau_i, \tau_j) \right)^2,$$

com

$$\hat{\gamma}_l^Z(\tau, \tau') = \frac{1}{N-l} \sum_{t=1}^{N-l} I(Z_t \leq \hat{q}_\tau) I(Z_{t+l} \leq \hat{q}_{\tau'}) - \tau \tau',$$

sendo  $I(\cdot)$  a função indicatriz e  $\hat{q}_\tau, \hat{q}_{\tau'}$  quantis empíricos.

D'Urso et al. (2014) e D'Urso e Maharaj (2012) introduziram um método de agrupamento em que o afastamento entre pares de séries temporais é obtido através de uma medida que pondera, por um lado, a distância entre as variâncias de onduletas associadas a pares bivariados de séries temporais, para um conjunto de escalas definidas *a priori*; e uma outra equivalente para as correspondentes covariâncias de onduletas. A dita medida entre o  $i$ -ésimo e o  $j$ -ésimo par  $\mathbf{Z}_i := (\mathbf{X}_{i,N}, \mathbf{Y}_{i,N})$  e  $\mathbf{Z}_j := (\mathbf{X}_{j,N}, \mathbf{Y}_{j,N})$ , é da forma

$$d^W(\mathbf{Z}_i, \mathbf{Z}_j) := \left\{ (a_{wV} \cdot d_{wv}(\mathbf{Z}_i, \mathbf{Z}_j))^2 + (a_{wC} \cdot d_{wc}(\mathbf{Z}_i, \mathbf{Z}_j))^2 \right\}^{\frac{1}{2}}, \quad (4)$$

em que os pesos  $a_{wV}, a_{wC} \geq 0$ , satisfazem  $a_{wV} + a_{wC} = 1$ , e  $d_{wv}$  e  $d_{wc}$  correspondem às distâncias entre as variâncias e as covariâncias de onduletas associadas, entre o  $i$ -ésimo e o  $j$ -ésimo par, sendo dadas pelas expressões

$$d_{wv}(\mathbf{Z}_i, \mathbf{Z}_j) = \sum_{r=1}^R \| \text{diag}(C_{\mathbf{Z}_i}(\mathbf{v}_r)) - \text{diag}(C_{\mathbf{Z}_j}(\mathbf{v}_r)) \|, \quad d_{wc}(\mathbf{Z}_i, \mathbf{Z}_j) = \sum_{r=1}^R \| \gamma_{\mathbf{Z}_i}(\mathbf{v}_j) - \gamma_{\mathbf{Z}_j}(\mathbf{v}_j) \|,$$

onde  $R$  representa o número de escalas,  $C$  é a matriz de variâncias/covariâncias de onduletas, sendo a componente da covariância representada pela função  $\gamma$ , no termo  $d_{wc}$ .

Finalmente, quando o propósito da classificação é agrupar séries temporais consoante o seu grau de dependência extremal conjunta, o adequado é considerar medidas de similaridade que incorporem informação sobre a relação de interdependência nas caudas. Com este intuito, Durante et al. (2015) e De Luca e Zuccolotto (2011) introduziram a medida

$$d^{TD}(\mathbf{X}_N, \mathbf{Y}_N) := -\log(\lambda_L),$$

sendo  $\lambda_L$  o coeficiente de dependência de cauda esquerda, definido como

$$\lambda_L = \lim_{x \rightarrow 0^+} P(U_1 \leq x | U_2 \leq x) = \lim_{x \rightarrow 0^+} \frac{C(x, x)}{x},$$

em que  $U_1 = F(X)$  e  $U_2 = F(Y)$ , e  $C(\cdot, \cdot)$  representa uma função cópula. A medida  $d^{TD} \geq 0$ , sendo que valores reduzidos implicam séries fortemente dependentes na cauda. Pelo contrário, valores elevados de  $d^{TD}$  implicam fraca dependência. De Luca e Zuccolotto (2011) propuseram utilizar a função cópula de Joe-Clayton

$$C(x_1, x_2) = 1 - \left\{ 1 - [(1 - (1 - x_1)^\kappa)^{-\theta} + (1 - (1 - x_2)^\kappa)^{-\theta} - 1]^{-1/\theta} \right\}^{\frac{1}{\kappa}}.$$

Neste caso é simples verificar que  $\lambda_L = 2^{-1/\theta}$ . De Luca e Zuccolotto (2015) consideraram uma extensão do modelo anterior em que  $\theta$  varia no tempo em função de um conjunto de covariáveis. De referir que de forma perfeitamente análoga, é possível definir a medida  $d^{TD}$  considerando o coeficiente de dependência de cauda direita

$$\lambda_U = \lim_{x \rightarrow 1^-} P(U_1 > x | U_2 > x).$$

### 2.3 Medidas baseadas em previsões

Nos casos em que o objetivo do agrupamento relaciona-se com o desempenho das previsões futuras das séries temporais, torna-se necessário definir medidas de distância para comparar densidades de previsão. Com este objetivo, Alonso et al. (2006) introduziram a medida

$$d^F(\mathbf{X}_N, \mathbf{Y}_N) := \int (f_{X_{N+h}}(x) - f_{Y_{N+h}}(x))^2 dx, \quad (5)$$

sendo  $f(\cdot)$  a função de densidade de previsão  $h$  passos à frente associada a um modelo autorregressivo de ordem  $p$ . Vilar et al. (2010) consideram também a distância  $d^F$  embora assumindo para  $X_t$  uma representação mais geral do tipo

$$X_t = m(X_{t-1}, \dots, X_{t-p}) + \varepsilon_t,$$

em que  $m(\cdot)$  é uma função dependente de  $X_{t-1}, \dots, X_{t-p}$ , satisfazendo um determinado conjunto de propriedades, e  $(\varepsilon_t)$  uma sucessão de variáveis aleatórias i.i.d. Além da distância  $d^F$ , estes autores também propuseram a medida

$$d^{|F|}(\mathbf{X}_N, \mathbf{Y}_N) := \int |f_{X_{N+h}}(x) - f_{Y_{N+h}}(x)| dx.$$

Por outro lado, quando o que está em causa é o agrupamento das séries em termos de previsões de longo prazo a partir, por exemplo, da comparação de distribuições de valores de retorno associadas a períodos de retorno definidos *a priori*, uma abordagem que entre em linha de conta com as propriedades extremas das séries torna-se mais adequada. Foi com este propósito que Scotto et al. (2010) introduziam um método de agregação que combina estatística de valores extremos e estatística bayesiana. Estes autores consideraram a distância  $L_2$  de Wasserstein ponderada

$$d^{EVT}(\mathbf{X}_N, \mathbf{Y}_N) := \left( \int_0^1 (F_{x_p}^{-1}(y|\mathbf{x}_N) - F_{y_p}^{-1}(y|\mathbf{y}_N))^2 y(1-y) dy \right)^{1/2}, \quad (6)$$

em que  $F_{z_p}$  representa a distribuição preditiva *a posteriori* do valor de retorno  $z_p$  associado a um período de retorno de  $1/p$  unidades de tempo. Para caracterizar o comportamento extremal das séries os autores utilizaram o métodos dos  $r$ -máximos. Uma abordagem semelhante (embora num contexto não bayesiano) foi proposta por Scotto et al. (2014) em que a distância entre  $\mathbf{X}_N$  e  $\mathbf{Y}_N$  é estimada também através da medida  $d^{EVT}$  sendo, no entanto, a caracterização do comportamento extremal das séries feita a partir do método POT (*Peak-Over-Threshold*).

### 3 Séries temporais geofísicas e ambientais

As séries temporais geofísicas e ambientais são, em geral, caracterizadas por uma componente sazonal bem definida, que é frequentemente o sinal dominante da série. Variáveis como a temperatura do ar ou do mar apresentam um ciclo anual muito regular, enquanto outras como a pressão atmosférica apresentam quer um ciclo anual quer um ciclo semi-anual, sendo a importância relativa dos dois sinais frequentemente da mesma ordem de grandeza, embora dependa fortemente da localização espacial (Barbosa et al., 2009). Ainda outras variáveis climáticas, como o nível do mar, são caracterizadas por uma componente sazonal complexa e de carácter não estacionário, cuja natureza reflete os diversos fatores (pressão atmosférica, temperatura, ventos e salinidade, por exemplo) que influenciam a variação do nível do mar à escala sazonal. Embora, tradicionalmente, a sazonalidade de dados climáticos tenha sido encarada como uma componente constante, a remover dos dados para posterior análise das correspondentes “anomalias”, no contexto de alterações climáticas a variação sazonal é ela própria de grande interesse (Barbosa, 2009). O ciclo sazonal tornou-se motivo de estudo per si, o que tem motivado a utilização de metodologias de análise de séries temporais que permitam ter em conta a periodicidade das séries, evitando o recurso ao cálculo de anomalias. Técnicas baseadas em onduletas, por exemplo, permitem caracterizar de forma flexível a sazonalidade de séries temporais e quantificar variações em amplitude e fase, sendo por isso utilizadas na definição de medidas de distância adequadas para o agrupamento de séries com componentes periódicas importantes.

Ao contrário da sazonalidade, que só recentemente começou a ser estudada com detalhe, a tendência de séries geofísicas e ambientais foi desde cedo um dos focos principais da análise de séries temporais, em particular no contexto de alterações climáticas. Quantificar e compreender a evolução a longo prazo de variáveis climáticas, quer a partir de dados observacionais, quer a partir de resultados de modelos numéricos, tem sido o objetivo principal de inúmeros estudos. O agrupamento de séries tendo, portanto, em conta o comportamento futuro é assim de particular importância, quer para as variáveis fundamentais cujo comportamento é relativamente previsível, como a temperatura, quer para variáveis integradas, como o nível do mar, cuja evolução a longo prazo depende de múltiplos fatores (temperatura, salinidade, fusão de glaciares e gelos continentais, entre outros) e é de muito difícil quantificação (Clark et al., 2015).

Por outro lado, uma das principais preocupações relacionadas com o impacto socio-económico das alterações climáticas está associada à ocorrência de fenómenos extremos. Tendo em conta a ainda reduzida extensão temporal dos dados existentes quando comparada com as escalas temporais dos fenómenos naturais, a quantificação de eventos extremos é bastante difícil. No entanto as observações disponíveis indicam uma alteração significativa na ocorrência de extremos associada a alterações climáticas (Heim, 2015), o que motiva o desenvolvimento de métodos de agrupamento de séries temporais com base no seu comportamento extremal, como os apresentados na secção anterior.

### 4 Aplicações dos métodos de agrupamento na classificação de séries temporais geofísicas e ambientais

Nesta última secção são referidos alguns trabalhos sobre agrupamento de séries temporais relativas a conjuntos de dados de agitação marítima, ambiente, incêndios e temperaturas.

Alonso et al. (2006) utilizaram a distância  $d^F$  em (5) para agrupar um conjunto de 24 países em função das emissões esperadas de CO<sub>2</sub> em 2012 (protocolo de Kyoto), a partir do conjunto de emissões observadas deste composto químico entre 1960 e 1999. Com recurso à mesma distância, Scotto et al. (2009) classificaram um conjunto de séries temporais do nível médio do mar obtidas em diferentes localizações do mar Báltico, em função de previsões para 3 e 6 meses. A recolha e análise de observações de marégrafos tem uma longa tradição no mar Báltico, sobretudo pelo interesse em estudar os efeitos do ajuste isostático

glacial na Fenoscândia. Os resultados obtidos pelos autores mostram a existência de 3 grupos com elevado grau de coerência espacial. Um primeiro conjunto de estações localizadas mais a norte, na região do mar de Bótnia e no Golfo da Finlândia; um segundo grupo composto pelas estações no Báltico central, e um terceiro grupo com as estações mais a sul e no mar Báltico ocidental. Os grupos obtidos refletem o efeito diferenciado na evolução a longo prazo do nível médio do mar Báltico de processos atmosféricos (principalmente vento zonal). Os ventos na direção este-oeste determinam a ocidente as trocas de água entre o mar Báltico e o mar do Norte, com salinidade muito mais elevada, que ocorrem pelos estreitos que ligam os dois mares, e a norte influenciam de forma significativa o sobre-elevamento do nível do mar em virtude do confinamento da bacia do mar Báltico.

Scotto et al. (2010) consideraram a medida  $d^{EVT}$  em (6) para classificar séries temporais do nível médio da superfície do mar obtidas a partir de marégrafos no Atlântico norte, em função dos valores de retorno esperados para 25, 50 e 100 anos. Os resultados mostraram, entre outras conclusões, uma clara distinção entre as estações com latitudes mais elevadas, para as quais os valores de retorno esperados são maiores, e o resto das estações. Scotto et al. (2011) fizeram o mesmo tipo de análise embora para classificar 32 séries temporais de temperaturas médias diárias obtidas na Europa entre 1901 e 2007. Neste caso os resultados evidenciaram que nas estações localizadas em elevada altitude são expectáveis valores de retorno mais baixos. Além disso, os resultados da classificação mostraram também uma clara distinção entre as estações mais setentrionais na Escandinávia e as estações na Europa central e do sul. Por outro lado, Scotto et al. (2014) compararam as distribuições preditivas associadas a valores de retorno entre 5 e 15 anos a partir da distância  $d^{EVT}$ , para classificar séries temporais de valores diários de área ardida nos 17 distritos de Portugal continental entre 1980 e 2010. Os resultados do estudo evidenciaram que existe uma clara diferença entre os distritos do norte do país (Viana do Castelo, Porto, Braga, Viseu, Vila Real e Bragança) e os restantes distritos.

Monteiro et al. (2012) aplicaram regressão quantílica e a distância  $L_2$  de Wasserstein ponderada, em que  $F$  representa a distribuição associada ao declive correspondente a um quantil fixado *a priori* (neste estudo 0.05, 0.50 e 0.95), na classificação de 11 séries de valores horários de  $O_3$  recolhidos em 11 estações de fundo espalhadas pela Península Ibérica (9 em Espanha e 2 em Portugal) durante o período 2000-2009. Os autores mostraram que a taxa de variação associada ao quantil de ordem 0.05 é bastante superior às verificadas nos quantis de ordem 0.5 e 0.95. A classificação das estações obtida considerando o quantil de ordem 0.05 mostrou a existência de 3 grupos, nomeadamente, um primeiro grupo contendo, unicamente, uma estação localizada no extremo mais oriental da Península Ibérica para a qual a variação média foi de  $\approx -28\mu g m^{-3}/década$ ; um segundo grupo formado por 5 estações, todas localizadas no norte da Península Ibérica, nas quais verifico-se uma variação média de  $\approx -10\mu g m^{-3}/década$ ; e um terceiro grupo contendo o resto das estações com uma variação média positiva de  $\approx +18\mu g m^{-3}/década$ . Os resultados obtidos na classificação das estações para o quantil de ordem 0.50 foram de todo idênticos. Finalmente, para o quantil de ordem 0.95 os resultados foram menos conclusivos, tendo a maioria das estações um comportamento bastante heterogéneo. A mesma metodologia foi aplicada por Barbosa et al. (2011) na classificação de séries temporais de temperaturas médias diárias recolhidas em 28 cidades Europeias no período 1901-2007.

Gouveia et al. (2015) aplicaram a distância  $d^W$  dada em (4) para agrupar séries temporais bivariadas de  $O_3$  e  $NO_2$ , recolhidas em 36 estações de monitorização da qualidade do ar espalhadas por Portugal continental, durante o período 2005-2013. A análise detalhada dos resultados revelou a existência de 3 grupos de estações. Um primeiro grupo composto, principalmente, por estações urbanas localizadas na grande Lisboa; um segundo grupo formado por estações urbanas localizadas no grande Porto, e também estações suburbanas. Finalmente, um terceiro grupo com prevalência de estações rurais. Nos dois primeiros grupos, as escalas  $v_3$  e  $v_4$  (ciclo intra-diário e diário) é que apresentam uma maior contribuição, no caso do  $NO_2$ . Relativamente ao  $O_3$ , é o ciclo diário que apresenta um peso maior, sobretudo na segunda classe. Em relação à covariância entre  $O_3$  e  $NO_2$ , esta é negativa nas duas primeiras classes sendo novamente as escalas  $v_3$  e  $v_4$  aquelas com um contributo maior. Para as estações na terceira

classe, a covariância é nula. D'Urso et al. (2014) fizeram uma aplicação semelhante mas para dados de CO, NO e NO<sub>2</sub> recolhidos em Roma (Itália). Uma outra aplicação para dados do nível médio do mar e pressão atmosférica recolhidos no mar Báltico entre 1979-2005, foi recentemente proposta por Barbosa et al. (2016). A pressão atmosférica influencia diretamente o nível do mar ao exercer uma força vertical sobre a superfície da água, mas o efeito hidrostático é alterado significativamente por outros fatores, como o vento, tornando a associação entre os dois parâmetros complexa espacial e temporalmente. A classificação das séries bivariadas de nível do mar e pressão atmosférica baseada em ondulas permitiu discriminar grupos de séries localizadas a sudoeste, junto à entrada do mar Báltico, em que a variabilidade é predominantemente de alta frequência e associada a vento zonal, de grupos de estações no interior da bacia do Báltico, caracterizadas por variabilidade de mais baixa frequência e associada à resposta a variações de pressão.

## Referências

- [1] Alonso, A.M., Berrendero, J.R., Hernández, A., Justel, A. (2006). Time series clustering based on forecast densities. *Computational Statistics and Data Analysis* **51**, 762-776.
- [2] Barbosa, S.M. (2009). Changing seasonality in Europe's air temperature. *European Physical Journal - Special Topics* **174**, 81-89.
- [3] Barbosa, S.M., Gouveia, S., Scotto, M.G., Alonso, A.M. (2016). Wavelet-based clustering of sea level records. *Mathematical Geosciences* **48**, 149-162.
- [4] Barbosa, S.M., Scotto, M.G., Alonso, A.M. (2011). Summarising changes in air temperature over Europe by quantile regression and clustering. *Natural Hazards and Earth System Sciences* **11**, 3227-3233.
- [5] Barbosa, S.M., Silva, M.E., Fernandes, M.J. (2009). Multi-scale variability patterns in NCEP/NCAR reanalysis sea-level pressure. *Theoretical and Applied Climatology* **96**, 319-326.
- [6] Brandmaier, A.M. (2015). **pdcc**: an R package for complexity-based clustering of time series. *Journal of Statistical Software* **62**, 1-23.
- [7] Caiado, J., Crato, N., Peña, D. (2006). A periodogram-based metric for time series classification. *Computational Statistics and Data Analysis* **50**, 2668-2684.
- [8] Cilibrasi, R., Vitányi, P.M.B. (2005). Clustering by compression. *IEEE Transactions on Information Theory* **51**, 1523-1545.
- [9] Clark, P.U., Church, J.A., Gregory, J.M., Payne, A.J. (2015). Recent progress in understanding and projecting regional and global mean sea level change. *Current Climate Change Reports* **1**, 224-246.
- [10] D'Urso, P., De Giovanni, L., Maharaj, E.A., Massari, R. (2014). Wavelet-based self-organizing maps for classifying multivariate time series. *Journal of Chemometrics* **28**, 28-51.
- [11] D'Urso, P., Maharaj, E.A. (2012). Wavelets-based clustering of multivariate time series. *Fuzzy Sets and Systems* **193**, 33-61.
- [12] D'Urso, P., Maharaj, E.A. (2009). Autocorrelation-based fuzzy clustering of time series. *Fuzzy Sets and Systems* **160**, 3565-3589.
- [13] De Luca, G., Zuccolotto, P. (2011). A tail dependence-based dissimilarity measure for financial time series clustering. *Advances in Data Analysis and Classification* **5**, 323-340.

- [14] De Luca, G., Zuccolotto, P. (2015). Dynamic tail dependence clustering of financial time series. *Statistical Papers* (no prelo).
- [15] Durante, F., Pappadà R., Torelli, N. (2015). Clustering of time series via non-parametric tail dependence estimation. *Statistical Papers* **56**, 701-721.
- [16] Galeano, P., Peña, D. (2000). Multivariate analysis in vector time series. *Resenhas do Instituto de Matemática e Estatística da Universidade de São Paulo* **4**, 383-403.
- [17] Gouveia, S., Scotto, M.G., Monteiro, A., Alonso, A. (2015). Wavelets-based clustering of air quality monitoring sites. *Environmental Monitoring and Assessment* **187**:694, 1-14.
- [18] Heim, Jr., R.R. (2015). An overview of weather and climate extremes - products and trends . *Weather and Climate Extremes* **10**, 1-9.
- [19] Kakizawa, Y., Shumway, R.H., Taniguchi, M. (1998). Discrimination and clustering for multivariate time series. *Journal of the American Statistical Association* **93**, 328-340.
- [20] Kalpakis, K., Gada, D., Puttagunta, V. (2001). Distance measures for effective clustering of ARIMA time-series. In Cercone, N., Lin, T.Y., Wu, X. editors, *Proceedings 2001 IEEE International Conference on Data Mining*, pp. 273-280.
- [21] Keogh, E., Lonardi, S., Ratanamahatana, C.A., Wei, L., Lee, S.H., Handley, J. (2007). Compression-based data mining of sequential data. *Data Mining and Knowledge Discovery* **14**, 99-129.
- [22] Lafuente-Rego, B., Vilar, J.A. (2015). Clustering of time series using quantile autocovariances. *Advances in Data Analysis and Classification* (no prelo).
- [23] Liu, S., Maharaj, E.A., Inder, B. (2014). Polarization of forecast densities: a new approach to time series classification. *Computational Statistics and Data Analysis* **70**, 345-361.
- [24] Maharaj, E.A. (1996). A significance test for classifying ARMA models. *Journal of Statistical Computation and Simulation* **54**, 305-331.
- [25] Maharaj, E.A. (2000). Clusters of time series. *Journal of Classification* **17**, 297-314.
- [26] Maharaj, E.A. (2002). Comparison of non-stationary time series in the frequency domain. *Computational Statistics and Data Analysis* **40**, 131-141.
- [27] Monteiro, A., Carvalho, A., Ribeiro, I., Scotto, M.G., Barbosa, S., Alonso, A., Baldasano, J.M., Pay, M.T., Miranda, A.I., Borrego, C. (2012). Trends in ozone concentrations in the Iberian Peninsula by quantile regression and clustering. *Atmospheric Environment* **56**, 184-193.
- [28] Piccolo, D. (1990). A distance measure for classifying ARIMA models. *Journal of Time Series Analysis* **11**, 153-164.
- [29] Scotto, M.G., Alonso, A.M., Barbosa, S.M. (2010). Clustering time series of sea levels: extreme value approach. *Journal of Waterway, Port, Coastal, and Ocean Engineering* **136**, 215-225.
- [30] Scotto, M.G., Barbosa, S.M., Alonso, A.M. (2009). Model-based clustering of Baltic sea-level. *Applied Ocean Research* **31**, 4-11.
- [31] Scotto, M.G., Barbosa, S.M., Alonso, A.M. (2011). Extreme value and cluster analysis of European daily temperature series. *Journal of Applied Statistics* **38**, 2793-2804.
- [32] Scotto, M.G., Barbosa, S.M., Alonso, A.M. (2013). Extreme-based clustering of environmental time series. *Boletín de Estadística e Investigación Operativa* **29**, 92-102.

- [33] Scotto, M.G., Gouveia, S., Carvalho, A., Monteiro, A., Martins, V., Flannigan, M., San Miguel-Ayanz, J., Miranda, A.I., Borrego, C. (2014). Area burned in Portugal over recent decades: an extreme value analysis. *International Journal of Wildland Fire* **23**, 812-824.
- [34] Vilar, J.A., Alonso, A.M., Vilar, J.M. (2010). Non-linear time series clustering based on non-parametric forecast densities. *Computational Statistics and Data Analysis* **54**, 2850-2865.



# A dinâmica TGARCH de potência na evolução temporal da série das manchas solares de Wolfer

Esmeralda Gonçalves, *esmerald@mat.uc.pt* Nazaré Mendes Lopes, *nazare@mat.uc.pt*

*CMUC, Departamento de Matemática, Universidade de Coimbra, Portugal*

## 1 Introdução

A procura de modelos estocásticos que permitam descrever, o mais adequadamente possível, sistemas evoluindo com o tempo tem sido objecto de investigação desde o início do século XX, em particular com a introdução em Yule ([12]) dos processos auto-regressivos. Até ao início dos anos 80 estes estudos foram dominados pelos modelos lineares, centrados nos processos auto-regressivos médias-móveis (ARMA).

A formulação linear presente nestes modelos revelou-se, no entanto, insuficiente para descrever alguns dados temporais, quer de natureza financeira ou monetária quer mesmo de natureza física. Com efeito, este tipo de séries temporais é em geral descrito por distribuições não Gaussianas e apresenta, entre outros factos estilizados, características de dinâmica não linear, assimetria e memória longa cuja tentativa de explicação levou ao grande desenvolvimento, registado nos últimos anos, desta área científica ([4], [11]).

Por exemplo, em muitas destas séries de dados constata-se que a sua volatilidade condicional depende fortemente do passado. Com vista a descrever este facto, vários modelos condicionalmente heteroscedásticos apareceram na literatura na sequência do trabalho de Engle ([3]), naturalmente generalizado por Bollerslev ([1]), dando lugar à classe geral de modelos GARCH. Um outro facto frequentemente encontrado naquele tipo de séries temporais é a reacção assimétrica da volatilidade condicional ao sinal das observações passadas, concretamente o seu comportamento diverso consoante decorre um período de crescimento ou de queda dos valores da série. Esta característica é tida em conta pelos modelos auto-regressivos condicionalmente heteroscedásticos com níveis, abreviadamente modelos TGARCH, nos quais o desvio padrão condicional do processo no instante  $t$  é uma função linear por intervalos dos valores positivos e negativos das observações passadas ([13]). Analogamente, a observação da presença de memória longa na variância condicional levou ao aparecimento de modelos condicionalmente heteroscedásticos de potência  $\delta$ ,  $\delta \in \mathbb{R}^+$ , inicialmente propostos por Ding, Granger e Engle ([2]) e recentemente estudados por Pan, Wang e Tong ([10]).

Uma extensão natural dos modelos GARCH que consegue incorporar quer a assimetria quer a propriedade de memória longa observadas na volatilidade condicional é obtida considerando os modelos TGARCH com potência real  $\delta$ , abreviadamente modelos  $\delta$ -TGARCH ([5]). Além disso, os modelos desta classe geral são também adequados para dados temporais não normais com caudas pesadas pois são em geral processos estocásticos leptocúrticos.

A análise probabilista da estrutura destes modelos, no que concerne em particular à sua estacionaridade e ergodicidade, é imprescindível para desenvolvimentos estatísticos e aplicações a dados reais. Neste sentido, começamos neste artigo por uma abordagem teórica aos modelos TGARCH de potência  $\delta$ , apresentando a sua definição e propriedades probabilistas relevantes. A sua importância prática será ilustrada com a procura de um modelo adequado para a descrição da série anual do número de manchas solares de Wolfer no período 1700-2009. O ajustamento de modelos ARMA a esta série temporal tem



tido objecto de vários estudos desde Yule ([12]) mas, tanto quanto é do nosso conhecimento, a análise sobre a presença da volatilidade condicional nos resíduos da estimação não foi ainda considerada. Ora, o nosso estudo indica, como veremos, que esta série de natureza física, composta por valores médios anuais, apresenta forte volatilidade condicional e que a sua evolução temporal é bem descrita por um modelo misto do tipo AR com erros  $\delta$ -TGARCH.

## 2 Processos TGARCH de potência $\delta$

### 2.1 Definição

Seja  $X = (X_t, t \in \mathbb{Z})$  um processo estocástico real e, para  $t \in \mathbb{Z}$ , consideremos  $X_t^+ = X_t \mathbb{I}_{\{X_t \geq 0\}}$ ,  $X_t^- = -X_t \mathbb{I}_{\{X_t < 0\}}$  e  $\underline{X}_t$  a  $\sigma$ -álgebra gerada por  $(X_{t-i}, i \geq 0)$ .

Dizemos que o processo estocástico  $X = (X_t, t \in \mathbb{Z})$  segue um modelo condicionalmente heteroscedástico com níveis generalizado de potência  $\delta$  ( $\delta$ -TGARCH) com ordens  $p$  e  $q$  ( $p, q \in \mathbb{N}$ ) se, para todo  $t \in \mathbb{Z}$ , se tem

$$\begin{cases} X_t = \sigma_t \varepsilon_t \\ \sigma_t^\delta = \omega + \sum_{i=1}^p \left[ \alpha_i (X_{t-i}^+)^{\delta} + \beta_i (X_{t-i}^-)^{\delta} \right] + \sum_{j=1}^q \gamma_j \sigma_{t-j}^{\delta} \end{cases}$$

para constantes reais  $\delta \neq 0$ ,  $\omega > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_i \geq 0$ ,  $i = 1, \dots, p$ ,  $\gamma_j \geq 0$ ,  $j = 1, \dots, q$ , e onde  $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$  é uma sucessão de variáveis aleatórias reais independentes e identicamente distribuídas tais que  $\varepsilon_t$  é independente de  $\underline{X}_{t-1}$ . Se  $\delta < 0$  tomamos, por convenção,  $(X_t^+)^{\delta} = 0$  se  $X_t \leq 0$  e  $(X_t^-)^{\delta} = 0$  se  $X_t \geq 0$ , para  $t \in \mathbb{Z}$ . O processo  $\varepsilon$  denomina-se processo gerador de  $X$ .

Se  $\gamma_j = 0$ ,  $j = 1, \dots, q$ , o modelo  $\delta$ -TGARCH( $p, q$ ) é denotado por  $\delta$ -TARCH( $p$ ).

Observamos que nestes processos, a equação de evolução envolve não só o desvio padrão condicional mas também os momentos de ordem  $\delta$  de  $X$ . Notemos ainda que esta formulação inclui os principais modelos condicionalmente heteroscedásticos, nomeadamente os modelos GARCH ([3], [1]), TGARCH ([13]),  $\delta$ -GARCH,  $\delta > 0$  ([9]) e APARCH ([2]).

### 2.2 Estrutura probabilista

Esta classe de modelos GARCH com níveis e de potência foi considerada por Pan, Wang and Tong (2008) com  $\delta > 0$ , tendo sido estabelecidas condições de estacionaridade estrita e de existência de momentos. Considerando condições mais gerais sobre a potência, os coeficientes e o processo gerador (sem hipóteses sobre os momentos e não necessariamente simétrico) estabelecemos ([5]) para este modelo:

- i) a existência de uma única solução estritamente estacionária e ergódica
- ii) uma condição necessária e suficiente de existência do momento de ordem  $\delta$  sob a qual a estacionaridade estrita é verificada,
- iii) a estacionaridade até à ordem  $\delta$ .

Para estabelecer a existência e unicidade de uma solução estritamente estacionária e ergódica foi fundamental encontrar uma representação Markoviana do modelo de coeficientes aleatórios, envolvendo um processo estritamente estacionário e ergódico a partir do qual tal solução é deduzida. Analogamente a Mittnik, Paolella and Rachev ([9]), obtivemos a seguinte representação vectorial

$$Y_{t+1} = A_t Y_t + B$$

considerando o processo estocástico vectorial  $Y = (Y_t, t \in \mathbb{Z})$  de dimensão  $m$ ,  $m = \max(p, q)$ , onde a componente de ordem  $k$ ,  $Y_t^{(k)}$ , é dada por

$$\begin{cases} Y_t^{(1)} = \sigma_t^\delta \\ Y_t^{(k)} = \sum_{i=k}^m \left[ \alpha_i (X_{t-i+k-1}^+)^{\delta} + \beta_i (X_{t-i+k-1}^-)^{\delta} + \gamma_i \sigma_{t-i+k-1}^\delta \right], k = 2, \dots, m, \end{cases}$$

com  $(A_t, t \in \mathbb{Z})$  uma sucessão de matrizes quadradas de ordem  $m$  independentes e identicamente distribuídas e  $B$  um vector determinista de  $\mathbb{R}^m$ , dados por

$$A_t = \begin{bmatrix} \sum_{i=1}^{m-1} \left[ \alpha_i (\varepsilon_t^+)^{\delta} + \beta_i (\varepsilon_t^-)^{\delta} + \gamma_i \right] e_i & I_{m-1} \\ \alpha_m (\varepsilon_t^+)^{\delta} + \beta_m (\varepsilon_t^-)^{\delta} + \gamma_m & 0_{m-1}^T \end{bmatrix}, \quad B = \begin{bmatrix} \omega e_1 \\ 0 \end{bmatrix};$$

e em que  $\{e_1, \dots, e_{m-1}\}$  é a base canónica de  $\mathbb{R}^{m-1}$ ,  $I_{m-1}$  a matriz Identidade de ordem  $m-1$  e  $0_{m-1}$  o vector nulo de  $\mathbb{R}^{m-1}$ .

No teorema seguinte resumimos os resultados acima referidos destacando a simplicidade das condições envolvidas e o facto de se exprimirem em termos dos coeficientes do modelo.

**Teorema.** Suponhamos que  $E(|\varepsilon_t|^\delta) < +\infty$  e  $P(\varepsilon_t = 0) \neq 1$ . Sendo  $E[(\varepsilon_t^+)^{\delta}] = \phi_{1,\delta}$  e  $E[(\varepsilon_t^-)^{\delta}] = \phi_{2,\delta}$ , tem-se:

1.  $E(|X_t|^\delta)$  existe e é independente de  $t$  se e só se  $\sum_{i=1}^m (\alpha_i \phi_{1,\delta} + \beta_i \phi_{2,\delta} + \gamma_i) < 1$ . Além disso  $X$  é fracamente estacionário até à ordem  $\delta$ .
2. Se  $\sum_{i=1}^m (\alpha_i \phi_{1,\delta} + \beta_i \phi_{2,\delta} + \gamma_i) < 1$ , então  $X$  é estritamente estacionário e ergódico.

Concluimos este brevíssimo resumo sobre a estrutura probabilista dos processos  $\delta$ -TGARCH referindo que, sob condições de regularidade não restritivas, se obtém uma representação da volatilidade condicional como função dos valores passados do processo, concretamente,

$$\sigma_t^\delta = c_0 + \sum_{i=1}^{+\infty} c_i (X_{t-i}^+)^{\delta} + \sum_{i=1}^{+\infty} \tilde{c}_i (X_{t-i}^-)^{\delta}, \text{ quase certamente,}$$

com coeficientes  $c_i$  e  $\tilde{c}_i$  que decrescem exponencialmente para zero quando  $i$  tende para  $+\infty$ . A representação é única se, adicionalmente,  $\varepsilon_0^+$  e  $\varepsilon_0^-$  são variáveis aleatórias não degeneradas. Assim, mediante tais condições, existe uma representação  $\delta$ -TARCH( $+\infty$ ) para  $\sigma_t$  e, como os coeficientes desta representação decrescem exponencialmente para zero, então  $\sigma_t$  pode ser aproximado, de modo consistente, por uma amostra finita de  $X$ .

### 3 Aplicação a uma série de natureza física

A compreensão do universo tem sido uma preocupação permanente da humanidade. Sabemos que os primeiros registos de representações gráficas de observações temporais estão ligadas à Astronomia e que, em particular, o gráfico da Figura 1, representando a inclinação das órbitas dos planetas em função

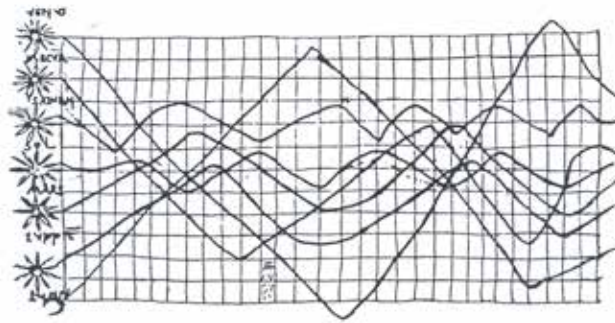


Figura 1: Inclinação das órbitas dos planetas do sistema solar (século X ou XI)

do tempo, data do século X ou XI e é considerado o mais antigo diagrama temporal do mundo ocidental ([8]).

Relembramos ainda que foi um conjunto de dados sobre o movimento dos planetas obtido por Brahe (1546-1601) que levou à formulação das leis de Kepler (1571-1630). A actividade solar tem também sido objecto de numerosos estudos entre os quais se destacam os ligados à evolução do número de manchas solares ou da área das regiões faculares ([7]) observadas em cada um dos hemisférios solares.

Para ilustrarmos a importância de considerar processos  $\delta-TGARCH$  para descrever a evolução dinâmica deste tipo de dados, consideramos a série anual das manchas solares no período 1700-2009. A análise será feita recorrendo a uma generalização da metodologia de Box-Jenkins que permitirá ter em conta as características de volatilidade condicional detectadas na série dos resíduos do modelo inicialmente retido com aquela metodologia.<sup>1</sup>

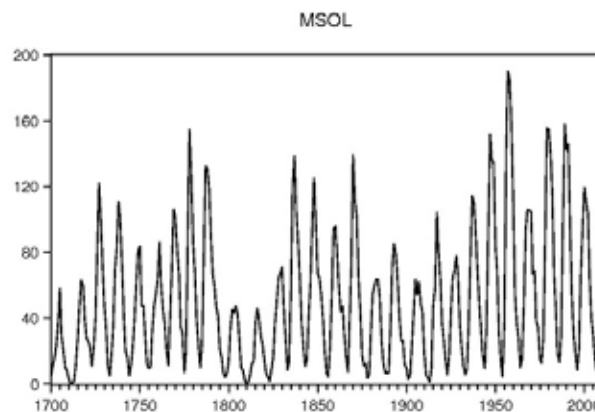


Figura 2: A série das manchas solares (1700 a 2009)

Com o objectivo de caracterizar a evolução temporal da série de dados em estudo, cuja trajectória se apresenta na Figura 2, começamos, de acordo com a metodologia de Box e Jenkins, por analisar as suas funções de autocorrelação e de autocorrelação parcial. Dos valores destas funções presentes na Figura 3, concluímos que a série é bem ajustada por um modelo AR(9).

<sup>1</sup>Na análise apresentada recorremos ao software estatístico Eviews.

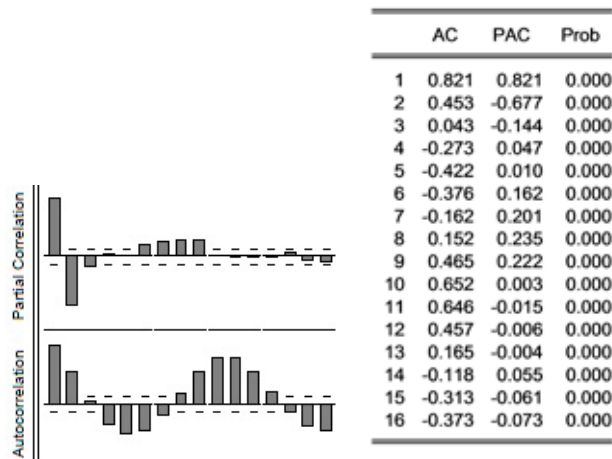


Figura 3: Autocorrelações e autocorrelações parciais empíricas da série das manchas solares

A estimação deste modelo AR produz uma série residual heteroscedástica, de acordo com o *output* presente na Figura 4. De facto, aplicando o teste ARCH-LM a este resíduo, a hipótese nula de homoscedasticidade é rejeitada com  $p$ -valor 0.0001.

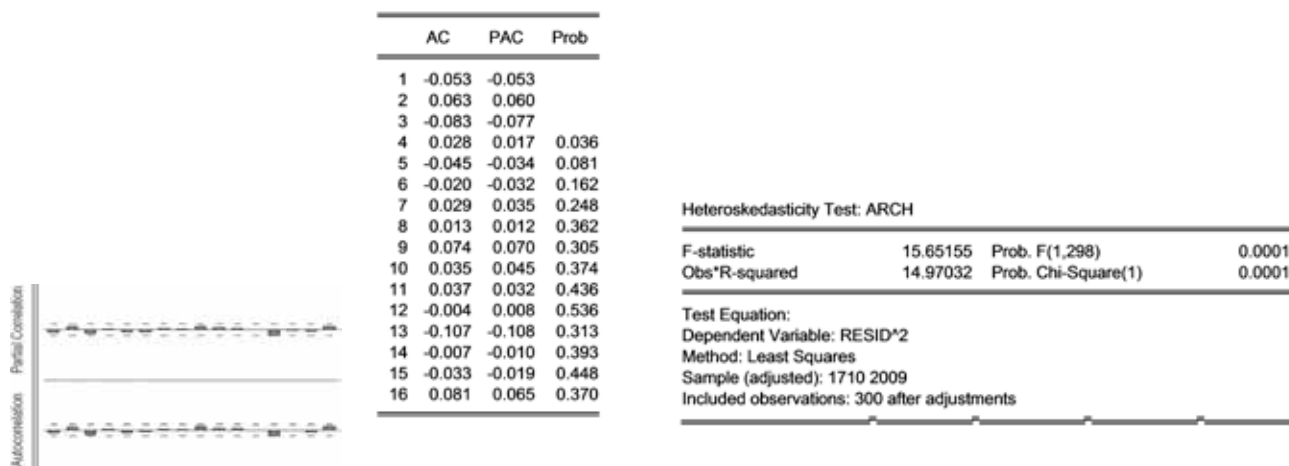


Figura 4: Autocorrelações e autocorrelações parciais empíricas do resíduo do modelo AR(9) estimado e *output* do teste ARCH-LM

Reanalisamos a série das manchas solares considerando a classe dos modelos  $AR(9)$  com erros  $\delta - TARCH(1)$ . Este valor da ordem condicionalmente auto-regressiva bem como a escolha de modelos com níveis são-nos sugeridos pelos valores das autocorrelações e autocorrelações parciais empíricas do módulo dos resíduos da estimação acima referida por um modelo  $AR(9)$  (presentes na Figura 5) que revelaram uma correlação ainda mais forte do que a observada nos quadrados dos resíduos da mesma estimação. Na Figura 6 encontram-se as estimativas dos parâmetros do novo modelo considerado.

A análise da série residual conduz à aceitação da hipótese de homoscedasticidade com  $p$ -valor igual a 0.8592, sendo o respectivo correlograma compatível com o de um ruído branco (Figura 7).

Apesar de as condições para a escolha de um bom modelo estarem razoavelmente cumpridas, entendemos procurar modelos do mesmo tipo mas com processo gerador  $\varepsilon$  não Gaussiano tendo as análises comparativas tido como base os critérios Akaike e Schwarz e a análise dos resíduos. Fixadas várias leis para  $\varepsilon$  fomos levadas a considerar, tendo em conta em particular a redução verificada nos critérios Akaike

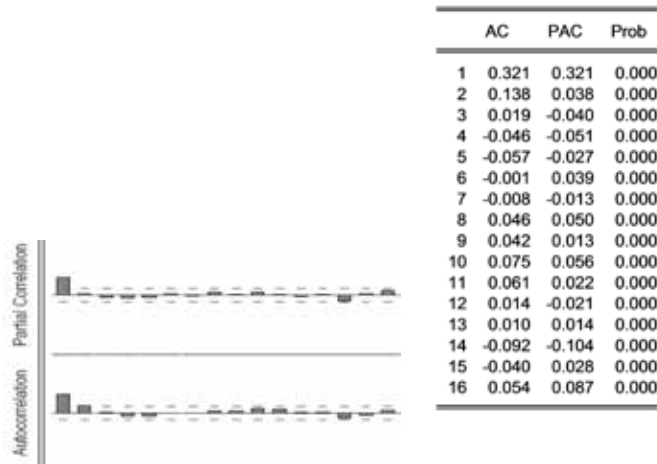


Figura 5: Autocorrelações e autocorrelações parciais empíricas do módulo dos resíduos do modelo AR(9) estimado

Dependent Variable: MSOL  
Method: ML - ARCH (Marquardt) - Normal distribution  
Sample (adjusted): 1709 2009  
Included observations: 301 after adjustments  
@SQRT(GARCH)\*C(8) = C(5) + C(6)\*(ABS(RESID(-1)) - C(7)\*RESID(-1))^C(8)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	48.62732	9.640444	5.044096	0.0000
AR(1)	1.215673	0.029010	41.90481	0.0000
AR(2)	-0.465296	0.034058	-13.66203	0.0000
AR(9)	0.180164	0.017169	10.49360	0.0000

Variance Equation				
C(5)	3.471423	3.772751	0.920130	0.3575
C(6)	0.334451	0.103940	3.217727	0.0013
C(7)	-0.607208	0.188537	-3.220628	0.0013
C(8)	0.579795	0.435436	1.331528	0.1830

R-squared	0.859250	Mean dependent var	50.39269
Adjusted R-squared	0.857828	S.D. dependent var	40.72569
S.E. of regression	15.35591	Akaike info criterion	8.148588
Sum squared resid	70033.82	Schwarz criterion	8.247116
Log likelihood	-1218.363	Hannan-Quinn criter.	8.188015
Durbin-Watson stat	2.043424		

Figura 6: Modelo AR(9) com erros TARCH(1) de potência  $\delta$

e Schwarz para 8.11 e 8.21 respectivamente, que a série de manchas solares  $S = (S_t)$  é bem descrita pelo modelo com a seguinte evolução:

$$S_t = 34.41 + 1.19S_{t-1} - 0.44S_{t-2} + 0.19S_{t-9} + \eta_t$$

onde o processo de erro  $\eta$  é tal que

$$\begin{cases} \eta_t = \sigma_t \varepsilon_t \\ \sigma_t^{0.81} = 5.72 + 0.58 (\eta_{t-i}^+)^{0.81} + 0.29 (\eta_{t-i}^-)^{0.81} \end{cases}$$

e  $(\varepsilon_t)$  são variáveis aleatórias reais independentes com distribuição de Student de parâmetro 10. A estacionaridade forte do modelo retido é assegurada pelo Teorema acima enunciado e pela análise das raízes do polinómio AR.

O modelo proposto conduz às três trajectórias presentes na Figura 8, concretamente a da série observada, da série estimada e do correspondente resíduo, sendo notória a qualidade do ajustamento obtido.

Para além da qualidade do ajustamento confirma-se, tal como em [7], a significativa volatilidade presente na actividade solar, sendo de realçar que são modelos de potência não inteira e processos geradores não

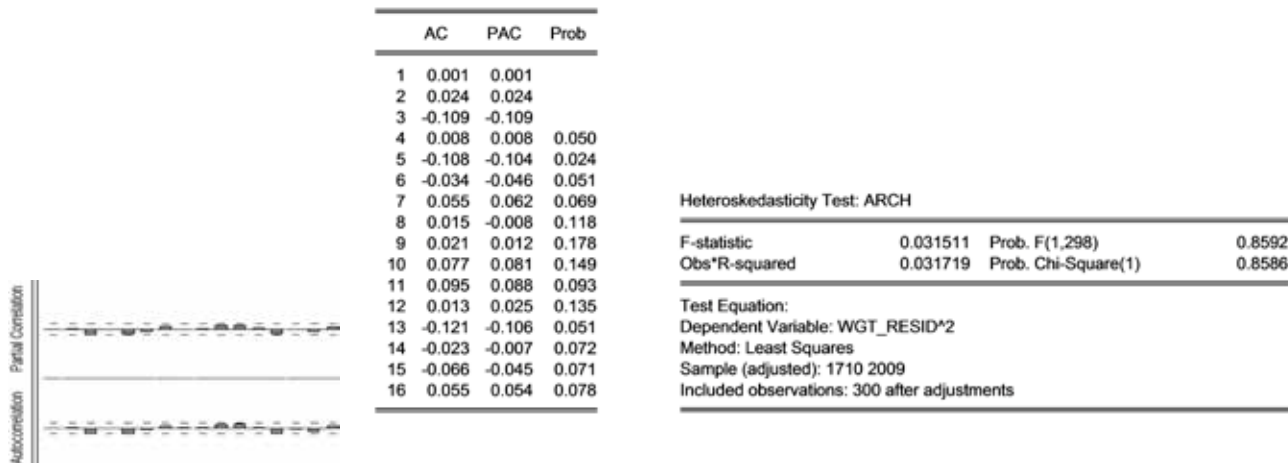


Figura 7: Autocorrelações e autocorrelações parciais empíricas do resíduo do modelo  $AR(9)$  com erros  $TARCH(1)$  de potência  $\delta$  e output do teste ARCH-LM

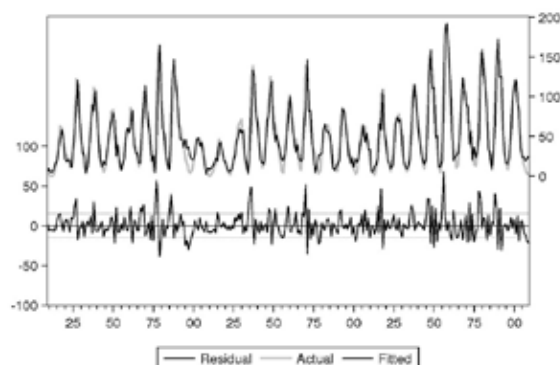


Figura 8: Trajectórias da série das manchas solares (cinza claro), da série ajustada (preto) e do resíduo

Gaussianos que melhor captam tal volatilidade. A presença natural de memória longa numa tal série não será porventura alheia a este facto.

Como curiosidade referimos finalmente um outro facto estilizado, inspirado no efeito de Taylor, frequentemente detectado em séries desta natureza e a que chamamos efeito de Taylor generalizado. Dizemos que uma série temporal estacionária  $X$  apresenta efeito de Taylor generalizado para a ordem  $h$ ,  $h \in \mathbb{Z}$ , se a autocorrelação empírica de ordem  $h$  da série  $X^2$  é, em módulo, menor do que o módulo da autocorrelação empírica de mesma ordem da série  $|X|$ :  $|\hat{\rho}_{X^2}(h)| < |\hat{\rho}_{|X|}(h)|$ ,  $h \in \mathbb{Z}$ . A Figura 9 revela a presença suave do efeito de Taylor na série das manchas solares, excepto nas ordens 8, 14, 19 e 30.

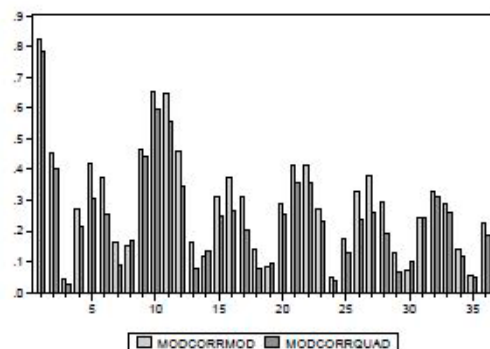


Figura 9: Módulos das autocorrelações empíricas das séries  $S$  (cinza claro) e  $S^2$  (cinza escuro)

## Referências

- [1] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307-327.
- [2] Ding, Z., Granger, C.W., Engle, R.F. (1993). A long memory property of stock market returns and a new model. *J. Empirical Finance* **1**, 83-106.
- [3] Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of the UK inflation. *Econometrica* **50**, 987-1008.
- [4] Francq, Ch., Zakoian, J.M. (2010). GARCH models, Structure, Statistical inference and Financial applications, Wiley.
- [5] Gonçalves, E., Leite, J., Mendes-Lopes, N. (2012). On the probabilistic structure of power threshold generalized ARCH stochastic processes. *Stat. Prob. Lett.* **82**, 1597-1609.
- [6] Gonçalves, E., Leite, J. and Mendes Lopes, N. (2014). On the probabilistic structure of power TGARCH models and applications to real data, *Book Series devoted to SMTDA2014*, Sally Mc-Clean, Raimondo Manca and Christos H. Skiadas Eds., 291-300.
- [7] Gonçalves, E., Mendes-Lopes, N., Dorotovic, I., Fernandes, J.M., Garcia, A. (2014). North and South Hemispheric Solar Activity for Cycles 21-23: Asymmetry and Conditional Volatility of Plage Region Areas. *Solar Physics* 289: 6, 2283-2296.
- [8] Kendall, G. (1973). Time Series, Griffin.
- [9] Mittnik, S., M.S. Paolella, S.T. Rachev (2002). Stationarity of stable power-GARCH processes. *Journal of Econometrics* **106**, 97-107.
- [10] Pan, J., H. Wang, H. Tong (2008). Estimation and tests for power-transformed and threshold GARCH models. *Journal of Econometrics* **142**, 352-378.
- [11] Turkman, K.F., Scotto, M.G., Bermudez P.Z. (2014). Non-linear time series, Extreme Events and Integer Value Problems, Springer.
- [12] Yule, G.U. (1927). On a method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers. *Philos. Trans. Royal Soc. London, Ser. A* **226**, 267-298.
- [13] Zakoian, J.M. (1994). Threshold heteroskedasticity models. *Journal of Economic Dynamics and Control* **18**, 931-955.



## • Artigos Científicos Publicados

- Bailey, R.A., Ferreira, S.S., Ferreira and D., Nunes, C. (2015). Estimability of Variance Components when all Model Matrices Commute. *Linear Algebra and its Applications*, 492, 144–160.
- Barbosa, S.M. Gouveia, S., Scotto, M.G. and Alonso, A.M. (2016). Wavelet-based clustering of sea level records. *Mathematical Geosciences* 48, 149-162.
- Bergel, A. e Egídio dos Reis, A. (2015) - Further developments in the Erlang(n) risk process, *Scandinavian Actuarial Journal*, 2015(1), 32-48, 2015
- Carvalho, F., Mexia, J. T., Santos, C. and Nunes, C. (2015). Inference for types and structured families of commutative orthogonal block structures. *Metrika*, 78, pp. 337–372.
- Ferreira, S.S., Nunes, C., Ferreira, D., Moreira, E., and Mexia, J.T. (2015). Estimation and Orthogonal Block Structure. *Hacettepe Journal of Mathematics and Statistics*, 45(58).
- Gouveia, S. Scotto, M.G., Pinna, G.D. Maestri, R., La Rovere, M.T. and Ferreira, P.J.S.G. (2015). Spontaneous baroreflex sensitivity for risk stratification of heart failure patients: optimal cut off and age effects. *Clinical Science* 129, 1163-1172.
- Kreer, M., Kizilersü, A., Thomas, A., e Egídio dos Reis, A. (2015) - Goodness-of-fit tests and applications for left-truncated Weibull distributions to non-life insurance. *European Actuarial Journal* 5, 139-163.
- Martinez, E. R., Cardoso, R. e Egídio dos Reis, A. (2015) - Some advances on the Erlang(n) dual risk model. *ASTIN Bulletin* 45, 127-150, 20125
- Menezes, R., Piai, H., Garcia-Soidán, P. and I.Sousa (2015). Spatial–temporal modellization of the NO<sub>2</sub> concentration data through geostatistical tools. *Journal of Statistical Methods & Applications*, 1-18.
- Scotto, M.G. Weiss, C.H. and Gouveia, S. (2015). Thinning-based models in the analysis of integer-valued time series: a review. *Statistical Modelling* 15, 590-618.
- Sepúlveda N, Paulino C.D., Drakeley C. (2015) - Sample size and power calculations for detecting changes in malaria transmission using antibody seroconversion rate. *Malaria Journal*, 14(1):529.
- Sepúlveda N, Stresman G, White MT, Drakeley CJ. (2015) - Current Mathematical Models for Analyzing Anti-Malarial Antibody Data with an Eye to Malaria Elimination and Eradication. *Journal of Immunology Research*.

## • Capítulos de Livros

- Caiado, J., E.A. Maharaj and P. D'Urso (2015). "Time Series Clustering", em *Handbook of Cluster Analysis*, C. Henning, M. Meila, F. Murtagh, R. Rocci (eds.), CRC Press, Taylor & Francis (em breve).

## • Livros

**Título:** Programa e Resumos. XXII Congresso SPE 2015

**Editores:** Clara Cordeiro, Conceição Ribeiro, Maria Helena Gonçalves, Nelson Antunes e Carlos Sousa.

Ano: 2015. Editora: Sociedade Portuguesa de Estatística. ISBN: 978-972-8890-36-0.

(Esta informação completa a que foi apresentada na página 67 do anterior Boletim SPE)



## • Teses de Mestrado

**Título:** Estimation of statistical cure from cancer using population-based data

**Autora:** Sara Elisabete Ferreira de Abreu, *sara\_elis5@hotmail.com*

**Orientadores:** Inês Sousa e Luís Antunes

**Título:** Avaliação de Conhecimentos sobre Cancro da Mama na região Porto Ocidental

**Autora:** João Firmino Domingues Barbosa Machado, *firmino.firminomachado@gmail.com*

**Orientadora:** Inês Sousa

**Título:** Os modelos lineares e em espaço de estados na análise dos dados de desemprego registados em Portugal

**Autora:** Benedita Raquel Fernandes Pereira, *benedita\_raquel@hotmail.com*

**Orientadores:** Arminda Manuela Gonçalves e Marco Costa

**Título:** Estimação da função de sobrevivência em dados de eventos recorrentes com censura pela direita

**Autora:** Andreia Catarina de Sousa Gonçalves, *andriacsgoncalves@hotmail.com*

**Orientador:** Luís Machado

## • Teses de Doutoramento

**Título:** Técnicas multivariadas de redução de dados: comparação, aplicabilidade e convergência de resultados

**Autor:** Fernando José do Nascimento Sebastião, *fsebast@ipleiria.pt*

**Orientadores:** Irene Oliveira e Jorge Cadima

Na minha tese de Doutoramento foram abordadas várias metodologias consagradas tais como a Análise em Componentes Principais (ACP), a Análise em Componentes Independentes (ACI), a Análise Espectral Singular (SSA) e a Análise Espectral Singular Multicanal (MSSA), assim como uma nova metodologia.

É muito comum a análise de dados recorrendo a técnicas multivariadas, nomeadamente para séries temporais. A ACI tem sido considerada mais adequada que a ACP na análise de séries temporais, principalmente quando se admite a não normalidade e se exige que as componentes a estimar sejam independentes. Neste sentido, como alternativa à MSSA, foi proposta uma nova abordagem baseada na ACI aplicada à matriz dos desfasamentos, que é designada *Lag*-Análise em Componentes Independentes (*Lag*ACI). Tal abordagem foi desenvolvida e suportada teoricamente pelos conceitos algébricos e algoritmos existentes das duas técnicas envolvidas. Foram dados exemplos de aplicação, com especial relevância para dados climáticos.

O principal objetivo da tese foi interligar o novo método com os restantes métodos estatísticos multivariados referidos, para permitir avaliar as semelhanças e diferenças entre estes. Analisaram-se diferentes abordagens em cada método e estudaram-se as implicações da escolha da matriz informativa das estruturas subjacentes aos dados, tendo em conta se estes estavam ou não sujeitos a algum tipo de pré-processamento, como por exemplo a centragem. Foram ainda considerados alguns indicadores comparativos de avaliação do desempenho de modelos para averiguar o grau de semelhança nos resultados transversais às técnicas em estudo.

Considerou-se que a nova abordagem representa uma contribuição para novos desenvolvimentos teóricos e aplicações no campo da análise de séries temporais.

Fernando Sebastião

**Título:** Estimação de parâmetros em modelos estocásticos de estruturas com comportamento dinâmico linear e quasi linear

**Autora:** Ana Prior, *anafpontes@gmail.com*

**Orientadora:** Paula Milheiro de Oliveira

A minha tese tem como objetivo a investigação de estimadores de parâmetros de equações diferenciais estocásticas que servem usualmente para modelar o comportamento dinâmico linear ou quasi linear de estruturas. Para além da obtenção de estimadores, pretende-se estudar as suas propriedades.

A dissertação está dividida em 3 partes. Na Parte I abordo o problema da estimação da matriz de deriva de um modelo linear estocástico homogéneo de dimensão  $2n$  com coeficientes constantes, observado em tempo contínuo e sendo a matriz de difusão singular. A matriz de deriva é uma matriz por blocos figurando nos blocos superiores,  $n \times n$ , a matriz nula e a matriz identidade  $e$ , nos blocos inferiores,  $n \times n$ , as matrizes  $M^{-1}K$  e  $M^{-1}C$ , sendo  $M$  a matriz de massa invertível,  $K$  a matriz de rigidez e  $C$  a matriz de amortecimento. Descrevo o estimador de máxima verosimilhança da matriz de deriva e apresento uma demonstração da sua propriedade da distribuição assintótica normal com recurso a técnicas da Transformada de Laplace. Estudo também a convergência da matriz de covariância deste estimador e explico a matriz de Informação de Fisher num caso em que se verifica uma condição de comutatividade da multiplicação de matrizes do modelo. Este estudo é acompanhado de simulações que ilustram e complementam os resultados teóricos obtidos.

Na Parte II trato o problema de estimação da matriz de deriva de dimensão 2 do mesmo modelo no caso em que ocorre uma mudança de regime nesta matriz. Esta mudança de regime acontece quando, num dado instante (conhecido ou desconhecido), o coeficiente de rigidez muda de um determinado valor  $k_1 > 0$  para  $k_2 > 0$  o que nos conduz a um modelo linear por troços. A equação diferencial estocástica em estudo, nestas condições, modela o movimento vibratório de estruturas sujeitas a ações aleatórias com mudança de regime. Em particular, descreve o cenário de uma estrutura que sofreu uma diminuição de rigidez após um determinado período de tempo. Considerando as observações em tempo discreto, descrevo uma forma de obter estimativas de máxima verosimilhança dos parâmetros  $k_1$ ,  $k_2$  e  $c$  (parâmetro de amortecimento) do modelo estocástico antes e depois da mudança de regime. No caso em que a mudança de regime ocorre num instante desconhecido, este passa a ser outro dos parâmetros a estimar. Mostro também como realizar um teste de hipóteses para a deteção de mudança de regime, no caso em que se desconhece se efetivamente esta ocorreu. Apresento estudos baseados em simulações que ilustram o procedimento e permitem analisar a influência dos parâmetros no desempenho dos testes e probabilidades de erro do teste.

Na Parte III, analiso o problema da estimação da matriz de deriva do modelo de dimensão 2 considerando um movimento browniano fracionário com parâmetro de Hurst  $H \in (0, \frac{1}{2})$ . Obtive o estimador de máxima verosimilhança dos parâmetros de rigidez e de amortecimento após transformação, assim como algumas das suas propriedades. O estudo é acompanhado da análise da matriz de covariância assintótica e são apresentadas simulações com o objetivo de ilustrar o comportamento do estimador obtido.

Ana Prior

**Título:** Análise de Variância com Amostras de Dimensão Aleatória e suas Aplicações.

**Autor:** Gilberto Capistrano Cunha de Andrade, *gilbertocapistrano@ubi.pt*

**Orientadores:** Célia Nunes e Dário Jorge da Conceição Ferreira

A minha tese apresenta uma extensão da análise de variância (ANOVA) ao caso em que as dimensões das amostras não são conhecidas.

A ANOVA é utilizada em muitas áreas de investigação, nomeadamente em investigação médica, agricultura ou psicologia, para citar apenas algumas, onde as dimensões das amostras podem não ser previamente conhecidas. Esta situação ocorre com frequência quando o intervalo de tempo para a recolha das observações é fixado à partida. Um bom exemplo corresponde à recolha de observações para um estudo onde se pretende comparar várias patologias de pacientes que chegam às urgências de um hospital num determinado período de tempo.

Neste trabalho estendeu-se a ANOVA, com um e mais fatores, ao caso em que as dimensões das amostras são desconhecidas, devendo por isso ser tratadas como realizações de variáveis aleatórias independentes. Defendemos que esta abordagem deve ser baseada na escolha adequada da distribuição destas variáveis, pelo que foram consideradas duas situações distintas:

- No primeiro caso assumimos que as variáveis aleatórias seguiam distribuições de Poisson, situação em que a ocorrência das observações corresponde a processos de contagem e não existem limites superiores para as dimensões das amostras (tal como ilustrado no exemplo anterior, referente à comparação de patologias);
- No segundo caso, considerámos a distribuição Binomial, quando existe um limite superior para as dimensões das amostras, que nem sempre é atingido uma vez que podem ocorrer falhas nas observações.

Como resultados, foram obtidas as estatísticas de teste e suas distribuições, condicional e não condicional assumindo as dimensões das amostras como aleatórias, para modelos de efeitos fixos, modelos de efeitos aleatórios e modelos mistos.

Adicionalmente, foram apresentadas várias aplicações com dados reais referentes a registros do cancro no Brasil, que nos permitiram ilustrar a utilidade da nossa abordagem assim como comparar os resultados obtidos com os da ANOVA usual.

Gilberto Capistrano Cunha de Andrade



## PRÉMIOS “ESTATÍSTICO JÚNIOR 2016”

Está aberto, até 1 de Junho de 2016, o concurso para atribuição de prémios “Estatístico Júnior 2016”, de acordo com o seguinte regulamento:

A atribuição de prémios “Estatístico Júnior 2016” é promovida pela Sociedade Portuguesa de Estatística (SPE), com o apoio da Porto Editora, e tem como objectivo estimular e desenvolver o interesse dos alunos do Ensino Básico e Secundário pelas áreas das Probabilidades e Estatística.

Os candidatos aos prémios “Estatístico Júnior 2016” devem ser alunos do 3.º Ciclo do Ensino Básico, do Ensino Secundário, dos Cursos de Educação e Formação (CEF) ou dos Cursos de Educação e Formação de Adultos (EFA), no ano letivo 2015-2016.

As candidaturas podem ser individuais ou em **grupo com um máximo de 3 alunos**. Do grupo pode ainda fazer parte um professor, do grau de ensino em que o trabalho se insere, ao qual caberá o papel de orientador.

Os candidatos devem apresentar um trabalho cuja temática deve estar relacionada com as Probabilidades ou a Estatística.

O **trabalho** deverá ser constituído por um texto escrito em Português com um máximo de 10 páginas A4 dactilografadas e um *poster* formato A2 que resuma os principais aspetos do trabalho.

Poderão ser atribuídos prémios “Estatístico Júnior 2016” a sete trabalhos: aos três primeiros classificados de entre os trabalhos candidatos do 3.º Ciclo do Ensino Básico, aos três primeiros classificados de entre os trabalhos candidatos do Ensino Secundário e um primeiro classificado de entre os trabalhos candidatos dos Cursos CEF ou EFA. Os prémios são constituídos por lotes de livros a selecionar das notas de encomenda da Porto Editora (à exceção de manuais escolares e livros auxiliares), no valor de 500€ para os classificados em primeiro lugar e de 200€ para os classificados em segundo e terceiro lugares.

Ao professor orientador do trabalho classificado em 1º lugar, em cada grau de ensino, é atribuída uma anuidade grátis como sócio da SPE, ajudas de custo para participação na Sessão de Entrega do Prémio e lotes de livros a selecionar das notas de encomenda da Porto Editora (à exceção de manuais escolares e livros auxiliares), no valor de 350€.

Aos grupos proponentes dos trabalhos classificados em 1º lugar será também oferecida uma ampliação do correspondente *poster* que será exposto na Sessão de Entrega do Prémio.

A candidatura é composta pelo **Boletim de Candidatura**, devidamente preenchido, e pelo **trabalho** (poster e texto). A candidatura, dirigida ao Presidente da SPE, deverá ser enviada para:

*impressa em papel para efeitos da avaliação* Sociedade Portuguesa de Estatística –  
Bloco C6, Piso 4 – Campo Grande – 1749-016 Lisboa  
*em formato digital (pdf) por e-mail:* [spe@fc.ul.pt](mailto:spe@fc.ul.pt)

O carimbo do correio validará a data de entrega do trabalho, sendo os autores notificados sobre a sua receção no prazo de uma semana.

A admissibilidade e apreciação dos trabalhos submetidos a concurso é da competência de um júri, cuja constituição e nomeação será da responsabilidade da Direção da SPE.

O júri é soberano nas decisões, não havendo lugar a impugnação ou recurso.

A atribuição dos prémios “**Estatístico Júnior 2016**” será anunciada logo que conhecida a decisão do júri e a sua entrega formal será realizada numa sessão expressamente dedicada a essa entrega.

Os prémios “**Estatístico Júnior 2016**” poderão não ser atribuídos.

O boletim de candidatura e este regulamento podem ser obtidos em

<http://www.spestatistica.pt/BoletimCandidaturaPEJ16.pdf>

<http://www.spestatistica.pt/RegulamentoPEJ16.pdf>

**Apoio**  Porto  
Editora

## Retrospectiva

O *Boletim SPE* através dos seus “Tema Central”

---

Outono de 2015 - Destaque: Estatística em Genética

---

Primavera de 2015 – Destaque: Estatística no Desporto

---

Outono de 2014 – Destaque: Estatística no Ensino Básico e Secundário

---

Primavera de 2014 – Destaque: (Um) Ano Internacional da Estatística

---

Outono de 2013 – Destaque: A "Escola Bayesiana" em Portugal

---

Primavera de 2013 – Destaque: Estatística não-paramétrica

---

Outono de 2012 – Destaque: Métodos Estatísticos em Medicina

---

Primavera de 2012 – Destaque: Estatística no Ensino Superior Politécnico

---

Outono de 2011 – Destaque: Análise de Sobrevivência

---

Primavera de 2011 – Destaque: Sondagens e Censos

---

Outono de 2010 – Destaque: Estatística Espacial

---

Primavera de 2010 – Destaque: Data Mining - Prospecção (Estatística) de Dados

---

Outono de 2009 – Destaque: Modelos Econométricos

---

Primavera de 2009 – Destaque: Investigação (em) Estatística

---

Outono de 2008 – Destaque: Processos Estocásticos

---

Primavera de 2008 – Destaque: ALEA - Um sítio do nosso mundo

---

Outono de 2007 – Destaque: Bioestatística

---

Primavera de 2007 – Destaque: A "Escola de Extremos" em Portugal

---

Outono de 2006 – Destaque: Ensino e Aprendizagem da Estatística

---

também disponíveis em <http://www.spestatistica.pt/index.php/publicacoes-57/boletins>



**SOCIEDADE PORTUGUESA  
DE ESTATÍSTICA**

**[www.spestatistica.pt](http://www.spestatistica.pt)**



**O MUNDO DA  
ESTATÍSTICA**

**ORGANIZAÇÃO PARTICIPANTE**

# Índice

Editorial .....	1
Mensagem da Presidente .....	3
Notícias .....	4
<i>Enigmística</i> .....	10
<b><i>SPE e a Comunidade</i></b>	
Sondagens e Estatística	
<i>Rui Oliveira Costa</i> .....	11
Estatística, inquéritos de opinião e cidadania	
<i>António Salvador</i> .....	13
Sondagens - Falham e acertam no que mostram, (...), e nos objetivos a que se propõem	
<i>João H. C. António</i> .....	17
Avaliação temporal do conhecimento matemático (...) à entrada do ensino superior	
<i>Milton Severo, Paulo Trindade e A. Rita Gaio</i> .....	24
<b><i>Séries Temporais e suas aplicações</i></b>	
Modelação de séries temporais de contagem	
<i>Maria Eduarda Silva</i> .....	33
Séries temporais em Economia	
<i>Paulo Teles e André Almeida</i> .....	41
Agrupamento de séries temporais e sua aplicação na análise de processos geofísicos e ambientais	
<i>Manuel G. Scotto e Susana M. Barbosa</i> .....	75
A dinâmica TGARCH de potência na evolução temporal da série das manchas solares de Wolfer	
<i>Esmeralda Gonçalves e Nazaré Mendes Lopes</i> .....	86
<b><i>Ciência Estatística</i></b>	
<i>Artigos Científicos Publicados</i> .....	94
<i>Livros e Capítulos de Livros</i> .....	94
<i>Teses de Mestrado</i> .....	95
<i>Teses de Doutoramento</i> .....	95
Prémios “Estatístico Júnior 2016” .....	98
Retrospectiva .....	100