

**Faculdade de Ciências da
Universidade de Lisboa**

Departamento de Matemática



**Instituto Superior de Ciências
do Trabalho e da Empresa**

Departamento de Finanças



A EVOLUÇÃO DO SETOR IMOBILIÁRIO

Bruno André Pereira de Sousa

Mestrado em Matemática Financeira

Dissertação orientada por:

Professora Doutora Diana Mendes, Professora Associada
ISCTE Business School, Departamento de Métodos Quantitativos para Gestão e
Economia

2020

Agradecimentos

Em primeiro lugar quero agradecer à professora Diana por toda a disponibilidade, acompanhamento e orientação que me deu ao longo deste projeto.

À minha namorada Cláudia por todo o carinho e apoio que me deu, e por estar presente nos momentos mais importantes da minha vida académica. Sem dúvida que tornou o meu percurso mais fácil. E feliz.

Quero agradecer também aos meus pais e à minha avó por nunca terem desistido do miúdo reguila, por toda a educação que me deram e por me motivarem todos os dias a ser um melhor ser humano. Acredito profundamente que sem a vossa ajuda nada disto era possível.

Por fim, agradeço também a todos os meus familiares, que estiveram sempre ao meu lado ao longo desta jornada.

Um obrigado a este grupo de pessoas, a vocês estarei eternamente grato.

Resumo

O presente projeto tem como objetivo compreender o comportamento histórico do setor imobiliário português e principalmente prever qual será o comportamento futuro do mesmo.

Partindo deste princípio, dividiu-se esta dissertação em três capítulos principais que, embora sejam independentes, têm o mesmo propósito, avaliar o comportamento do preço das casas em Portugal.

No primeiro capítulo é realizada a revisão de literatura que aborda a evolução da economia portuguesa, os fatores macroeconómicos que afetam o setor imobiliário e, por fim, o contexto atual de Portugal face aos efeitos/consequências da pandemia mundial.

No segundo capítulo são abordados conceitos estatísticos e econométricos considerados fundamentais para a compreensão da série temporal analisada.

No terceiro capítulo, com o auxílio de representações gráficas e de análises estatísticas, é elaborado o estudo empírico da série temporal – a média trimestral do preço das casas portuguesas. Esta análise consiste na avaliação da estacionariedade da série, seguindo-se a modelação da mesma através do modelo autorregressivo de médias móveis, terminando com a previsão dos preços que serão praticados em Portugal nos próximos dois anos.

Finalmente, é realizada uma síntese do trabalho desenvolvido através da interpretação dos resultados obtidos de maneira a cumprir o objetivo acima mencionado.

Palavras-Chave: Economia Portuguesa; Preço das casas em Portugal; Setor Imobiliário; Série Temporal; Previsão.

Abstract

The main goal of this project is to understand the historical behavior of the Portuguese real estate sector, and primarily, predict its future behavior.

From this premise, this thesis was split in three main chapters that, although they are independent, have the same purpose, the assessment of house prices in Portugal.

In the first chapter, a literature review regarding the evolution of the Portuguese economy, macroeconomic factors that affect the real estate sector and, finally, the current state of Portugal as a country due to the effects/consequences of the global pandemic was performed.

The second chapter addresses statistical and econometrics concepts considered fundamental to the understanding of the time series in discussion.

In the third chapter, resorting to graphical representations and statistical analysis, is developed the empirical study of the time series – the quarterly average of the Portuguese house prices. This analysis consists in the evaluation of the series stationarity, followed by its modeling through the autoregressive moving average model, ending with the forecast of prices that will be charged in Portugal over the next two years.

Finally, a synthesis of the work developed through the interpretation of the results obtained to meet the above-mentioned objective is carried out.

Keywords: Portuguese economy; House prices in Portugal; Real estate sector; Time Series; Forecast.

Índice

Introdução	2
1. Revisão de Literatura	4
1.1 Enquadramento Teórico.....	4
1.1.1 Mercado de Capitais Portugêses (PSI-20)	4
1.1.2 Principais catalisadores para o crescimento do Setor Imobiliário.....	7
1.1.3 Panorama Atual da Economia – Efeitos Pandemia	13
2. Metodologia.....	16
2.1 Conceitos Gerais das Séries Temporais.....	16
2.2 O que descreve uma série temporal?	18
2.3 Estacionariedade	21
2.3.1 Natureza da Série Temporal.....	24
2.4 Testes de Raiz Unitária.....	25
2.5 Modelos Lineares Univariados.....	28
2.5.1 Modelos Autorregressivos de ordem p – AR(p)	28
2.5.2 Modelos de Média Móvel de ordem q – MA(q)	30
2.5.3 Modelos Autorregressivos de Médias Móveis– ARMA(p,q)	31
2.5.4 Modelos Autorregressivos Integrados de Médias Móveis– ARIMA(p,d,q).....	31
2.5.5 Critérios de Seleção de Modelos	32
2.6 Verificação dos Pressupostos – Análise dos Resíduos	33
2.6.1 Ruído Branco.....	34
2.7 Previsão	34
2.7.1 Erro de Previsão	35
2.7.2 Previsão com Modelos ARIMA.....	36
3. Estudo Empírico	37
3.1 Análise Gráfica e Estatísticas Descritivas	37
3.2 Estacionariedade da série	40
3.3 Análise Gráfica e Estatísticas Descritivas – DIFLOG / Séries de Retornos.....	42
3.4 Estacionariedade da série dos retornos – Diferença dos Logaritmos (DIFLOG).....	44
3.5 Estimacão dos Modelos Econométricos.....	45
3.5.1 Identificacão do Modelo	45
3.5.2 Seleçãõ do Modelo	46
3.6 Análise dos resíduos dos modelos ARMA/ARIMA.....	47
3.7 Previsão da série.....	50

3.7.1 Previsão In-Sample.....	51
3.7.2 Análise Out-Of-Sample.....	52
Conclusão	55
Bibliografia	56

Índice de Figuras

Figura 1: Evolução do Mercado de Capitais Português PSI-20.....	4
Figura 2: Evolução da Euribor a 12 meses nos últimos 10 anos.....	9
Figura 3: Número de dormidas nos alojamentos turísticos por 100 habitantes	10
Figura 4: Taxa de Desemprego em Portugal nos últimos 10 anos	11
Figura 5: Taxa de Crescimento real do PIB nos últimos 10 anos (Taxa de Variação).....	12
Figura 6: Valor médios dos prédios urbanos transacionados.....	13
Figura 7: Exemplo de uma série estacionária.....	22
Figura 8: Exemplo de uma série em níveis não estacionária	23
Figura 9: Exemplo de uma série em diferenças sucessivas – 1ª diferença.....	23
Figura 10: Exemplo de uma série em diferenças sucessivas – 2ª diferença	23
Figura 11: Série preço real das casas em Portugal (Dados Trimestrais)	37
Figura 12: Histograma – Série do preço das casas em Portugal	39
Figura 13: Série dos retornos do preço real das casas - DIFLOG.....	42
Figura 14: Histograma – Série dos retornos do preço real das casas	43
Figura 15: Análise das funções ACF e PACF da série dos retornos.....	45
Figura 16: Comportamento dos resíduos.....	48
Figura 17: Previsão In-Sample da série dos retornos do preço real das casas.....	51
Figura 18: Previsão Out-of-Sample da série dos retornos do preço das casas	53

Índice de Tabelas

Tabela 1: Resultado das estatísticas descritivas obtidas na série do preço real das casas em Portugal.....	39
Tabela 2: Teste da Normalidade aplicado à série do preço real das casas em Portugal.....	39
Tabela 3: Resultado das estatísticas descritivas obtidas na série dos retornos do preço real das casas.....	43
Tabela 4: Teste da Normalidade aplicado à série dos retornos.....	43
Tabela 5: Testes de Estacionariedade aplicados à série dos retornos.....	44
Tabela 6: Critérios de Seleção Modelos ARMA/ARIMA aplicados à série dos retornos.....	46
Tabela 7: Primeiro Pressuposto dos Resíduos – Média Nula.....	48
Tabela 8: Segundo Pressuposto dos Resíduos – Homocedasticidade dos Resíduos.....	49
Tabela 9: Terceiro Pressuposto dos resíduos – Independência dos Resíduos Ljung-Box.....	50
Tabela 10: Quarto Pressuposto dos resíduos – Independência dos Resíduos Durbin-Watson.....	50
Tabela 11: Erro de previsão obtido no período in-sample.....	52
Tabela 12: Retornos preditos pelo modelo ARMA(1,1) para os próximos 2 anos.....	53
Tabela 13: Erro de previsão obtido na previsão out-of-sample.....	54

Introdução

Na última década, observou-se nos mercados financeiros um forte crescimento em praticamente todos os setores. De facto, após a crise de 2008, os Bancos Centrais como o BCE e a FED têm optado por políticas expansionistas de maneira a estimular o consumo, aumentar a liquidez e rejuvenescer a economia. Estas medidas resultaram num aumento significativo da procura e os mercados em geral lucraram com isso. A título ilustrativo, o índice Norte Americano S&P500 deste o início de 2009 até ao final de 2019 teve um crescimento notável de aproximadamente 265%.

Uma das grandes apostas dos bancos centrais para expandir a economia centra-se no corte de taxas de juro. Este estímulo concede às empresas e particulares um acesso mais facilitado ao dinheiro e, conseqüentemente, um aumento do consumo.

Contudo, não foi apenas o corte das taxas de juro que levou ao crescimento da economia global. Na verdade, este fator foi apenas a ponta de um iceberg que, ao longo desta década, tornou os mercados financeiros atrativos para os investidores.

O mercado imobiliário foi um dos mercados que mais aumentou na última década, nomeadamente em Portugal. Se se juntar ao baixo nível das taxas de juro, a diminuição da taxa de desemprego e, portanto, o aumento do rendimento das famílias portuguesas percebe-se porque é que este setor tem sido um dos que mais cresceu nos últimos anos. Mais, o aumento substancial do turismo e do número de estrangeiros a residir no país também são fatores importantes ao crescimento deste setor.

Não há dúvida que o cenário macroeconómico otimista no consumo, agregado aos fatores previamente mencionados, provocou um aumento da inflação no preço do imobiliário. Muitos analistas acreditam que o setor imobiliário Português se encontra numa bolha especulativa e que, em 2020, este mercado não irá abrandar uma vez que Portugal está a passar por uma fase de estabilidade política e recuperação económica. Do outro lado da moeda, a mais recente pandemia que se instalou no Mundo – o covid-19 – abrandou drasticamente as maiores economias mundiais tanto em termos de produção como de consumo. Os efeitos secundários que este vírus pode provocar na economia mundial ainda não estão a ser sentidos, contudo, muitos especialistas acreditam que podemos estar a caminhar para o início de uma nova recessão.

Será que o mercado imobiliário vai estagnar? Será que os preços das casas vão descer? O presente trabalho pretende dar resposta a estas questões e avaliar se, hipoteticamente, poderemos vir a observar uma quebra de rendimentos generalizada neste setor.

Deste modo, dividiu-se este projeto em três capítulos principais: a revisão de literatura, onde se faz um enquadramento teórico da economia portuguesa e da evolução do setor imobiliário; um capítulo dedicado às definições, onde são abordados os conceitos estatísticos e os modelos econométricos abordados; e, por fim, dedicou-se um capítulo ao estudo empírico, onde é estudada a série temporal do preço real das casas em Portugal, com o auxílio de métodos econométricos.

1. Revisão de Literatura

1.1 Enquadramento Teórico

1.1.1 Mercado de Capitais Português (PSI-20)

O PSI-20 ou *Portuguese Stock Index 20* é o principal índice de referência do mercado de capitais português. Este índice é composto pelas 18 maiores empresas cotadas na bolsa de valores de Lisboa e reflete a evolução do preço das ações dessas empresas ao longo do tempo.

O PSI-20 teve início a 31 de dezembro de 1992 com 3.000 pontos de cotação. Atualmente o índice está cotado ligeiramente acima dos 4.000 pontos sugerindo que, desde a sua criação, as empresas apresentaram um desempenho positivo nas suas contas (embora pouco expressivo). Contudo, importa referir que o índice português já atingiu 14.882 pontos em março de 2000.

A título meramente informativo - visto que não faz sentido comparar uma superpotência económica, como é o caso da América, com a realidade portuguesa - no horizonte temporal de existência do PSI-20, o índice S&P500 (que representa os 500 maiores ativos cotados na bolsa americana) subiu dos 379.10 pontos para perto dos 3.500 pontos de índice, dando aos investidores retornos acima dos 600%.

Dada a volatilidade registada nos pontos de índice do PSI-20 ao longo do tempo, torna-se importante realizar uma análise histórica à série de maneira a compreender melhor a evolução da economia portuguesa.



Figura 1: Evolução do Mercado de Capitais Português PSI-20

No início da década de 2000 a bolha especulativa causada em redor das empresas tecnológicas de internet teve o seu término. Consequentemente, os mercados mundiais inverteram a sua tendência de crescimento económico passando por uma fase de *bear market* (período de queda acentuada da economia suportado por um pessimismo generalizado dos investidores). O PSI-20 não foi exceção e, até meados de 2003, desvalorizou perto dos 10.000 pontos de índice, assinalando a primeira grande queda no mercado português. Esta crise que teve consequências no mundo inteiro e ficou conhecida por “ponto com”.

Após 2003 a economia portuguesa registou um crescimento constante até 2007. O índice português, a meados de 2007 conseguira recuperar quase na totalidade os pontos de índice que perdera com a crise do “ponto com”.

Contudo, em julho de 2007, nos Estados Unidos, surgiu uma nova crise. A crise do “*subprime*” que, posteriormente, deu origem à crise económico-financeira internacional de 2008, teve origem precisamente no mercado imobiliário e é, por este motivo, um tema incontornável nesta dissertação.

A Crise do “*Subprime*”

Na verdade, a crise de 2008 foi consequência de uma valorização continua do mercado imobiliário que teve início em 1994, ganhando a confiança dos investidores públicos e privados.

Em 1994, o governo americano criou duas entidades com o objetivo de garantir a sustentabilidade e a liquidez do mercado imobiliário norte americano. Estas entidades, ou *lenders* como eram conhecidas nos Estados Unidos, embora privadas, eram financiadas diretamente pelo estado americano e visavam claramente o lucro.

Com a forte aposta do estado americano no crescimento deste setor o leque de potenciais clientes a comprarem casa própria e a financiarem-na estendeu-se a todas as classes sociais, englobando minorias étnicas que não tinham condições para ter uma hipoteca e/ou uma casa própria (de frisar que o governo queria permitir que todos os indivíduos tivessem direito a uma casa própria).

Não obstante à venda de imóveis a pessoas com alto risco de incumprimento, os investidores a nível mundial acreditavam no mercado imobiliário americano – porque, historicamente este mercado tende sempre a crescer – e continuavam a comprar produtos financeiros constituídos por várias hipotecas.

Por consequência do colapso das empresas de internet em 2000 e dos acontecimentos do 11 de Setembro de 2001, a Reserva Federal dos Estados Unidos viu-se obrigada a baixar drasticamente as taxas de juro - o próximo subcapítulo irá aprofundar quais as consequências das alterações das taxas de juro no mercado imobiliário - aumentando a liquidez e o acesso facilitado a créditos. Desta forma, embora involuntariamente, a FED aumentou a confiança dos cidadãos nos pedidos mais facilitados de crédito onde qualquer um poderia financiar uma hipoteca com taxas de juro mais reduzidas, dando origem a uma bolha cada vez maior no setor imobiliário.

Em 2006, com uma economia solidificada que já não precisava de ser estimulada, a FED voltou a aumentar as taxas de juro. Consequentemente, o número de novas hipotecas baixou e o valor dos imóveis desvalorizou drasticamente. Nesta altura, as casas já não valiam aquilo que tinha sido concebido em termos de crédito e havia muito menos compradores para essas mesmas casas.

De notar que, neste momento a crise já se tinha espalhado por todo o mundo e os investidores públicos e privados que investiam direta ou indiretamente nestes produtos baseados em hipotecas de alto risco, estavam a ver os seus investimentos a perder o seu valor no mercado.

Em 2008, aquando da queda do grande banco americano *Lehman Brothers* e dos resgates aos ativos “tóxicos” na ordem dos 700 mil milhões de dólares que os Estado Unidos aprovaram, já se tomava como certo que o mundo estava perante uma grande crise financeira.

A avalanche de desaires estava imparável e atingiu sobretudo a Europa, evidenciando a debilidade bancária e os desequilíbrios das finanças públicas, com os Estados, famílias e empresas muito endividados.

O que de início se parecia com uma crise financeira transformou-se numa crise económica e, a partir de 2010, Portugal sofreu uma crise na dívida soberana levando mesmo à necessidade de recorrer ao resgate financeiro do país.

Em 2012, após intervenção da troika (e ainda sofrendo as consequências da crise de 2008), foi a vez dos bancos portugueses recapitalizarem-se e o Estado viu-se obrigado a fazer injeções de capital nos maiores bancos nacionais. Desde então, seguiram-se anos de mais injeções e resgates na banca portuguesa.

Atualmente, o país ainda não recuperou totalmente da crise de 2008 apesar da banca portuguesa ter tido um processo de transformação profundo. Os pontos de índices do PSI-20 apresentam valores abaixo dos que foram observados na queda de 2008, o que comprova a debilidade da economia portuguesa num contexto contemporâneo

1.1.2 Principais catalisadores para o crescimento do Setor Imobiliário

O mercado imobiliário tem vindo a aumentar a sua popularidade ao longo dos anos, simbolizando uma parte integral da economia mundial nos dias de hoje. Mais, os elevados retornos obtidos através deste mercado têm provocado um aumento do número de investidores neste tipo de ativos e conseqüentemente um aumento do número de compradores – note-se que os investidores são, regra geral, constantes compradores e vendedores de imóveis. Assim sendo, pode-se admitir que o mercado imobiliário é um mercado em constante expansão e crescimento.

Quando os cenários macroeconómicos são benéficos, i.e., otimistas, existem alguns fatores que podem impulsionar ainda mais o crescimento deste setor e tornar o preço dos imóveis sobrevalorizados.

Pode dizer-se que o mercado imobiliário não depende apenas da oferta e da procura, mas que a oferta e a procura deste mercado podem aumentar devido a diversos catalisadores económicos.

Neste sentido, para tentar desmistificar o aumento exponencial dos preços dos imóveis em Portugal, nesta secção pretende-se analisar os impactos que os estímulos provocados na economia têm sobre o mercado imobiliário.

Taxas de Juro

Em Portugal, o número de transações de imóveis continua a crescer a um bom ritmo e importa analisar os fatores que provocam este crescimento. Um dos fatores mais determinantes para manter este crescimento são as taxas de juro.

As taxas de juro desempenham um papel fundamental no sistema financeiro, bancário e económico.

Após a crise imobiliária de 2007, o Banco Central Europeu reviu as políticas monetárias de forma a dar novos estímulos à economia. Entre estas revisões esteve a redução das taxas de juro nas instituições e nos privados. Como resultado, o acesso a crédito tornou-se mais fácil e o poder de compra aumentou, impulsionando assim a economia.

A redução de taxas de juro, juntando ao facto do valor da Euribor (indicador da taxa de juros média dos empréstimos interbancários sem garantia da zona euro) se encontrar em mínimos históricos, permite a concessão de crédito a taxas variáveis negativas.

O baixo nível das taxas de juro provoca um aumento na procura de novas habitações dada a redução substancial do valor da hipoteca a pagar ao banco, e provavelmente, é um

dos fatores macroeconómicos que mais impulsiona o crescimento do número de imóveis transacionados.

A acrescentar às baixas taxas de juro o mercado de investimento estrangeiro, catalisados por fundos de investimento, ainda encontra em Portugal um mercado acessível para investir. Como resultado, o preço dos imóveis tende a subir.

As famílias portuguesas, com a impossibilidade de competir contra grandes capitais privados vê-se obrigada a procurar imóveis na periferia. Desta forma, num efeito de contágio, também o preço dos imóveis nas periferias tende a subir, levando a um aumento generalizado dos preços.

Deve-se ter em consideração que o aumento da procura é impulsionado por taxas de juro baixas e o receio que os imóveis continuem a subir os seus preços.

Sabendo que a redução das taxas de juro leva a uma maior facilidade na concessão de crédito e conseqüentemente maior disponibilidade financeira para pagar a hipoteca das casas, o que acontecerá quando os bancos centrais decidirem subir essas taxas de juro, como aconteceu em 2008?

Com a subida das taxas de juro o preço do dinheiro vai subir e a mensalidade das hipotecas com taxa variável vai aumentar. Com isto, o investimento em novas habitações por parte das famílias portuguesas vai baixar, baixando conseqüentemente a procura de novos investimentos no mercado imobiliário. Nesta altura, quem quiser vender os seus imóveis terá de baixar o preço de venda levando a uma redução generalizada dos preços no mercado.

Em suma, o aumento da taxa de juro gera um aumento dos preços e uma menor liquidez na economia, ao par que, a diminuição das taxas de juro leva à diminuição dos preços e uma maior liquidez aos investidores que, desta forma, irão estimular o aumento da economia e o aumento dos mercados em geral.

Contudo, importa referir que existem imóveis com menos sensibilidade à variação das taxas de juro. Isto porque existem imóveis diferenciados dos demais, seja por se situarem no centro histórico, serem imóveis de “luxo” ou continuarem a atrair investimento estrangeiro.

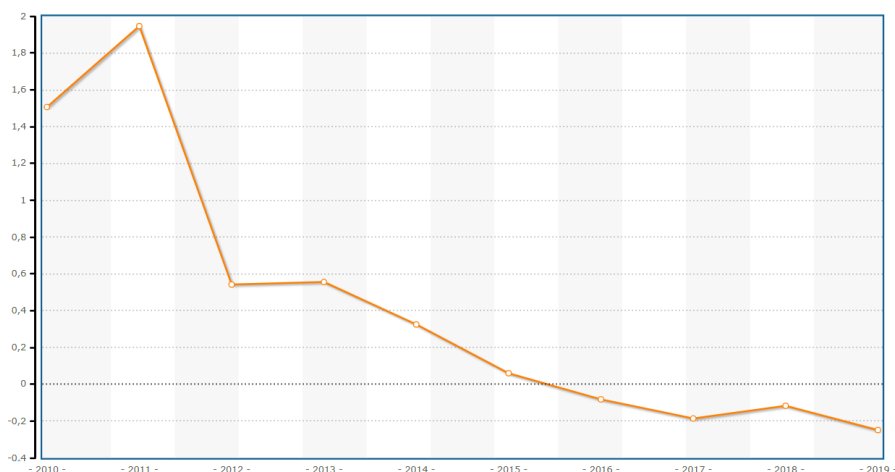


Figura 2: Evolução da Euribor a 12 meses nos últimos 10 anos

Turismo

Não obstante ao facto de haver diversos fatores a impulsionar o crescimento do mercado imobiliário, o mesmo não significa que o façam de forma independente ou não tenham ligação entre si.

Posto isto, a juntar às baixas taxas de juro, Portugal encontra-se, neste momento, na rota do turismo, das estadias de curta e longa duração, com especial incidência em Lisboa. Este crescimento exponencial do turismo que se faz sentir em todo o país motiva muitos investidores imobiliários a criarem alojamentos locais e hoteleiros, promovendo uma natural subida dos preços de mercado. Quanto maior for o número de turistas a visitar Portugal maior será a procura dos imóveis para venda.

De facto, o fluxo turístico tem implicações em diferentes setores de atividade e não apenas no mais visível, como a hotelaria e a restauração. A construção de hotéis e a recuperação de edifícios para arrendamento de curta duração para cidadãos estrangeiros, faz com a que a procura de novos imóveis para novos investimentos aumente.

Todavia, sabe-se que o setor turístico é cíclico e numa altura em que o ciclo esteja negativo poderá observar-se uma diminuição da procura que porventura originar uma queda no preço dos imóveis que sustentam este setor.

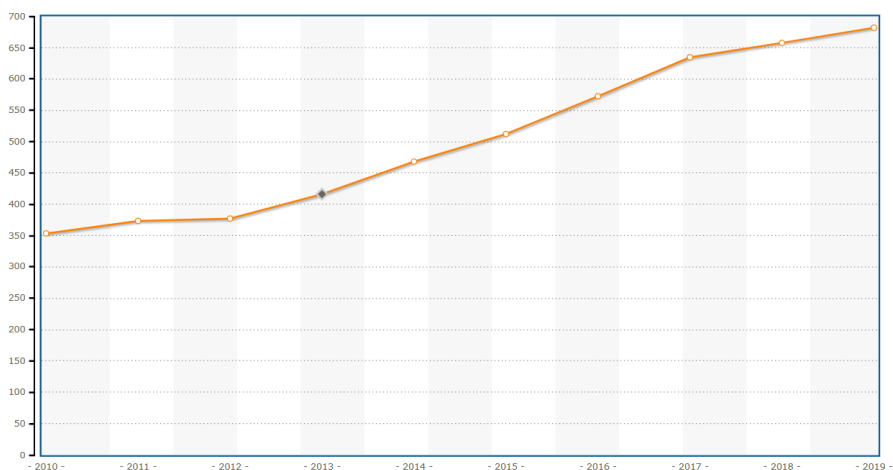


Figura 3: Número de dormidas nos alojamentos turísticos por 100 habitantes

Taxa de Desemprego

De acordo com o Instituto Nacional de Estatística (INE), a taxa de desemprego permite definir a relação entre a população desempregada e a população ativa (que corresponde à soma da população empregada com a população desempregada), sendo calculada da seguinte forma: $(\text{População desempregada} / \text{População ativa}) \times 100$.

Naturalmente, quanto menor for a taxa de desemprego de um país maior será o poder de compra e o consumo que, conseqüentemente, implicará a requisição de mais mão-de-obra a fim de oferecer mais e melhores serviços e produtos à população. Este ciclo resulta numa tendência de descida na taxa de desemprego, ou seja, ao longo do tempo espera-se que o número de empregados de um país tenha tendência a aumentar.

Mais, a taxa de desemprego e a inflação têm uma forte correlação negativa, i.e., quanto menor a taxa de desemprego maior será a inflação dos preços elevando a necessidade de as empresas serem mais aliantes para os seus trabalhadores aumentando os respetivos salários. Caso contrário, o aumento da inflação resultaria num menor poder de compra estagnando a evolução da economia do país.

Em Portugal, a taxa de desemprego tem vindo a diminuir desde o pico da crise em 2013. Este indicador, como já tinha sido mencionado, aumenta o poder de compra e, em resultado, diversos setores aumentam as suas vendas e os seus lucros, como é o caso do setor imobiliário que desde 2013 tem aumentado exponencialmente as suas receitas.

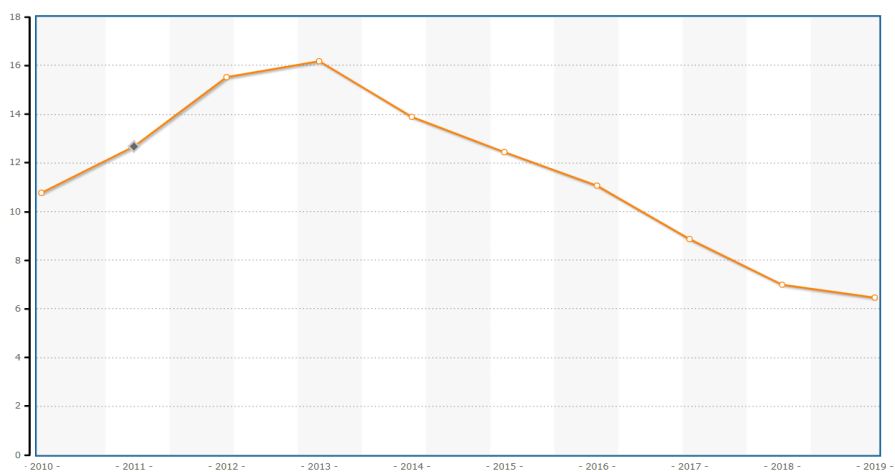


Figura 4: Taxa de Desemprego em Portugal nos últimos 10 anos

PIB

O Produto Interno Bruto (PIB) é um indicador global que representa a capacidade produtiva e de geração de rendimento de um país. Mais concretamente, quantifica a atividade económica do país em valores monetários e representa a soma de todos os bens e serviços produzidos durante um determinado período (mensal, trimestral, anual, etc.).

O PIB tem um impacto direto na confiança dos consumidores uma vez que mede o nível geral de riqueza de um país. Ora, quanto maior o PIB maior a confiança dos consumidores na obtenção de mais bens e serviços e, por consequência, haverá mais investidores nos setores que geram maior riqueza no país, de maneira a otimizarem as suas rentabilidades.

Em março de 2019, de acordo com os dados divulgados pelo INE, o mercado imobiliário valia 12% do PIB português. Estes dados revelam o grande impacto que o setor imobiliário tem no rendimento do país, reforçando a confiança por parte dos investidores que acreditam num crescimento constante deste setor.

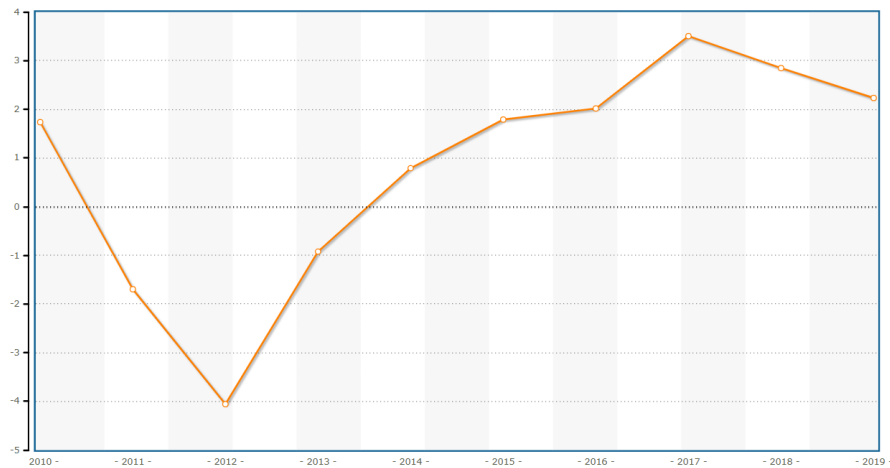


Figura 5: Taxa de Crescimento real do PIB nos últimos 10 anos (Taxa de Variação)

Em suma, o mercado imobiliário vive dias de magnificência.

Na visão dos investidores, este mercado tem uma tendência positiva de crescimento, mesmo que esse crescimento não seja tão expressivo como foi nestes últimos anos. Para ajudar a compreender esta confiança, esta secção apresentou de que forma os cenários macroeconómicos regalam este setor, permitindo que o mesmo continue a ter um crescimento sustentável.

Embora o mercado imobiliário esteja a passar por tempos de euforia, é natural que os lucros apresentados e o seu crescimento comecem a estagnar. Isto porque a população portuguesa vive de rendimentos limitados e a inflação tem tendência a aumentar, diminuindo o poder de compra. Assim, se os investidores continuarem a aumentar o preço dos seus imóveis na esperança de aumentar os seus lucros, podem contribuir para uma diminuição da procura por falta de rendimento disponível pelas famílias portuguesas para sustentar tais encargos.

Não obstante, enquanto os principais catalisadores deste mercado continuarem a ser favoráveis para o seu crescimento, nada aparenta que o mesmo quebre ou saia da sua “bolha”.

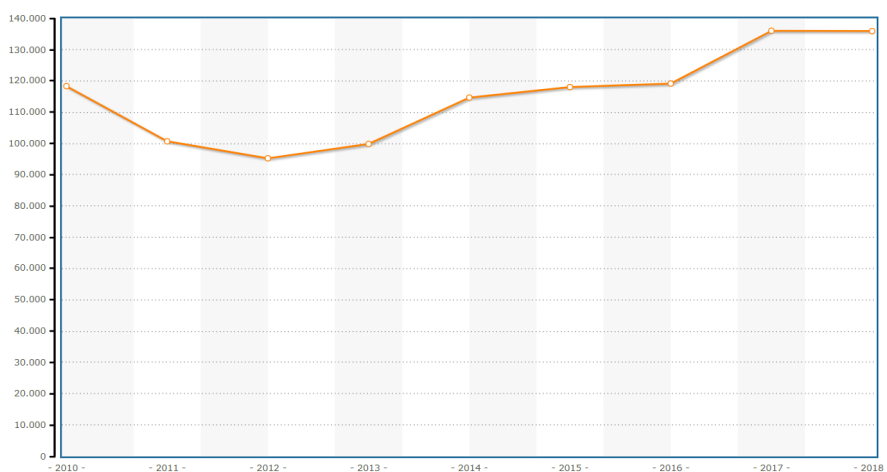


Figura 6: Valor médios dos prédios urbanos transacionados

1.1.3 Panorama Atual da Economia – Efeitos Pandemia

Vivem-se tempos de incerteza onde as consequências que a pandemia pode gerar na economia ainda se encontram por terras desconhecidas. No entanto, é indiscutível que o vírus Covid-19 veio afetar a economia global e todos os setores que a compõe, sem exceção.

Tanto os bancos centrais como as políticas orçamentais dos países vão ter um papel fundamental na moderação dos efeitos do vírus na economia. No entanto, as medidas que forem tomadas pelos mesmos, terão de ter em consideração os impactos na vida da população e, por outro lado, a evolução que o vírus terá em cada país. Desta forma, embora se tomem medidas de mitigação/prevenção para uma próxima crise económica, é irrefutável que se irá passar por uma, provocando uma nova recessão em alguns países.

Durante os primeiros meses de pandemia, onde ainda subsistia um clima de medo e incerteza, o Estado da maioria dos países decidiu cessar o crescimento da sua economia de maneira a proteger a saúde dos seus cidadãos. Naturalmente que houve setores onde rapidamente se sentiu uma quebra drástica da sua faturação e muitas famílias tiveram um corte nos seus rendimentos. A título exemplificativo, em Portugal, o setor da restauração que, com as portas fechadas perdeu a totalidade dos rendimentos, o setor do turismo também registou um grande corte no que eram os seus lucros habituais para a época e o setor de aviação que se viu obrigado a recorrer a despedimentos coletivos. Regra geral, a economia portuguesa levou um choque nas suas contas e todos os setores tiveram uma quebra nos rendimentos.

Porém, de que forma a “crise pandémica” que se instaurou no país desde março afetou o setor imobiliário? Ao contrário do que seria de esperar, em plena crise pandémica, a procura das casas não se retraiu e o setor imobiliário não quebrou. Embora pareça

contraditório, o racional para a sustentação do mercado imobiliário perante uma crise pandémica resulta não só dos fatores macroeconómicos como das medidas preventivas adotadas pelo Estado português. As historicamente baixas taxas de juro, aliadas às moratórias de crédito praticadas pela banca e o regime simplificado de *lay-off* tornam eloquente tal explicação.

Tal como foi previamente mencionado, as baixas taxas de juro aumentam o poder de compra, aliciando a compra de novos imóveis.

Relativamente ao regime simplificado de *lay-off*, o mesmo foi concebido para ajudar as empresas a suportar os encargos que tinham dada a baixa produção. No entanto, este regime também “ajudou” os trabalhadores que de um modo geral, conseguiram suportar as dificuldades financeiras que iam surgindo. Desta forma, mesmo com as reduções observadas nos salários das famílias portuguesas, estas conseguiram suportar as dívidas e as despesas que possuíam.

Por fim, as moratórias de crédito permitiram que as empresas (e particulares) suspendessem o pagamento dos seus créditos à banca dando espaço no seu capital para pagar os salários e continuar com a sua atividade. Com isto, muitas das empresas continuaram a trabalhar a 100% não sofrendo quebras severas na sua atividade e os seus trabalhadores continuaram a receber o seu salário por inteiro. Também os particulares puderam beneficiar de moratórias de crédito e prorrogar o pagamento dos créditos à banca. De um modo geral, estas moratórias permitiram que os principais agentes físicos que estimulam a atividade imobiliária não tivessem uma quebra nos seus rendimentos e, como resultado, a procura de novos imóveis manteve-se nos níveis que se encontrara antes da pandemia.

No entanto, muitas das micro empresas e das empresas em nome individual, com menos capacidade financeira foram obrigadas a fecharem atividade, aumentando a taxa de desemprego em Portugal – em agosto de 2020 a taxa de desemprego era de 8.1%, sendo o valor mais elevado desde agosto de 2018. No término das moratórias, quanto as empresas voltarem a pagar o crédito concedido pelo banco, perante os baixos lucros e a incapacidade de pagar todas as despesas, deverão reduzir o número de postos de trabalho contribuindo para que a taxa de desemprego continue a aumentar. Com o aumento da taxa de desemprego a prestação das casas pode tornar-se insustentável e as pessoas poderão não ter capacidade para pagar os seus créditos, vendendo os seus imóveis, ou em última instância, entregando as casas aos bancos. Também é expectável que o próprio turismo, enquanto não existir a cura para o Covid-19, seja severamente afetado, devido não só às inseguranças da população mundial, como também à redução drástica do número de voos praticados pelas companhias aéreas.

Sintetizando, não esquecendo outros fatores macroeconómicos, o aumento da taxa de desemprego e a redução do turismo podem resultar numa diminuição da procura de novos imóveis e conseqüentemente na redução do valor das casas. Contudo, deve-se ter em consideração que as conseqüências desta crise pandémica ainda não estão a ser refletidas, na sua totalidade, na economia.

2. Metodologia

No presente capítulo são descritos os conceitos teóricos que serão abordados no decorrer do estudo empírico da presente dissertação. Neste sentido, serão abordados todos os conceitos necessários para a interpretação dos resultados às análises efetuadas no capítulo 3.

Primeiramente serão introduzidas as noções básicas sobre séries temporais: definição e classificação das mesmas.

De seguida, com o auxílio de alguns testes econométricos será explicado a importância da existência de raízes unitárias nas séries temporais.

Posteriormente segue-se uma breve descrição dos modelos de séries temporais univariadas, mais concretamente os modelos ARMA/ARIMA.

Posto isto, serão introduzidos os conceitos sobre a utilidade e a validação dos modelos acima referidos.

Por fim, seguem-se os conceitos de previsão que será o tema central deste projeto.

O desenvolvimento deste capítulo será acompanhado pelo livro “*Applied Time Series Analysis: A Practical Guide to Modeling and Forecasting*, Academic Press, Elsevier Inc.” (Mills, T.C., 2019).

2.1 Conceitos Gerais das Séries Temporais

O registo de fenómenos a variar irregularmente com o tempo tem o nome de série temporal (Kitagawa, 2010).

Uma série temporal é uma sequência de dados obtidos por um conjunto de observações ordenadas durante um intervalo de tempo.

De um modo geral, uma série temporal pode caracterizar-se da seguinte maneira:

$$X_t \text{ (discreta)}, t = 1, 2, \dots, n \quad \text{ou} \quad X(t) \text{ (contínua)}, t \in T \subseteq \mathbb{R}$$

Tipicamente uma série temporal é medida em intervalos de tempo iguais (minutos, horas, dias, meses, anos, etc., ...), portanto, a ordem dos dados é um dos fatores mais importantes no estudo deste tipo de séries.

A definição de série temporal pode repartir-se em duas ramificações. A série temporal **univariada** onde apenas se considera uma observação em cada intervalo de

tempo, e a série temporal **multivariada** que regista em cada intervalo de tempo o valor respetivo de cada série que entrar no estudo.

A tese apresentada tem como objetivo o estudo de uma série temporal **univariada**.

Os movimentos característicos desta sequência de dados podem ser classificados nas seguintes quatro componentes:

Sazonalidade – Movimentos ou Variações Por Estações

São referentes a movimentos padronizados a que uma série temporal obedece por períodos sucessivos no tempo.

Movimentos de Tendência

Conforme o nome indica são movimentos que seguem uma tendência/direção geral segundo o qual a série se desenvolve num intervalo de tempo. A determinação deste tipo de série pode ser feita através de regressão linear.

Movimentos ou variações cíclicas

São referentes a oscilações a longo prazo em torno de uma curva de tendência que pode ou não ser periódica.

Movimentos irregulares ou aleatórios

São referentes a movimentos esporádicos provocados por fatores externos (normalmente aleatórios) como por exemplo um surto epidémico. Admite-se que os efeitos destes acontecimentos provocam variações somente num curto período, contudo, tais efeitos podem ser tão severos que acarretam novos movimentos cíclicos ou de outra natureza.

Após definidos os principais movimentos que uma série temporal pode descrever é importante detalhar as principais etapas que devem ser percorridas para fazer um estudo rigoroso deste tipo de séries. Desta forma, podemos resumir o estudo das séries temporais nos seguintes passos:

Descrição: É nesta etapa que se explora os primordiais da série. Tipicamente é usual desenhar-se cronogramas, histogramas e identificarem-se *outliers*. Também é costume analisar as estatísticas descritivas dos dados, como a média aritmética, a mediana, a variância, os máximos e mínimos, os quartis, a taxa de variação no intervalo total ou intervalos parciais, etc. (Murteira et al., 1993).

Modelação: Nesta etapa pretende-se identificar a estrutura (estocástica ou determinística) e adaptar os dados ao modelo mais apropriado. Deste modo, dado que existem vários tipos de séries temporais, é necessário escolher uma classe que se adeque ao modelo e estimar os parâmetros que nele se incluem. (Kitagawa, 2010).

Previsão: O passo da previsão permite-nos prever (com maior ou menor precisão) o comportamento futuro das séries temporais com base nas correlações ao longo do tempo entre as variáveis usando as observações conhecidas da série temporal.

Controlo: O estudo da série temporal não fica concluído sem se passar por esta fase. O controlo permite avaliar se as características quantitativas da série em produção continuam dentro dos *thresholds* previamente definidos. Caso os limites estabelecidos tenham sido ultrapassados devem-se corrigir os fatores responsáveis pelo comportamento invulgar.

2.2 O que descreve uma série temporal?

Após apresentados os conceitos gerais das séries temporais, esta secção foca-se na apresentação das principais estatísticas e testes que descrevem este tipo de séries.

As definições que serão apresentadas permitem ao analista um maior controlo sobre os dados e uma melhor interpretação do tipo de série que o mesmo está a analisar.

Estatísticas

Seja $X(t)$ um processo estocástico tal que $E(X(t)^2) < +\infty$ para todo o $t \in T$, definem-se como as suas principais estatísticas:

1. Valor Médio, $\mu(t)$, como sendo:

$$\mu(t) = E(X(t))$$

2. Variância, σ^2 , como sendo:

$$\sigma^2 = \text{Var}(X(t)) = E[(X(t) - \mu(t))^2]$$

3. Covariância, $\gamma(t_1, t_2)$, como sendo:

$$\begin{aligned}\gamma(t_1, t_2) &= \text{cov}[X(t_1), X(t_2)] = E\{[X(t_1) - \mu(t_1)][X(t_2) - \mu(t_2)]\} \\ &= E[X(t_1)X(t_2)] - \mu(t_1)\mu(t_2)\end{aligned}$$

4. Função de Correlação, $\rho(t_1, t_2)$, como sendo:

$$\rho(t_1, t_2) = \frac{\gamma(t_1, t_2)}{\sigma(t_1)\sigma(t_2)} = \frac{\text{cov}[X(t_1), X(t_2)]}{\sqrt{\text{Var}[X(t_1)]\text{Var}[X(t_2)]}}$$

Assimetria

Em estatística, a assimetria é uma medida que permite estudar a distribuição de um conjunto de dados. Mais concretamente, a assimetria analisa o grau de afastamento que uma distribuição apresenta relativamente ao eixo de simetria da distribuição normal.

O coeficiente de simetria é:

$$S_k = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\sqrt{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^3}}$$

Daqui resultam 3 cenários possíveis:

1. $s_k = 0$, a distribuição é simétrica, ou seja, média = mediana = moda;
2. $s_k > 0$, a distribuição é assimétrica positiva (à direita) e, como resultado, média \geq mediana \geq moda;
3. $s_k < 0$, a distribuição é assimétrica negativa (à esquerda) e, como resultado, média \leq mediana \leq moda.

Curtose

Em estatística, a curtose, tal como a simetria, também é uma medida que permite estudar a distribuição de uma amostra. Mais concretamente, a curtose é uma medida que estuda o grau de achatamento de uma distribuição em relação a uma curva normal de referência.

A curtose é:

$$k = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\sqrt{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^2}}$$

Daqui resultam 3 cenários possíveis:

1. $k = 3$, a distribuição é mesocúrtica, ou seja, a distribuição é normal;
2. $k > 3$, a distribuição é leptocúrtica, ou seja, a distribuição é mais alta que a distribuição normal;
3. $k < 3$, a distribuição platicúrtica, ou seja, a distribuição é mais achatada que a distribuição normal.

Em suma, na estatística descritiva de um conjunto de dados a assimetria e a curtose são utilizadas para inferências estatísticas sobre a normalidade da variável aleatória em estudo.

Não obstante, é usual recorrer-se a testes estatísticos para tirar conclusões mais fidedignas. Desta forma o teste de *Jarque-Bera* permite, estatisticamente, concluir se uma distribuição segue, ou não, uma distribuição normal.

Teste de Normalidade de *Jarque-Bera*

Este teste estuda a normalidade de uma variável aleatória através da média, variância, simetria e curtose da mesma sendo por isso um dos testes mais comuns a aplicar a séries temporais aquando estudada a normalidade.

H_0 : A variável é normalmente distribuída

versus

H_1 : A variável não segue uma distribuição normal

Assim, se a série tiver uma distribuição normal o *p-value* associado será superior aos níveis de significância usuais (1%, 5% e 10%), levando à não rejeição da hipótese nula.

2.3 Estacionariedade

É dedicada uma secção exclusiva à estacionariedade das séries temporais dada a importância que esta propriedade possui na maioria dos modelos econométricos. Mais concretamente, a maioria dos modelos econométricos só é válido para séries temporais estacionárias.

Quando uma série não é estacionária diz-se que a mesma possui uma raiz unitária (vide subcapítulo 2.4) contudo, existem algumas técnicas que permitem obter a estacionariedade de uma série.

Na teoria das probabilidades, um processo estocástico representa a evolução no tempo de um conjunto ou de uma família de variáveis aleatórias. A distribuição de probabilidades destes processos é estável ao longo do tempo, ou seja, todos os valores de uma sequência de observações têm a mesma distribuição.

Estacionariedade Fraca

Um processo estocástico $\{X(t), t \in T\}$ diz-se fracamente estacionário se $\forall t \in T$ tem-se que:

1. Média constante ao longo do tempo

$$E(X_t) = E(X_{t-i}) = \mu$$

2. Variância constante e finita

$$E[(X_t - \mu)^2] = E[(X_{t-i} - \mu)^2] = \sigma^2 < \infty$$

3. Covariância constante

$$E[(X_t - \mu)(X_{t-i} - \mu)] = E[(X_{t-i} - \mu)(X_{t-j-i} - \mu)] = \gamma_i$$

onde i é o *lag* entre duas observações.

Se o terceiro critério de estacionariedade for respeitado então garante-se que os termos de covariância são estáveis e que a informação da série permite prever valores futuros.

De frisar que a estacionariedade implica estacionariedade fraca, mas o contrário não é necessariamente verdadeiro. Normalmente quando se garante estacionariedade refere-se, em geral, a estacionariedade fraca.

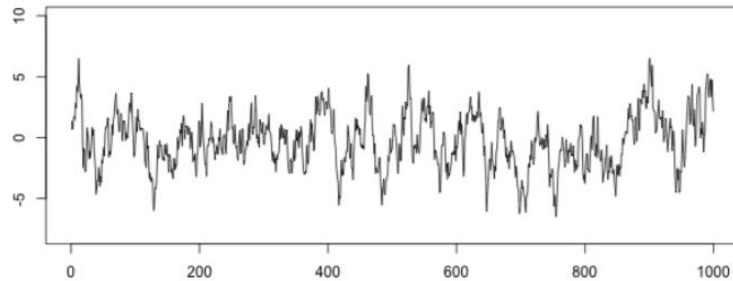


Figura 7: Exemplo de uma série estacionária

Processo para estabilizar uma série

Regra geral, uma série temporal não é estacionária. Geralmente é necessário realizar alterações ao formato original da série de forma a obtermos a estacionariedade. De seguida, apresenta-se dois dos processos mais comuns para estabilizar uma série temporal:

1. Diferenças Sucessivas

$$\Delta x_t = x_t - x_{t-1} \text{ (1ª diferença)}$$

$$\Delta\Delta x_t = \Delta^2_{x_t} = x_t - 2x_{t-1} + x_{t-2} \text{ (2ª diferença)}$$

2. Diferenças Logaritmizadas/Retornos

$$\Delta \log(x_t) = \log(x_t) - \log(x_{t-1})$$

Nas figuras abaixo é apresentado um exemplo de diferenças sucessivas que transforma uma série em níveis não estacionária numa série estacionária. Estes exemplos foram extraídos do livro Mills, T.C. (2019, pp. 2, 22).

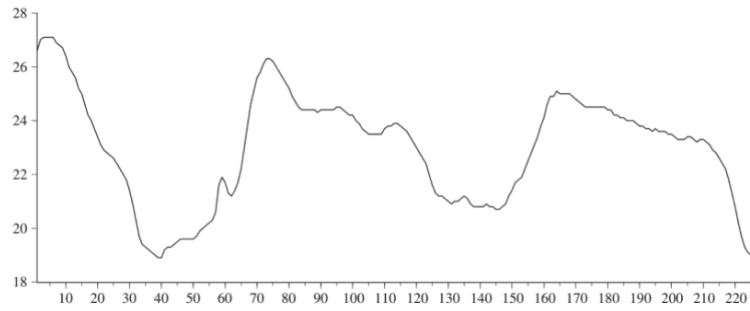


Figura 8: Exemplo de uma série em níveis não estacionária

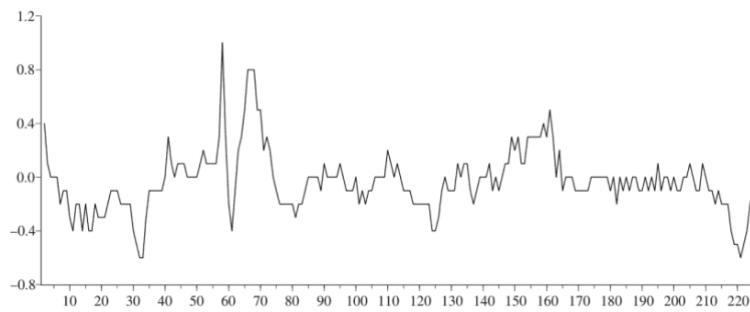


Figura 9: Exemplo de uma série em diferenças sucessivas – 1ª diferença

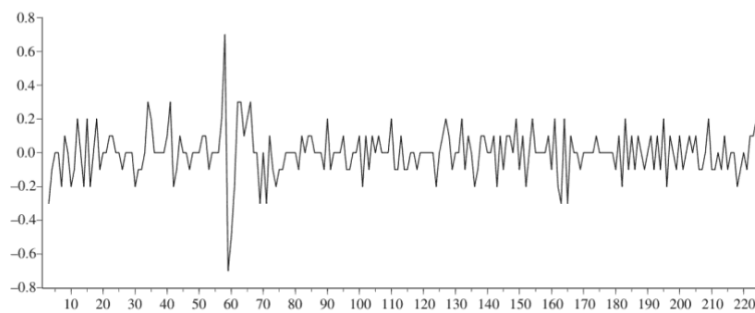


Figura 10: Exemplo de uma série em diferenças sucessivas – 2ª diferença

Se uma série não estacionária for diferenciada d vezes antes de ser estabilizada, é dita integrada de ordem d , e denotamo-la por $X_t \sim I(d)$. Por sua vez, uma série estacionária diz-se integrada de ordem zero, $I(0)$.

No exemplo apresentado o processo de estacionarizar a série foi até à segunda diferença. Deste modo, pode dizer-se que esta série é integrada de ordem 2, i.e., $I(2)$.

2.3.1 Natureza da Série Temporal

As séries temporais podem apresentar uma tendência que, por sua vez, pode ser determinística ou estocástica.

Uma série apresenta uma tendência determinista se apenas variar em função do tempo. Por outro lado, uma série apresenta tendência estocástica se a componente aleatória afetar a trajetória/comportamento da série a longo prazo.

De forma a clarificar os tipos de tendência e, como os mesmos se refletem na estacionariedade de uma série, considere-se o seguinte modelo de regressão linear:

$$X_t = \rho X_{t-1} + \beta_0 + \beta_1 t + \varepsilon_t$$

1. Se $\rho = 0$, o processo estocástico tem tendência pura e desenvolve-se em torno de uma tendência linear, apresentando, desta forma, uma tendência determinística – **série não estacionária**;
2. Se $\rho = 1$ e $\beta_1 = 0$, obtém-se um processo estacionário puro em primeiras diferenças (designado por passeio aleatório) e estamos na presença de uma tendência estocástica – **série não estacionária**;
3. Se $\rho = 1$ e $\beta_0, \beta_1 \neq 0$, o processo estocástico segue um passeio aleatório com *drift* e tendência determinística – **série não estacionária**;
4. Se $0 < \rho < 1$, existe correção serial em X, mas não existe tendência estocástica – **série estacionária**.

De modo geral, quando testamos a não estacionariedade de uma série temporal estamos a averiguar se $\rho = 1$, ou seja, se a série temporal possui uma raiz unitária e, como resultado, não é estacionária.

O próximo subcapítulo foca-se no estudo dos testes de raízes unitárias mais frequentes aplicados a séries temporais.

2.4 Testes de Raiz Unitária

Para uma série ser estacionária requer-se que o processo esteja num estado de “equilíbrio estatístico”. Um processo estocástico diz-se estritamente estacionário se as suas propriedades não são afetadas ao longo do tempo (Mills, T.C., 2019).

Os testes de raiz unitária foram criados para detetar se a série é, ou não, estacionária. Ou seja, quando uma série não é estacionária diz-se que esta possui uma raiz unitária. Para o efeito, a generalidade dos testes afere como hipótese nula H_0 se o processo é não estacionário (possui uma raiz unitária) *versus* a hipótese de que a série é estacionária (não possui raiz unitária).

A unidade da raiz é uma característica das séries temporais que podem causar problemas na inferência estatística. Na verdade, maioria das séries temporais reprova nestes testes, sendo usual transformar a série (conforme abordado no subcapítulo anterior) antes de se proceder a este teste.

Neste subcapítulo serão abordados quatro dos testes mais usados no estudo da raiz unitária: *Dickey-Fuller* (DF) e *Dickey-Fuller Aumentado* (ADF), *Phillips-Perron* (PP) e *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS).

Teste de Raiz Unitária

A ideia deste teste é testar se:

$H_0: \rho = 1$ (processo não estacionário)

versus

$H_1: \rho < 1$ (processo estacionário)

Fazendo a regressão de X_t em X_{t-1} , verifica-se se o coeficiente ρ é estatisticamente igual a 1. Se for, não rejeitamos a hipótese nula, ou seja, o processo é não estacionário e estamos na presença de uma raiz unitária. Caso contrário, o processo é estacionário.

Teste de Dickey-Fuller (DF)

O desenvolvimento deste teste parte do pressuposto que os erros são independentes e identicamente distribuídos (i.i.d), ou seja, não apresentam qualquer correlação.

Partindo do modelo já conhecido, $X_t = \rho X_{t-1} + \beta_0 + \beta_1 t + \varepsilon_t$, e, subtraindo ambos os lados da equação por X_{t-1} , vem:

$$X_t - X_{t-1} = (\rho - 1)X_{t-1} + \beta_0 + \beta_1 t + \varepsilon_t$$

As hipóteses a serem testadas para o teste DF são,

H0: $\rho = 1$ (a série não é estacionária)

versus

H1: $\rho < 1$ (a série é estacionária)

Desta forma, para valores de teste superiores aos níveis de significância usuais o teste não rejeita H0 admitindo que a série não é estacionária, ou seja, a série tem uma raiz unitária.

No caso dos erros aleatórios ε_t apresentarem correlação serial entre si o teste de *Dickey-Fuller* não pode ser aplicado. Nestes casos deve ser utilizado o teste com correção paramétrica *Dickey-Fuller Aumentado* (ADF).

Teste de Dickey-Fuller Aumentado (ADF)

Este teste sugere que quando os resíduos apresentam correlação deve-se aumentar a regressão em termos suficientes ΔX_{t-j} até eliminar a autocorrelação.

O modelo adotado neste teste é o seguinte:

$$\Delta X_t = (\rho - 1)X_{t-1} + \sum \beta_i \Delta x_{t-i} + \varepsilon_t$$

Usualmente é utilizada a regra empírica $[12(N/100)^{1/4}]$, onde N é o tamanho da série, para determinar o número de *lags* a utilizar. Este tipo de tratamento resulta geralmente na adição de:

- 1 a 2 *lags* para dados anuais;
- ≥ 4 *lags* para dados trimestrais;
- ≥ 12 *lags* para dados mensais;
- ≥ 30 *lags* para dados diários.

As hipóteses a serem testadas para o teste ADF são,

H0: $\rho = 1$ (a série não é estacionária)

versus

H1: $\rho < 1$ (a série é estacionária)

As conclusões deste teste são idênticas às apresentadas no teste anterior.

Teste de Phillips-Perron (PP)

Este teste é construído na base do teste de DF, contudo, ao contrário do teste de ADF, o teste de PP incorpora uma correção automática no teste de *Dickey-Fuller*, permitindo que os erros sejam autocorrelacionados.

Para o modelo de regressão linear original, $X_t = \rho X_{t-1} + \beta_0 + \beta_1 t + \varepsilon_t$, as hipóteses a serem testadas são:

H0: $\rho = 1$ (a série não é estacionária)

versus

H1: $\rho < 1$ (a série é estacionária)

Sendo este teste semelhante aos demais apresentados, as conclusões daqui retiradas também são as mesmas dos testes anteriores.

De acordo com Davidson and MacKinnon (2004), a performance do teste de Phillips-Perron para amostras finitas de dados apresenta um pior desempenho face ao ADF.

Teste de Kwiatkowski-Phillips-Schmidt-Shin (KPSS)

Por sua vez, o teste de KPSS foi concebido de uma forma diferente dos anteriores. Neste teste, a presença de uma raiz unitária não está na hipótese nula, mas na alternativa.

Seja $S_t = \sum_{i=1}^t \varepsilon_i$ e σ^2 um estimador para a variância dos erros. A estatística de teste é dada por:

$$LM = \sum_{i=1}^t \frac{S_t^2}{n^2 \hat{\sigma}^2}$$

As hipóteses a serem estudadas neste teste são,

H0: $\rho < 1$ (a série é estacionária)

versus

H1: $\rho = 1$ (a série não é estacionária)

Neste teste rejeitamos H_0 se a estatística de teste $LM_{KPSS} >$ valores críticos habituais (1%, 5% e 10%), indicados por KPSS (1992).

O teste de KPSS é usado como complemento dos testes anteriormente apresentados uma vez que apresenta uma variação na argumentação das hipóteses de teste.

Habitualmente este teste é abordado como critério de desempate quando os testes ADF e PP são discordantes.

2.5 Modelos Lineares Univariados

Os modelos lineares univariados são usados com o objetivo de modelar e realizar previsões de uma série temporal, Box e Jenkins (1970).

Na base dos modelos univariados, destaca-se o modelo ARMA e seus precedentes (AR – Autorregressivo e MA – Médias Móveis). Este tipo de modelo foi concebido para séries temporais estacionárias, ao par que, o modelo ARIMA é aplicado a séries temporais não estacionárias, sendo I correspondente ao processo de integração (diferenciação).

2.5.1 Modelos Autorregressivos de ordem p – AR(p)

O modelo autorregressivo AR admite que a variável dependente seja linearmente descrita pelos seus valores passados somados a um termo estocástico. Assim, um modelo autorregressivo AR é uma representação de um processo aleatório que pode ser expresso da seguinte forma:

$$X_t = a_0 + \sum_{i=1}^p a_i X_{t-i} + \varepsilon_t, \text{ com } \varepsilon_t \sim WN(0, \sigma^2)$$

onde p é a ordem de defasamento temporal, a_0, a_1, \dots, a_p são os parâmetros do modelo e ε_t é um processo aleatório que assinala a nova informação recebida no instante t .

Seja L o operador de defasamento, representado por $L(X_t) = X_{t-1}$, a equação anterior pode ser escrita de forma equivalente, como sendo:

$$\Phi(L)X_t = a_0 + \varepsilon_t, \quad \text{onde } \Phi(L) = 1 - \sum_{i=1}^p a_i L^i$$

Diz-se que um modelo autorregressivo de ordem p é estável/estacionário se e só se todas as raízes $z = \frac{1}{\alpha_i}$ do polinómio característico

$$\Phi(z) = 1 - \sum_{i=1}^p a_i z^i$$

estão situadas no exterior do círculo unitário, i.e., $|a_i| < 1$.

Funções de Autocorrelação e Autocorrelação Parcial

A autocorrelação é uma ferramenta matemática usada para encontrar padrões de repetição. A função de autocorrelação de um modelo $AR(p)$ estacionário com média nula $E(X_t) = 0$ e covariância $\gamma_h = E(X_t, X_{t+h})$ é dada por:

$$\rho_k = a_1 \rho_{k-1} + a_2 \rho_{k-2} + \dots + a_k \rho_{k-p} = \sum_{i=1}^p a_i \rho_{k-i}$$

Onde $\gamma_0 = \text{var}(X_t)$ e $\rho_k = \frac{\gamma_h}{\gamma_0}$.

Por sua vez, a variância de um processo $AR(p)$ enquanto função das autocorrelações é dada por:

$$\text{Var}(X_t) = \frac{\sigma^2}{a_1 \rho_1 + a_2 \rho_2 + \dots + a_k \rho_k} = \frac{\sigma^2}{\sum_{i=1}^k a_i \rho_i}$$

Por fim, a função de autocorrelação parcial (PACF) de um processo $AR(p)$ é dada por:

$$r(p) = \begin{cases} a_h \neq 0, & \text{para } h < p \text{ (existe PACF)} \\ 0, & \text{para } h > p \text{ (não existe PACF)} \end{cases}$$

denota-se uma convergência brusca para zero ao fim de p passos. A função PACF faz a diferença entre um processo AR e um processo ARMA.

2.5.2 Modelos de Média Móvel de ordem q - MA(q)

Uma média móvel é um estimador calculado a partir de amostras sequenciais de uma população. Mais concretamente, um processo de médias móveis de ordem q define-se, para cada instante t, como a média ponderada das últimas q observações de um modelo de ruído branco.

Seja ε_t um processo aleatório - ruído branco - com $E(\varepsilon_t) = 0$ e $var(\varepsilon_t) = \sigma^2$. Então o processo MA(q) pode ser representado por:

$$X_t = \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} = \theta_i L^i \varepsilon_t$$

Onde $\theta_q(L) = 1 + \sum_{i=1}^q \theta_i L^i$ e θ_i são constantes reais. Um processo MA(q) é fracamente estacionário para todos os valores de $\{\theta_u\}$.

Dado que os erros ε_t são não correlacionados, então a variância é dada por:

$$Var(X_t) = \left(1 + \sum_{i=1}^q \theta_i^2 \right) \sigma^2$$

No que respeita à função de auto covariância a mesma é dada por:

$$\gamma_h = \sigma^2 (\theta_h + \theta_{h+1} \theta_h + \dots + \theta_q \theta_{q-h}), \quad 0 \leq h \leq q$$

Com

$$\gamma_0 = v(x_t) = \sigma^2 (1 + \theta_1^2 + \dots + \theta_q^2)$$

Por sua vez, para a função de autocorrelação, vem:

$$\rho_k = \frac{(\theta_k + \theta_{k+1} \theta_1 + \dots + \theta_q \theta_{q-k})}{(1 + \theta_1^2 + \dots + \theta_q^2)}, \quad 0 \leq h \leq q$$

A função de autocorrelação (ACF) de um processo MA(q) anula-se no lag(q+1) e a função de autocorrelação parcial converge suavemente para zero.

2.5.3 Modelos Autorregressivos de Médias Móveis- ARMA(p,q)

Os modelos autorregressivos de médias móveis ARMA(p,q) resultam da junção do modelo autorregressivo de ordem p, AR (p), com o de média móvel de ordem q, MA(q).

Estes modelos oferecem uma descrição frugal de um processo estocástico fracamente estacionário em termos de dois polinómios (Percival, Donald B.; Walden, Andrew T., 1993).

O modelo ARMA é uma ferramenta muito usada no estudo das séries temporais para prever valores futuros de uma série temporal, através dos valores já registados em momentos anteriores (AR) e de uma combinação linear de erros aleatórios que ocorreram no passado (MA).

Posto isto, o modelo ARMA(p,q) pode ser escrito como sendo:

$$X_t = a_0 + \varepsilon_t + \sum_{i=1}^p a_i X_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$$

Por sua vez, o valor médio do modelo ARMA(p,q) é dado por:

$$E(X_t) = \frac{a_0}{1 - \sum_{i=1}^p a_i}$$

De notar que as combinações relativas aos comportamentos derivados dos modelos AR e MA apresentam a função de autocorrelação de um processo ARMA. Contudo, para lags inferiores a q, a função ACF vai ser idêntica ao de um modelo AR(p).

Não obstante ao relatado, este modelo tem algumas limitações no que respeita a séries temporais com algumas irregularidades como, por exemplo, assimetrias, volatilidade irregular, intervalos de tempo variável e quebras de estrutura.

2.5.4 Modelos Autorregressivos Integrados de Médias Móveis-ARIMA(p,d,q)

Os modelos ARMA só podem ser utilizados para séries estacionárias. Quando este princípio não é observado mas é possível estacionarizar a série após d diferenças sucessivas, obtém-se uma nova versão do modelo ARMA, denominado por ARIMA(p,d,q) onde p representa a ordem do modelo AR, q representa a ordem do modelo MA e d representa o número de vezes que a série foi integrada até se obter a estacionariedade.

Seja X_t uma série não estacionária e W_t uma série estacionária obtida através do processo de diferenciação da série X_t d vezes, podemos definir W_t como:

$$W_t = \Delta^d X_t = (1 - L)^d X_t$$

Onde L é o operador de desfasamento.

A partir daqui pode-se tratar W_t como sendo um modelo ARMA(p,q).

Através do modelo ARIMA(p,d,q) podemos descrever todos os modelos apresentados anteriormente através das seguintes equivalências:

- ARIMA(p,0,q) = ARMA(p,q);
- ARIMA(p,0,0) = AR(p);
- ARIMA(0,0,q) = AR(q).

2.5.5 Critérios de Seleção de Modelos

A importância de se escolher o “melhor modelo” parte do pressuposto que não existem modelos verdadeiros. Contudo, existem modelos aproximados da realidade causados pela perda de informação. “*All models are wrong, but some are useful*” George E. P. Box.

Assim sendo, é necessário, através de critérios de seleção apurar o melhor modelo, dentro daqueles que foram ajustados. O ACF e PACF são capazes de, através de análises gráficas, sugerir o modelo que deve ser usado. Todavia, as séries temporais podem ser muito voláteis acabando por enviesar as ordens dos resultados observados nestas funções. Neste sentido, é comum completar o estudo da seleção do melhor modelo com critérios de informação mais robustos e capazes de mensurar a qualidade estatística de um modelo.

Estes critérios culminam da soma de dois fatores distintos. Um dos fatores é a função da soma dos resíduos ao quadrado e o outro é o resultado de uma penalidade pela perda de graus de liberdade ao adicionarmos parâmetros extra.

Os modelos que apresentarem critérios de informação menores são, por sua vez, melhores modelos. Porém, a adição de um termo extra só resulta numa redução do valor do critério se a soma dos resíduos ao quadrado diminuir o suficiente de modo a compensar o aumento no valor do termo de penalidade.

Os três critérios mais usuais e que irão ser abordados nesta dissertação são: o critério de informação Akaike (1974) – *AIC*, o critério de informação Schwarz Bayesian (1978) – *BIC* e por fim o critério Hannan-Quinn – *HQIC*.

Estes três critérios podem ser descritos através das seguintes expressões matemáticas:

- $AIC = \ln(\hat{\sigma}^2) + \frac{2m}{T}$
- $BIC = \ln(\hat{\sigma}^2) + \frac{m}{T} \ln(T)$
- $HQIC = \ln(\hat{\sigma}^2) + \frac{2m}{T} \ln(\ln(T))$

Onde, $m = p + q + 1$ corresponde ao número total de parâmetros estimados e $\hat{\sigma}^2$ diz respeito à variância residual obtida através do quociente entre a soma dos resíduos ao quadrado e o número de observações T .

2.6 Verificação dos Pressupostos – Análise dos Resíduos

A validação do modelo é feita aquando seleção do melhor modelo que descreve os dados. Esta validação incide, essencialmente, na análise dos resíduos do modelo onde o mesmo deve verificar os seguintes pressupostos para ser validado:

- $E(\varepsilon_t) = 0 \Rightarrow$ Os erros têm média zero
Se o termo de regressão tem um termo constante não nulo, então este pressuposto nunca vai ser violado. Caso contrário, pode originar o enviesamento nas estimativas dos coeficientes.
- $Var(\varepsilon_t) = \sigma^2 \Rightarrow$ A variância é constante e finita (Homocedasticidade)
Quando a variância dos erros não é constante, recorre-se, na maioria dos casos, a transformações logarítmicas de modo a reduzir-se os problemas de heterocedasticidade.
- $Cov(\varepsilon_i, \varepsilon_j) = 0, i \neq j \Rightarrow$ Os erros são linearmente independentes
Na presença de independência linear, os resíduos não apresentam nenhum padrão.

- $Cov(\varepsilon_t, y_t) = 0 \Rightarrow$ Não há relação entre os erros e a variável independente
Se a média dos resíduos for nula, este pressuposto pode ser reescrito na fórmula $E(\varepsilon_t y_t) = 0$.
- $\varepsilon_t \sim N(0, \sigma^2) \Rightarrow$ Os erros são normalmente distribuídos
Esta verificação é importante para a realização de testes de hipóteses e intervalos de confiança sobre os parâmetros.

Nota-se que os resíduos de um modelo de regressão são tipicamente assumidos como sendo algum tipo de ruído branco.

2.6.1 Ruído Branco

Até esta fase foi abordado o conceito de ruído branco sem nunca se explicar tal conceito. Nesta secção, e após definição dos pressupostos dos resíduos que, na sua generalidade, seguem um ruído branco, iremos abordar a sua definição.

Um modelo de ruído branco é o bloco base na construção de séries temporais. Este tipo de modelos pode ser apresentado por:

$$x_t = \varepsilon_t, \quad t \in T$$

Seja $\{\varepsilon_t\}$ uma sequência, define-se ruído branco se cada valor da série temporal tiver média zero, variância constante e não apresentar correlação serial. Na prática, um modelo para se denominar de ruído branco tem de validar os três primeiros pressupostos apresentados anteriormente.

O ruído branco é escrito matematicamente por:

$$\varepsilon_t \sim WN(0, \sigma^2) \quad \text{ou} \quad x_t \sim WN(0, \sigma^2)$$

e é um processo estacionário.

2.7 Previsão

Uma característica importante dos modelos univariados é a habilidade de fornecer uma previsão sobre os valores futuros de uma série de observações (Mills, T.C. 2019).

Existem dois pontos importantes quando se prevê os dados de uma série temporal:

1. A previsão dos valores futuros da série;
2. O erro de previsão indexado a cada ponto de previsão.

De notar que qualquer modelo é capaz de prever os valores futuros de uma série temporal, mas como previamente mencionado, existem modelos que dão melhores valores de previsão que outros. Neste sentido, é comum realizar-se uma previsão *in-sample* e respetivo erro de previsão face aos valores reais da série antes de se proceder à realização da previsão *out-of-sample*.

2.7.1 Erro de Previsão

Como foi mencionado, o erro de previsão é realizado para as observações *in-sample* e é dado pela diferença entre o valor atual de uma observação i (X_i), e o previsto (\hat{X}_i), sendo representado da seguinte forma:

$$e_i = X_i - \hat{X}_i$$

Onde e_i corresponde ao erro de previsão observado no instante i .

De forma a avaliar o desempenho global de um modelo de previsão foram criadas medidas de erro capazes de indicar se a previsão efetuada pelo modelo é boa ou má.

As medidas mais utilizadas são, o erro médio (ME – *Mean Error*), o erro absoluto médio (MAE – *Mean Absolute Error*), o erro quadrático médio (MSE – *Mean Squared Error*) e a raiz do erro quadrático médio (RMSE – *Root Mean Squared Error*).

Matematicamente estas medidas podem ser escritas da seguinte forma:

- Erro médio (ME) = $\frac{\sum_{t=1}^n e_t}{n}$
- Erro absoluto médio (MAE) = $\frac{\sum_{t=1}^n |e_t|}{n}$
- Erro quadrático médio (MSE) = $\frac{\sum_{t=1}^n e_t^2}{n}$
- Raiz do Erro quadrático médio (RMSE) = \sqrt{MSE}

Na prática, quanto menores forem os erros de previsão melhores serão as previsões efetuadas pelo modelo eleito. No entanto, quanto maior a janela de previsão menor será a taxa de acerto.

2.7.2 Previsão com Modelos ARIMA

Em termos gerais, a previsão tem como objetivo prever um valor X_{t+h} , $h \geq 1$ admitindo que se conhece o valor de todas as observações até ao instante t .

Seja $\hat{X}_t(h)$ a previsão de uma observação no instante $t+h$. Então, a equação de previsão em função de h para um modelo ARIMA(p,d,q) é dada por:

$$\hat{X}_t(h) = \sum_{i=1}^{p+q} \phi_i \hat{X}_t(h-1), \quad h > q$$

A solução geral da equação é da forma:

$$\hat{X}_t(h) = \sum_{i=1}^{p+q} c_i^{(t)} f_i(h), \quad h > q - p - d$$

onde $f_i(h)$ são funções de h e $c_i^{(t)}$ são coeficientes adaptados que dependem da origem da previsão t determinados por $\hat{X}_t(1), \hat{X}_t(2), \dots, \hat{X}_t(p+q)$.

3. Estudo Empírico

Após definidos e apresentados os conceitos teóricos que estarão presentes no estudo da série (vide capítulo 2), este capítulo pretende abordar estes conceitos de uma forma prática, tendo como objetivo o desenvolvimento de um modelo econométrico que melhor se ajuste aos dados da série e que minimize o erro de previsão.

A série que irá ser analisada relata, com dados trimestrais, a evolução do preço real das casas por metro quadrado em Portugal ao longo do tempo. Os dados estão compreendidos entre 1988 e 2019 num conjunto de 128 observações.

De referir que os dados referentes a esta série foram retirados da Organização para a Cooperação e Desenvolvimento Económico – OCDE.

Todos os gráficos que serão apresentados nesta secção foram obtidos através do *software Python*. Por sua vez, todos os *outputs* exibidos em cada subcapítulo foram extraídos do *Python* e, posteriormente, trabalhados em *Excel*.

3.1 Análise Gráfica e Estatísticas Descritivas

Nesta secção é apresentado graficamente a evolução trimestral do preço das casas entre 1988 e 2019, bem como, a análise descritiva da série em estudo.

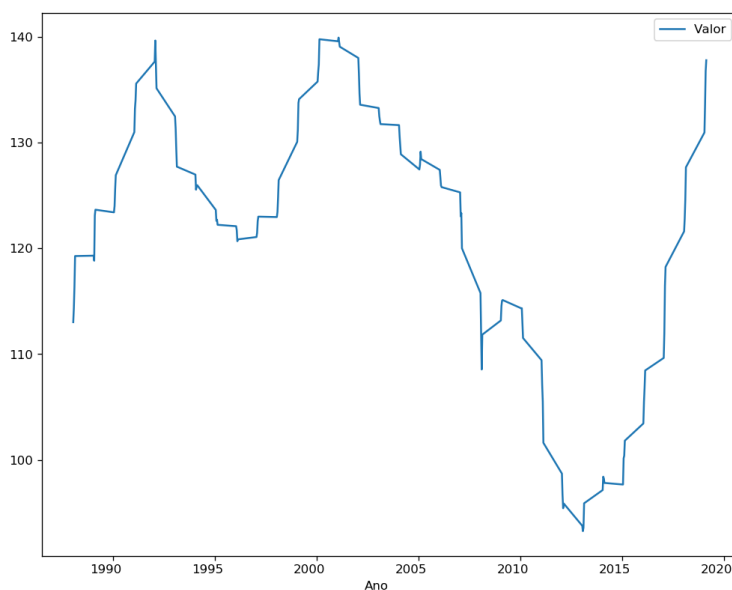


Figura 11: Série preço real das casas em Portugal (Dados Trimestrais)

Na figura acima apresentada, consegue-se observar uma série não linear sem uma tendência visível ao longo do tempo.

Desta maneira, importa frisar 2 momentos do tempo cruciais, onde é possível observar uma possível tendência na evolução dos preços. São estes:

1. **Finais de 2002 até 2013:** Existe uma tendência de baixa nos preços ao longo deste período, impulsionado pela recessão que era sentida em Portugal no início da década de 2000 provocada pela crise mundial, conhecida como o “ponto com”. Todavia, observa-se uma ligeira subida de preços em 2006 originada por uma ligeira retoma da economia portuguesa, contudo, a Grande Recessão de 2008 continuou a catapultar o preço para mínimos históricos;
2. **De 2013 até ao momento:** Num Portugal pós-crise, observa-se a maior subida dos preços à habitação alguma vez registada em Portugal, impulsionado, principalmente, por fatores macroeconómicos favoráveis ao crescimento deste setor.

Em suma, observa-se uma grande volatilidade no preço ao longo do tempo, o que dá fortes indicações que a série analisada pode não ser estacionária.

Feita a análise gráfica, procedeu-se ao estudo empírico da série através do auxílio de estatísticas descritivas e do teste de normalidade de *Jarque-Bera*, de forma a conhecer melhor a série temporal em estudo.

Estatísticas	Valor
Observações	128.00
Média	121.21
Desvio-Padrão	12.69
Mínimo	93.25
Q25	113.14
Q50	123.38
Q75	131.05
Máximo	139.96
Assimetria	-0.61
Curtose	-0.56

Tabela 1: Resultado das estatísticas descritivas obtidas na série do preço real das casas em Portugal.

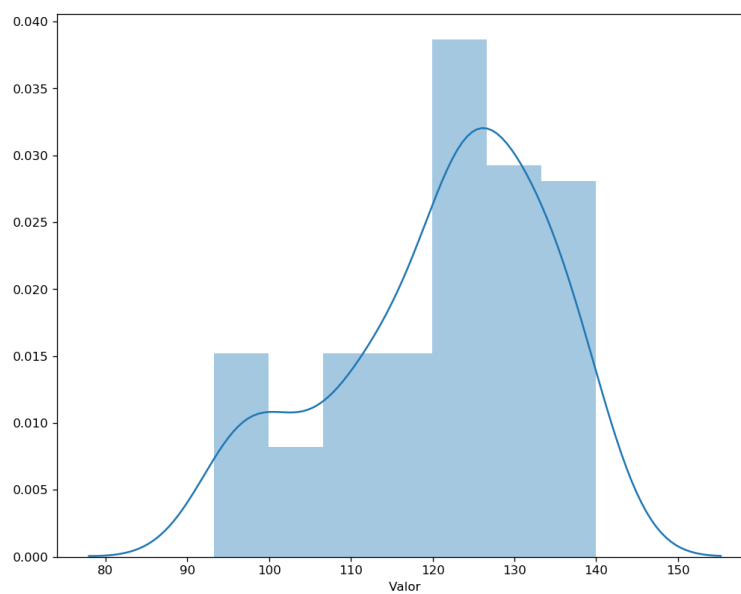


Figura 12: Histograma – Série do preço das casas em Portugal

Teste	Valor	P-value
Jarque-Bera	9.54	0.0008

Tabela 2: Teste da Normalidade aplicado à série do preço real das casas em Portugal.

Os resultados obtidos na estatística descritiva da série confirmam aquilo que era observado no gráfico da evolução do preço ao longo do tempo.

Os níveis de curtose sugerem uma série com uma distribuição platicúrtica dado que $k = -0.56 < 3$;

Por sua vez, a assimetria $s = -0.61 < 0$ sugere uma distribuição assimétrica negativa onde Média \leq Mediana \leq Moda;

Também o histograma sugere que a série tem uma distribuição dispare da distribuição normal;

Por fim, recorreu-se ao teste de *Jarque-Bera* com o intuito de aferir a normalidade da série. Como o $p\text{-value} = 0.0008$ é inferior a todos os níveis de significância (1%, 5% e 10%) rejeitamos a hipótese nula, concluindo assim que a série não segue uma distribuição normal.

3.2 Estacionariedade da série

Um processo estacionário tem a propriedade de que a média, variância e a covariância são constantes ao longo do tempo. O estudo de uma série não estacionária pode ter consequências económicas severas no comportamento do modelo.

Neste subcapítulo iremos estudar a estacionariedade da série do preço real das casas portuguesas com o auxílio dos 3 testes seguintes:

1. *Augmented Dickey-Fuller* (ADF), indicando, ou não, a presença de uma raiz unitária no modelo;
2. *Phillips-Perron* (PP) que, tal como o teste ADF, testa a hipótese de o modelo ter uma raiz unitária e, deste modo, a série ser integrada de primeira ordem;
3. *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS) que, ao contrário dos testes referidos nos pontos anteriores, testa a hipótese nula da série temporal ser estacionária ao longo do tempo.

Na tabela abaixo apresenta-se os resultados obtidos nos 3 testes de estacionariedade realizados:

Teste	Valor	P-value
ADF	-1.69	0.4355
PP	-1.65	0.4590
KPSS	0.41	0.0734

Tabela 3: Testes de Estacionariedade aplicados à série do preço real das casas em Portugal.

A suspeita da não estacionariedade da série temporal referida no subcapítulo anterior é confirmada com o resultado destes testes.

Tanto para o teste de *Augmented Dickey-Fuller* como para o teste de *Phillips-Perron* não rejeitamos a hipótese nula, visto que o *P-value* é superior a todos os níveis de significância. Deste modo, conclui-se, por via destes 2 testes, que a série em análise não é estacionária.

Também para o teste *Kwiatkowski-Phillips-Schmidt-Shin* conclui-se que a série não é estacionária para o nível de significância de 10%. Por outro lado, este teste admite a estacionariedade da série para os níveis de significância 1% e 5%.

Em resumo, dado que o teste do KPSS é um complemento dos restantes e que o mesmo atribui a não estacionariedade da série para o nível crítico de 10%, ao par que os testes ADF e PP atribuem a não estacionariedade da série para todos os níveis de significância, podemos concluir que a série temporal analisada não é estacionária.

Neste sentido, o foco das próximas duas secções recai sobre a estacionarização da série temporal até aqui analisada.

3.3 Análise Gráfica e Estatísticas Descritivas – DIFLOG / Séries de Retornos

Tendo em vista a estacionarização da série temporal, aplicou-se o processo das diferenças logaritmizadas à mesma, obtendo os seguintes resultados:

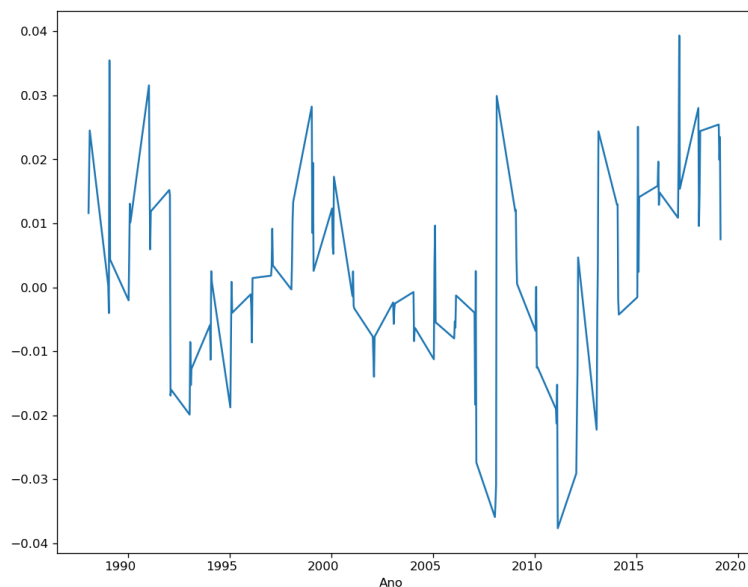


Figura 13: Série dos retornos do preço real das casas - DIFLOG

Pela figura apresentada acima, consegue-se observar uma estabilização do valor médio em níveis próximos de zero. Com a transformação do preço em retornos, também é possível observar uma menor volatilidade nos resultados, dando fortes indícios que esta série é estacionária.

Após representado graficamente a série dos retornos, é realizado um estudo empírico sobre a mesma, de forma a conhecer melhor o tipo de série que será avaliada.

Estatísticas	Valor
Observações	127.00
Média	0.00
Desvio-Padrão	0.02
Mínimo	-0.04
Q25	-0.01
Q50	0.00
Q75	0.01
Máximo	0.04
Assimetria	-0.10
Curtose	0.02

Tabela 3: Resultado das estatísticas descritivas obtidas na série dos retornos do preço real das casas.

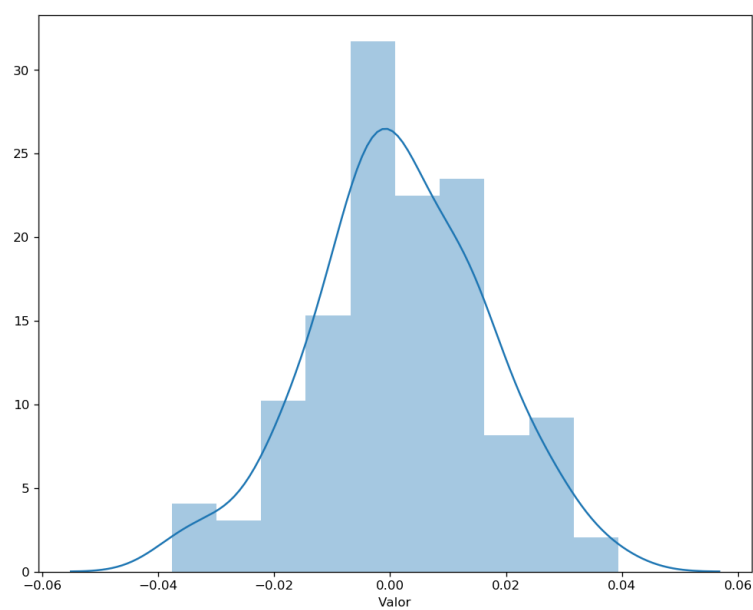


Figura 14: Histograma – Série dos retornos do preço real das casas

Teste	Valor	P-value
Jarque-Bera	0.23	0.8911

Tabela 4: Teste da Normalidade aplicado à série dos retornos.

Os resultados obtidos na estatística descritiva da série dos retornos estão congruentes com os resultados obtidos na análise gráfica.

Os níveis de curtose, à semelhança da série original, indicam que a série segue uma distribuição platicúrtica dado que $k = 0.02 < 3$;

Também o valor da assimetria, $s = -0.10$, mantém o registo da série original, indicando uma distribuição assimétrica (ligeiramente) negativa;

Contudo, o histograma apresenta um comportamento normal na sua distribuição;

Para finalizar a análise, e de maneira a averiguar a normalidade da série, recorreu-se novamente ao teste de *Jarque-Bera*. Desta vez, tem-se que o $p\text{-value} = 0.8911$, ou seja, o $p\text{-value}$ é superior aos níveis de significância usuais (1%, 5% e 10%) levando à não rejeição da hipótese nula, garantindo a normalidade da série.

3.4 Estacionariedade da série dos retornos – Diferença dos Logaritmos (DIFLOG)

Aquando satisfeita a condição de normalidade para a série dos retornos, reúnem-se as condições suficientes para estudar a estacionariedade da série (*DIFLOG*). Este estudo pretende apurar se é possível modelar a série.

Desta forma, tal como foi feito para os dados originais, aplicou-se os testes ADF, PP e KPSS à série dos retornos obtendo o seguinte *output*:

Teste	Valor	P-value
ADF	-3.94	0.0017
PP	-5.92	0.0000
KPSS	0.15	0.1000

Tabela 5: Testes de Estacionariedade aplicados à série dos retornos.

Da análise aos testes ADF e PP – que, como já vimos, têm como hipótese nula a presença de uma raiz unitária na série – conclui-se que os níveis de significância (1%, 5% e 10%) são superiores aos $p\text{-values}$ garantindo que a série não tem raiz unitária.

Por sua vez, o $p\text{-value}$ do teste KPSS é igual ao maior nível de significância, garantindo que a hipótese nula não é rejeitada, ou seja, a estacionariedade para a série DIFLOG também é garantida por este teste.

Satisfeitos os três testes de estacionariedade existem condições para afirmar que a série dos retornos é estacionária, atingindo o objetivo proposto no final da secção 3.2.

Posto isto, todas as análises que se seguem são suportadas pela série dos retornos do preço das casas em Portugal.

3.5 Estimação dos Modelos Econométricos

Como vimos nos subcapítulos anteriores, a série original dos dados segue um modelo ARIMA integrada de ordem 1, dado que, foi necessário diferenciar a série uma vez para obter-se uma série estacionária. Por sua vez, a série dos retornos segue um modelo do tipo ARMA.

Com o objetivo de encontrar o melhor modelo ARMA que se adequa à série das diferenças logaritmizadas, iremos dividir esta secção nas seguintes subsecções:

1. Estudo da autocorrelação (ACF) e autocorrelação parcial (PACF) do modelo, com a finalidade de identificar a ordem do mesmo através de análises gráficas;
2. Seleção do melhor modelo ARMA que se adequa à série.

3.5.1 Identificação do Modelo

Conforme o mencionado, esta subsecção pretende, através de representações gráficas do ACF e do PACF, determinar a ordem do modelo ARMA.

Abaixo apresentam-se os outputs obtidos de cada uma das funções.

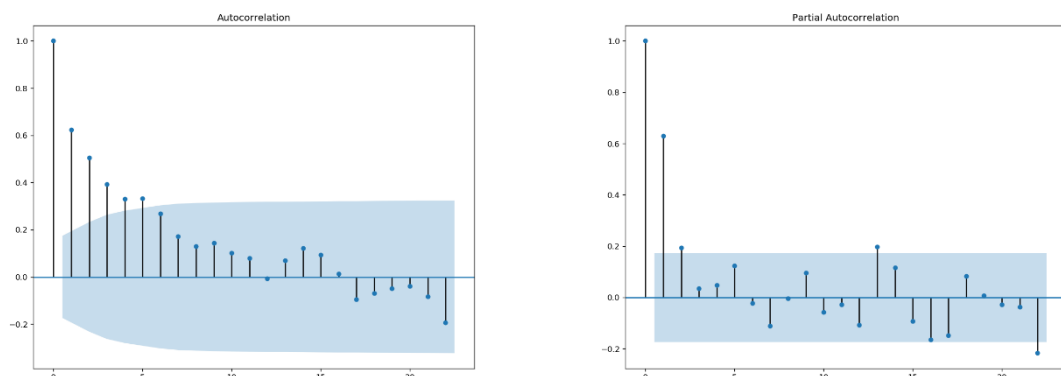


Figura 15: Análise das funções ACF e PACF da série dos retornos

A análise gráfica da função ACF apresenta um declínio geométrico para 0 sugerindo um modelo autorregressivo.

No que concerne à função PACF, a análise gráfica apresenta 2 picos significativos sugerindo um modelo autorregressivo de ordem 2.

Desta feita, na subsecção seguinte iremos testar diferentes tipos de modelos até ordem 2 de forma a encontrar o modelo que melhor se ajusta aos dados.

3.5.2 Seleção do Modelo

Nesta subsecção iremos escolher, qual o modelo que melhor se ajusta à série em análise. A decisão vai recair sobre os seguintes critérios de informação:

1. *Akaike's Information Criterion* (AIC);
2. *Bayesian Information Criterion* (BIC);
3. *Hannan-Quinn Information Criterion* (HQIC).

Assinale-se que, quanto menor forem estes critérios melhor será o modelo.

De todos os modelos testados apenas os modelos AR(1), AR(2), MA(1), MA(2) e ARMA(1,1) serviram de objeto de estudo porque os betas dos restantes modelos tinham *p-values* superiores aos valores críticos usuais – sendo estes automaticamente eliminados.

No quadro abaixo apresentam-se os valores obtidos para estes critérios de cada um dos modelos testados:

Modelo	AIC	BIC	HQIC
AR (1)	-762.832	-754.299	-759.365
AR (2)	-765.929	-754.552	-761.307
MA (1)	-738.508	-729.975	-735.041
MA (2)	-752.833	-741.456	-748.211
ARMA (1,1)	-766.553	-755.176	-761.931

Tabela 6: Critérios de Seleção Modelos ARMA/ARIMA aplicados à série dos retornos.

Após a identificação sugestiva do melhor modelo a aplicar à série dos retornos através da análise gráfica do ACF e PACF (subsecção 3.5.1) consegue-se concluir com o auxílio do quadro acima que o modelo que minimiza os 3 critérios de seleção é o ARMA(1,1) sendo, desta forma, o escolhido para modelar a série temporal.

Considere-se o modelo de regressão linear analisado no capítulo 2:

$$X_t = \rho X_{t-1} + \beta_0 + \beta_1 t + \varepsilon_t$$

O modelo ARMA(1,1) escolhido para modelar a série temporal dos retornos, pode ser escrito da seguinte forma:

$$X_t = 0.0023 + 0.8261 \times X_{t-1} + (-0.3503) \times t + \varepsilon_t$$

O modelo acima apresentado mostra que, tanto o ρ como β_1 apresentam valores diferentes de zero. Desta forma, pode-se concluir que o modelo apresentado tem 1 *lag* em X_t ($\rho \neq 0$) e 1 *lag* no termo de erro ($\beta_1 \neq 0$), confirmando-se que se trata de um modelo autorregressivo de médias móveis ARMA(1,1)

Como $\beta_0 \neq 0$, a série está perante um passeio aleatório com *drift* e, deste modo, o valor previsto para X_t no futuro não depende apenas do valor presente de X aliado a uma combinação de valores passados de um processo de ruído branco (termo de erro).

Por outro lado, como $0 \leq \rho \leq 1$, existe correlação serial em X e a série diz-se estacionária.

3.6 Análise dos resíduos dos modelos ARMA/ARIMA

Neste subcapítulo pretende-se estudar os pressupostos dos resíduos do melhor modelo obtido na secção anterior - ARMA(1,1). Esta análise tem como objetivo validar se o modelo passa pelos critérios que permitam a sua previsão.

A avaliação do modelo que permite avaliar se o mesmo pode modelar a série dos retornos recairá nos 3 critérios seguintes:

1. Média Nula;
2. Homocedasticidade dos resíduos;
3. Independência dos resíduos.

Não obstante aos testes que serão realizados, é importante perceber como se comportam os resíduos ao longo do tempo e a sua distribuição.

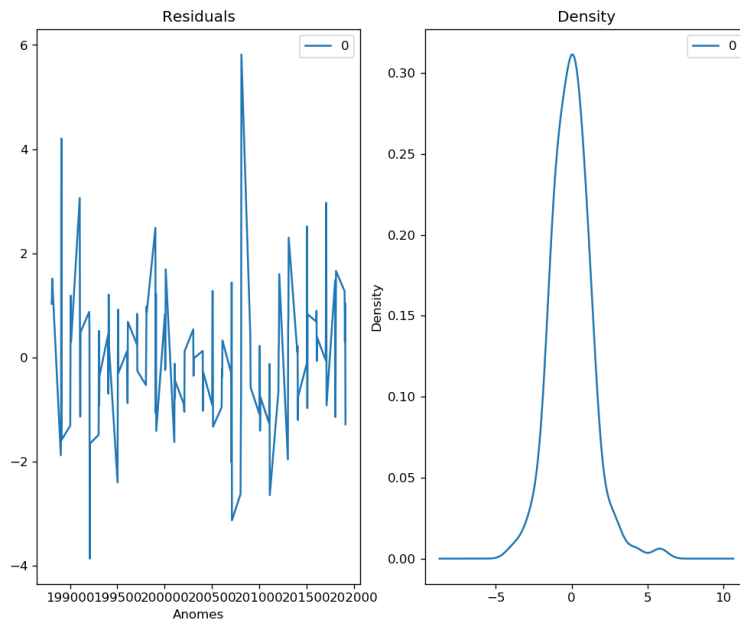


Figura 16: Comportamento dos resíduos

Pela análise gráfica observamos que os resíduos evoluem ao longo do tempo por valores próximos de zero e que seguem uma distribuição normal.

Ambos os indicadores dão fortes indícios que o modelo poderá dar boas previsões sobre a série.

Média Nula

O primeiro pressuposto que iremos validar centra-se na média dos resíduos. Para este pressuposto ser validado a média dos resíduos do modelo ARMA(1,1) para a série dos retornos deve ser nula.

Estatísticas	Valor
Média	0.00

Tabela 7: Primeiro Pressuposto dos Resíduos – Média Nula.

Conforme apresentado na tabela o primeiro pressuposto para os resíduos é cumprido.

Homocedasticidade dos Resíduos

A homocedasticidade dos resíduos pretende apurar se a variância é estável ao longo do tempo. Para validar este pressuposto recorreremos ao teste ARCH que tem como hipótese nula a não existência de efeitos ARCH nos resíduos, ou seja, a variância da série é constante.

Teste	Valor	P-value
ARCH	1.82	0.8738

Tabela 8: Segundo Pressuposto dos Resíduos – Homocedasticidade dos Resíduos.

Pela tabela acima pode-se observar que o *p-value* do teste ARCH é superior a todos os níveis de significância usuais levando à não rejeição da hipótese nula. Deste modo, conclui-se que não existem efeitos ARCH nos resíduos e, conseqüentemente, a variância é constante ao longo do tempo.

Validado o segundo pressuposto, estão reunidas as condições para avaliar o terceiro e último requisito de avaliação dos resíduos.

Independência dos Resíduos

A fim de verificar que o modelo se adequa corretamente aos dados da série dos retornos, os resíduos devem, além de ter média nula e variância constante, ser independentes.

Para validar este pressuposto recorreu-se ao teste *Ljung-Box* até ordem 10 que tem como hipótese nula a não-correlação dos resíduos. Paralelamente também se recorreu ao teste *Durbin-Watson* que testa a correlação dos resíduos de ordem 1.

Abaixo apresenta-se os resultados obtidos em ambos os testes:

Teste	Valor	P-value
Ljung-Box	0.015	0.8999
	0.017	0.9915
	0.476	0.9240
	0.791	0.9396
	2.432	0.9867
	3.149	0.7900
	3.737	0.8095
	4.235	0.8353
	4.688	0.8606
	4.720	0.9091

Tabela 9: Terceiro Pressuposto dos resíduos – Independência dos Resíduos Ljung-Box.

Teste	Valor
Durbin-Watson	1.97

Tabela 10: Quarto Pressuposto dos resíduos – Independência dos Resíduos Durbin-Watson.

No teste de *Ljung-Box*, como os *p-values* são superiores aos níveis de significância usuais não se rejeita a hipótese nula, ou seja, os resíduos são independentes entre si – até ao lag de ordem 10.

Relativamente ao teste de *Durbin-Watson* verifica-se que o mesmo apresenta um valor muito próximo de 2 indicando que não há, praticamente, correlação de primeira ordem entre os resíduos.

O modelo ARIMA(1,1,1), para a série original dos dados, passou nos 3 testes aplicados aos resíduos. Desta forma, valida-se o modelo e conclui-se que o mesmo pode ser usado como objeto de previsão.

3.7 Previsão da série

Na última secção desta dissertação irá ser abordado o tema da previsão da série temporal. Após escolhido o melhor modelo, reúne-se as condições necessárias para avaliar se o mesmo apresenta uma boa previsão sobre a série dos dados.

Neste sentido, a previsão irá ser realizada sobre a série dos retornos dos dados, i.e., avaliar-se-á se o modelo ARMA(1,1) se ajusta à série dos retornos através da análise gráfica

da previsão *in-sample* e do erro de previsão obtido com o teste MSE – erro quadrático médio. Posteriormente, com o auxílio deste modelo, será feita uma avaliação sobre preço futuro das casas em Portugal através da previsão *out-of-sample*. Esta análise incide sobre os próximos 8 retornos estimados pelo modelo, indicando qual o comportamento futuro que é esperado para série em análise.

3.7.1 Previsão In-Sample

Tal como mencionado anteriormente, nesta secção pretende-se avaliar como é que o modelo ARMA(1,1) prevê a série dos retornos nas observações *in-sample* - recorde-se que, o modelo ARMA(1,1) para a série dos retornos é equivalente ao modelo ARIMA(1,1,1) para a série original dos dados.

Mais concretamente, será avaliado, através de observação gráfica, o comportamento do modelo ao prever as observações passadas da série, i.e., os dados da série que são previamente conhecidos.

Caso o modelo apresente uma boa previsão da série, existem condições para prever as suas observações futuras, *out-of-sample*.

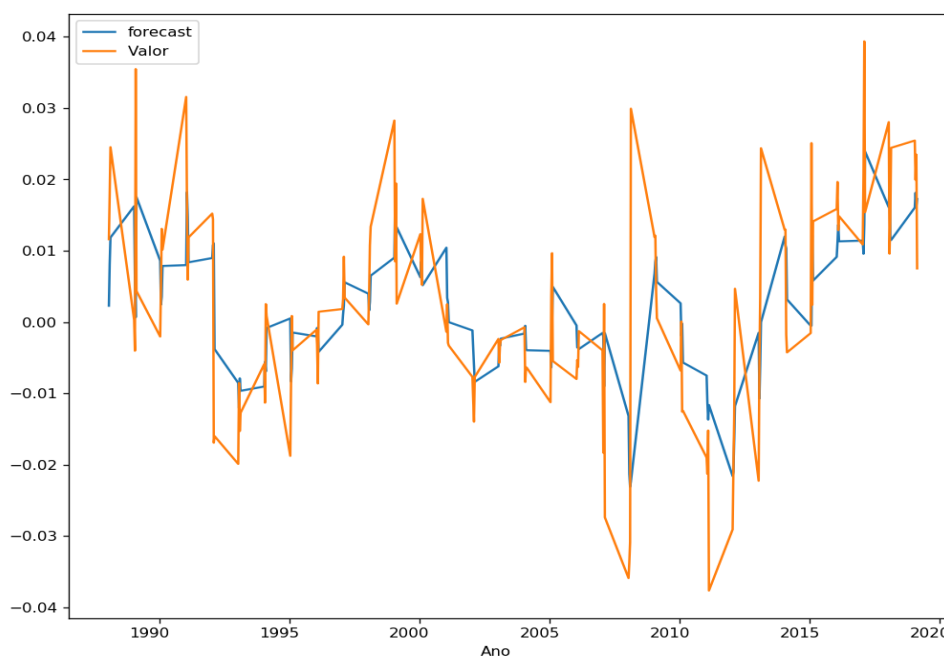


Figura 17: Previsão *In-Sample* da série dos retornos do preço real das casas

Através da observação gráfica, pode-se observar que o modelo escolhido para modelar os dados da série apresenta bons resultados de previsão, dado que, embora

apresente menor oscilação, consegue representar os momentos de maiores retornos no mercado imobiliário e os momentos onde o setor sofreu uma contração dos preços.

Teste	Valor
MSE	1.31E-04

Tabela 11: Erro de previsão obtido no período *in-sample*

Recorrendo ao teste do erro quadrático médio (MSE), observa-se que o mesmo toma um valor muito pequeno, indicando que o modelo ARMA (1,1) para a série dos retornos tem uma boa capacidade de previsão.

De acordo com as conclusões retiradas para a mostra *in-sample*, verifica-se que o modelo é adequado para realizar uma previsão *out-of-sample* da série.

3.7.2 Análise Out-Of-Sample

Por último, e de acordo com a ordem de trabalhos estabelecida para este subcapítulo, segue-se a análise out-of-sample da previsão dos retornos do preço dos imóveis para os próximos 6 períodos, com base no modelo ARMA(1,1).

Neste caso serão preditos os 8 valores dos trimestres futuros desde a última observação, que correspondem aos anos 2020 e 2021.

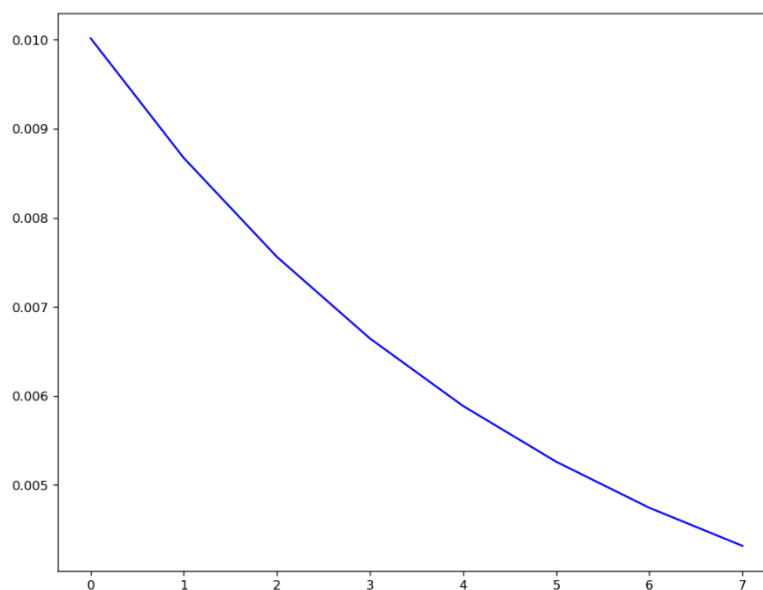


Figura 188: Previsão *Out-of-Sample* da série dos retornos do preço das casas

Como observado no gráfico acima, o modelo retorna valores cada vez menores para os sucessivos trimestres - note-se que, cada observação de [0,7] corresponde aos 8 trimestres desde 2020.

Da análise gráfica pode-se concluir que, embora o modelo aponte para um mercado em crescimento, os seus retornos ao longo do tempo serão cada vez menores, aproximando-se progressivamente de 0. Desta forma, o modelo indica que os investidores podem contar com lucros cada vez menores neste setor.

Data a Prever	Valor Predito
2020-Q1	0.010
2020-Q2	0.009
2020-Q3	0.008
2020-Q4	0.007
2021-Q1	0.006
2021-Q2	0.005
2021-Q3	0.005
2021-Q4	0.004

Tabela 12: Retornos preditos pelo modelo ARMA(1,1) para os próximos 2 anos

Com o intuito de perceber qual o erro de previsão obtido para amostra *out-of-sample*, iremos, com o auxílio do teste MSE, avaliar o erro de previsão obtido nos primeiros

2 trimestres de 2020, ou seja, irá ser comparado os valores preditos com os valores reais dos retornos obtidos nos primeiros 2 trimestres de 2020.

Data a Prever	Valor Registrado	Valor Predito	Erro de Previsão
2020-Q1	0.015	0.010	2.90E-05
2020-Q2	0.000	0.009	7.54E-05

Tabela 13: Erro de previsão obtido na previsão *out-of-sample*

De acordo com a tabela 13, conclui-se que os retornos preditos são semelhantes aos reais. De notar que existiu uma quebra nos retornos mais a bruta do que o era esperado pelo modelo para o segundo trimestre de 2020. No entanto, o erro de previsão de cada observação toma um valor pequeno, dando sinais de uma boa previsão sobre os retornos futuros da série.

Não obstante aos resultados apresentados, deve-se ter em conta que os mesmos são meros indicadores de preços e não têm em consideração a evolução da economia e dos cenários macroeconómicos do país. Estes resultados são meros indicadores do que se pode esperar do mercado imobiliário português, tendo apenas em consideração a série temporal histórica que o mesmo apresenta.

Conclusão

O objetivo desta dissertação passou pelo estudo do mercado imobiliário português que, até à última observação da série estudada, tem uma tendência positiva no preço muito influenciado pelos fatores macroeconómicos.

Neste sentido, a série do preço das casas em Portugal foi tratada como uma série temporal, onde o objetivo foi centrado na seleção do modelo que melhor se ajustasse aos seus dados e, conseqüentemente, tivesse um bom poder preditivo no comportamento futuro do preço dos imóveis portugueses.

Para se obter o melhor modelo recorreu-se a vários critérios de informação que serviram de comparabilidade dos modelos “aptos” a estimar/prever a série. Identificou-se que o modelo ARIMA(1,1,1) era o melhor modelo de ajustamento a estes dados. Posteriormente realizou-se a previsão do preço futuro sobre este modelo onde se concluiu que, embora se continue a verificar a tendência crescente no mercado imobiliário, estima-se que os retornos esperados sejam cada vez menores, levando possivelmente a uma estagnação dos preços no futuro, ou a uma possível queda.

De notar que existiram algumas limitações no desenvolvimento deste estudo e os resultados apresentados devem ser considerados como meros indicadores, que podem não corresponder aos resultados mais elucidativos. Por um lado, tratando-se de uma série de dados trimestrais, a janela de amostras observada é bastante reduzida (128 observações), o que pode contribuir para uma menor robustez dos resultados apresentados. Por outro lado, a série tem dados até 2019, data em que Portugal ainda não tinha registado nenhum caso positivo de Covid-19. Desta forma, nem os indicadores macroeconómicos portugueses que influenciam o comportamento da série estudada, nem a própria série, tinham sofrido os impactos resultantes da pandemia mundial. Assim, importa ressaltar o leitor que as previsões *out-of-sample* não se encontram ajustadas à realidade económica e que a redução dos preços poderá ser mais fugaz do que o que é apresentado neste projeto, tendo em conta os resultados mais recentes da taxa de desemprego, do turismo e do PIB português.

Bibliografia

- Andrade, V. (2020). *Porque é que os preços das casas não caem apesar da pandemia?*. Jornal Expresso.
- Aníbal, S. (2018). *Novo alerta de bolha especulativa imobiliária em Portugal*. Jornal Público.
- Boura, T. (2015). *Risco de estrutura temporal de taxas de juro em financiamentos indexados*. ISCTE Business School.
- Box, G e Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*, 2ª Edição. San Francisco, Holden-Day.
- Campos, F. (2020). Covid-19: como a pandemia afeta o mercado imobiliário?. Available at: https://www.deco.proteste.pt/investe/investimentos/imobiliario/comprar-casa/noticias/2020/03/covid-19-como-afeta-mercado-imobiliario?fbclid=IwAR1GdNKVOpULLV0DtYxD2zey-0ewb_W9OoIr21eeCberIc-Qooa9_NrC1ko
- Davidson, R. e MacKinnon, J. (2004). *Econometric Theory and Methods*. Oxford University Press.
- Fanico, M. (2009). *Mercado Imobiliário Português: Performance, Ciclos e Tendências*. ISCTE Business School
- Fernandes, C. (2018). *A Influência da Presidência de Trump sobre as Séries Financeiras Americanas*. ISCTE Business School.
- Gonçalves, N. (2020). *A economia e a Covid-19*. A Visão.
- Henriques, J. (2017). *A bolha imobiliária de 2008: uma análise acerca da evolução do valor dos imóveis em Portugal*. CEGOT.
- Kitagawa, G. (2010). *Introduction to time series modeling*. Florida, CRC Press.
- Mariano, A. (2018). *A Crise de 2008 (a crise do “subprime”) explicada de forma simples*. Available at: <https://arrowplus.pt/a-crise-de-2008-a-crise-do-subprime-exlicada-de-forma-simples>
- Mateus, C (2020). *Taxa de desemprego volta a subir e atinge 8.1%*. Jornal Expresso
- Marques, A. (2017). *Desafios na Previsão de Séries Temporais Financeiras: O caso da Taxa de Câmbio EUR/USD*. ISCTE Business School.
- Mills, T.C. (2019). *Applied Time Series Analysis: A Practical Guide to Modeling and Forecasting*, Academic Press, Elsevier Inc. Loughborough University, United Kingdom.
- Milheiro, J. (2019). *Novo Recorde: mercado imobiliário já vale 12% do PIB*. TSF Rádio Notícias.
- Murteira, B.; Muller, D. e Turkman, K. (1993). *Análise de Sucessões Cronológicas*. McGraw Hill.

- Passy, J. (2020). *America's housing market is showing the first signs of trouble from the coronavirus pandemic*. Available at:
https://www.deco.proteste.pt/investe/investimentos/imobiliario/comprar-casa/noticias/2020/03/covid-19-como-afeta-mercado-imobiliario?fbclid=IwAR1GdNKVOpULLV0DtYxD2zey-0ewb_W9OoIr21eeCberIc-Qooa9_NrC1ko
- Percival, D. e Walden, T. (1993). *Spectral Analysis for Physical Applications*. Cambridge University Press.
- Sousa, J. (2020). *Mercado imobiliário em Portugal dá sinais de abrandamento e “futuro incerto”*. *Jornal Económico*.
- Tavares, E. (2018). *O Lehman caiu há dez anos. E Portugal ainda recupera*. *Diário de Notícias*.
- Tavares, M. (2019). *Are There Signals of a Housing Bubble? An Empirical Analysis of The Portuguese Case*. ISCTE Business School.
- Tiago, R. (2018). *Taxa de juros: como a sua variação influencia a economia do país?*. *Suno Artigos*.
- Viana, S. (2019). *As taxas de juro e o seu impacto no mercado imobiliário*. *Jornal Capitalmag*.
- Vieira, A. (2016). *O impacto económico do turismo – uma abordagem espacial*. Faculdade de Economia do Porto.