

Maria Isabel Vaz Pitacas

**ESTUDO COMPARATIVO DE MÉTODOS DE PREVISÃO
APLICAÇÃO À SÉRIE DO NÚMERO SEMANAL
DE ADULTOS ALOJADOS NUM HOTEL**



Departamento de Matemática Aplicada
Faculdade de Ciências da Universidade do Porto
Dezembro / 1999

TM

**ESTUDO COMPARATIVO DE MÉTODOS DE PREVISÃO
APLICAÇÃO À SÉRIE DO NÚMERO SEMANAL
DE ADULTOS ALOJADOS NUM HOTEL**



*Tese submetida à Faculdade de Ciências da Universidade do Porto
para a obtenção do grau de Mestre em Estatística*

Departamento de Matemática Aplicada
Faculdade de Ciências da Universidade do Porto
Dezembro / 1999

Dissertação orientada por

Maria Eduarda da Rocha Pinto Augusto da Silva

Professora Auxiliar do
Departamento de Matemática Aplicada da
Faculdade de Ciências da Universidade do Porto

*Ao meu marido Vitor
e à minha filha Patrícia*

Agradecimentos

Gostaria de começar por agradecer à minha orientadora, Professora Doutora Maria Eduarda Silva pelo apoio científico dado a este projecto. Os seus conselhos e críticas foram muito construtivos, pois conduziram-me pelas linhas deste trabalho de uma forma mais segura e esclarecida.

Aos colegas e amigos que me apoiaram e encorajaram.

Ao Instituto Politécnico de Tomar por todo o apoio concedido, principalmente no acesso ao *software* indispensável a este trabalho.

Ao hotel e em especial ao seu director, que pediu o anonimato, pelo fornecimento incondicional dos dados.

À D. Ana Bela Vieira pela dedicação e paciência que demonstrou nestes meses ao ajudar-me, principalmente com a Patrícia.

À mãe e à irmã do Vitor por cuidarem da Patrícia.

Aos meus pais pelo carinho e apoio que me deram no decorrer deste trabalho.

Ao meu marido, Vitor, pelo carinho, compreensão e paciência incondicionais.

À minha filha pelo seu sorriso.

A todos, o meu muito obrigado.

Resumo

Neste trabalho o nosso objectivo é obter previsões para as séries temporais do número semanal de adultos alojados num hotel, para o que utilizámos três métodos diferentes, nomeadamente: previsão Holt-Winters, modelos estruturais e modelação Box-Jenkins ARIMA.

Concluimos que, embora o melhor ajustamento em termos de variância residual seja fornecido pelo método de decomposição clássica, as melhores previsões são dados pelo Modelo Estrutural Básico (MEB).

Palavras chave: séries temporais; decomposição clássica; Holt-Winters; modelos estruturais; filtro de Kalman; Box-Jenkins; modelos ARIMA

Abstract

Our aim is to obtain predictions for the time series of the weekly occupancy rate of adult customers in an hotel, we use three different methodologies namely Holt-Winters prediction, structural models and Box-Jenkins ARIMA approach.

We conclude that although the best fit in terms of residual variance is given by the classical decomposition method, the best predictions are given by the Basic Structural Model (BSM).

Key words: time series; classical decomposition; Holt-Winters; structural models; Kalman filter; Box-Jenkins; ARIMA models

Índice

Dedicatória	3
Agradecimentos	4
Resumo	5
Abstract	6
Índice	7
Capítulo 1. Introdução	9
1.1. Motivação e Estrutura do Trabalho	9
1.2. Conceitos Fundamentais	10
Capítulo 2. Método de Análise de Séries Temporais	13
2.1. Introdução	13
2.2. Método de Decomposição Clássica	13
2.2.1. Modelo Aditivo	14
2.2.2. Previsão	17
2.3. Modelação Box-Jenkins	18
2.3.1. Identificação do Modelo	20
2.3.2. Estimção e Diagnóstico	21
2.3.3. Previsão	22
2.4. Modelos Estruturais	23
2.4.1. Modelo Estrutural Simples	24
2.4.2. Modelo Estrutural Básico (MEB)	25
2.4.3. Modelo Cíclico	27
2.4.4. Os Modelos Estruturais na Sua Forma Reduzida	28
2.4.5. Selecção do modelo	32
2.4.6. Análise da Componente Residual	32
2.4.7. Modelos em Espaço de Estados	32
2.4.8. Filtro de Kalman	36

Capítulo 3. Modelação do Número Semanal de Adultos Alojados num Hotel	40
3.1. Apresentação da Série	40
3.2. Diagnóstico Preliminar dos Dados	41
3.3. Modelação da Série Através da Decomposição Clássica	43
3.3.1. Estimação das Componentes	44
3.3.2. Estudo da Componente Residual	45
3.4. Modelação Box-Jenkins	48
3.4.1. Identificação do Modelo	48
3.4.2. Estimação dos Parâmetros	49
3.4.3. Estudo dos Resíduos	50
3.5. Modelo Estrutural	53
3.5.1. Estimação das Componentes	53
3.5.2. Estudo da Componente Residual	54
3.6. Comparação de Resultados	55
3.7. Previsão	57
3.8. Conclusões	60
Bibliografia	62
Anexo I	63
Anexo II	65
Anexo III	66
Anexo IV	68
Anexo V	70
Anexo VI	71
Anexo VII	72
Anexo VIII.	73

Capítulo 1. Introdução

1.1. Motivação e Estrutura do Trabalho

Define-se uma série temporal como um conjunto de observações associadas a determinado fenómeno aleatório, efectuadas em períodos sucessivos de tempo e estatisticamente relacionadas.

Ao analisar uma série temporal pretende-se, de uma forma geral, alcançar dois objectivos, os quais se encontram fortemente relacionados. O primeiro, a *modelação*, consiste em encontrar um modelo que tenha em conta as relações existentes entre as observações, permitindo a descrição da série temporal. O segundo, diz respeito à *previsão* de valores futuros, pois como observa Murteira (1994:5) “É a própria existência do tempo que leva as pessoas a fazer previsões e a tentar antecipar a evolução no futuro das sucessões que estudam”.

Um dos métodos possíveis para a análise e previsão das séries temporais é a decomposição clássica da série em várias componentes básicas não directamente observáveis. São elas: a *tendência*, *ciclo*, *sazonalidade* e *erro*, que por não serem observáveis, poderão ser estimadas a partir da informação contida nos dados que constituem a série.

Quando estuda determinada série temporal, o analista confronta-se, por vezes, com o problema da não estacionaridade. A metodologia desenvolvida por Box e Jenkins, baseia-se na ideia de que tais séries podem tornar-se estacionárias através de operações de diferenciação. Este método segue essencialmente três etapas: na primeira faz-se a escolha do modelo ARMA, na segunda etapa estimam-se os parâmetros do modelo e na terceira diagnostica-se o comportamento dos resíduos. Na terceira etapa, se o modelo se apresentar como satisfatório, este é usado na previsão de valores futuros. Caso não se apresente satisfatório voltamos à primeira etapa.

A metodologia de Box e Jenkins, relativamente à abordagem decomposicionista, apresenta rigor e controlo, mas não tem uma interpretação e execução óbvias.

Para ultrapassar as dificuldades apresentadas nas duas metodologias anteriores usam-se os modelos estruturais de séries temporais, desenvolvidos por Harvey (1996), que aliam a facilidade de interpretação ao rigor formal e controlo estatístico.

Estes modelos são formulados em termos das componentes tradicionais sendo cada uma destas formalizada como um processo estocástico.

Neste trabalho pretende-se modelar e obter previsões para a série do número semanal de adultos alojados num hotel aplicando as três metodologias atrás referidas. Iniciamos o nosso estudo tentando ajustar à série das observações um modelo de decomposição, prosseguimos com um modelo estrutural e concluímos com o ajustamento de um modelo ARIMA. Os três melhores modelos obtidos com as diferentes metodologias serão depois comparados quer em termos de ajustamento, quer em termos de previsão.

Assim, no Capítulo 2 começamos por descrever a decomposição clássica de uma série temporal e apresentamos o método utilizado na previsão. De seguida descrevemos a metodologia de Box-Jenkins, que se divide nas fases: (i) identificação, (ii) estimação e diagnóstico e (iii) previsão. E por último centramo-nos nos principais modelos estruturais, na sua representação em espaço de estados e no filtro de Kalman.

Finalmente, no Capítulo 3 aplicaremos as diferentes metodologias estudadas na construção de modelos para o número semanal de adultos alojados num hotel. Na sequência dos modelos construídos faremos as previsões para esta série temporal.

1.2. Conceitos Fundamentais

Sendo o objectivo desta introdução abordar alguns conceitos básicos relacionados com a análise de séries temporais, vamos de seguida definir os conceitos de processo estocástico, série temporal, estacionaridade, função autocovariância, função de autocorrelação, função de autocorrelação parcial e ruído branco.

Seja $(\Omega, \mathcal{A}, \mathcal{P})$ um espaço de probabilidades e T um conjunto arbitrário, define-se *processo estocástico* como sendo uma função $Y(t, \omega)$, definida em $T \times \Omega$, que para cada $t \in T$ é uma variável aleatória. E representa-se por $\{Y(t, \omega): \omega \in \Omega, t \in T\}$. Para cada ω_0 fixo, $Y(t, \omega_0)$ é uma realização ou trajectória do processo.

Uma *série temporal*, ou sucessão cronológica, define-se como um conjunto de observações $Y_{t_1}, Y_{t_2}, \dots, Y_{t_N}$ feitas em pontos ou períodos t_1, t_2, \dots, t_N , e representa-se por

$\{Y_t, t = 1, 2, \dots, N\}$. A série temporal também pode ser registada em tempo contínuo, e neste caso representa-se por $\{Y(t), 0 \leq t \leq T\}$.

Uma série temporal é apenas uma das realizações de determinado processo estocástico; no entanto, o termo série temporal é empregue para designar a realização observada do processo estocástico.

O processo $\{Y_t\}$ diz-se estritamente estacionário sse a distribuição conjunta de $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_n})$ é igual à distribuição conjunta de $(Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_n+k})$ para todo o n -úplo (t_1, t_2, \dots, t_n) e todo o k , ou seja, $F_{t_1, t_2, \dots, t_n}(Y_1, Y_2, \dots, Y_{t_n}) = F_{t_1+k, t_2+k, \dots, t_n+k}(Y_1, Y_2, \dots, Y_{t_n})$ em todos os pontos (y_1, y_2, \dots, y_n) e em que $F(\cdot)$ representa a função distribuição conjunta de $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_n})$. De uma forma resumida significa que um processo estritamente estacionário goza da propriedade de que a distribuição de um qualquer conjunto de margens se mantém a mesma quando estas são sujeitas a uma translação no tempo. Na prática esta propriedade não se verifica com facilidade, recorrendo-se por isso à estacionaridade fraca.

Um processo diz-se fracamente estacionário até à ordem m , ou simplesmente estacionário, se os vectores n -dimensionais $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_n})$ e $(Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_n+k})$ possuírem momentos conjuntos até à m -ésima ordem e esses momentos forem iguais.

O processo $\{Y_t, t \in T\}$ diz-se *fracamente estacionário de 1ª ordem* se $E(y_t) = \mu, \forall t$ e *fracamente estacionário de 2ª ordem* ou para a covariância sse for tal, que para todo o t :

$$1) E(Y_t) = \mu, \text{ constante, } \forall t$$

$$2) E(Y_t^2) = \mu_2', \text{ constante, } \forall t$$

$$\text{e como } (Y_t) = E(Y_t^2) - E^2(Y_t) = \mu_2' - \mu^2 = \sigma^2$$

$$3) E(Y_t Y_s) = E[Y_t Y_s] - \mu^2, \text{ i.e., função apenas da diferença } t-s.$$

Seja $\{Y_t, t \in T\}$ um processo estacionário de 2ª ordem em que $E(Y_t) = \mu$ e $\text{var}(Y_t) = \sigma^2$, $t = 0, 1, 2, \dots$ a *função de autocovariância* é dada por $\gamma_k = E\{(Y_t - \mu)(Y_{t+k} - \mu)\} = \text{cov}(Y_t, Y_{t+k})$.

A *função de autocorrelação* (FAC) é dada por $\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \text{corr}(Y_t, Y_{t+k})$ onde para cada k a função ρ_k mede a correlação entre pares de valores do processo separados por um intervalo k .

$$\text{O estimador da FAC é dado por } \hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}, \quad k = 0, 1, \dots \text{ onde } \bar{Y}$$

é a média amostral $\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t$. A estimação da FAC efectua-se para $k > 0$.

As funções de autocovariância e autocorrelação gozam das propriedades:

$$1) \gamma(0) = \text{var}(Y_t); \quad \rho_0 = 1$$

$$2) |\gamma_k| \leq \gamma(0) \text{ e em consequência } |\rho_k| \leq 1 \text{ em virtude da desigualdade de Cauchy-Schwarz: } |E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$$

$$3) \gamma_k = \gamma_{-k}; \quad \rho_k = \rho_{-k}$$

4) As funções γ_k e ρ_k são semidefinidas positivas.

A função de autocorrelação parcial (FACP) é dada por $\phi_{kk} = \text{corr}(Y_t, Y_{t+k})$, retirando os efeitos de Y_t e Y_{t+k-1} . Para cada k mede a correlação entre Y_t e Y_{t+k} , depois de eliminar o efeito produzido pelas variáveis $Y_{t+1}, \dots, Y_{t+k-1}$.

O cálculo do estimador da FACP é feito por meio recursivo partindo de :

$$\hat{\phi}_{11} = \hat{\rho}_1 \text{ e utilizando } \hat{\phi}_{kk} = \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_j} \text{ em que } \hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \phi_{kk} \phi_{k-1,k-j},$$

$$j = 1, 2, \dots, k-1.$$

A representação gráfica da FAC e da FACP designam-se, respectivamente, por correlograma e correlograma parcial estimados, pois na prática geralmente apenas dispomos de uma realização do processo.

O processo $\{e_t\}$ constitui ruído branco se for formado por uma sucessão de variáveis aleatórias independentes e identicamente distribuídas com média constante $E(e_t) = \mu$, geralmente assumida como sendo zero, $\text{var}(e_t) = \sigma^2$ e covariância $\gamma_k = 0$ para $k \neq 0$, ou seja, é um processo estacionário de 2ª ordem

$$\text{com } \rho_k = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \text{ e } \gamma_k = \begin{cases} \sigma^2, & k = 0 \\ 0, & k \geq 1 \end{cases}$$

Quando a amostra é significativamente grande, assume-se que $\tilde{\rho}_k$ tem distribuição $N(0, 1/n)$. Os valores (95%), tanto da FAC como da FACP, deverão situar-se dentro dos limites de confiança: $\pm 1.96/\sqrt{n}$.

Capítulo 2. Método de Análise de Séries Temporais

2.1. Introdução

Neste capítulo faremos uma abordagem aos três métodos de análise de séries temporais utilizados no estudo da série das observações. Começamos pelo método de decomposição clássica, prosseguimos com a modelação Box-Jenkins e finalizamos com os modelos estruturais.

2.2. Método de Decomposição Clássica

A decomposição das séries temporais começou a ser estudada nos finais do século passado e desenvolveu-se nos princípios deste século. Este foi um dos primeiros métodos utilizados na análise de séries temporais (Makridakis, Wheelwright & Hyndman, 1998).

Este método consiste numa descrição matemática dos movimentos presentes nas séries, os quais após serem identificados, podem ser utilizados para facilitar a interpretação e projecção da série temporal. Os movimentos, ou componentes, que se distinguem nas séries temporais são usualmente classificados em quatro tipos:

Tendência (T_t) - Variação em média ou variância ao longo do tempo, ou ainda, mudança de nível. Movimentos que se manifestam suavemente e consistentemente durante períodos longos.

Componente Sazonal (S_t) - Representa uma variabilidade periódica em relação à tendência.

Componente Cíclica (C_t) - Esta componente associa-se às fases alternadas de expansão e depressão que afectam a série, mas não apresenta qualquer periodicidade definida. Os ciclos longos são dificilmente separáveis da tendência.

Componente Residual (e_t) - Esta componente representa tudo o que não pode ser atribuído às outras três componentes.

Embora todas as séries temporais contenham a componente residual, as outras três componentes, que reflectem um certo padrão comportamental, podem não estar presentes na sua totalidade.

O objectivo da análise das séries temporais é identificar as componentes presentes de modo a identificar as suas causas e prever valores futuros.

Então, sendo y_t a série dos valores observados,

$$y_t = f(T_t, S_t, C_t, e_t) \quad (2.1)$$

é a expressão genérica do modelo de decomposição.

Este modelo pode ser do tipo aditivo

$$y_t = T_t + S_t + C_t + e_t, \quad (2.2)$$

do tipo multiplicativo

$$y_t = T_t S_t C_t e_t \quad (2.3)$$

ou ter uma forma mista

$$y_t = T_t S_t C_t + e_t \quad (2.4)$$

A opção entre (2.2) e (2.3) baseia-se no tipo de sazonalidade presente, e uma das formas de perspectivar as características desta passa pela observação do cronograma. Se se observar que com o aumento (ou diminuição) do nível da tendência há um aumento (ou diminuição) da amplitude dos movimentos periódicos, então é aconselhável um modelo do tipo (2.3). Por outro lado, se os movimentos periódicos permanecerem estáveis em torno da tendência, deve-se optar por um modelo aditivo.

Por ser a que mais relevância parece ter para o nosso trabalho, na secção que se segue vamos-nos restringir à construção de um modelo de decomposição do tipo aditivo, e dentro deste tipo de modelos consideraremos apenas os que contêm na sua estrutura a tendência, a componente sazonal e, claro, a componente residual.

2.2.1. Modelo Aditivo

Como referimos anteriormente, o modelo aditivo é especificado do seguinte modo:

$$y_t = T_t + S_t + e_t, \quad t = 1, 2, \dots, n \quad (2.5)$$

onde T_t é a tendência, S_t a componente sazonal e e_t a componente residual.

Nesta perspectiva de análise do comportamento da série temporal começa-se por dessazonalizar a série com médias móveis, prossegue-se com a estimação da componente da tendência e finalmente estima-se a componente residual.

♦ Dessazonalização com Médias Móveis

Considerando então que a componente sazonal está presente, o objecto inicial é estimar os coeficientes sazonais.

Admita-se que $\sum_{t=k+1}^{k+S} S_t = 0$, para qualquer inteiro k , sendo S o período sazonal. Isto significa que o valor dos efeitos depressivos compensa o valor dos efeitos expansivos. Os diferentes valores que a função S_t toma, isto é, a $S_{k+1}, S_{k+2}, \dots, S_{k+S}$ chamamos coeficientes sazonais.

A estimação dos coeficientes sazonais compreende duas fases: na primeira estimam-se as suas componentes não normalizadas e na segunda estas são normalizadas. Na primeira fase, o primeiro passo é construir um processo de médias móveis associado a y_t .

Se S for par⁽¹⁾, então definimos o processo de médias móveis centrado como:

$$M_t = \frac{1}{S} \left(\left(\frac{1}{2} \right) y_t + y_{t-S/2+1} + \dots + y_t + \dots + y_{t+S/2-1} + \left(\frac{1}{2} \right) y_{t+S/2} \right), \quad (2.6)$$

para $t = \frac{S}{2} + 1, \dots, N - \frac{S}{2}$. Realce-se que t agora é o instante médio das observações incluídas na média móvel.

Se T_t for constante ou linear, M_t é considerado um bom estimador do nível T_t da série, no momento t .

Num segundo passo obtemos as estimativas das componentes sazonais e residuais através das diferenças

$$S_t^* = y_t - M_t, \quad (2.7)$$

pois

$$y_t - M_t = T_t + S_t + e_t - M_t \approx S_t + e_t, \quad t = \frac{S}{2} + 1, \dots, N - \frac{S}{2} \quad (2.8)$$

Admita-se que, sem perda de generalidade, $N = mL$, e represente-se t com $i = S(j-1)$, onde $i = 1, 2, \dots, S$ e $j = 1, 2, \dots, m$, com m o número de anos de observações disponíveis. Constroem-se para os, anteriormente obtidos, $S_t^* = S_{i+S(j-1)}^*$ as S médias

$$\bar{S}_i = \frac{1}{m-1} \sum_{j=1}^m S_{i+S(j-1)}^*, \quad i = 1, 2, \dots, S \quad (2.9)$$

Os \bar{S}_i são as estimativas não normalizadas dos coeficientes sazonais no instante i de cada ano.

Passando à segunda fase, as estimativas \bar{S}_i serão normalizadas, para assegurar que

$$\sum_{t=k+1}^{k+S} S_t = 0, \quad \forall k \text{ inteiro, tomando então}$$

⁽¹⁾ Se S for ímpar as médias simples já são centradas.

$$\hat{S}_i = \bar{S}_i - \frac{1}{S} \left(\sum_{i=1}^S \bar{S}_i \right), \quad i = 1, 2, \dots, S \quad (2.10)$$

e $\hat{S}_i = \hat{S}_{i-S}$, $i > S$, onde a estimativa da componente sazonal é dada por (2.10).

A série dessazonalizada y_t^D é então definida pela remoção da componente sazonal na série original y_t , isto é,

$$y_t^D = y_{i+S(j-1)} - \hat{S}_i, \quad t = 1, \dots, N \quad (2.11)$$

Está feita, portanto, a dessazonalização por médias móveis de y_t .

Prossegue-se, nesta perspectiva de decomposição, com a estimação da componente de tendência.

◆ Estimação da Tendência

É importante olhar atentamente para o gráfico da série para assim identificar a forma da tendência a estimar. Pois a tendência, nas suas abordagens determinísticas ou estocásticas, pode ser (i) constante, (ii) linear ou ter (iii) outras formas (quadrática, exponencial, logarítmica, etc.).

Se concluirmos que a tendência é linear, isto é,

$$T_t = \beta_0 + \beta_1 t + e_t \quad (2.12)$$

a regressão linear simples é uma técnica que fornece uma boa estimativa da tendência. Pois fornece estimadores lineares centrados de variância mínima, através do método dos mínimos quadrados.

Também podemos utilizar a regressão múltipla a fim de ajustar, se for esse o caso, uma tendência do tipo polinomial, ou seja,

$$y_t = \beta_0 + \beta_1 t + \dots + \beta_p t^p + e_t \quad (2.13)$$

Mas nem sempre é claro o tipo de tendência que mais se adequa às observações. Nesse caso constrói-se uma sucessão de médias móveis associada, em que se substitui cada observação pela média da própria observação com as que lhe estão próximas.

◆ Estimação da Componente Residual

Temos vindo a analisar a estimação das componentes de um modelo do tipo aditivo (2.5), nesta altura falta estimar a componente residual e_t . A estimativa desta componente é dada por

$$\hat{e}_t = y_t - \hat{T}_t - \hat{S}_t, \quad t = 1, \dots, N \quad (2.14)$$

Após a decomposição espera-se obter uma série estacionária, se tal suceder resta modelar a série dos resíduos utilizando a metodologia Box-Jenkins (Brockwell & Davis, 1996).

2.2.2. Previsão

A decomposição das séries, em componentes não observáveis, além de ser uma ferramenta que permite compreender o comportamento das séries temporais, também é um método de previsão. Neste trabalho a decomposição será utilizada para descrever e compreender a estrutura da série e assim ajudar na escolha de um método de previsão.

Quando a série y_t (como em 2.5) apresenta tendência e sazonalidade com período S usa-se o método Holt-Winters (que é uma extensão do método de Holt) (Brockwell & Davis, 1996). Este método permite construir as previsões pontuais da série. Baseia-se em três equações de alisamento: a do nível, a da tendência e a da sazonalidade, e apresenta a vantagem de não se supor um padrão sazonal fixo.

Considerando então $y_t = a + bt + S_t + e_t$ o modelo subjacente, a função de previsão é

$$\hat{y}_{t+h} = \hat{a}_t + \hat{b}_t h + \hat{S}_{t+h-kS} \quad (2.15)$$

com $h = 1, 2, \dots$ e $k = 1$ se $0 < h \leq L$, $k = 2$ se $L < h \leq 2L$.

Assim

$$\hat{a}_t = \alpha(y_t - \hat{S}_{t-S}) + (1 - \alpha)(\hat{a}_{t-1} + \hat{b}_{t-1}), \quad 0 < \alpha < 1 \quad (2.16)$$

$$\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1 - \beta)\hat{b}_{t-1}, \quad 0 < \beta < 1 \quad (2.17)$$

$$\hat{S}_t = \gamma(y_t - \hat{a}_t) + (1 - \gamma)\hat{S}_{t-S}, \quad 0 < \gamma < 1 \quad (2.18)$$

Para aplicar este método é necessário definir os valores iniciais para $\hat{a}(0)$, $\hat{b}(0)$ e \hat{S}_{i-S} , $i = 1, \dots, S$.

Começando em y_1 , tem-se:

$$(i) \hat{b}(0) = \frac{\hat{y}_m - \bar{y}_1}{(m-1)S}, \quad m = 2 \text{ ou } m = 3 \text{ e } mS < N \quad (2.19)$$

$$(ii) \hat{a}(0) = \bar{y}_1 - \frac{S+1}{2} b(0) \quad (2.20)$$

com \bar{y}_1 a média do primeiro ano e \bar{y}_m a média do ano m ;

(iii) para $t = 1, \dots, mS$ os índices sazonais

$$S_{i+S(j-1)}^* = y_{i+S(j-1)} - \left[\bar{y}_j - \left(\frac{S+1}{2} - i \right) \hat{b}(0) \right] \quad (2.21)$$

onde $i = 1, \dots, S$; $j = 1, \dots, m$ e \bar{y}_j constitui a média do ano j e ao qual pertence

$$y_t = y_{i+S(j-1)};$$

(iv) controem-se as médias

$$\bar{S}_i = \frac{1}{m} \sum_{j=1}^m S_{i+S(j-1)}^*, \quad i = 1, \dots, S \quad (2.22)$$

que de seguida se normalizam.

As constantes de alisamento α , β e γ escolhem-se minimizando uma função dos erros de previsão.

Os coeficientes sazonais estimados, por vezes ao fim de S períodos, terão de ser renormalizados, devido à sucessiva aplicação de (iii).

Até aqui vimos como obter previsões pontuais para a série, mas muitas vezes interessamos construir intervalos de confiança para as previsões, os chamados intervalos de previsão. Isto é, pretende-se obter intervalos que se crê conterem o valor futuro da variável com certa probabilidade.

No entanto, fazer previsão por intervalos depende da especificação de certas hipóteses relativas ao comportamento da variável, do seu modelo teórico e em particular do seu termo residual. O método de Holt-Winters não é o mais adequado à dedução de intervalos de previsão, sendo que um dos principais motivos é o facto de o modelo estatístico equivalente não estar teoricamente fundamentado.

2.3. Modelação Box-Jenkins

Neste capítulo debruçar-nos-emos sobre a metodologia Box-Jenkins utilizada na análise de séries temporais, com a qual se pretende escolher modelos ARMA⁽²⁾ (*auto regressive moving average*) para modelar a série das observações.

Um processo $\{y_t\}$ diz-se um modelo ARMA (p, q) se a equação às diferenças estocásticas é satisfeita

$$y_t - \phi_1 y_{t-1} - \dots - \phi_p y_{t-p} = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$$

ou escrevendo de uma forma mais concisa

$$\phi_p(B)y_t = \theta_q(B)e_t \quad (2.23)$$

onde B é o operador atraso $B^i y_t = y_{t-i}$, $i = 0, \pm 1, \dots$ e $\{e_t\}$ é um processo de ruído branco com $E(e_t) = 0$ e $\text{var}(e_t) = \sigma^2$.

As raízes do polinómio $\theta_q(B) = 0$ devem situar-se fora do círculo unitário para que o processo seja invertível e para o processo ser estacionário as raízes de $\phi_p(B) = 0$ devem estar

⁽²⁾ Os modelos ARIMA e SARIMA são uma generalização dos modelos ARMA.

fora do círculo unitário. Assume-se daqui para a frente que o processo é estacionário e invertível e que os polinómios $\theta_q(B) = 0$ e $\phi_p(B) = 0$ não têm raízes comuns.

A FAC de um processo ARMA (p, q) decai gradualmente para zero (ou sob a forma de sinusoidal amortecida), pois o seu comportamento é como o de um processo autoregressivo (AR). A FACP, cujo comportamento é semelhante à de um processo de médias móveis (MA), também decai gradualmente para zero (ou sob a forma de sinusoidal amortecida).

Segundo Box-Jenkins (1994) na construção de um modelo podemos distinguir três fases essenciais, conforme é representado esquematicamente na Figura 2.1: (i) identificação, (ii) estimação e diagnóstico e (iii) utilização do modelo na previsão.

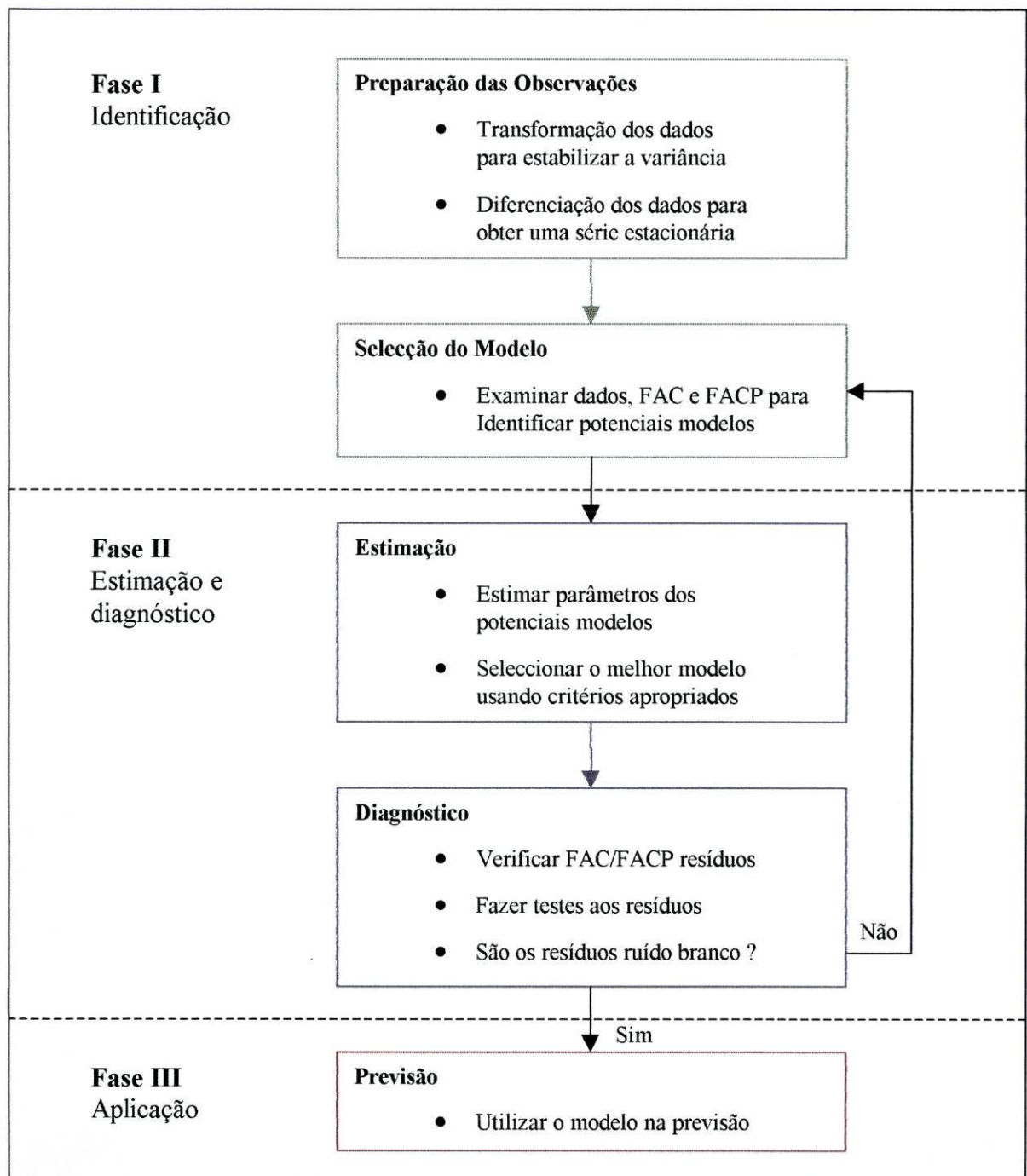


Figura 2.1 - Metodologia Box-Jenkins (Makridakis, Wheelwright & Hyndman, 1998, p. 314)

Na análise da nossa série serão estas as fases por nós utilizadas e que neste capítulo descrevemos sucintamente.

2.3.1. Identificação do Modelo

Na fase de identificação dos modelos dois desafios se nos colocam: (i) preparação das observações e (ii) determinar os possíveis modelos ARMA (p, q) para a série em estudo.

(i) Preparação das Observações

Na prática, na maior parte dos casos procuramos ajustar um modelo a uma série $\{y_1, y_2, \dots, y_n\}$ que é não estacionária (em média, em variância ou em ambas). Isto levanta um problema uma vez que na modelação Box-Jenkins, pretende-se escolher modelos ARMA, que correspondem a processos estacionários. Portanto se lidamos com uma série não estacionária é preciso transformá-la de modo a que se torne estacionária.

Quanto à estabilização da variância deverá ser feita antes das transformações estabilizadoras da média, utilizando a transformação de Box-Cox (Box & Cox, 1964).

A transformação Box-Cox é uma família paramétrica de transformação de y_t para $y_t^{(\lambda)}$ definida da seguinte maneira

$$T(y_t) = y_t^{(\lambda)} = \begin{cases} \frac{y_t^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log y_t, & \lambda = 0 \end{cases} \quad (2.24)$$

Deverão então experimentar-se vários valores de λ até chegar ao valor de λ mais apropriado. Ao “melhor” λ corresponderá a menor soma de quadrados residuais, isto é, para o qual

$$S(\lambda) = \sum_{t=1}^N (y_t^{(\lambda)} - \hat{\mu}_t^\lambda)^2 \quad (2.25)$$

é menor.

No caso da série ser não estacionária em média torna-se necessário diferenciar a série.

A técnica de diferenciação é usada, consoante os casos, para remover quer a tendência quer a sazonalidade. A diferenciação para eliminar a tendência tem a forma

$$\nabla^d y_t = (1 - B)^d y_t \quad (2.26)$$

onde o inteiro d , na maior parte das situações práticas, toma os valores 0, 1 ou 2.

Torna-se no entanto necessário recordar que, embora o operador d possa ser qualquer inteiro, $d \geq 0$, se deve ter em atenção que se a ordem da diferenciação for muito elevada, introduzem-se variâncias espúrias.

Quando estamos perante dados sazonais poderá ser apropriado aplicar a diferenciação sazonal, para estacionar a série, através do operador diferença sazonal

$$\nabla_s^D y_t = (1 - B^S)^D y_t \quad (2.27)$$

onde o inteiro D é a ordem da diferenciação sazonal e S é a periodicidade.

(ii) Escolha de possíveis modelos ARMA (p, q)

Após a aplicação das transformações necessárias (se for esse o caso) para estabilizar a série, examinamos o cronograma, a FAC e FACP amostrais da série convenientemente transformada e tentamos encontrar as ordens de p e q com vista a encontrar os modelos ARMA (p, q) que parecem ajustar-se à série dos dados. De seguida passamos à segunda fase da modelação Box-Jenkins.

2.3.2. Estimação e Diagnóstico

Nesta fase vamos decidir qual, de entre os vários modelos propostos, será utilizado para previsão. Para isso temos de estimar os parâmetros dos vários modelos e, usando critérios adequados, seleccionar o melhor modelo. Por fim, e ainda nesta fase, fazemos a análise dos resíduos e assim decidimos se o modelo vai, ou não, ser utilizado para previsão.

(i) Estimação dos parâmetros dos potenciais modelos

Já temos no horizonte um leque de modelos que poderão ajustar-se à série em estudo, mas para chegar à forma final destes modelos é preciso estimar os seus parâmetros, os quais são estimados pelo método da máxima verosimilhança. Os estimadores $\hat{\beta}$ de $\beta = (\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q)'$ são aproximadamente $N(\beta, n^{-1}V(\beta))$, onde $V(\beta)$ é a matriz de covariância assintótica (Brockwell & Davis, 1996 e 1991).

Estando em poder das estimativas dos parâmetros de cada modelo é preciso analisar a sua significância. Assim as estimativas que apresentem um *p-value* superior a 0.05 (para um nível de significância de $\alpha = 5\%$) deverão ser excluídos do modelo. Por outro lado os parâmetros não devem ser correlacionadas, sob pena de produzirem um modelo instável.

(ii) Seleção do modelo

Os diferentes modelos ajustados são comparados com base nos critérios AIC e BIC. Segundo Murteira, Müller e Turkman (1994) existem dois importantes critérios de selecção de modelos: Akaike (AIC) e critério de informação de Bayes (BIC).

No critério AIC, a regra de decisão é seleccionar o modelo que minimiza

$$AIC(m) = -2 \log L(\tilde{\Psi}) + 2m$$

onde $L(\tilde{\Psi})$ é o valor da função de máxima verosimilhança e m é o número de parâmetros (Harvey, 1993).

O critério BIC dado por

$$BIC(m) = -2 \log L(\tilde{\Psi}) + m \log T,$$

também selecciona o modelo ao qual corresponde o valor de m que minimiza esta função.

(iii) Diagnóstico

Nesta segunda sub-etapa vamos verificar se o modelo seleccionado fornece uma descrição adequada dos dados através da análise dos resíduos.

Se os resíduos se comportarem como um processo de ruído branco, deverão ser uma sequência de variáveis aleatórias não correlacionadas, provenientes de uma distribuição fixa, com média constante, normalmente assumida como sendo zero, e variância constante. Estas hipóteses sobre a componente residual poderão ser verificadas através de representações gráficas e testes estatísticos adequados.

Assim, constrói-se o cronograma dos resíduos, o qual deverá revelar uma média próxima de zero e não apresentar movimentos periódicos e analisamos os gráficos das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) amostrais. A normalidade da componente residual poderá ser verificada graficamente através do histograma dos resíduos e da representação em papel de probabilidades. Além disso deveremos analisar também os testes adequados para verificar se esta componente tem um comportamento compatível com uma amostra de ruído branco. Estes testes são utilizados para avaliar a aleatoriedade dos resíduos e estão resumidamente apresentados no Anexo I.

Caso os resíduos tenham todas as características de ruído branco, o modelo em causa é utilizado para previsão.

2.3.3. Previsão

Após a selecção do modelo que melhor se ajusta à série dos dados, este é utilizado para previsão de valores futuros da série: a partir do passado da série $y_t, y_{t-1}, y_{t-2}, \dots$ pretendemos prever um valor não observado da série $y_{t+m}, m > 0$.

O objectivo é escolher $y_t(m) = f(y_t, y_{t-1}, \dots)$ que minimize $E[(y_{t+m} - y_t(m))^2]$. E a melhor função para f é o valor médio condicional de y_{t+m} em $y_t, y_{t-1}, y_{t-2}, \dots$ ou seja, o preditor de erro quadrático médio mínimo.

Seja \hat{y}_{t+m} a previsão de $\{y_t\}$ em $t + m$:

$$\hat{y}_{t+m} = y_t(m) = E(y_{t+m} | y_t, y_{t-1}, y_{t-2}, \dots). \quad (2.28)$$

Se $\{y_t\}$ satisfaz um ARIMA (p, d, q) da forma:

$$\Phi_p(B)(1-B)^d y_t = \Theta_q(B)e_t, \quad (2.29)$$

onde as raízes do polinómio $\Phi(B)$ são em módulo estritamente superiores à unidade e o valor das raízes de $\Theta(B)$ não interfere com a estacionaridade do processo.

Se

$$\Psi(B) = \Phi_p(B)(1-B)^d = 1 - \psi_1 B - \dots - \psi_{p+d} B^{p+d} \quad (2.30)$$

tem-se

$$(1 - \psi_1 B - \dots - \psi_{p+d} B^{p+d}) y_{t+m} = \Theta_q(B) e_t \quad (2.31)$$

então

$$y_{t+m} = \psi_1 y_{t+m-1} + \dots + \psi_{p+d} y_{t+m-p-d} + e_{t+m} - \theta_1 e_{t+m-1} - \dots - \theta_q e_{t+m-q} \quad (2.32)$$

Tomando então o valor médio condicionado a $y_t, y_{t-1}, y_{t-2}, \dots$ em ambos os membros da equação anterior, obtém-se a equação de previsão

$$\hat{y}_t(m) = \psi_1 \hat{y}_t(m-1) + \dots + \psi_{p+d} \hat{y}_t(m-p-d) + \hat{e}_t(m) - \theta_1 \hat{e}_t(m-1) - \dots - \theta_q \hat{e}_t(m-q) \quad (2.33)$$

onde

$$\hat{y}_t(j) = E(y_{t+j} | y_t, y_{t-1}, y_{t-2}, \dots) \quad (2.34)$$

e

$$\hat{e}_t(j) = 0, \quad j \geq 1 \quad (2.35)$$

quando $j \leq 0$

$$\hat{y}_t(j) = y_{t+j} \quad (2.36)$$

e

$$\hat{e}_t(j) = e_{t+j}. \quad (2.37)$$

No caso de pretendermos calcular um intervalo de confiança a $(1-\alpha)$ 100% para a previsão do valor futuro Y_{t+m} , admitindo que os resíduos têm distribuição gaussiana:

$$e_t(m) \cap N\left(0, \sigma_e^2 \sum_{j=0}^{m-1} \psi_j^2\right), \text{ constrói-se então o intervalo pretendido } Y_{t+m} \pm Z_{1-\alpha/2} \sigma_e \sqrt{\sum_{j=0}^{m-1} \psi_j^2}.$$

2.4. Modelos Estruturais

Nas séries temporais é frequente observar que estas apresentam tendência, sazonalidade, ciclos e perturbações aleatórias.

Por vezes a tendência não pode ser ajustada por linhas rectas, a menos que o período de tempo seja razoavelmente pequeno. Isto significa que a maior parte das vezes a tendência determinística limita o estudo das séries temporais. O ideal seria que a tendência não fosse tão restritiva, mas sim mais flexível. A forma de o conseguir é permitir que os parâmetros, do nível e do declive, evoluam ao longo do tempo. Surge, assim, necessidade de construir uma tendência estocástica, em que o nível e o declive se alteram com o decorrer do tempo. De modo semelhante, torna-se necessário que as outras componentes sejam suficientemente flexíveis de modo a que reflectam os padrões do comportamento da série.

Os modelos compostos por componentes com tais características, são muitas vezes vistos como modelos de regressão, nos quais as variáveis são funções do tempo e cujos parâmetros variam ao longo do tempo. Tais modelos designam-se por modelos estruturais de séries temporais e foram desenvolvidos por Harvey (1996).

Os modelos estruturais de séries temporais são uma classe de modelos formulados em termos de componentes estocásticas, que têm por isso uma interpretação directa. Se para além da tendência e da sazonalidade, se incluírem variáveis exploratórias, o modelo resulta numa mistura de séries temporais e regressão. A regressão é um caso especial, na qual não existem componentes estocásticas, para além do termo da perturbação aleatória. A combinação das variáveis exploratórias com as componentes não observáveis, abre um vasto leque de possibilidades na modelação. De seguida iremos abordar os principais modelos estruturais.

2.4.1. Modelo Estrutural Simples

O modelo estrutural simples é definido por duas componentes: a tendência e a perturbação aleatória. Este modelo é definido por:

$$y_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.38)$$

Se a série se caracteriza por um *nível localmente constante*, então μ_t é a tendência sem declive e ε_t é uma perturbação aleatória com características de ruído branco de média zero e variância σ_ε^2 , que se assume ser não correlacionado com μ_t . A tendência define-se por

$$\mu_t = \mu_{t-1} + \eta_t \quad (2.39)$$

Se a tendência for considerada linear, introduz-se em 2.39 o declive β_t e tem-se

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.40a)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.40b)$$

onde η_t e ξ_t são perturbações aleatórias com características de ruído branco, com média zero e variâncias σ_η^2 e σ_ξ^2 , respectivamente. A sucessão é portanto caracterizada por um nível localmente linear.

2.4.2. Modelo Estrutural Básico (MEB)

Este modelo difere do anterior pois além de incluir as componentes de nível de tendência μ_t e aleatória ε_t , inclui a componente sazonal γ_t , de período s . Assim, o modelo estrutural básico representa-se do seguinte modo:

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.41a)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.41b)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.41c)$$

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \quad (2.41d)$$

Os termos ε_t , η_t , ξ_t e ω_t são perturbações aleatórias com características de ruído branco que se assume serem não correlacionadas entre si.

As equações (2.41b) e (2.41c) juntas definem o modelo da tendência: a equação (2.41b) representa o modelo do nível da tendência e β_t (equação (2.41c)) é o modelo do respectivo declive, que tem comportamento de passeio aleatório.

A tendência tem diferentes modelos estocásticos consoante se admite que cada uma das variâncias σ_η^2 e σ_ξ^2 é nula, ou até que ambas são nulas.

- Se $\sigma_\eta^2 = \sigma_\xi^2 = 0$, o modelo da tendência ((2.41b) e (2.41c)) é uma função determinística do tempo em que $\beta_t = \beta_{t-1} = \dots = \beta$ e $\mu_t = \mu_{t-1} + \beta t$, $t = 1, \dots, T$
- Se $\sigma_\xi^2 = 0$ e $\sigma_\eta^2 > 0$, o declive é constante $\beta_t = \beta_{t-1} = \dots = \beta$ e a tendência fica

$$\mu_t = \mu_{t-1} + \beta + \eta_t \Leftrightarrow$$

$$\Leftrightarrow (1 - B)\mu_t = \beta + \eta_t,$$

que apesar de continuar a ser um processo estocástico tem uma componente determinística: β , sendo B o operador atraso.

No caso de $\beta = 0$ a tendência é um passeio aleatório.

- Se $\sigma_{\xi}^2 > 0$ e $\sigma_{\eta}^2 = 0$, então

$$\begin{aligned}\mu_t &= \mu_{t-1} + \beta_{t-1} \Leftrightarrow \\ \Leftrightarrow \mu_t - \mu_{t-1} &= \beta_{t-1} \Leftrightarrow \\ \Leftrightarrow (1 - B)\mu_t &= \beta_{t-1}.\end{aligned}$$

e como

$$\begin{aligned}\beta_t &= \beta_{t-1} + \xi_t \Leftrightarrow \\ \Leftrightarrow \beta_t - \beta_{t-1} &= \xi_t \Leftrightarrow \\ \Leftrightarrow (1 - B)\beta_t &= \xi_t \Leftrightarrow \\ \Leftrightarrow \beta_t &= \frac{\xi_t}{1 - B}\end{aligned}$$

Temos então que $\beta_{t-1} = \frac{\xi_{t-1}}{1 - B}$ e substituindo esta em $(1 - B)\mu_t = \beta_{t-1}$ fica

$$\begin{aligned}(1 - B)\mu_t &= \frac{\xi_{t-1}}{1 - B} \Leftrightarrow \\ \Leftrightarrow (1 - B)^2 \mu_t &= \xi_{t-1}\end{aligned}$$

que é ruído branco nas segundas diferenças.

- Se $\sigma_{\xi}^2 > 0$ e $\sigma_{\eta}^2 > 0$, temos que

$$\begin{aligned}\beta_t &= \beta_{t-1} + \xi_t \Leftrightarrow \\ \Leftrightarrow \beta_t - \beta_{t-1} &= \xi_t \Leftrightarrow \\ \Leftrightarrow (1 - B)\beta_t &= \xi_t \Leftrightarrow \\ \Leftrightarrow \beta_t &= \frac{\xi_t}{1 - B}\end{aligned}$$

e conseqüentemente $\beta_{t-1} = \frac{\xi_{t-1}}{1 - B}$.

Quanto à equação da tendência, vem:

$$\begin{aligned}\mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t \Leftrightarrow \\ \Leftrightarrow \mu_t - \mu_{t-1} &= \beta_{t-1} + \eta_t \Leftrightarrow \\ \Leftrightarrow (1 - B)\mu_t &= \frac{\xi_{t-1}}{1 - B} + \eta_t \Leftrightarrow \\ \Leftrightarrow (1 - B)^2 \mu_t &= \xi_{t-1} + (1 - B)\eta_t \Leftrightarrow \\ \Leftrightarrow (1 - B)^2 \mu_t &= \eta_t - \eta_{t-1} + \xi_{t-1}.\end{aligned}$$

Portanto a tendência no caso, $\sigma_{\xi}^2 > 0$ e $\sigma_{\eta}^2 > 0$, é um processo ARIMA (0, 2, 1).

A componente sazonal é dada pela equação $\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \Leftrightarrow \sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t$.

Esta componente pode ser actualizada quer sob a forma de variável muda

$$\sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t$$

quer como função trigonométrica

$$\gamma_t = \sum_{j=1}^{\lfloor s/2 \rfloor} \gamma_{jt}$$

A representação trigonométrica de γ_{jt} é um ciclo não estacionário, com $\lambda_j = 2\pi j/s$, $j = 1, 2, \dots, \lfloor s/2 \rfloor$:

$$\begin{bmatrix} \gamma_{jt} \\ \gamma_{jt}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_j & \sin \lambda_j \\ -\sin \lambda_j & \cos \lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{jt} \\ \omega_{jt}^* \end{bmatrix} \quad (2.42)$$

onde ω_t e ω_t^* são processos de ruído branco com média zero e variância σ_j^2 .

$$\text{De (2.42) tem-se } \begin{cases} \gamma_{jt} = \gamma_{j,t-1} \cos \lambda_j + \gamma_{j,t-1}^* \sin \lambda_j + \omega_{jt} \\ \gamma_{jt}^* = -\gamma_{j,t-1} \sin \lambda_j + \gamma_{j,t-1} \cos \lambda_j + \omega_{jt}^* \end{cases}$$

Se s for par então $\gamma_{jt}^* = 0$, logo $\gamma_{jt} = \gamma_{j,t-1} \cos \lambda_j + \omega_{jt}$.

2.4.3. Modelo Cíclico

Os modelos cíclicos apresentam uma componente cíclica, ψ_t , para além das componentes da tendência e aleatória.

Podem definir-se três tipos de modelos com componente cíclica: o modelo ciclo mais ruído, o modelo de tendência mais ciclo e o modelo de tendência cíclica.

Define-se ciclo estocástico ψ_t da seguinte forma

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} k_t \\ k_t^* \end{bmatrix} \quad (2.43)$$

e utilizando o operador atraso B , tem-se que

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \begin{bmatrix} 1 - \rho \cos \lambda_c B & -\rho \sin \lambda_c B \\ \rho \sin \lambda_c B & 1 - \rho \cos \lambda_c B \end{bmatrix}^{-1} \begin{bmatrix} k_t \\ k_t^* \end{bmatrix} \quad (2.44)$$

é um vector de um processo AR(1), com $0 \leq \rho < 1$ e $0 \leq \lambda_c \leq \pi$.

◆ Ciclo Mais Ruído

Este modelo é formalizado da seguinte forma:

$$y_t = \mu + \psi_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.45)$$

Mas substituindo (2.43) em y_t obtém-se o modelo:

$$y_t = \mu + \frac{(1 - \rho \cos \lambda_c B)k_t + (\rho \sin \lambda_c B)k_t^*}{1 - 2\rho \cos \lambda_c B + \rho^2 B^2} + \varepsilon_t, \quad t = 1, \dots, T \quad (2.46)$$

◆ Tendência Mais Ciclo

Neste modelo considera-se a componente cíclica como estacionária e representa-se por

$$y_t = \mu_t + \psi_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.47a)$$

onde μ_t é o modelo da tendência como no modelo estrutural básico:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.47b)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.47c)$$

◆ Tendência Cíclica

Tal como no modelo anterior a componente cíclica deste modelo assume-se como estacionária e

$$y_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.48a)$$

com $\mu_t = \mu_{t-1} + \psi_{t-1} + \beta_{t-1} + \eta_t$, em que o ciclo é incorporado na tendência. (2.48b)

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.48c)$$

Então, com $\Delta = 1 - B$, tem-se

$$y_t = \frac{\eta_t}{\Delta} + \frac{(1 - \rho \cos \lambda_c B)k_{t-1} + (\rho \sin \lambda_c B)k_{t-1}^*}{\Delta(1 - 2\rho \cos \lambda_c B + \rho^2 B^2)} + \frac{\xi_{t-1}}{\Delta^2} + \varepsilon_t \quad (2.49)$$

2.4.4. Os Modelos Estruturais na Sua Forma Reduzida

Os modelos estruturais de séries temporais contêm vários termos de perturbação, mas uma vez que são lineares, estas perturbações podem ser combinadas de modo a que exista apenas uma perturbação. Quando os modelos estruturais de séries temporais se encontram nesta nova forma diz-se que estão na sua forma reduzida, cujas características probabilísticas são equivalentes a modelos, com restrições, da família ARIMA.

Uma das vantagens dos modelos estruturais na sua forma reduzida, é que são formulados em termos das componentes não observáveis dos modelos ARIMA. Pois a classe dos modelos ARIMA abrange muitos modelos e parâmetros, os quais não têm uma interpretação sensível.

Na prática, em termos de previsão, os modelos ARIMA até funcionam bastante bem, mas tal só acontece porque os modelos estruturais de séries temporais têm a forma ARIMA reduzida (Harvey, 1997).

Apresentamos de seguida a forma reduzida do modelo estrutural simples do modelo estrutural básico.

◆ Modelo Estrutural Simples

◆ Nível Localmente Constante

O modelo é dado pelas seguintes equações:

$$y_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.50)$$

$$\mu_t = \mu_{t-1} + \eta_t \quad (2.51)$$

Sendo

$$\begin{aligned} \mu_t &= \mu_{t-1} + \eta_t \Leftrightarrow \\ \Leftrightarrow \mu_t - \mu_{t-1} &= \eta_t \Leftrightarrow \\ \Leftrightarrow (1 - B)\mu_t &= \eta_t \Leftrightarrow \\ \Leftrightarrow \mu_t &= \frac{\eta_t}{1 - B} \end{aligned}$$

Então, substituindo μ_t em $y_t = \mu_t + \varepsilon_t$, vem

$$\begin{aligned} y_t &= \frac{\eta_t}{1 - B} + \varepsilon_t \Leftrightarrow \\ \Leftrightarrow (1 - B)y_t &= \eta_t + (1 - B)\varepsilon_t \Leftrightarrow \\ \Leftrightarrow (1 - B)y_t &= \eta_t + \varepsilon_t - \varepsilon_{t-1} \end{aligned}$$

Então a forma reduzida do modelo estrutural simples com nível localmente constante é

$$(1 - B)y_t = \eta_t + \varepsilon_t - \varepsilon_{t-1}.$$

No segundo membro temos um processo ARMA (0, 0) e um processo ARMA (0, 1), logo y_t é um processo ARIMA (0, 1, 1), isto é, satisfaz uma equação da forma $(1 - B)y_t = Z_t + \theta Z_{t-1}$, onde Z_t é o ruído branco e $\theta = \left[(q^2 + y_q)^{1/2} - 2 - q \right] / 2$, onde $q = \rho_\eta^2 / \rho_\varepsilon^2$ é a razão das variâncias (Harvey, 1996). Verifica-se que sob estas condições $-1 \leq \theta \leq 0$.

Assim, o modelo estrutural simples com nível localmente constante é um processo ARIMA (0, 1, 1), mas com restrições.

◆ Nível Localmente Linear

O modelo estrutural simples com nível localmente linear é definido pelo seguinte conjunto de equações

$$y_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.52a)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.52b)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.52c)$$

Consequentemente temos

$$\begin{aligned} \beta_t &= \beta_{t-1} + \xi_t \Leftrightarrow \\ \Leftrightarrow \beta_t - \beta_{t-1} &= \xi_t \Leftrightarrow \\ \Leftrightarrow (1 - B)\beta_t &= \xi_t \Leftrightarrow \\ \Leftrightarrow \beta_t &= \frac{\xi_t}{1 - B} \end{aligned}$$

e também

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t \Leftrightarrow \\ \Leftrightarrow \mu_t - \mu_{t-1} &= \beta_{t-1} + \eta_t \Leftrightarrow \\ \Leftrightarrow (1 - B)\mu_t &= \beta_{t-1} + \eta_t \Leftrightarrow \\ \Leftrightarrow \mu_t &= \frac{\beta_{t-1}}{1 - B} + \frac{\eta_t}{1 - B} \Leftrightarrow \\ \Leftrightarrow (1 - B)\mu_t &= \beta_{t-1} + \eta_t \end{aligned}$$

Como $\beta_{t-1} = \frac{\xi_{t-1}}{1 - B}$ e substituindo em $\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$, temos que

$$(1 - B)\mu_t = \frac{\xi_{t-1}}{1 - B} + \eta_t \Leftrightarrow$$

$$\Leftrightarrow \mu_t = \frac{\xi_{t-1}}{(1 - B)^2} + \frac{\eta_t}{1 - B} \quad (2.53)$$

Substituindo esta última em $y_t = \mu_t + \varepsilon_t$, vem:

$$\begin{aligned} y_t &= \frac{\xi_{t-1}}{(1 - B)^2} + \frac{\eta_t}{1 - B} + \varepsilon_t \Leftrightarrow \\ \Leftrightarrow (1 - B)^2 y_t &= \xi_{t-1} + (1 - B)\eta_t + (1 - B)^2 \varepsilon_t \Leftrightarrow \\ \Leftrightarrow (1 - B)^2 y_t &= \xi_{t-1} + \eta_t - \eta_{t-1} + \varepsilon_t - 2\varepsilon_{t-1} + B\varepsilon_{t-1} \Leftrightarrow \\ \Leftrightarrow (1 - B)^2 y_t &= \xi_{t-1} + \eta_t - \eta_{t-1} + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2} \Leftrightarrow \\ \Leftrightarrow (1 - B)^2 y_t &= \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2} + \eta_t - \eta_{t-1} + \xi_{t-1} \end{aligned}$$

Esta última expressão representa um processo ARIMA (0, 2, 2) e é a forma reduzida do modelo estrutural simples com nível localmente linear.

◆ Modelo Estrutural Básico

O modelo é dado por:

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.54a)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.54b)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.54c)$$

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \quad (2.54d)$$

E no caso em que $\sigma_\xi^2 > 0$ e $\sigma_\eta^2 > 0$, a equação (2.54b) fica como em (2.53)

$$\begin{aligned} \mu_t &= \frac{\xi_{t-1}}{(1-B)^2} + \frac{\eta_t}{1-B} \Leftrightarrow \\ \Leftrightarrow (1-B)^2 \mu_t &= \eta_t - \eta_{t-1} + \xi_{t-1} \Leftrightarrow \\ \Leftrightarrow \mu_t &= \frac{\eta_t - \eta_{t-1} + \xi_{t-1}}{(1-B)^2} \end{aligned}$$

Quanto à componente da sazonalidade também pode ser escrita assim:

$$\sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t \Leftrightarrow \gamma_t + \gamma_{t-1} + \gamma_{t-2} + \dots + \gamma_{t-s+1} = \omega_t$$

com $S(B) = 1 + B + B^2 + \dots + B^{s-1}$, vem $S(B)\gamma_t = \omega_t$ que é ruído branco, ou então $\gamma_t = \frac{\omega_t}{S(B)}$.

$$\text{Como } S(B) = \frac{1-B^s}{1-B} \text{ vem } \gamma_t = \frac{\omega_t}{\frac{1-B^s}{1-B}} \Leftrightarrow \gamma_t = \frac{(1-B)\omega_t}{1-B^s}$$

E no modelo $y_t = \mu_t + \gamma_t + \varepsilon_t$ substituímos μ_t e γ_t pelas expressões que determinámos, ficando

$$\begin{aligned} y_t &= \frac{\eta_t - \eta_{t-1} + \xi_{t-1}}{(1-B)^2} + \frac{(1-B)\omega_t}{1-B^s} + \varepsilon_t \Leftrightarrow \\ \Leftrightarrow (1-B)^2(1-B^s)y_t &= \\ &= (1-B^s)\eta_t - (1-B^s)\eta_{t-1} + (1-B^s)\xi_{t-1} + (1-B)^3\omega_t + (1-B)^2(1-B^s)\varepsilon_t \end{aligned}$$

a qual é a forma reduzida do modelo estrutural básico, que é um MA (s+1) (Harvey, 1996).

No Anexo II apresentamos de um modo resumido os principais modelos estruturais, suas componentes e a sua forma reduzida.

2.4.5. Selecção do Modelo

Deparamo-nos agora com o problema de escolher qual o modelo, de entre os analisados, que melhor se ajusta à série. Tal como na Secção 2.3.2. utilizam-se os critérios adequados de selecção de modelos.

2.4.6. Análise da Componente Residual

Após a selecção do modelo que parece ajustar-se melhor à série dos dados observados, faz-se a análise dos resíduos, como na Secção 2.3.2..

2.4.7. Modelos em Espaço de Estados

Na teoria do controlo, a representação de um sistema em espaço de estados é essencial pois permite que os parâmetros dos modelos sejam ajustados de modo recursivo à medida que surgem novas observações.

Define-se espaço de um sistema como a informação mínima, do passado e do presente, necessária para que o comportamento futuro de um sistema possa ser "traçado" com base no presente estado e nas observações futuras.

Assim a propriedade Markoviana é o fundamento da representação de um sistema em espaço de estados: dado um estado presente, o futuro do sistema é independente do passado.

Suponhamos que temos um processo y_t , $t = 1, \dots, T$ caracterizado por evoluir no tempo e cujos valores futuros pretendemos prever.

Seja x_t um vector de estados ($m \times 1$), cujas componentes são as variáveis de estado, em geral não observáveis. Num dado instante t , os valores assumidos por estas variáveis definem o estado do processo.

A relação entre as observações, y_t , e o vector de estados x_t é traduzido pela seguinte equação:

$$y_t = z_t x_t + d_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.55)$$

que se designa por equação das observações, equação de medida ou equação de "output".

Em modelos univariados, z_t é um vector $1 \times m$, d_t é vector 1×1 e ε_t é o erro com distribuição $N(0, \sigma^2 h_t)$, com h_t função conhecida do tempo.

O vector de estados, x_t , embora não directamente observável é actualizado ao longo do tempo através da equação de transição que o define:

$$x_t = T_t x_{t-1} + C_t + R_t \eta_t, \quad t = 1, \dots, T \quad (2.56)$$

onde T_t é uma matriz $m \times m$, C_t é um vector $m \times 1$, R_t é uma matriz $m \times g$ (cuja inclusão é arbitrária) e η_t é um vector $g \times 1$. Sendo este último um vector de perturbações aleatórias não correlacionadas com média zero e matriz de covariância Q_t , i.e., $E(\eta_t) = 0$ e $var(\eta_t) = Q_t$.

A representação do modelo em espaço de estados é dada pelas equações (2.55) e (2.56).

◆ Representação em Espaço de Estados dos Principais Modelos Estruturais de Séries Temporais

Os modelos estruturais de séries temporais, tal como outros modelos (após manipulações algébricas), admitem representação em espaço de estados, o que é uma vantagem, pois a partir daqui podemos fazer previsões através do filtro de Kalman.

A seguir damos exemplos da representação em espaço de estados de alguns modelos estruturais.

◆ Modelo Estrutural Simples

a) Nível Localmente Constante

O modelo estrutural simples com nível localmente constante, como vimos no capítulo anterior é

$$y_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.57a)$$

$$\mu_t = \mu_{t-1} + \eta_t \quad (2.57b)$$

Já se encontra representado em espaço de estados com $z = 1$ e $T = 1$ e

$$x_t = \mu_t = \mu_{t-1} + \eta_t.$$

b) Nível Localmente Linear

O modelo estrutural simples com nível localmente linear representa-se, na sua forma habitual, da seguinte maneira:

$$y_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.58a)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.58b)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.58c)$$

A sua representação em espaço de estados é

$$y_t = [1 \ 0]x_t + \varepsilon_t, \quad (2.59a)$$

A equação de transição, matricial, é:

$$x_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \end{bmatrix} \quad (2.59b)$$

com $x_t = [\mu_t \ \beta_t]'$, $Z = [1 \ 0]$ e $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

◆ Modelo Estrutural Básico

Este modelo é representado pelas seguintes equações

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.60a)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.60b)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.60c)$$

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \quad (2.60d)$$

O modelo estrutural básico é adequado para muitas séries mensais e trimestrais ($s = 4$).

E a componente da sazonalidade para $s = 4$ é

$$\gamma_t = -\sum_{j=1}^3 \gamma_{t-j} + \omega_t \quad (2.61)$$

A sua representação em espaço de estados, com a componente da sazonalidade de período $s = 4$, é a seguinte

$$y_t = [1 \ 0 \ 1 \ 0 \ 0]x_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.62a)$$

Com equação matricial de transição

$$x_t = \begin{bmatrix} \mu_t \\ \beta_t \\ \gamma_t \\ \gamma_{t-1} \\ \gamma_{t-2} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ \gamma_{t-1} \\ \gamma_{t-2} \\ \gamma_{t-3} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \\ \omega_t \\ 0 \\ 0 \end{bmatrix} \quad (2.62b)$$

em que o vector de estados é

$$x_t = [\mu_t \ \beta_t \ \gamma_t \ \gamma_{t-1} \ \gamma_{t-2}]' \text{ e } Z = [1 \ 0 \ 1 \ 0 \ 0] \text{ e } T = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

◆ Modelo Cíclico

a) Tendência Mais Ciclo

O modelo cíclico com as componentes da tendência e do ciclo representa-se por

$$y_t = \mu_t + \psi_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.63a)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (2.63b)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.63c)$$

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} k_t \\ k_t^* \end{bmatrix} \quad (2.63d)$$

Admite a seguinte representação em espaço de estados:

$$y_t = [1 \ 0 \ 1 \ 0]x_t + \varepsilon_t, \quad t = 1, \dots, T \text{ que é a equação de medida.} \quad (2.64a)$$

A equação de transição em termos matriciais é

$$x_t = \begin{bmatrix} \mu_t \\ \beta_t \\ \psi_t \\ \psi_t^* \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho \cos \lambda_c & \rho \sin \lambda_c \\ 0 & 0 & -\rho \sin \lambda_c & \rho \cos \lambda_c \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \\ k_t \\ k_t^* \end{bmatrix} \quad (2.64b)$$

O vector de estados é

$$x_t = [\mu_t \ \beta_t \ \psi_t \ \psi_t^*]' e \quad Z = [1 \ 0 \ 1 \ 0] e$$

$$T = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho \cos \lambda_c & \rho \sin \lambda_c \\ 0 & 0 & -\rho \sin \lambda_c & \rho \cos \lambda_c \end{bmatrix}$$

b) Tendência Cíclica

O modelo estrutural com tendência cíclica é

$$y_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, T \quad (2.65a)$$

$$\mu_t = \mu_{t-1} + \psi_{t-1} + \beta_{t-1} + \eta_t \quad (2.65b)$$

$$\beta_t = \beta_{t-1} + \xi_t \quad (2.65c)$$

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} k_t \\ k_t^* \end{bmatrix} \quad (2.65d)$$

Representando-o em espaço de estados, a equação de medida é

$$y_t = [1 \ 0 \ 0 \ 0]x_t + \varepsilon_t, \quad (2.66a)$$

a equação de transição matricial é

$$x_t = \begin{bmatrix} \mu_t \\ \beta_t \\ \psi_t \\ \psi_t^* \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho \cos \lambda_c & \rho \sin \lambda_c \\ 0 & 0 & -\rho \sin \lambda_c & \rho \cos \lambda_c \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \\ k_t \\ k_t^* \end{bmatrix} \quad (2.66b)$$

Sendo

$$x_t = [\mu_t \quad \beta_t \quad \psi_t \quad \psi_t^*]' \text{ e o vector de estados } Z = [1 \quad 0 \quad 0 \quad 0] \text{ e}$$

$$T = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho \cos \lambda_c & \rho \sin \lambda_c \\ 0 & 0 & -\rho \sin \lambda_c & \rho \cos \lambda_c \end{bmatrix}$$

2.4.8. Filtro de Kalman

A representação de um modelo em espaço de estados permite a aplicação do filtro de Kalman para calcular o estimador do vector de estados no instante t , com base na informação disponível nesse instante.

O filtro de Kalman é um processo de estimação recursiva óptima, que permite construir as previsões dos modelos estruturais (Harvey, 1997). Dadas estimativas iniciais, este permite obter os parâmetros do modelo e ajustar essas estimativas sempre que surge uma nova observação, dando uma estimativa do erro em cada actualização.

Os modelos utilizados para previsões em séries temporais dividem-se em lineares e não lineares. Segundo Harvey (1993:266) “Qualquer modelo que tem uma forma linear em espaço de estados e perturbações gaussianas é certamente linear”. Ora, os modelos estruturais, objecto deste nosso estudo, têm todas as características de modelos lineares. Assim, enquanto nos modelos lineares se utiliza o filtro de Kalman discreto, nos modelos não lineares aplica-se o filtro de Kalman estendido (Welch & Bishop, s/d). Portanto nós só iremos abordar o filtro de Kalman discreto.

◆ Filtro de Kalman Discreto

O filtro de Kalman discreto aplica-se quando um processo contínuo é obtido em intervalos de tempo discretos. Este filtro é definido por um conjunto de equações, as quais por

sua vez se dividem em dois grupos: equações de previsão e equações de actualização. As primeiras permitem fazer uma previsão inicial (estimativa a priori) para o processo e para o erro, no instante seguinte. O segundo grupo de equações actualiza as estimativas a priori.

Resumidamente, o filtro de Kalman aplica-se da seguinte forma:

- A partir da informação inicial das estimativas dos parâmetros iniciais e do erro associado, calcula-se a matriz de ganho.
- O erro entre a estimativa do parâmetro e a observação é determinado e multiplicado pela matriz de ganho para actualizar os parâmetros e a covariância do erro estimados.
- Estas actualizações são depois usadas para se obter informação inicial e segue-se todo o procedimento descrito para actualizar, os parâmetros e o erro, no próximo instante.

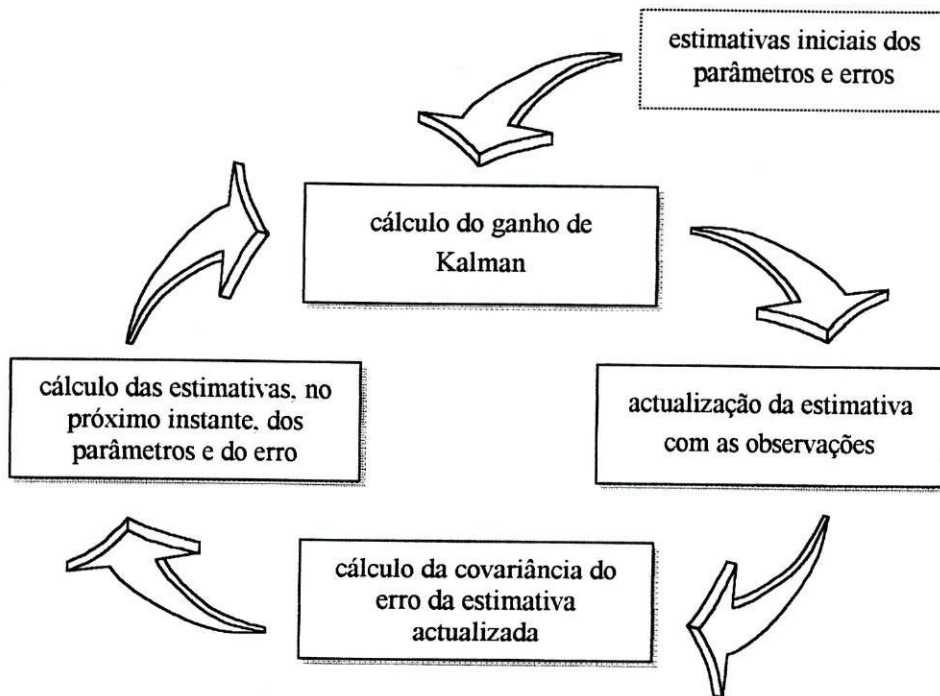


Figura 2.2 - Esquema do processo do filtro de Kalman

◆ Descrição dos Elementos do Filtro de Kalman

- $\hat{x}_{t|t}$ → vector dos parâmetros actuais, no instante t .
- $\hat{x}_{t|t-1}$ → vector da estimativa actual dos parâmetros no instante anterior à actualização.
- $\hat{x}_{t+1|t}$ → vector da estimativa do parâmetro no instante $t + 1$.
- $P_{t|t}$ → matriz da covariância do erro dos parâmetros actuais.
- $P_{t|t-1}$ → matriz da covariância do erro dos parâmetros estimados.

- $P_{t+1|t}$ → matriz da covariância do erro da estimativa do parâmetro no próximo instante, $t + 1$.
 Y_t → vector contendo as observações.
 Z_t → matriz (não contendo ruído) que dá a relação entre x_t e y_t .
 $\varepsilon_t \varepsilon_t'$ → matriz que denota o erro nas observações, que é conhecida ou que terá de ser estimada a priori.
 T_t → matriz que contém o modelo que relaciona x_t com x_{t+1} .
 K_t → matriz de ganho que relaciona a quantidade de influência do erro entre $x_{t|t-1}$ e Y_t .

◆ Algoritmo do Filtro de Kalman Discreto

Consideremos o modelo

$$y_t = Z_t x_t + \varepsilon_t \quad , t = 1, \dots, T \quad (2.67a)$$

$$x_t = T_t x_{t-1} + R_t \eta_t \quad (2.67b)$$

onde $\varepsilon_t \sim N(0, \sigma^2 h_t)$, com h_t função conhecida do tempo e R_t também conhecida.

Seja x_{t+1} o processo que pretendemos estimar com as condições iniciais $\hat{x}_0 = \mu_0$ e $P(0) = P_0$ e Y_t o processo das observações.

Suponhamos que inicialmente temos $\hat{x}_0 = \hat{x}_{0|0}$, $P_0 = P_{0|0} = P_{1|0}$ e que pretendemos calcular $\hat{x}_{1|1}$.

Começamos por calcular a função de previsão a um passo $\hat{x}_{1|0} = T_0 \hat{x}_{0|0}$ e depois calculamos a matriz de ganho K_1 , que é dada por $K_1 = P_{1|0} Z_1' [Z_1 P_{1|0} Z_1' + \varepsilon_1 \varepsilon_1']^{-1}$. De seguida actualizamos a variância através de $P_{1|1} = [I - K_1 Z_1] P_{1|0}$ e finalmente actualizamos a estimativa através de $\hat{x}_{1|1} = \hat{x}_{1|0} + K_1 [Y_1 - Z_1 \hat{x}_{1|0}]$.

No cálculo de $\hat{x}_{2|2}$, tal como acima, começamos por obter $\hat{x}_{2|1} = T_1 \hat{x}_{1|1}$ e a covariância do erro é $P_{2|1} = T_1 P_{1|1} T_1' + R_2 R_2'$.

E depois calculamos a matriz de ganho $K_2 = P_{2|1} Z_2' [Z_2 P_{2|1} Z_2' + \varepsilon_2 \varepsilon_2']^{-1}$ para podermos actualizar o estado $\hat{x}_{2|1}$, e temos: $\hat{x}_{2|2} = \hat{x}_{2|1} + K_2 [Y_2 - Z_2 \hat{x}_{2|1}]$, que é a estimativa actualizada.

Por fim actualiza-se a covariância do erro com: $P_{2|2} = [I - K_2 Z_2] P_{2|1}$.

Estas duas últimas actualizações são depois as estimativas iniciais para a estimativa no instante seguinte e assim sucessivamente para as restantes estimativas.

Um possível esquema ilustrativo do processo de actualização através do filtro de Kalman encontra-se na Figura 2.3.

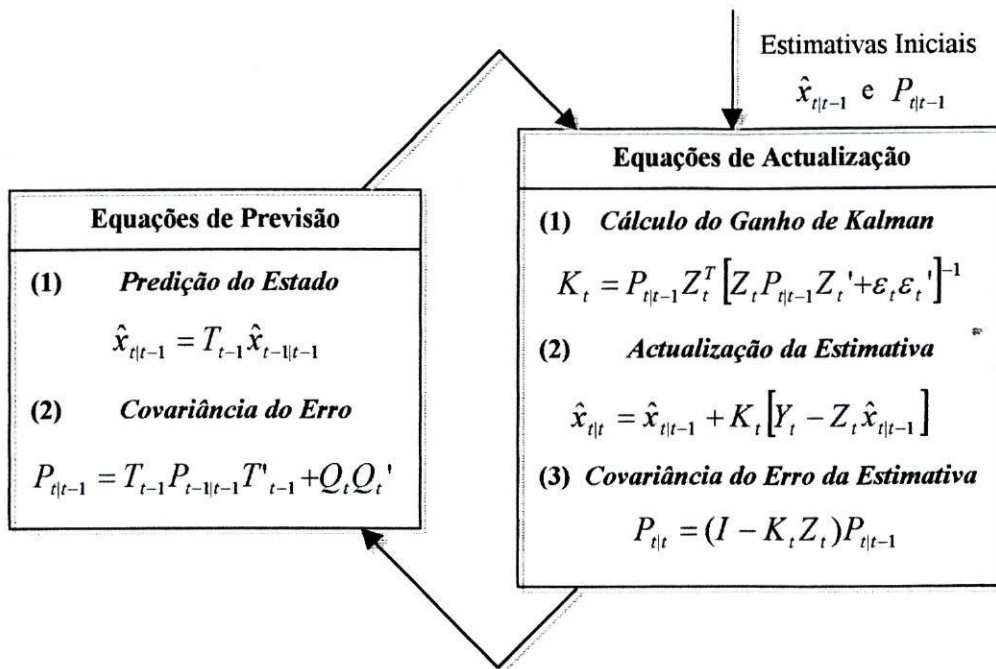


Figura 2.3 - Esquema do algoritmo do filtro de Kalman

◆ Intervalos de Confiança para Previsões

Se o modelo se encontra na forma (2.67), o predictor a m passos para o processo das observações Y_t é $\hat{Y}_{T+m|T} = Z' \hat{x}_{T+m|T} = Z' T^m \hat{x}_T$, $m = 1, 2, \dots$ e um intervalo de confiança a $(1-\alpha)100\%$ para Y_{T+m} é um tipo $\tilde{Y}_{T+m|T} \pm t_{T-d}^{(\alpha/2)} (EQM(\tilde{Y}_{T+m|T}))^2$.

3. Modelação do Número Semanal de Adultos Alojados num Hotel

3.1. Apresentação da Série

Neste capítulo vamos analisar a série do número semanal de adultos alojados num hotel, utilizando para tal as três metodologias estudadas nos dois capítulos anteriores e iremos comparar as previsões.

Os dados disponíveis são provenientes de um hotel de quatro estrelas com capacidade diária para 354 adultos, e referem-se ao número semanal de adultos alojados no mesmo, de 1 de Janeiro de 1995 a 31 de Dezembro de 1998, num total de 208 observações. A série passará a ser designada por série Adultos.

Iniciamos o estudo tentando ajustar aos nossos dados um modelo de decomposição, prosseguimos com um modelo estrutural e finalizamos com um modelo ARMA.

Os três melhores modelos obtidos nas diferentes abordagens são depois comparados conjuntamente quer em termos de ajustamento quer em termos de previsões.

Para a realização deste trabalho foi indispensável a utilização de instrumentos informáticos apropriados.

Na análise da nossa série utilizámos os *packages* estatísticos STAMP 5.0 (*Structural Time Series Analyser, Modeller and Predictor*), SPSS 9.0 (*Statistical Program for Social Sciences*) e ITSM for Windows (mais concretamente o programa PEST – *Parameter Estimation*).

Também utilizámos a folha de cálculo *Microsoft Excel 97*, sobretudo para ajustar o modelo de decomposição aditivo.

3.2. Diagnóstico Preliminar dos Dados

Neste capítulo pretendemos fazer uma análise dos dados e o primeiro passo é olhar para os gráficos dos mesmos, juntamente com as estatísticas resumo simples. Por se tratarem de observações efectuadas ao longo do tempo, os gráficos que se revelam mais importantes no estudo são o cronograma e os correlogramas analisados nas alíneas c) e d) seguintes.

a) Estatísticas Resumo

Na Tabela 3.1 apresentamos o *output* das estatísticas resumo dos dados.

N	208
Média	836,04
Erro Padrão da Média	22,77
Mediana	852,00
Moda	149 ^a
Desvio Padrão	328,32
Variância	107795,70
Enviesamento	-,112
Erro Padrão do Enviesamento	,169
Curtose	-,446
Erro Padrão da Curtose	,336
Amplitude	1510
Mínimo	101
Máximo	1611
Soma	173896

^a Existem várias modas. É apresentada a de menor valor.

Tabela 3.1 - Estatísticas resumo da série Adultos

b) Histograma

Tal como se pode observar na Figura 3.1, a série não se afasta muito da distribuição normal, apresenta uma certa assimetria positiva e a curva tem um aspecto leptocúrtico.

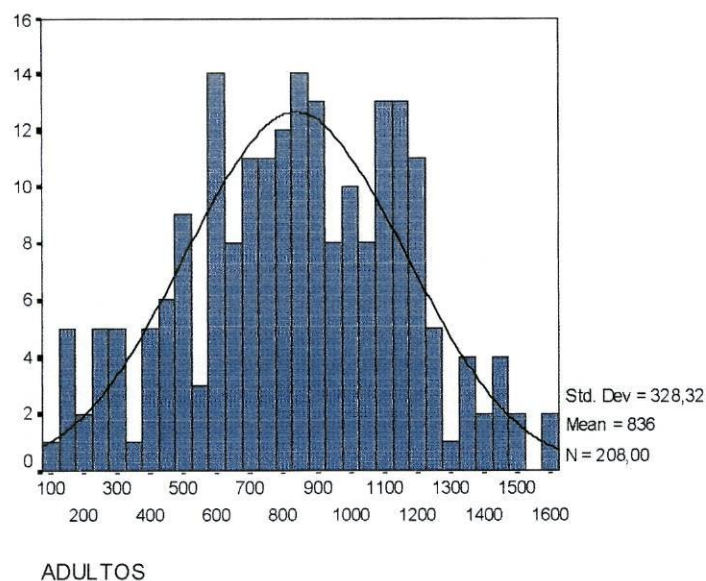


Figura 3.1 - Histograma da série Adultos

c) Cronograma

Na Figura 3.2 encontram-se representados os valores da série Adultos de 1 de Janeiro de 1995 a 31 de Dezembro de 1998.

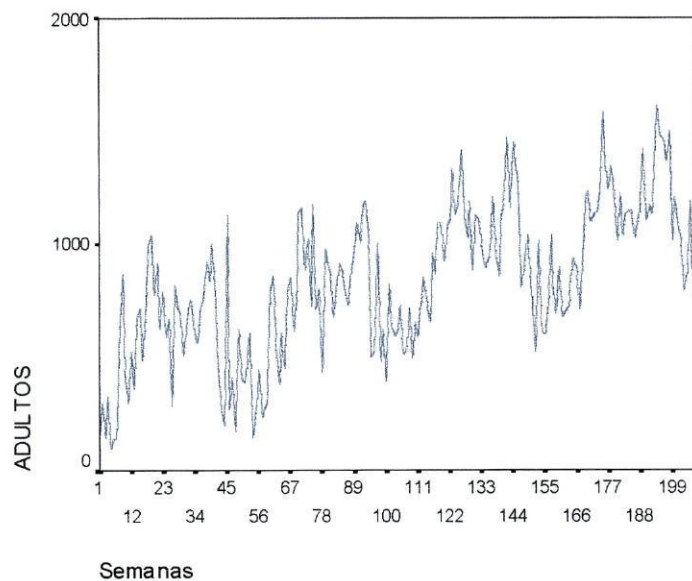


Figura 3.2 - Cronograma da série Adultos

Como se pode observar na Figura 3.2, a variabilidade não parece sofrer alterações com o nível da série, o que é um sinal de homocedasticidade. Fomos no entanto verificar se era preciso estabilizar a série através de várias transformações de Box-Cox, e concluímos que $\lambda = 1$, pelo que não foi necessário aplicar qualquer transformação.

No entanto a série parece evoluir em torno de uma tendência linear, acompanhada de um efeito sazonal, sendo portanto não estacionária

d) Correlograma

Nas Figuras 3.3 e 3.4 apresentamos os correlogramas da série em estudo, constatando que estamos perante uma série não estacionária.

A FAC estimada decresce lentamente para zero, indicando a não estacionaridade e a FACP decai para zero de forma sinusoidal.

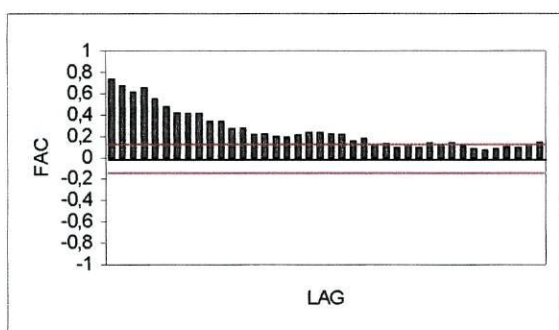


Figura 3.3 - FAC da série Adultos

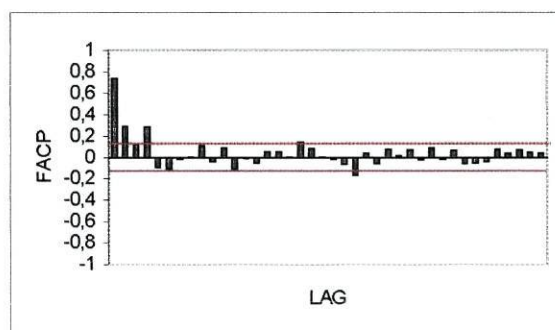


Figura 3.4 - FACP da série Adultos

e) Periodograma

Para examinar as periodicidades que regem a variação dos dados calculámos o periodograma da série Adultos, cuja representação gráfica e análise se encontra no Anexo IV.

3.3. Modelação da Série Através da Decomposição Clássica

A Decomposição Clássica da série Adultos tem como objectivo compreender melhor o seu comportamento e auxiliar na escolha de um modelo de previsão.

Por observação da Figura 3.2 verificamos que estão presentes a tendência e a componente sazonal de amplitude constante, o que sugere o modelo do tipo aditivo como sendo o mais adequado (Murteira, Müller & Turkman, 1994).

Seja

$$Y_t = S_t + T_t + \varepsilon_t, \quad t = 1, \dots, 208$$

o modelo representativo dos dados, definido como em (2.5).

3.3.1. Estimação das Componentes

A sazonalidade e a tendência são componentes não observáveis, tendo que ser estimadas a partir da informação contida na série, pelo que o objectivo inicial é estimar os índices sazonais S_t para dessazonalizar a série Y_t .

Deste modo, começamos por aplicar uma média móvel centrada, M_t , de 52 semanas aos dados da nossa série, para assim obter as estimativas não normalizadas \bar{S}_t , da componente sazonal.

Numa segunda etapa, de modo a assegurar que $\sum_{t=K+1}^{K+S} S_t = 0, \quad \forall_K$ (com S como o número de observações anuais), calculamos as estimativas normalizadas \hat{S}_t , dos índices sazonais, com base nas \bar{S}_t . No Anexo V encontra-se a tabela destas estimativas, as quais estão representadas na Figura 3.5.

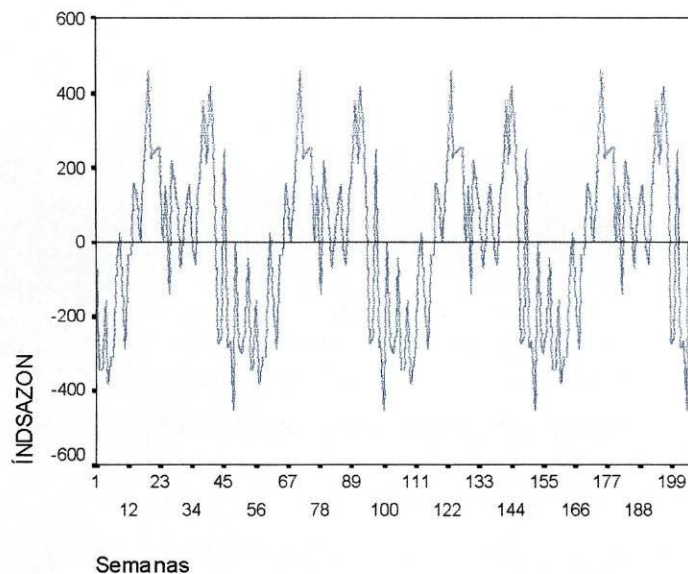


Figura 3.5 - Índices sazonais da série Adultos

Finalmente dessazonalizamos a série Y_t subtraindo a esta os efeitos sazonais.

Na Figura 3.6 apresentam-se os gráficos da série dos valores observados Y_t e a série dos valores dessazonalizados Y_t^D .

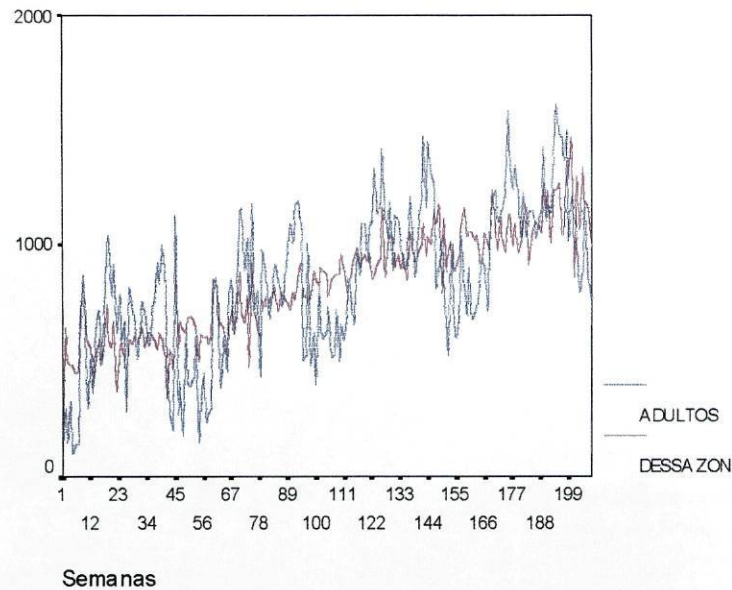


Figura 3.6 - Série Adultos original e série Adultos dessazonalizada

Uma vez que as flutuações sazonais foram amaciadas, verificamos que a série dessazonalizada apresenta apenas uma tendência linear crescente.

Na perspectiva da decomposição da série, prosseguimos a análise das componentes estimando a componente da tendência $\hat{T}_t = a + bt$ pelo método dos mínimos quadrados, obtendo como estimativa da componente tendência $\hat{T}_t = 473.1033 + 3.473061t$.

A componente residual foi obtida por subtração de \hat{T}_t à série de valores dessazonalizados: $\hat{e}_t = y_t - \hat{y}_t$.

3.3.2. Estudo da Componente Residual

Nesta seção vamos fazer a análise dos resíduos.

Como podemos observar na Figura 3.7 os resíduos não têm um comportamento padrão.

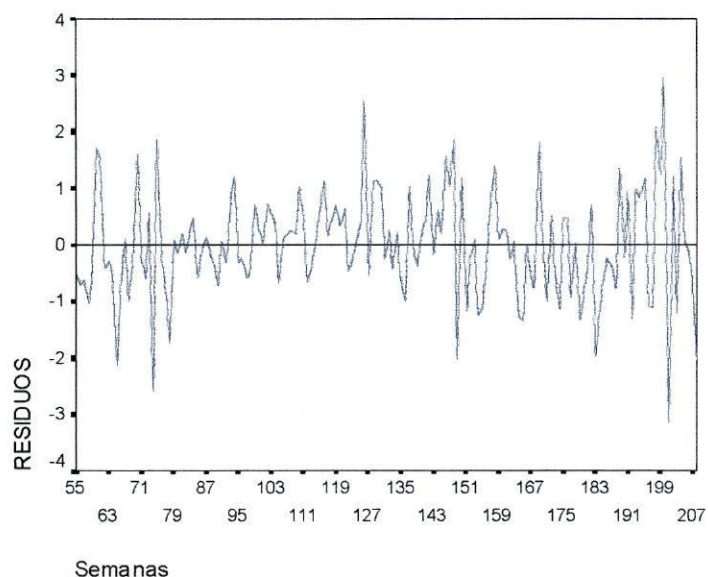


Figura 3.7 - Cronograma da componente residual

O modelo estimado ajusta-se bem à série em estudo se os resíduos sugerirem que são ruído branco (Murteira, Müller & Turkman, 1994).

Nas Figuras 3.8 e 3.9 estão representados os valores estimados das funções FAC e FACP.

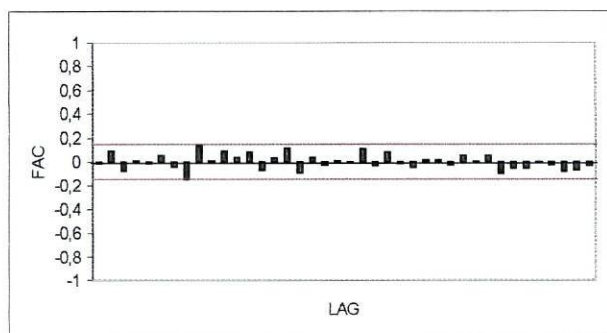


Figura 3.8 - FAC da componente residual

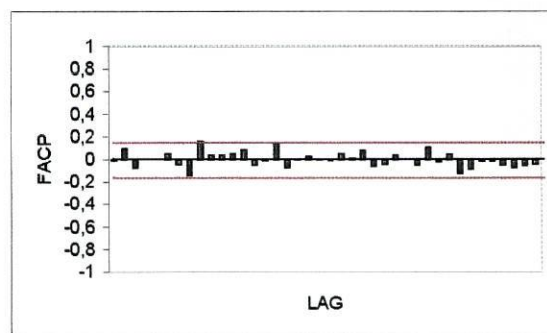


Figura 3.9 - FACP da componente residual

De facto os valores estimados não saem das bandas de confiança, demonstrando, por isso, um comportamento muito semelhante ao do ruído branco (Figuras 3.8 e 3.9).

Para verificarmos se de facto se comportam como um processo de ruído branco, de seguida vamos averiguar se se verificam as hipóteses subjacentes:

I - Normalidade

Para verificar a normalidade dos resíduos vamos em primeiro lugar observar o seu histograma e de seguida a sua representação em papel de probabilidades.

a) Histograma dos Resíduos

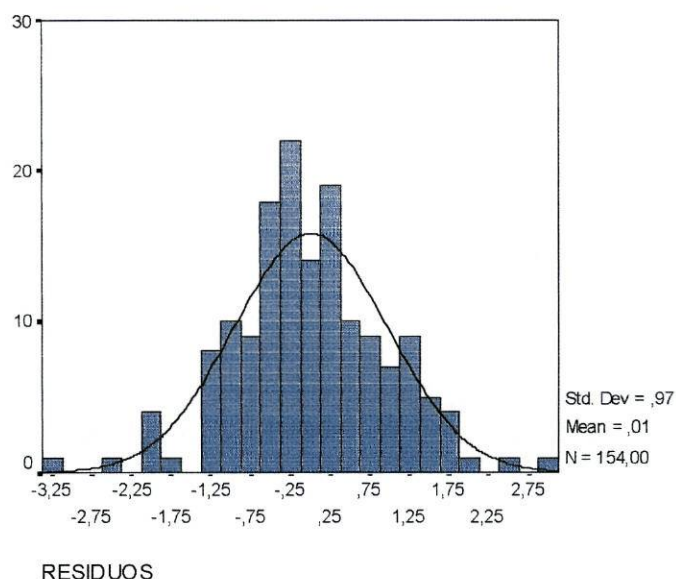


Figura 3.10 - Histograma da componente residual

A forma deste histograma sugere uma distribuição aproximadamente Gaussiana com resíduos standartizados.

b) Representação em Papel de Probabilidades

Para concluirmos da normalidade dos resíduos, estes foram representados em papel de probabilidades.

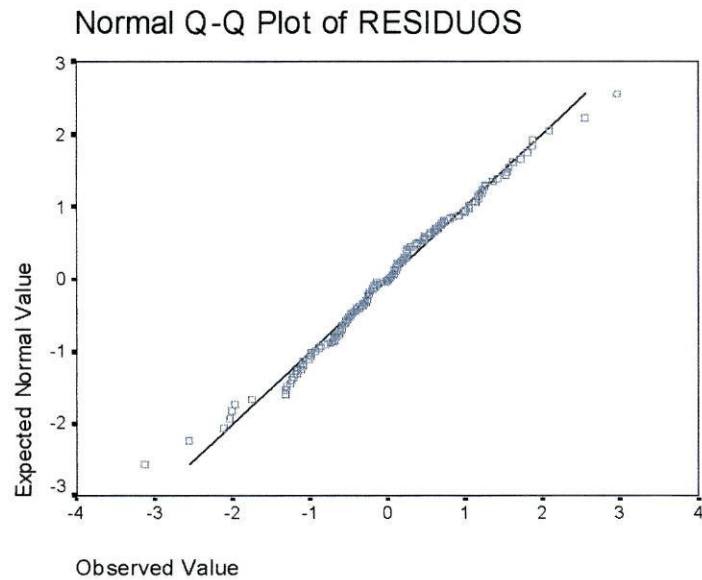


Figura 3.11 - Representação em papel de probabilidades da componente residual

Os pontos do gráfico seguem praticamente a linha recta (notando-se apenas um afastamento nas caudas), sugerindo que seguem uma distribuição normal.

II - Variância Constante

Não é evidente que exista violação desta hipótese, pois os resíduos apresentam-se com um comportamento aleatório, não seguindo nenhum padrão (Figura 3.12).

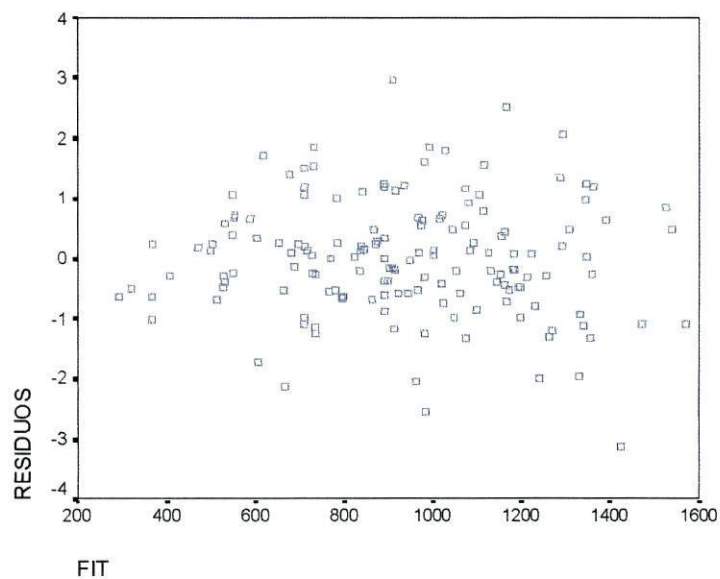


Figura 3.12 - Componente residual vs. Valores Estimados

III - Independência

Na prática, violações desta hipótese são muito difíceis de distinguir graficamente. Assim, uma forma de analisar esta hipótese é voltarmos a olhar para a Figura 3.12. Como não temos evidência do tipo de resíduos positivos se seguem a resíduos positivos e resíduos negativos se seguem a resíduos negativos, então não nos parece que esta hipótese seja violada.

Os testes (descritos no Anexo I) efectuados aos resíduos levam à não rejeição das hipóteses de gaussianidade e independência. Assim a série é completamente explicada pelo modelo $y_t = 473,1 + 3,47t + S_t + e_t$.

Teste	Valores	Distribuição	Hipótese a Testar
Ljung-Box Portmanteau	23.18	χ^2_{20}	Os resíduos são Observações Gaussianas i.i.d.
McLeod-Li Portmanteau	24.27	χ^2_{20}	
Pontos de Viragem	131	$N(137.33; 6.05^2)$	
Rank	10764	$N(10764; 1505.26^2)$	
Ordem do Modelo com AICC Mínimo = 0			

Tabela 3.2 - Testes efectuados à componente residual

3.4. Modelação Box-Jenkins

Tentámos ajustar à série dos dados um modelo ARMA, como tal seguimos a metodologia de Box-Jenkins apresentada na Figura 2.1.

Deste modo, numa primeira fase tentámos identificar o modelo.

3.4.1. Identificação do Modelo

Para neutralizar a tendência diferenciámos a série observada, aplicando o operador diferença, com $d = 1$:

$$X_t = \nabla Y_t = Y_t - Y_{t-1}$$

De facto, diferenciando a série conseguimos estacioná-la como se pode ver na Figura 3.13.

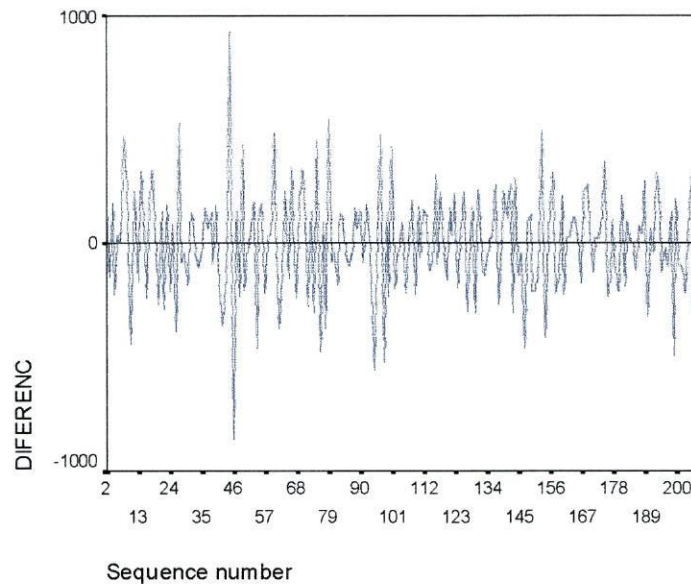


Figura 3.13 - Série Adultos diferenciada a 1

Nas Figuras 3.14 e 3.15 apresentamos os gráficos da FAC e FACP da série diferenciada.

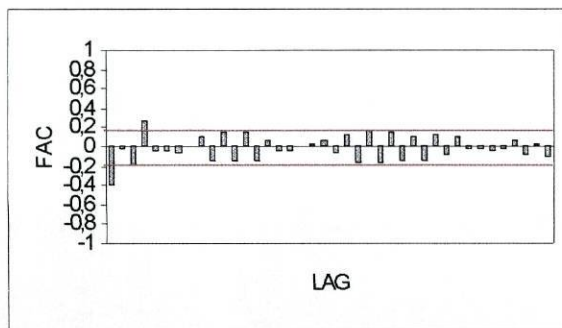


Figura 3.14 - FAC dos resíduos

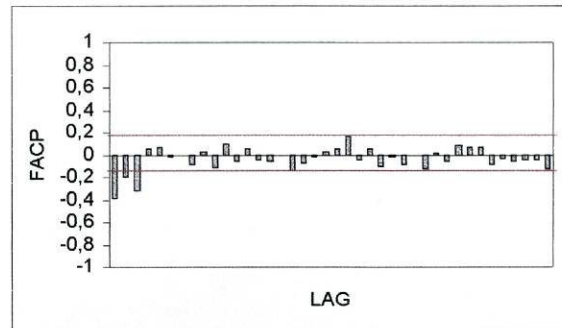


Figura 3.15 - FACP dos resíduos

Os gráficos da FAC e da FACP da série diferenciada sugeriam vários modelos ARMA. Fomos então testar vários modelos ARMA, cujo resumo da análise apresentamos no Anexo VI, concluindo que o mais adequado parece ser um AR (3).

3.4.2. Estimação dos Parâmetros

O modelo que iremos estudar é, portanto, um ARIMA (3, 1, 0):

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - B)y_t = e_t$$

Estimámos os parâmetros pelo método da máxima verosimilhança, e os valores encontrados são os que figuram na Tabela 3.3.

Parâmetro	Estimativa	Desvio do Erro	T	Valor de p
ϕ_1	-0.5277295	0.0660738	-8.0130862	0.0000
ϕ_2	-0.3474348	0.0718723	-4.8608065	0.0000
ϕ_3	-0.3307658	0.0664675	-5.0349918	0.0000

Tabela 3.3 - Estudo dos parâmetros

Pela matriz de correlações vemos que os parâmetros estão pouco correlacionados, o que sugere a razoável qualidade do modelo (Murteira, Müller & Turkman, 1994).

	AR1	AR2	AR3
AR1	1,0000000	,4311042	,1981055
AR2	,4311042	1,0000000	,4336456
AR3	,1981055	,4336456	1,0000000

Tabela 3.4 - Matriz de correlações dos parâmetros

Então o modelo estimado é

$$(1+0,5277B+0,3474B^2+0,3308B^3)(1-B)y_t = e_t$$

3.4.3. Estudo dos Resíduos

Esta é a segunda sub-etapa da segunda fase, na qual vamos verificar as hipóteses sobre os resíduos.

◆ Avaliação da Qualidade do Ajustamento

Pela análise visual do cronograma dos resíduos, Figura 3.16, concluímos que estes não revelam padrões de comportamento.

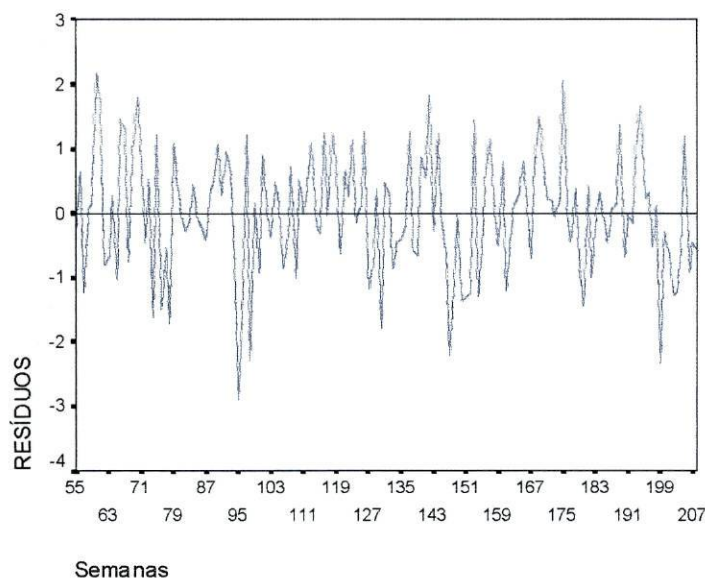


Figura 3.16 - Cronograma dos Resíduos

Nas Figuras 3.17 e 3.18, temos representadas a FAC e a FACP e verificamos que são compatíveis com as do ruído branco.

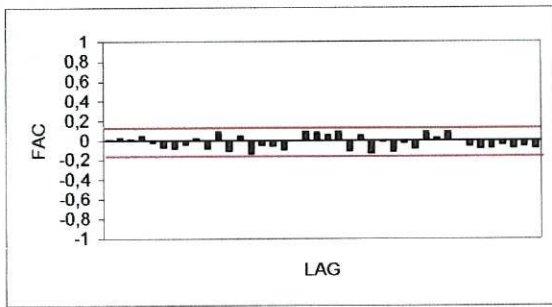


Figura 3.17 - FAC dos resíduos

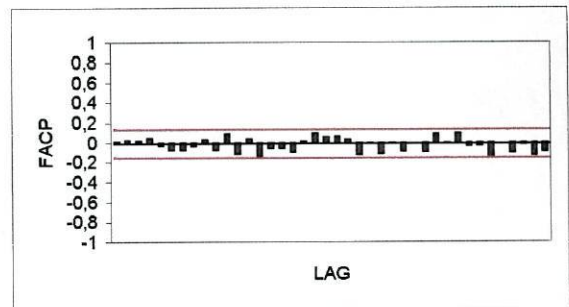


Figura 3.18 - FACP dos resíduos

Mas para validar esta afirmação temos de constatar se de facto são uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média próxima de zero e variância constante.

I - Normalidade

a) Histograma dos Resíduos

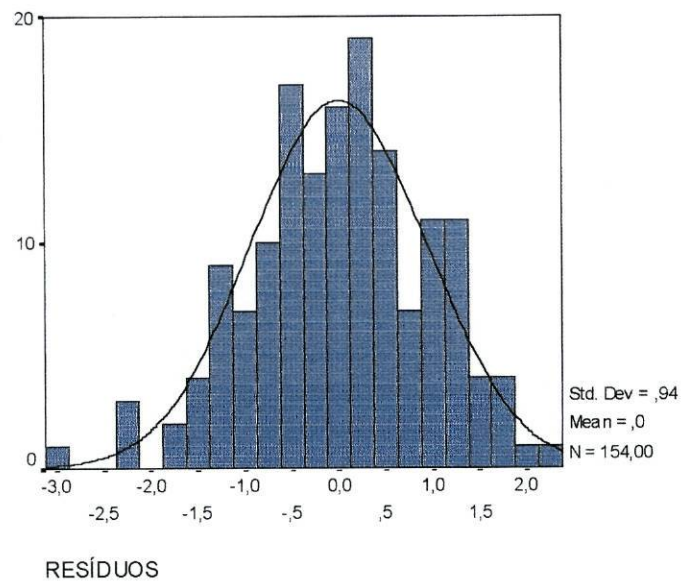


Figura 3.19 - Histograma dos Resíduos

O histograma dos resíduos revela um comportamento destes muito próximo da distribuição Gaussiana (Figura 3.19).

b) Representação em Papel de Probabilidades

Na Figura 3.20 temos a representação dos resíduos em papel de probabilidades na qual, confirmando o que já tínhamos dito anteriormente, podemos observar que seguem de perto a linha recta, o que sugere que seguem uma distribuição Gaussiana.

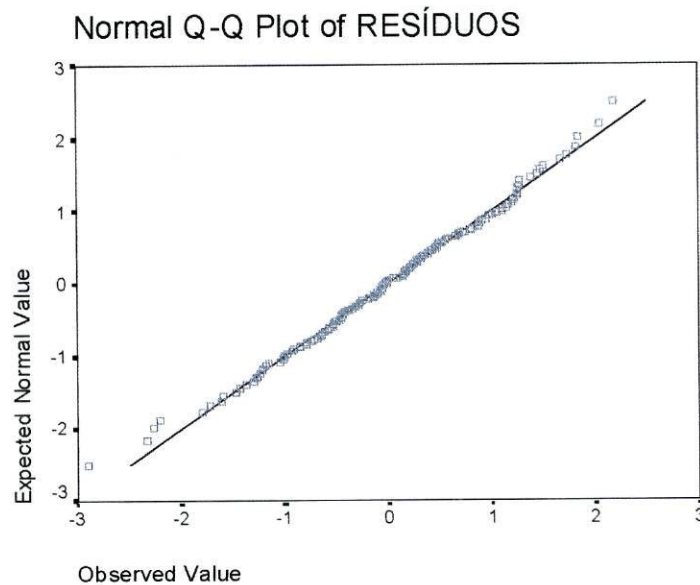


Figura 3.20 - Representação em Papel de Probabilidades dos resíduos

II - Variância Constante

Com base na representação dos resíduos do modelo estimado versus os valores estimados, podemos identificar claras violações da hipótese da variância constante.

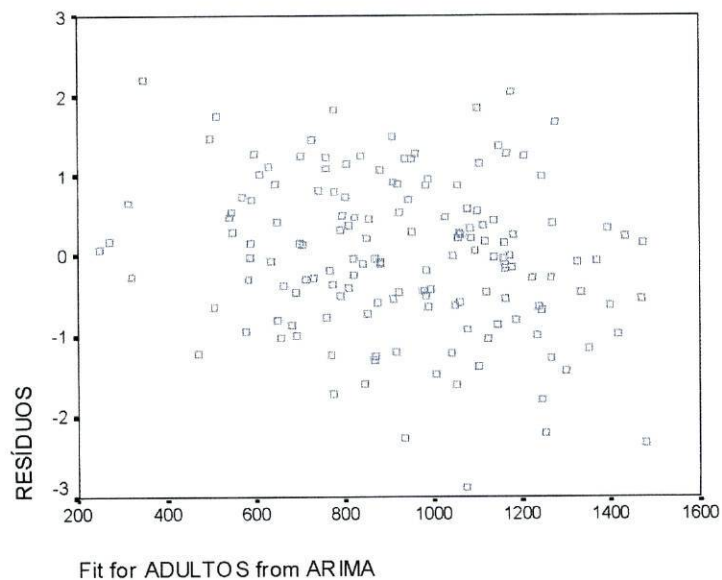


Figura 3.21 - Resíduos vs. valores estimados

No nosso caso, de acordo com a Figura 3.21, não é evidente que haja heterocedasticidade, pois não existe nenhum tipo de padrão no seu aspecto.

III - Independência

Finalmente temos que verificar se os resíduos são independentes e pela Figura 3.21 observamos que de facto têm um comportamento aleatório, não revelando falta de independência.

Para finalizar a análise dos resíduos temos os valores obtidos, com o ITSM, para os vários testes na Tabela 3.5.

Teste	Valores	Distribuição	Hipótese a Testar
Ljung-Box Portmanteau	23.30	χ^2_{20}	Os resíduos são observações Gaussianas i.i.d.
McLeod-Li Portmanteau	31.94	χ^2_{20}	
Pontos de Viragem	128	$N(136.67; 6.04^2)$	
Teste Rank	10224	$N(10660.50; 1494.45^2)$	
Ordem do Modelo com AICC Mínimo = 0			

Tabela 3.5 - Testes efectuados aos Resíduos

Concluimos que os valores obtidos levam à não rejeição da hipótese nula.

3.5. Modelo Estrutural

Visando ajustar um modelo estrutural à série Adultos, a selecção das componentes a incluir no modelo é baseada nas características observáveis da série.

Uma vez que, tal como já tínhamos visto, a série apresenta tendência e sazonalidade, então fomos testar alguns modelos que incluíssem estas componentes de previsão, dos quais o modelo estrutural básico (MEB) com sazonalidade trigonométrica foi o que apresentou melhores resultados (Anexo VII).

3.5.1. Estimação das Componentes

Dado que este modelo já foi definido na Secção 3.2., de seguida apenas apresentamos as estimativas significativas das componentes (Tabela 3.6).

μ_t	β_t	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_9
1192.0	3.3624	-192.47	-60.466	-125.51	-84.169	39.751	90.779	35.073
γ_{17}	γ_{19}	γ_{21}	γ_{26}	γ_{28}	γ_{34}	γ_{35}	γ_{45}	γ_{47}
-54.412	-50.855	32.234	64.391	43.693	34.575	-33.464	-43.414	-39.457

Tabela 3.6 - Componentes do MEB

Como podemos observar na Tabela 3.6, e devido ao facto de termos utilizado a sazonalidade trigonométrica, obtivemos poucos coeficientes sazonais, o que corrobora a afirmação de Andrews (1994:130): “*The primary advantage of trigonometric seasonality is that, in some cases, the higher-order frequencies may be removed, resulting in fewer seasonal coefficients. This could be particularly useful when modeling Weekly or daily data.*”

Ainda dando relevo a esta afirmação, realçamos que o MEB com sazonalidade muda apresentou 33 coeficientes sazonais, ou seja, mais do dobro do que com a sazonalidade trigonométrica.

3.5.2. Estudo da Componente Residual

É através da análise dos resíduos que verificamos se o modelo estimado descreve bem a sucessão em estudo. O modelo pode ser considerado adequado se os resíduos se apresentarem como ruído branco. Observando o cronograma dos resíduos (Figura 3.22) vemos que a sua média é próxima de zero.

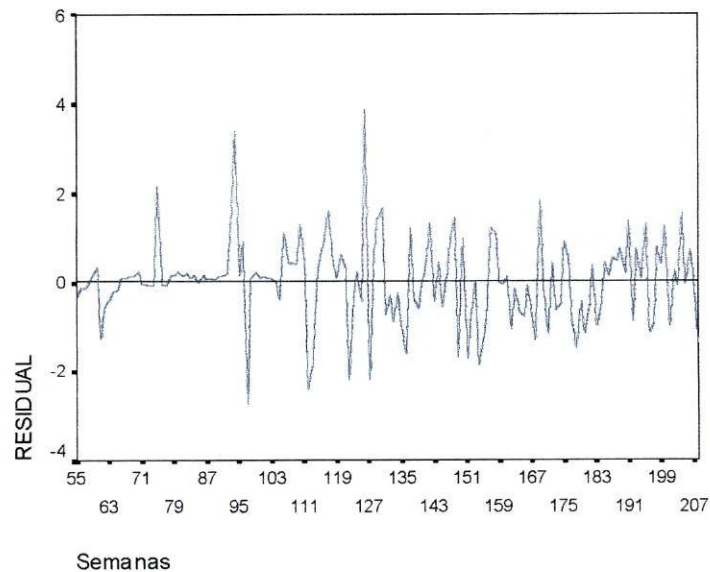


Figura 3.22 - Cronograma da componente residual

Os valores estimados da FAC e FACP, Figuras 3.23 e 3.24, apresentam características de ruído branco.

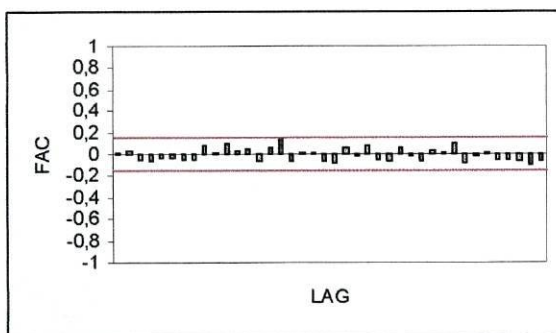


Figura 3.23 – FAC da componente residual

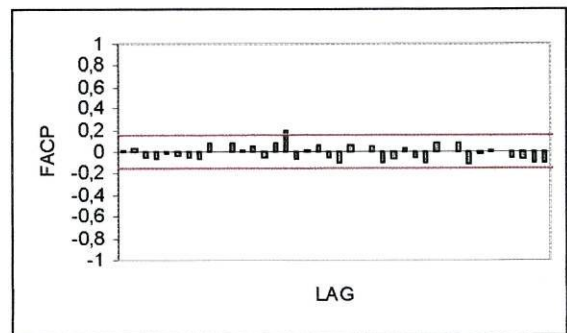


Figura 3.24 - FACP da componente residual

Fizemos outras representações gráficas dos resíduos (Anexo VIII), através das quais constatámos que estes se afastam ligeiramente da normalidade, não nos parecendo no entanto que exista heterocedasticidade e que estejam correlacionados.

Na Tabela 3.7 apresentamos os resultados obtidos, com o ITSM, para os vários testes efectuados aos resíduos. Os resultados agradam-nos particularmente, pois levam à não rejeição da hipótese nula, confirmando que o MEB com sazonalidade trigonométrica é o modelo adequado para a série Adultos.

Teste	Valores	Distribuição	Hipótese a Testar
Ljung-Box Portmanteau	16.69	χ^2_{20}	Os resíduos são observações Gaussianas i.i.d.
McLeod-Li Portmanteau	24.21	χ^2_{20}	
Pontos de Viragem	101	$N(101.33; 27.04)$	
Teste Rank	5901	$N(5890.5; 12.89)$	
Ordem do Modelo com AICC Mínimo = 0			

Tabela 3.7 - Testes efectuados à componente residual

3.6. Comparação de Resultados

Nesta fase do estudo a comparação dos resultados obtidos pelos três modelos passa a ser o nosso principal objectivo. Começamos portanto por avaliar a qualidade do ajustamento de cada um dos modelos, utilizando para tal o desvio padrão da série residual (Müller, 1997).

Deste modo, apesar de os três modelos se ajustarem bem à série Adultos, verificamos que o que apresenta melhores resultados, em termos de desvio padrão da série residual, é o modelo de decomposição (Tabela 3.8).

Modelo	$\hat{\sigma}_e$
Decomposição	100.2
MEB	105.1
ARIMA	199.6

Tabela 3.8 - Desvio padrão da componente residual

Passando agora à análise gráfica do ajustamento dos vários modelos, desde a 3ª semana de 1996 até à 52ª semana de 1998, começamos pelo modelo de Decomposição.

Podemos observar na Figura 3.25 que apesar de o modelo de decomposição não apresentar um mau ajustamento, mesmo assim verificamos que em alguns picos a aderência é fraca.

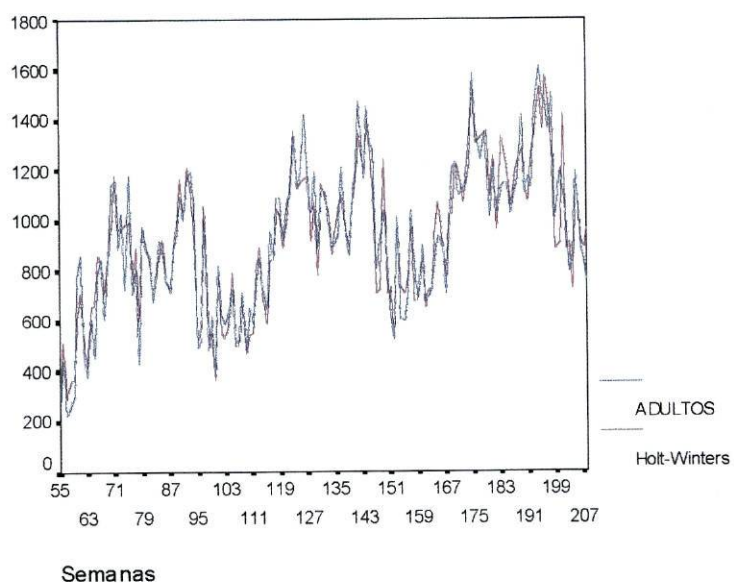


Figura 3.25 - Série adultos e valores ajustados pelo modelo de decomposição

Por seu lado, o MEB (Figura 3.26) também apresenta um ajustamento razoável à série Adultos, embora não tão bom quanto o modelo anterior.

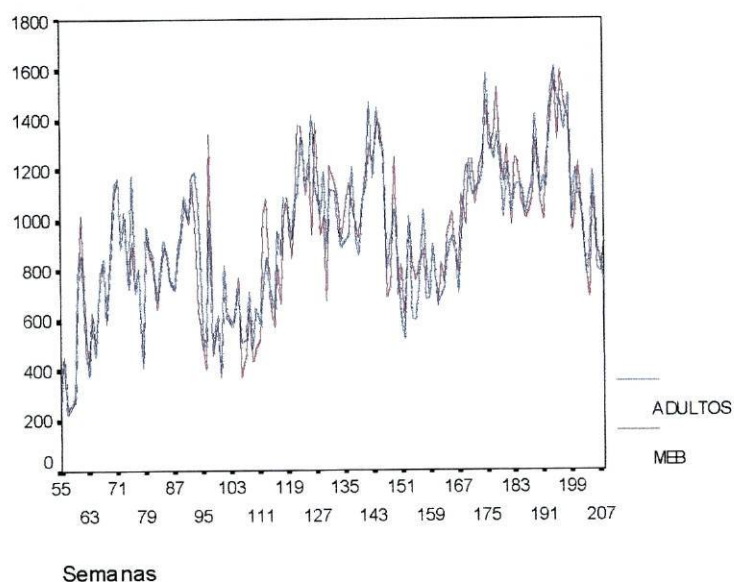


Figura 3.26 - Série Adultos e valores ajustados pelo MEB

Por último, no que diz respeito ao modelo ARIMA, verificamos que este é o que apresenta maiores dificuldades em reproduzir as observações, como se pode constatar na Figura 3.27.

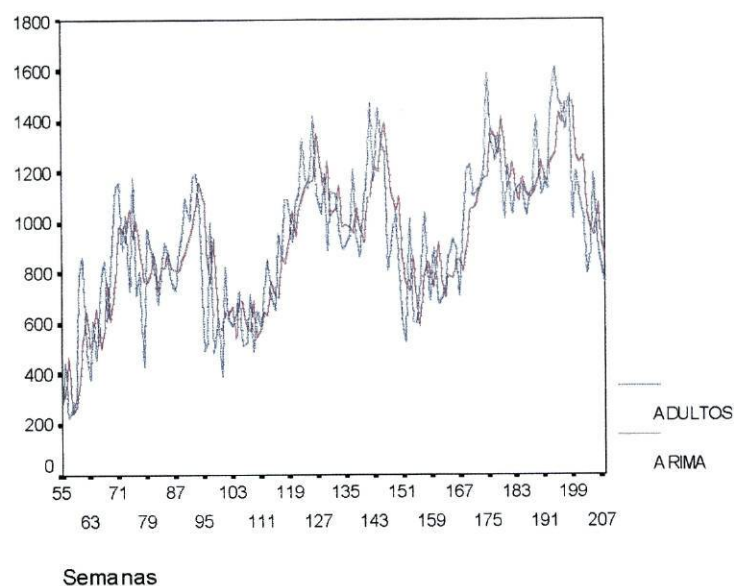


Figura 3.27 - Série Adultos e valores ajustados pelo modelo Arima

Até aqui vimos o ajustamento dos vários modelos à série Adultos em termos do seu passado, mas o mais interessa no nosso estudo é o comportamento dos modelos em termos futuros. Pois nem sempre o modelo que melhor se ajusta ao conjunto das observações é o que produz as melhores previsões.

3.7. Previsão

Nesta Secção do nosso estudo apresentamos as previsões que obtivemos com os três modelos para as primeiras 17 semanas de 1999 (as primeiras subsequentes às da nossa série), as quais apresentamos na Tabela 3.9.

	Semanas	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225
Valores Observados		1203	979	868	1072	821	850	866	978	1057	998	818	1068	1330	1348	1233	1205	1253
Previsões	Decomp.	1157	859	863	1055	833	909	910	1159	1252	1072	948	1204	1207	1404	1380	1252	1376
	MEB	1207	986	873	1081	859	894	919	1043	1132	1076	906	1147	1378	1419	1300	1310	1343
	ARIMA	926	891	882	850	889	890	894	885	895	899	903	903	908	911	915	917	920

Tabela 3.9 - Valores observados e previstos para a série Adultos

Para podermos avaliar rigorosamente a qualidade das previsões e saber a credibilidade que lhes podemos conferir, importa determinar o erro quadrático médio (Makridakis, Wheelwright & Hyndman, 1998): $EQM = \frac{1}{17} \sum_{t=1}^{17} e_t^2$ e como se pode ver na Tabela 3.10 o MEB é o modelo que apresenta melhores previsões.

Modelo	Holt-Winters	MEB	ARIMA
EQM	11228.4	3941.5	52001.6

Tabela 3.10 - EQM para os três modelos estudados

Nas Figuras 3.28, 3.29 e 3.30 podemos ver isoladamente as previsões e os intervalos de previsão de cada um dos modelos para a série Adultos.

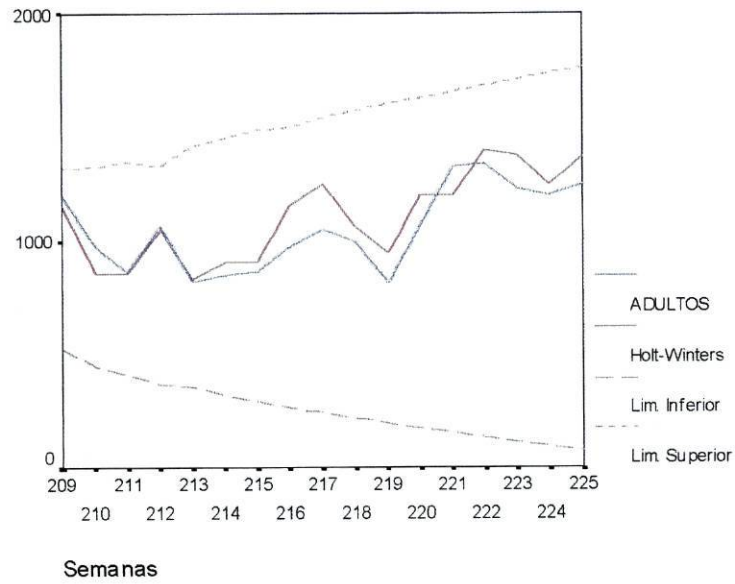


Figura 3.28 - Série Adultos, previsões e intervalos de previsão do modelo de Decomposição

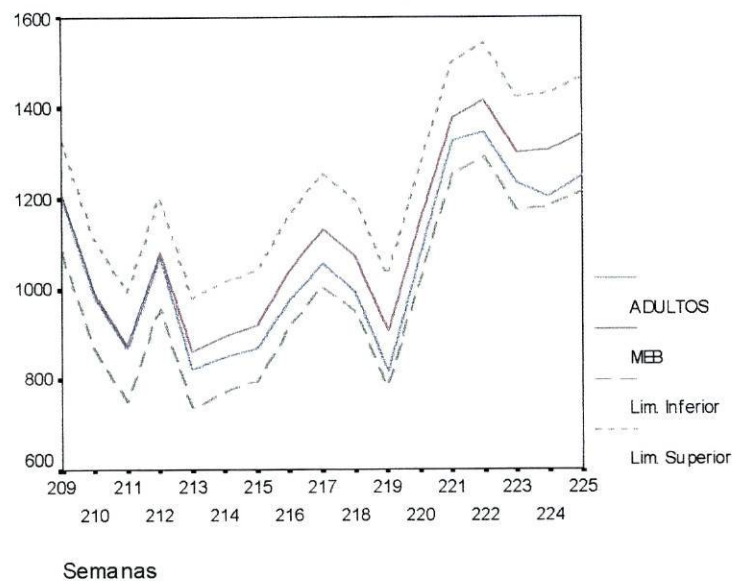


Figura 3.29 - Série Adultos, previsões e intervalos de previsão obtidas pelo modelo MEB

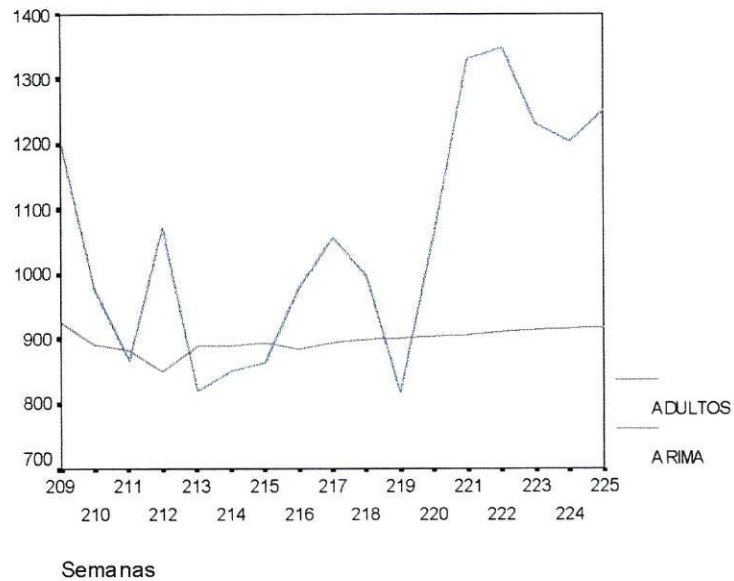


Figura 3.30 - Série Adultos e previsões do modelo ARIMA (3, 1,0)

Para melhor podermos comparar as previsões obtidas pelos três modelos representámo-las conjuntamente na Figura 3.31.

Como já havíamos observado, o modelo que melhor se ajusta à série Adultos é o MEB, e o modelo que mais se afasta é o modelo ARIMA (3, 1, 0).

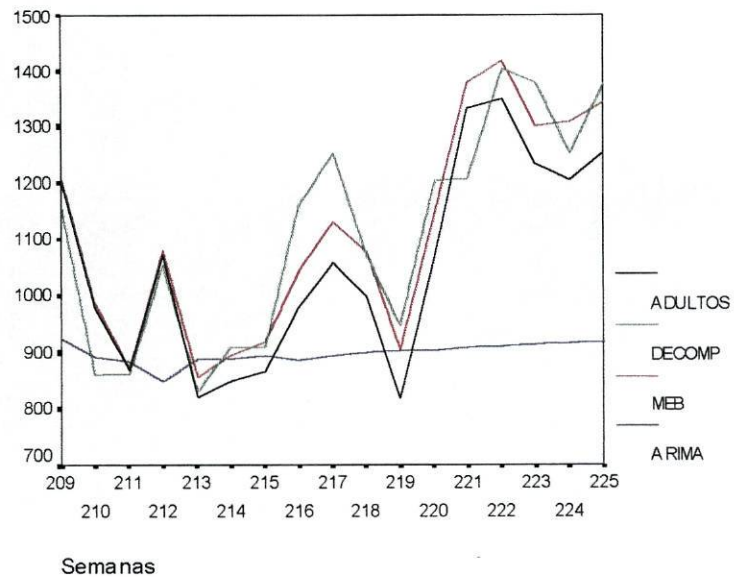


Figura 3.31 - Série Adultos, modelo de decomposição, MEB e ARIMA

Constatamos que as previsões do MEB e as obtidas pelo método Holt-Winters são semelhantes, mas nas duas últimas semanas as previsões deste último são um pouco melhores do que as do MEB. Relativamente aos outros dois, o modelo ARIMA é o modelo que oferece as previsões menos ajustadas.

Por observação da Figura 3.31 vemos que nas primeiras cinco semanas as previsões do método Holt-Winters e do MEB são praticamente coincidentes umas com as outras e com a série Adultos. A partir da 7ª semana até à 15ª, o MEB é o que melhor acompanha o comportamento da série Adultos. Deste modo, em termos globais, consideramos que o modelo que melhores previsões nos fornece é o MEB.

3.8. Conclusão

Neste trabalho estudámos três métodos diferentes de análise de séries temporais: decomposição clássica, modelos estruturais e Box-Jenkins. As duas primeiras são muito semelhantes, no sentido em que ambas decompõem as séries nas componentes que as compõem: tendência, sazonalidade, ciclo e ruído.

Nos modelos estruturais estas componentes são formalizadas como processos estocásticos. O que é óptimo do ponto de vista da previsão, pois permite a representação destes modelos em espaço de estados, que por sua vez permite a utilização do filtro de Kalman. Este filtro é um processo de estimação recursivo óptimo que actualiza as estimativas à medida que vão surgindo novas observações.

Na decomposição clássica também se definem e se constróem as componentes, que são não observáveis, permitindo uma certa sistematização do trabalho. Mas a estimação das componentes está sujeita a alguma arbitrariedade dos métodos a seguir, justificando-se a sua aplicação pela presença da componente residual com comportamento de ruído branco. Daí que teoricamente este método esteja menos explicado que na prática.

A aplicação da metodologia Box-Jenkins apresenta mais rigor que a decomposição clássica, mas menos facilidade em ser executada e interpretada. O rigor da sua aplicação advém do facto de ser sustentada por uma família de modelos teóricos eventualmente capazes de reproduzir o comportamento da série. A sua execução é dificultada por exigir algumas condições (e.g., estacionaridade) da série temporal antes da sua modelização, e porque o modelo que se obtém não oferece a possibilidade de inferência imediata das características da série.

Na parte final deste trabalho modelámos a série do número semanal de adultos alojados num hotel, utilizando para tal as três metodologias referidas. Verificou-se que o modelo de decomposição clássica foi o que apresentou melhor ajustamento à série em termos de variância residual.

Obtivemos as previsões para as primeiras dezassete semanas subsequentes às nossas observações, com o intuito de observar qual o método que conseguia obter melhores resultados. Constatámos que o modelo estrutural básico (MEB) foi o que forneceu melhores previsões. Pelo método Holt-Winters (na sequência da decomposição clássica da série) também se obtiveram boas previsões, não tanto quanto pelos modelos estruturais, mas bastante melhores do que as obtidas com o modelo ARIMA (utilizando o método Box-Jenkins). Este último revelou uma quase total falta de aderência ao comportamento da série.

Como projecto para trabalho futuro seria interessante analisar uma série semelhante, mas com observações diárias, em que fossem considerados os feriados, os fins-de-semana e outras datas relevantes, pois para a área da hotelaria, caso se conseguissem resultados animadores, seria um trabalho de grande interesse.

Julgamos que considerando estes aspectos num futuro estudo, o trabalho se revelaria muito mais conclusivo e valioso.

Bibliografia

- Andrews, R. L. (1994). Forecasting Performance of Structural Time Series Models. *Journal of Business & Economic Statistics*, 12 (1), 129-133.
- Box, G. E. P. & Cox, D. R. (1964). An Analysis of Transformation. *Journal of Royal Statistical Society, Series B*, (26), 211-252.
- Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control*. New Jersey: Prentice Hall.
- Brockwell, P. J. & Davis, R. A. (1991). *Time Series: Theory and Methods*. New York: Springer-Verlag.
- Brockwell, P. J. & Davis, R. A. (1996). *Introduction to Time Series and Forecasting*. New York: Springer.
- Harvey, A. C. (1993). *Time Series Models*. New York: Harvester Wheatsheaf.
- Harvey, A. C. (1996). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Harvey, A. C. (1997). Trends, Cycles and Autoregressions. *The Economic Journal*, (107), 192-201.
- Makridakis, S., Wheelwright, S. C. & Hyndman, R. J. (1998). *Forecasting: Methods and Applications*. New York: John Wiley & Sons, Inc..
- Müller, D. (1997). *Como Lidar com as Mudanças de Estrutura em Séries Temporais ?* Textos de Apoio nº 10. CEMAPRE.
- Murteira, B. J. F., Müller, D. A. & Turkman, K. F. (1994). *Análise de Sucessões Cronológicas*. Lisboa: McGraw-Hill.
- Welch, G. & Bishop, G.(s/d). *An Introduction to the Kalman Filter*.
<http://www.cs.unc.edu/~welch/kalman/kalman.html> (17 Mai.1999).

Anexo I

Testes estatísticos efectuados aos resíduos (Brockwell & Davis, 1994 e 1991)

- **Teste de Ljung-Box Portmanteau** – Neste teste é utilizada a estatística

$$Q = n(n+2) \frac{\sum_{k=1}^h \hat{\rho}_{Rk}^2}{(n-k)}$$

onde $\hat{\rho}_{Rk}$ é a autocorrelação amostral dos resíduos, no lag k e h deve ser da ordem de \sqrt{n} onde n é a dimensão da amostra. Este teste rejeita o modelo ARMA (p, q) proposto, no nível α , se o valor observado de Q for maior do que o quantil $(1-\alpha)$ da distribuição χ_{h-p-q}^2 .

- **Teste de McLeod-Li Portmanteau** – Apesar de se basear na mesma estatística que o teste anterior, este teste em vez de utilizar as autocorrelações amostrais utiliza as autocorrelações amostrais dos quadrados dos resíduos $\hat{\rho}_{RRk}$, obtendo-se:

$$\tilde{Q} = n(n+2) \frac{\sum_{k=1}^h \hat{\rho}_{RRk}^2}{(n-k)}$$

A hipótese dos resíduos serem normais e i.i.d. será rejeitada no nível α , se o valor de \tilde{Q} for maior do que o quantil $(1-\alpha)$ da distribuição χ_{h-p-q}^2 .

- **Teste baseado nos “Pontos de Viragem” (Turning Points)** – A estatística T utilizada no presente teste é o número de pontos de viragem na sucessão dos resíduos. Mostra-se que para uma sucessão i.i.d., T é assintoticamente normal com média $\mu_T = 2(n-2)/2$ e variância $\sigma_T^2 = (16n-29)/90$. A hipótese dos resíduos serem uma sucessão de observações i.i.d. é então rejeitada se $|T - \mu_T| / \sigma_T > \Phi_{1-\alpha/2}$ onde $\Phi_{1-\alpha/2}$ é o quantil $(1-\alpha/2)$ da distribuição standard.

- **Teste Rank** – Este é um teste muito útil para detectar alguma tendência linear nos resíduos. Seja P o número de pares (i, j) tais que $\hat{R}_j > \hat{R}_i$ e $j > i, i = 1, \dots, n-1$. Se os resíduos forem i.i.d., então a média de P é $\mu_P = n(n-1)/4$, a variância é $\sigma_P^2 = n(n-1)(2n+5)/8$ e P é assintoticamente normal. A hipótese dos resíduos serem uma sucessão de observações i.i.d. é consequentemente rejeitada se $|P - \mu_P| / \sigma_P > \Phi_{1-\alpha/2}$ onde $\Phi_{1-\alpha/2}$.
- **Modelo AR com AICC mínimo** – Se os resíduos forem compatíveis com uma sequência de observações i.i.d., então a autorregressão com AICC mínimo ajustável aos resíduos deverá ter ordem $p = 0$. No ITSM (PEST) através das equações de Yule-Walker são calculados os valores de AICC para modelos AR de ordens 0, 1, ..., 26 ajustados aos resíduos, dando-nos depois a ordem p do modelo com AICC mínimo. Um valor de $p > 0$ sugere a existência de alguma correlações nos resíduos, contradizendo a hipótese dos resíduos serem observações de um ruído branco.

Anexo II Tabela A.1 – Principais Modelos e Componentes Estruturais de Séries Temporais

Modelo	Componentes	Especificação	Operador de Estacionariedade	Forma Reduzida
	(1a) Passeio aleatório	$\mu_t = \mu_{t-1} + \eta_t$	$\Delta = 1 - B$	—
	(1b) Passeio aleatório com declive	$\mu_t = \mu_{t-1} + \beta + \eta_t$	Δ	—
A	Nível local/ passeio aleatório com ruído	$\gamma_t = \mu_t + \varepsilon_t$ com μ_t como em (1a)	Δ	ARIMA (0, 1, 1)
	(2) Tendência estocástica	$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$ $\beta_t = \beta_{t-1} + \xi_t$	Δ^2	—
B	Tendência Localmente linear	$\gamma_t = \mu_t + \varepsilon_t$ com μ_t como em (2)		ARIMA (0, 2, 2)
	(3) Ciclo estocástico	$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} k_t \\ k_t^* \end{bmatrix}$ onde ψ_t é o ciclo $0 \leq \rho < 1$, e $0 \leq \lambda_c \leq \pi$	1	ARMA (2, 1)
C	Ciclo com ruído	$\gamma_t = \mu_t + \psi_t + \varepsilon_t$, onde $0 \leq \rho < 1$	1	Constante + ARMA (2, 2)
D	Tendência mais ciclo	$\gamma_t = \mu_t + \psi_t + \varepsilon_t$, com μ_t como em (2)	Δ^2	ARIMA (2, 2, 4)
E	Tendência cíclica	$\gamma_t = \mu_t + \varepsilon_t$, $\mu_t = \mu_{t-1} + \psi_{t-1} + \beta_{t-1} + \eta_t$ com β_t como em (2)	Δ^2	ARIMA (2, 2, 4)
	(4) Ciclo não estacionário	Como (3) mas $\rho = 1$	$1 - 2\cos\lambda_c B + B^2$	$(1 - 2\cos\lambda_c B + B^2)\psi_t \sim \text{MA}(1)$
	(5a) Variável muda da sazonalidade	$\gamma_t = \sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t$	$S(B) = 1 + B + B^2 + \dots + B^{s-1}$	$S(B)\gamma_t \sim \text{WN}$
	(5b) Sazonalidade trigonométrica	$\gamma_t = \sum_{j=1}^{\lfloor s/2 \rfloor} \gamma_{t,j}$ onde $\gamma_{t,j}$ é um ciclo não estacionário, (4) com $\lambda_c = \lambda_j = 2\pi j/s$, $j=1, 2, \dots, \lfloor s/2 \rfloor$	$S(B)$	$S(B)\gamma_t \sim \text{MA}(s-2)$
F	Modelo estrutural básico	$\gamma_t = \mu_t + \gamma_t + \varepsilon_t$ onde μ_t é como em (2) e γ_t é como em (5a) ou (5b)	$\Delta_s = (1 - B)(1 - B^s)$	$\Delta_s \gamma_t \sim \text{MA}(s+1)$

Anexo III

• Exemplo de Aplicação do Filtro de Kalman

Consideremos o modelo em espaço de estados

$$\begin{aligned}y_t &= s_t + w_t + \varepsilon_t \\s_t &= 0.6s_{t-1} + 0.5w_{t-1} + u_t \\w_t &= 0.3s_{t-1} + 0.8w_{t-1} + v_t\end{aligned}$$

onde $\varepsilon_t \sim N(0, 1)$, $u_t \sim N(0, 2)$, $v_t \sim N(0, 3)$ e $\text{cov}(u_t, v_t) = 0.7$.

Dispomos das observações y_1, y_2, \dots, y_{t-1} e pretendemos obter \hat{y}_{T+1} .

Como o modelo 4.3.1 se encontra representado em espaços de estados, temos que

$$Z = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad T = \begin{bmatrix} 0.6 & 0.5 \\ 0.3 & 0.8 \end{bmatrix}, \quad R = \begin{bmatrix} 2 & 0.7 \\ 0.7 & 3 \end{bmatrix} \quad \text{E } h_t = 1.$$

Supondo que se tem $\hat{x}_{T-1|T-1} = \begin{bmatrix} 10 \\ 5 \end{bmatrix}$ e matriz de covariâncias estimadas $P_{T-1|T-1} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$.

• Previsão

Através do filtro de Kalman temos $\hat{x}_{T|T-1} = T\hat{x}_{T-1|T-1} = \begin{bmatrix} 0.6 & 0.5 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 10 \\ 5 \end{bmatrix} = \begin{bmatrix} 8.5 \\ 7 \end{bmatrix}$

e a matriz de covariância do erro é $P_{T|T-1} = TP_{T-1|T-1}T' + RR' = \begin{bmatrix} 2.670 & 1.555 \\ 1.555 & 4.314 \end{bmatrix}$.

O preditor a um passo é dado por $\hat{y}_T = Z\hat{x}_{T|T-1} = 15.5$.

Se dispomos, agora, de $y_T = 16$ então o erro de previsão é $\xi_T = y_T - \hat{y}_T = 0.5$.

• Actualização

Recorrendo às equações de actualização do filtro de Kalman, a estimativa actualizada é

$$\hat{x}_{T|T} = \hat{x}_{T|T-1} + K_T[y_T - Z\hat{x}_{T|T-1}] = \begin{bmatrix} 8.690 \\ 7.265 \end{bmatrix} \quad \text{em que } K_T = \begin{bmatrix} 0.380 \\ 0.529 \end{bmatrix}.$$

A matriz de covariâncias do erro é $P_{T|T} = \begin{bmatrix} 1.060 & -0.675 \\ -0.675 & 1.210 \end{bmatrix}$.

Finalmente, como pretendíamos $\hat{y}_{T+1} = Z\hat{x}_{T|T} = 16.75$.

Anexo IV

Da análise da representação gráfica da série (Figura 4.2), constatamos que a série Adultos apresenta tendência linear.

Ajustámos então uma tendência linear à série Adultos, efectuando uma regressão linear simples, em que a variável independente é o tempo, que permitiu obter os resultados sumariados na Tabela A.2.

	Estimativas	Erro Padrão	Estatística t	Valor- p
β_0	465,27	34,793	13,373	0,000
β_1	3,548	0,289	12,290	0,000

Tabela A.2 - Resultados da regressão linear simples

A partir dos resultados obtidos construímos a série Adultos sem tendência, cuja representação gráfica pode ser analisada na Figura A.1.

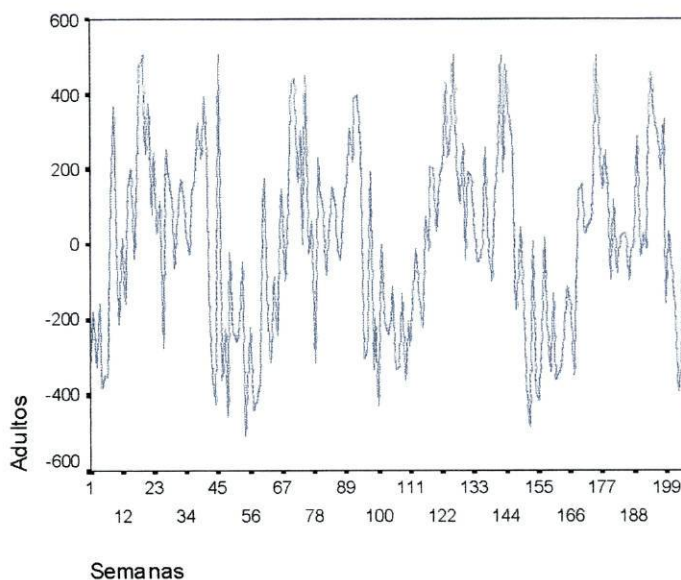


Figura A.1 - Cronograma da série Adultos sem tendência

Esta série foi posteriormente utilizada para examinar quais as periodicidades, se as houver, que regem a variação dos dados. Para isso calculámos o periodograma da série Adultos, depois de retirada a tendência, cuja representação gráfica se apresenta na Figura A.2. Como dispomos de 208 observações, o periodograma consta de 104 ordenadas.

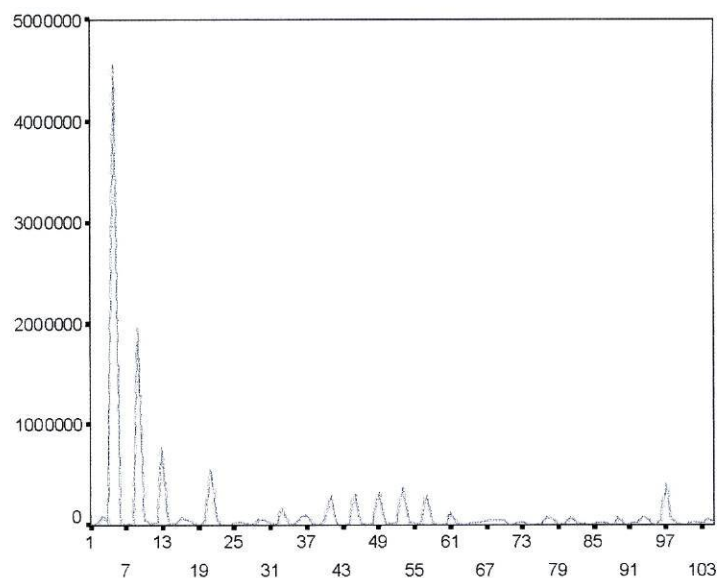


Figura A.2 - Periodograma da série Adultos sem tendência

Para avaliar a significância dos picos observados utilizamos os testes de Walker e de Whittle, cujos resultados apresentamos na Tabela A.3.

Período	Frequência	Periodograma	g^*	Valor- p
52	0,12083	4564622	73,96084	0,000
26	0,241661	1954216	0,236232	0,000
17,33333	0,362491	770623,1	0,121968	0,000
10,4	0,604152	549706,6	0,099089	0,003
2,166667	2,899932	414140,7	0,082863	0,019
4	1,570796	376889,7	0,082223	0,022
4,333333	1,449966	337759,4	0,080287	0,029
4,727273	1,329135	315144,1	0,081451	0,028
5,2	1,208305	303394,5	0,085368	0,020
3,714286	1,691627	297577,9	0,091546	0,011

Tabela A.3 - Resultados da regressão linear simples

Verificamos que a série apresenta várias componentes periódicas bastante significativas, sendo algumas delas múltiplas de outras. Destacam-se, no entanto, as correspondentes aos períodos 52, 26, 17 e 10 semanas.

Anexo V

<i>S</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
	-42,4	-343,1	-343,3	-154,4	-379,6	-307,7	-310,2	-64,1	25,0	-158,6	-285,3	-33,7	-33,3
<i>S</i>	14	15	16	17	18	19	20	21	22	23	24	25	26
	160,1	131,9	0,9	121,6	263,6	458,8	222,5	245,9	254,0	258,9	-1,1	151,2	-139,6
<i>S</i>	27	28	29	30	31	32	33	34	35	36	37	38	39
	221,8	157,3	82,9	-69,1	23,2	102,6	156,5	-3,5	-57,2	122,8	198,6	378,9	212,5
<i>S</i>	40	41	42	43	44	45	46	47	48	49	50	51	52
	418,3	315,0	134,0	-275,9	-257,1	251,4	-284,2	-264,1	-452,5	-1,6	-275,3	-300,5	-233,0

Tabela A.4 - Estimativas normalizadas, \hat{S}_i , dos índices sazonais

Anexo VI

ARMA (p, q)	$\hat{\sigma}_e^2$	AIC	BIC	Parâmetros Significativos*
(0, 1)	42795,1	2797,3	2803,9	1
(0, 2)	42882,5	2798,7	2808,7	1
(0, 3)	43090,6	2800,7	2814,1	2
(1, 1)	42886,2	2798,8	2808,8	1
(1, 2)	41209,6	2794,0	2807,3	2
(1, 3)	42205,2	2797,5	2814,2	3
(2, 3)	42170,7	2801,3	2821,3	2
(3, 3)	40705,8	2792,2	2815,5	3
(3, 0)	40235,0	2785,7	2795,7	3
(3, 1)	40528,5	2789,2	2805,9	3
(3, 2)	40509,6	2790,2	2810,2	3

* Apresentamos alguns dos modelos com pelo menos um parâmetro significativo

Tabela A.5 - Modelos ARMA testados

Anexo VII

		Testes	
		AIC	BIC
Sazona- lidade	Muda	9,894	10,793
	Trigonométrica	9,849	10,748

Tabela A.6 - Tipos de sazonalidade testados

Anexo VIII

• Normalidade

a) Histograma dos Resíduos

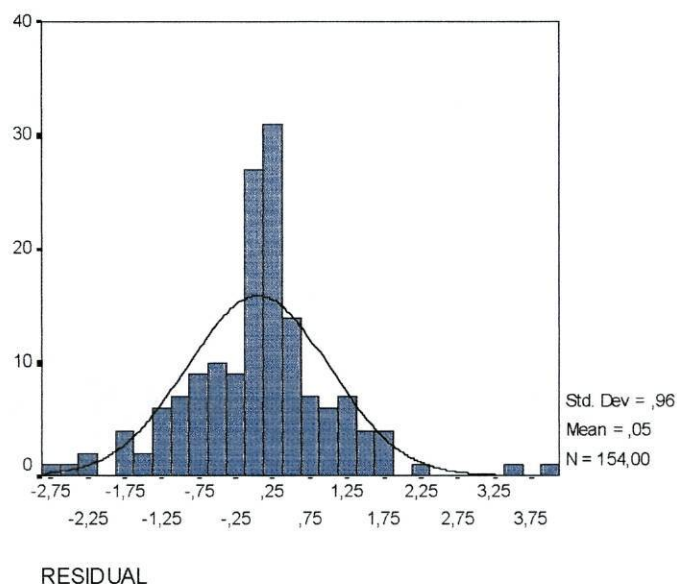


Figura A.3 - Histograma da componente residual

b) Representação em Papel de Probabilidades

Uma das formas mais evidentes para averiguar a normalidade dos resíduos é fazer a sua representação em papel de probabilidades.

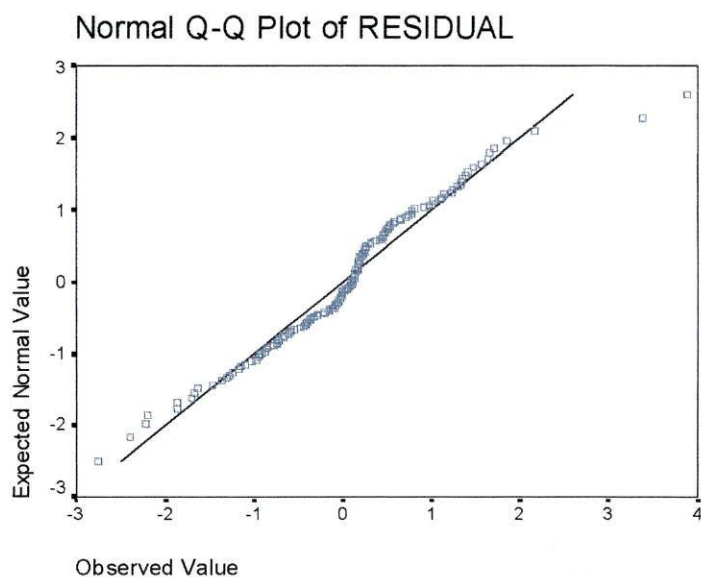


Figura A.4 - Representação em papel de probabilidades da componente residual

- **Variância Constante e Independência**

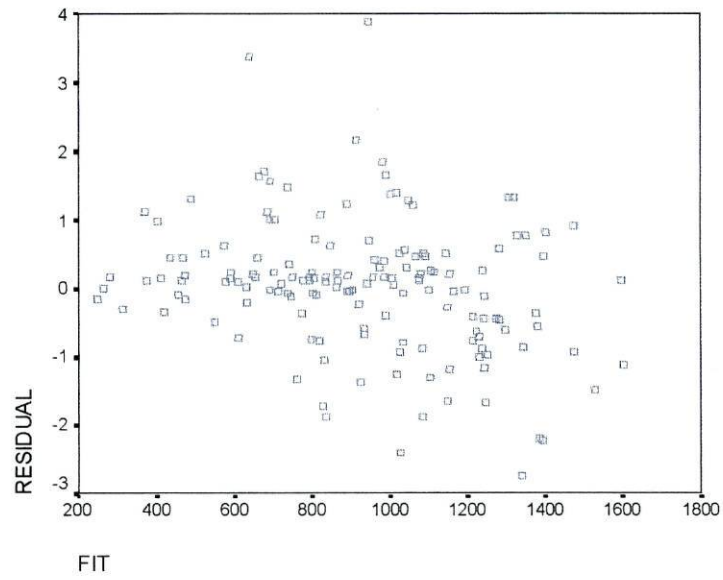


Figura A.5 - Componente residual vs. valores estimados