



NOVA
NOVA SCHOOL OF
SCIENCE & TECHNOLOGY

DEPARTAMENTO
DE MATEMÁTICA

MARIA QUIRINO ROSA
Licenciada em Matemática

PREVISÕES USANDO SÉRIES TEMPORAIS E MODELOS DE REGRESSÃO PARA CRIAÇÃO DE POLÍTICAS DE GESTÃO DE STOCKS

MESTRADO EM MATEMÁTICA E APLICAÇÕES

Universidade NOVA de Lisboa
Novembro, 2021



PREVISÕES USANDO SÉRIES TEMPORAIS E MODELOS DE REGRESSÃO PARA CRIAÇÃO DE POLÍTICAS DE GESTÃO DE STOCKS

MARIA QUIRINO ROSA

Licenciada em Matemática

Orientador: Nelson Fernando Chibeles Pereira Martins
Professor, Universidade NOVA de Lisboa

Coorientadora: Maria Isabel Azevedo Gomes
Professora, Universidade NOVA de Lisboa

Previsões usando séries temporais e modelos de regressão para criação de políticas de gestão de stocks

Copyright © Maria Quirino Rosa, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

AGRADECIMENTOS

Gostaria de agradecer em primeiro lugar aos meus orientadores, o professor Nelson Chibeles e a professora Maria Isabel Gomes, pelo apoio e acompanhamento contínuo durante a elaboração da dissertação.

Gostaria também de agradecer às empresas Processware e ANA Aeroportos, pela disponibilização dos dados necessários para desenvolvimento da dissertação.

Agradeço ainda aos meus familiares, namorado e amigos.

RESUMO

Durante 2020 e 2021, vários setores foram fortemente impactados pela pandemia, incluindo a indústria aeronáutica. Surge então a necessidade de reforçar toda a estrutura de empresas que contribuem para o bom funcionamento dos aeroportos. Nomeadamente a otimização de processos da cadeia de abastecimento nas empresas fornecedoras dos aeroportos. A gestão de stocks do fornecimento de catering do aeroporto de Lisboa não é exceção. Atualmente são reportados elevados valores de rotura de stock, devido à gestão manual existente.

Este trabalho tem como objetivo modelar o comportamento da procura através do número de passageiros do aeroporto, para construir políticas de gestão de armazéns. Surge assim a oportunidade de estudar o comportamento da procura de material e averiguar se este é afetado pelo tráfego aéreo.

Para alcançar os objetivos, recorreu-se à análise de séries temporais e a métodos de previsão através de modelos de regressão, uma aplicação de *machine learning*. Tendo como ponto de partida as séries temporais, é encontrada uma relação entre duas variáveis ao longo do tempo, o fluxo de passageiros no aeroporto e a procura de matéria prima. São propostos vários modelos de regressão, incluindo o modelo de regressão linear por troços, para prever a procura, tendo como variável explicativa o número de passageiros do aeroporto de Lisboa. Através desta previsão, estudaram-se duas políticas de gestão de stocks. Para tal, foram estimados os custos associados à cadeia de abastecimento. A primeira política é um modelo determinístico, onde a procura é assumida como constante. A segunda tem em conta a variabilidade da procura e a aplicação do algoritmo de Hadley-Whitin.

Os resultados mostram como a procura de material está fortemente relacionada com o número de passageiros do aeroporto de Lisboa. Esta relação é descrita no modelo de regressão linear por troços. A construção de uma política de gestão de stocks que considera a variabilidade da procura e admite um stock de segurança considerável descreve uma solução adequada para a resolução do problema.

Palavras-chave: Séries temporais, Machine Learning, Regressão linear por troços, Gestão de stocks, Algoritmo de Hadley-Whitin

ABSTRACT

During 2020 and 2021, several sectors were heavily impacted by the pandemic, including the aviation industry. There is a need to strengthen the entire structure of companies which contribute to the proper functioning of airports. In particular, the optimization of supply chain processes at airport supplier companies. Lisbon airport catering supply stock management is no exception. Currently, high of stock-out values are reported, due to the existing manual management.

This work aims to model demand behaviour through the number of passengers at Lisbon airport, in order to build warehouse management policies. This gives rise to the opportunity to study demand behaviour and find out if it is affected by air traffic.

To achieve these objectives, time series analysis and prediction methods were used through regression models, an application of machine learning. Taking time series as a starting point, a relationship between the two variables over time is found. Several regression models are proposed, including the linear regression model by sections, to forecast demand, using the number of passengers at Lisbon airport as the explanatory variable. Through this forecast, two stock management policies were studied. To this end, the costs associated with supply chain were estimated. The first policy is a deterministic model, where demand is assumed to be constant. The second takes into account the variability of demand and the application of the Hadley-Within algorithm.

The results show how demand is strongly related to the number of passengers at Lisbon airport. This relationship is described in the piece-wise linear regression model. A stock management policy which considers the variability of demand and admits a considerable safety stock describes an adequate solution to solve the problem.

Keywords: Time series, Machine learning, Piece-wise linear regression, Inventory control, Hadley-Within algorithm

ÍNDICE

Índice de Figuras	xv
Índice de Tabelas	xvii
Siglas	xix
1 Introdução	1
1.1 Âmbito e motivação	1
1.2 Objetivos e organização da dissertação	2
2 Revisão da literatura	5
2.1 Previsão e séries temporais	5
2.2 Modelos de regressão	6
2.3 Gestão de stocks	7
3 Metodologias	11
3.1 Análise de séries temporais	11
3.1.1 Definição	11
3.1.2 Processo de análise	11
3.2 Modelação através de regressão	14
3.2.1 Definição	14
3.2.2 Avaliação de modelos de regressão	15
3.2.3 Regressão linear	16
3.2.4 Regressão linear Piece-wise	16
3.2.5 Regressão linear generalizada	17
3.2.6 Regressão quadrática	17
3.2.7 Regressão logarítmica	18
3.3 Gestão de stocks	19
3.3.1 Definição	19
3.3.2 Processo de construção	19

3.3.3	Modelo determinístico	20
3.3.4	Modelo de revisão periódica para procura variável	21
4	Caso de estudo	23
4.1	Caracterização do problema	23
4.2	Caracterização dos dados	25
4.2.1	Dados cedidos por ProcessWare	25
4.2.2	Dados cedidos por ANA Aeroportos	27
4.2.3	Dados finais	28
5	Resultados e discussão	29
5.1	Análise da série temporal da procura	29
5.1.1	Agregação dos dados semanalmente	30
5.2	Análise da série temporal de passageiros	31
5.3	Modelação da procura	34
5.3.1	Agregação dos dados semanalmente	35
5.3.2	Dados diários	41
5.3.3	Resultados finais	46
5.4	Políticas de gestão de stocks	47
5.4.1	Modelo determinístico	48
5.4.2	Modelo de revisão periódica	50
6	Conclusões	53
6.1	Conclusões	53
6.2	Desenvolvimentos futuros	55
	Referências Bibliográficas	57

ÍNDICE DE FIGURAS

1.1	Abordagem e organização da tese	2
2.1	Diagrama da cadeia de abastecimento e distribuição de produto	7
3.1	Série temporal com sazonalidade semanal	12
3.2	Ajuste a uma regressão linear	16
3.3	Ajuste a uma regressão quadrática	18
3.4	Ajuste a uma regressão logarítmica	18
4.1	Nível de stock ao longo dos meses	25
5.1	Série temporal da procura de material de setembro de 2020 a julho de 2021	29
5.2	Autocorrelação da procura ao longo do tempo	30
5.3	Série temporal da procura semanal de material	31
5.4	Série temporal do número de PEAL de setembro de 2020 a julho de 2021 .	32
5.5	Autocorrelação do número de PEAL	32
5.6	Desvio de valores obtidos entre sexta-feiras consecutivas, sábados e domingos, respetivamente	33
5.7	Autocorrelação do número de PEAL, excluindo os fins de semana	33
5.8	Relação entre a procura mensal e PEAL mensal	34
5.9	Gráfico de dispersão entre a procura semanal e número de PEAL semanal .	35
5.10	Ajuste do segundo método de regressão linear Piece-wise	37
5.11	Histograma de D quando $x < 54712$	38
5.12	Gráfico de dispersão entre a procura diária e o número de PEAL diário . .	41
5.13	Ajuste do 2º método de regressão linear Piece-wise	43
5.14	Histograma de D, quando $x < 8579$	44
5.15	Histograma de D, quando $x \geq 8579$	44

ÍNDICE DE TABELAS

4.1	Cinco observações dos dados de entradas no armazém	25
4.2	Cinco observações dos dados de inventário	26
4.3	Cinco observações dos dados de consumo de produção	26
4.4	Dez observações da cronologia do processo	27
4.5	Primeiras observações dos dados ANA Aeroportos	27
4.6	Dez observações da compilação final	28
5.1	Compilação de resultados dos quatro melhores modelos	46
5.2	Distribuição do número de PEAL pelo mês de julho de 2021	47
5.3	Previsão da procura durante o mês de julho de 2021	47
5.4	Valores ótimos do modelo determinístico	49

SIGLAS

DSGE Dynamic Stochastic General Equilibrium

IATA International Air Transport Association

INE Instituto Nacional de Estatística

PEAL Passageiros embarcados no aeroporto de Lisboa

INTRODUÇÃO

1.1 Âmbito e motivação

Atualmente o setor de aviação tem sofrido grandes alterações. O impacto da pandemia entre 2020 e 2021 foi especialmente drástico neste ramo de negócio. De acordo com um artigo publicado pela *International Air Transport Association (IATA)* a 28 de janeiro de 2021, estima-se que cerca de 4.8 milhões de postos de trabalho ligados diretamente à aviação serão extintos até ao início de 2022, o que traduz uma redução de 43% em comparação com a situação pré-COVID, [1]. Estes incluem postos de trabalho em companhias aéreas, aeroportos, fabricantes e gestores de tráfego aéreo. É necessária uma enorme capacidade de adaptação por parte dos aeroportos e empresas abastecedoras para que o bom funcionamento dos mesmos prevaleça. A sua logística e gestão de cadeias de abastecimento são os principais problemas do tráfego aéreo, seja este de passageiros ou de carga, [2]. Toda a logística de abastecimento está dependente de rigorosas previsões de tráfego aéreo para prever as quantidades de matéria prima a encomendar de modo a abastecer os aviões que passam nos aeroportos.

A gestão e controlo de inventário têm sido tópicos de grande relevância ao longo dos anos, uma vez que a gestão eficiente das cadeias de abastecimento minimiza custos em todos os processos envolventes. A principal causa do aumento constante desses custos são os níveis excessivos de stock nos armazéns, [3]. Por isso é essencial haver rigorosas previsões da procura de mercado.

Esta dissertação surge no âmbito de um estágio curricular realizado na Processware, uma empresa de software portuguesa que soluciona problemas operacionais em diversos setores tais como indústria de manufatura, catering e indústria agrícola, [4]. Um dos seus clientes é a principal empresa abastecedora do aeroporto de Lisboa, Cateringpor. Atualmente a gestão de matéria prima da Cateringpor é auxiliada pela Processware. Durante o ano de 2020, a gestão de armazéns e inventário tem sido condicionada, resultando em grandes roturas de stock.

1.2 Objetivos e organização da dissertação

O objetivo desta dissertação é construir uma política de gestão de stocks, através da modelação da procura. De modo a sugerir uma política de gestão de stocks baseada na procura de matéria prima, é necessário conhecer e poder prever o comportamento da mesma. Sendo este um material utilizado na produção de bens alimentares abastecidos ao aeroporto de Lisboa e respetivos aviões que por ele passam, é preciso averiguar se a sua procura se relaciona com o número de passageiros embarcados no aeroporto de Lisboa.

Para atingir estes objetivos, o primeiro passo consiste em estudar o comportamento da procura do material através de séries temporais. Deve-se averiguar se as observações da série estão relacionadas ao longo do tempo e estudar possíveis padrões. Depois, averiguar se existe alguma relação entre a procura e o tráfego aéreo do aeroporto de Lisboa. Deve-se estudar quais as variáveis explicativas da procura, por exemplo, se a procura do material se relaciona com o número de Passageiros embarcados no aeroporto de Lisboa (PEAL). É ainda necessário criar um modelo de previsão da procura. Paralelamente, estimar custos associados à cadeia de abastecimento e construir um modelo de gestão de stocks adequado. A partir deles, é possível prever a procura e níveis de stock para o período pretendido. A figura 1.1 ilustra esta abordagem.

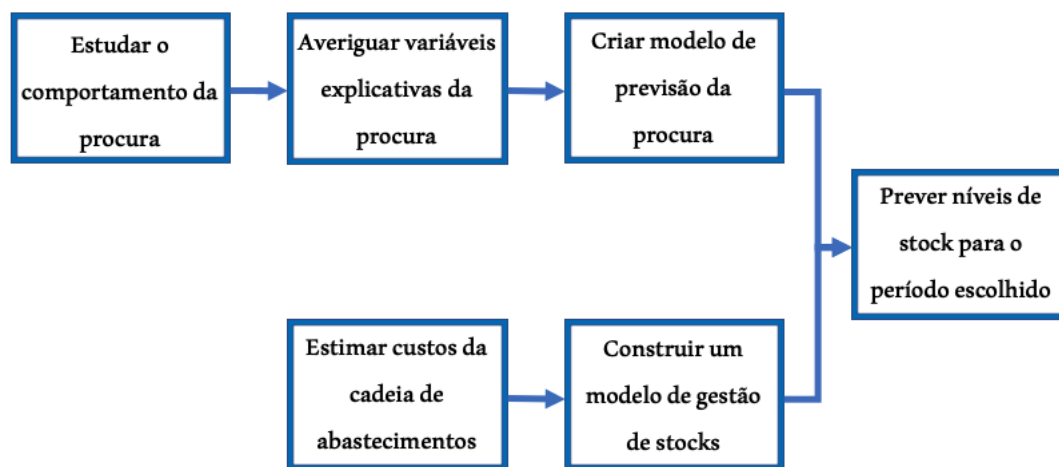


Figura 1.1: Abordagem e organização da tese

Excluindo a introdução, a dissertação está organizada em cinco fases de trabalho diferentes. A primeira fase diz respeito à Revisão da Literatura. A Revisão da Literatura está organizada pelos temas que serão abordados ao longo da dissertação. Estes serão aplicados ao caso de estudo e os seus métodos matemáticos serão detalhados na metodologia. O primeiro tema é previsão e séries temporais. O segundo e terceiro temas abordam modelos de regressão e gestão de stocks, respetivamente.

A segunda fase de trabalho descreve as metodologias necessárias para a resolução do problema. Esta secção apresenta metodologias para análise de séries temporais, modelação através de regressão e gestão de stocks. Para além da descrição dos processos necessários, são ainda apresentados conceitos e modelos matemáticos, respetivas equações e introdução de variáveis.

A terceira fase de trabalho descreve o caso de estudo, caracterizando o problema e os dados concedidos por duas entidades – Processware e ANA Aeroportos. As abordagens de resolução do problema também são descritas nesta secção.

A quarta fase apresenta os resultados obtidos e a sua análise. Estes estão organizados na análise das séries temporais da procura e passageiros, na modelação da procura e nas políticas de gestão de stocks.

A última fase resume todo o trabalho desenvolvido, dá ênfase aos resultados mais importantes e propõe futuros desenvolvimentos. Esta informação pode ser encontrada na secção de conclusões.

REVISÃO DA LITERATURA

2.1 Previsão e séries temporais

Ao longo do tempo a humanidade tem demonstrado interesse em conseguir prever vários fenómenos, tais como o comportamento dos animais ou a relação entre o Sol e a Terra, [5]. Estas previsões eram baseadas em observação e experiência. Assim surgiram leis e regras baseadas nestes conceitos que asseguravam um maior rigor das previsões. A previsibilidade de um certo evento pode depender de vários fatores, tais como: a natureza de fatores externos que influenciam a variável resposta, o número de observações e a diversidade da amostra, [6]. O conceito de fazer prognósticos baseados em observações históricas designa-se por previsão de séries temporais, sendo que um dos seus grandes desafios é escolher o período temporal para realizar o estudo.

Os estudos podem ser classificados segundo a sua hierarquização temporal - previsões a curto, médio e longo prazo. Previsões a longo prazo podem ser aplicadas a problemas de decisão macro, onde é necessário conhecer, por exemplo, o mercado global e ter acesso a informação macro da área de negócio. Em comparação, as previsões a curto prazo são utilizadas quando é necessário maior rigor, como é o caso da previsão diária de energia eólica gerada em parques eólicos na China, [7]. Previsões a curto prazo são frequentemente utilizadas para prever a procura de determinados processos, sendo estes logísticos ou financeiros. É necessária a utilização de métodos estatísticos rigorosos que, através de dados recolhidos durante um período prévio, estimam tais previsões. Assim, previsões a curto prazo tendem a ser mais exatas e fidedignas quando comparadas com previsões a longo prazo, [8]. Dependendo da natureza dos dados do problema, as previsões podem ser realizadas com recurso a diferentes métodos estatísticos, como redes neuronais artificiais ou *machine learning*, [9].

Existem três grandes objetivos para a utilização de séries temporais: analisar o comportamento da amostra, prever comportamentos futuros e controlar o resultado de determinados processos para que se torne mais favorável. As séries temporais podem ser influenciadas por diversos fatores exteriores, como variáveis aleatórias explicativas. Alguns dos grandes desafios de trabalhar com séries temporais são determinar o número de variáveis explicativas, o período temporal a considerar e a interpretação dos resultados obtidos, [10].

2.2 Modelos de regressão

Os modelos preditivos baseiam-se na existência de uma variável de estudo - variável resposta - e respetivas variáveis explicativas - preditores, [11]. Ou seja, o objetivo é prever valores da variável resposta, tendo em conta que existe uma relação entre esta variável e respetivas variáveis explicativas. Ao longo dos anos, a análise de regressão tem-se estabelecido como uma ferramenta poderosa que pode ser aplicada a várias áreas de negócio, [12].

Os algoritmos de regressão são considerados métodos supervisionados de *machine learning*, um ramo de Inteligência Artificial, [13]. *Machine learning* inclui vários tipos de aprendizagem. As aprendizagens supervisionadas visam construir algoritmos através de dados de experiências passadas, onde estão presentes as variáveis independentes e dependentes. Através deles, é possível perceber a relação entre as variáveis e a sua significância, [14]. Para além dos métodos de supervisão, existem também os métodos não supervisionados. Estes também requerem uma base de dados inicial, no entanto os seus algoritmos correm sem supervisão humana. Não existe controlo humano, uma vez que os algoritmos apresentam a capacidade de criar e memorizar padrões, através de tentativa e erro, de modo a atingir o melhor resultado para o objetivo proposto, [14]. *Machine learning* engloba outros tipos de aprendizagem como aprendizagem por reforço e redução de dimensionalidade, entre outros, [15].

Dependendo das relações existentes entre a variável resposta e possíveis variáveis explicativas, podem ser utilizados modelos de relação linear entre as variáveis resposta e explicativa, é construído um modelo de regressão linear ou multilinear, caso a variável resposta se relacione com mais que um preditor, [16]. Muito embora os modelos lineares sejam menos complexos, o seu ajuste por vezes não é o ideal. Quando a relação entre as variáveis não é linear, existem outras alternativas como modelos lineares generalizados, análises de regressão quadráticas, logarítmicas, entre outras. Por vezes, estes modelos podem ser mais adequados, uma vez que proporcionam um ajuste de valores mais preciso. Contudo, a sua complexidade é muito maior, o que implica uma maior complexidade do problema de programação matemática.

Um bom compromisso entre a complexidade e a precisão dos modelos é conseguido através de regressão linear piece-wise. Este tipo de regressão consiste em avaliar a existência de diferentes tendências e padrões entre as variáveis resposta e explicativa e separar

em partições, através de pontos de rotura, [17]. Assim, admitindo que existem n relações lineares entre as variáveis resposta e explicativa separadas por $n - 1$ pontos de rotura, um modelo de regressão linear piece-wise constrói n regressões lineares que se situam em n espaços diferentes. Este modelo apresenta maior precisão quando comparado com modelos lineares, uma vez que estabelece relações lineares a conjuntos de amostras com menor número de observações.

Um artigo publicado em 2016 pela revista *Expert Systems With Applications* descreve a sua aplicação a sete bases de dados de problemas reais entre os quais o desempenho hidrodinâmico das velas de iates, a resistência da compressão do betão, a eficiência energética de edifícios e o impacto do design das lâminas de aerofólio e da velocidade do vento no nível de pressão sonora, [18]. A vasta variedade da natureza e o elevado número de observações dos dados permitem afirmar que este método pode ser aplicado a variados problemas com eficácia.

2.3 Gestão de stocks

O tema de gestão de stocks e controlo de inventário tem sido alvo de grande e rápido desenvolvimento ao longo dos anos. O sucesso de muitas empresas e fábricas depende da capacidade e excelência logística. O seu grande objetivo é disponibilizar produtos com qualidade a um preço acessível e cumprindo com os prazos estabelecidos, [19]. Gestão de stocks é a atividade responsável por organizar a disponibilidade dos produtos aos seus consumidores, que inclui todo o processo da cadeia de abastecimento, [20]. Esta é responsável pela gestão de fornecedores, organização de encomendas e seus custos e transporte dos produtos até ao consumidor, [21]. Os seus grandes objetivos incluem compatibilizar os processos de abastecimento e de procura do produto, minimizar o risco de falha de abastecimento e minimizar os custos relacionados com a cadeia de abastecimento. A figura 2.1 mostra todo o processo da cadeia de abastecimento e distribuição de produto. É de notar que esta dissertação centra-se essencialmente na primeira parte do diagrama, referente ao processo de fornecimento das matérias primas e sua manufatura.

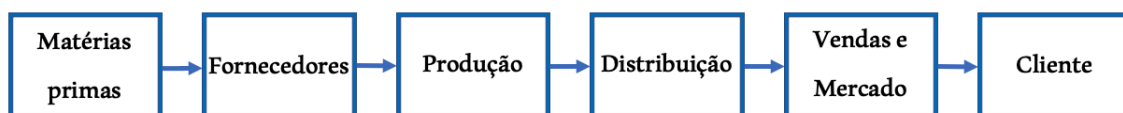


Figura 2.1: Diagrama da cadeia de abastecimento e distribuição de produto

Os custos associados à cadeia de abastecimento correspondem a cerca de 55% do custo total do produto, [22]. Estes incluem o custo de processamento da encomenda, custos de transporte, custos de inventário e custos associados à procura do produto, [23]. Assim, o controlo de inventário é muito vantajoso para as empresas, uma vez que uma boa política de gestão de stocks proporciona uma redução de custos para as empresas.

Vários estudos têm surgido na área de gestão de stocks para alcançar uma maior eficiência de gestão de recursos e otimizar os processos de armazenamento de matérias, [24–26]. A vasta natureza dos problemas e do tipo de indústria a que se referem promoveram o estudo de diferentes modelos. Um estudo publicado em 2016, [27], realça a importância das estratégias de controlo de inventário e o seu impacto em diferentes setores comerciais como a indústria de manufatura.

Se as ordens de encomenda têm como objetivo manter o nível de stock existente em armazém constante ao longo do tempo, de modo a manter um equilíbrio no stock existente em armazém, então a quantidade a encomendar é frequentemente fixa, [28]. Este tipo de política é designado por política do nível de stock ótimo (*re-order level policy*) e consiste em determinar o período ideal entre encomendas. Por outro lado, se as encomendas são feitas regularmente e o seu tamanho é variável devido a uma grande variabilidade dos níveis de stock dos armazéns resultante de uma procura variável, então este tipo de política designa-se por política do ciclo de stock ótimo (*re-order cycle policy*).

Independentemente do tipo de política de gestão de stocks, é sempre necessário conhecer a natureza e características da procura. Quando a procura é fixa – constante – ao longo do tempo, então o modelo de gestão de stocks é possível recorrer a modelos determinísticos para construir a melhor política possível. Nos modelos determinísticos, são utilizados lotes de quantidade económica, que determinam a quantidade ideal de matéria prima a encomendar. Os parâmetros são conhecidos e as variáveis possíveis de calcular. Quando a procura é aleatória, são utilizados modelos probabilísticos ou estocásticos, [29].

Quando a variabilidade da procura é considerada, então é necessário estudar o seu comportamento estatístico, [30]. Vários autores estudaram modelos e algoritmos para encontrar equilíbrio estocástico em problemas de procura variável, [31, 32]. Este equilíbrio, também conhecido por Dynamic Stochastic General Equilibrium (DSGE), é frequentemente utilizado em modelos clássicos de macroeconomia, [33]. Na área de gestão de stocks, sendo o objetivo encontrar o equilíbrio entre a procura do produto e o abastecimento da sua matéria prima, é importante reduzir a complexidade dos modelos.

Por isso, sugere-se que a procura tome um comportamento seguindo a distribuição Normal. O objetivo desta generalização do comportamento da procura é facilitar a determinação dos melhores pontos de reencomenda, [34]. Esta aproximação foi considerada a melhor aproximação contínua e, de modo geral, é através dela que os modelos estocásticos apresentam melhores resultados.

Um artigo em particular sugere uma interessante abordagem ao assumir um comportamento Gamma da procura e avaliar se este modelo é tão eficaz e preciso quando é assumido que a procura tem comportamento Normal, [35]. Estudando o comportamento da procura de uma determinada amostra e após verificação que esta é, de facto, proveniente de uma distribuição Gamma, faz sentido averiguar se um modelo construído para uma procura com esta particularidade apresenta maior rigor. Este modelo foi baseado nos modelos existentes para procura Normal, sendo por isso considerado uma extensão do anterior. Este modelo considera procuras mais complexas. Assim, os seus resultados

são positivos, uma vez que apresenta resultados ligeiramente melhores que os obtidos através dos modelos tradicionais. Contudo, a grande limitação é a sua complexidade. Não só a sua implementação é mais lenta como também é composto por muitas restrições.

Existem duas formas de abordar a variabilidade da procura em modelos de gestão de stocks: sistemas de revisão contínua e sistemas de revisão periódica. O sistema mais comum e de onde grande parte da investigação de controlo de stocks é proveniente é o sistema de revisão periódica, [36]. Isto deve-se ao facto de grande parte dos sistemas de stock serem controlados diária ou semanalmente. Contudo, sistemas de revisão contínua têm-se tornado cada vez mais recorrentes devido à automação da gestão de armazéns, [37].

A grande vantagem dos sistemas de revisão periódica é a sua simplicidade de implementação. Tem-se como hipótese que o nível de stock é verificado em intervalos de tempo regulares, [38]. O tamanho da encomenda varia para cada ciclo e é definido com auxílio a um nível de stock máximo em armazém. O maior desafio apresentado por este tipo de sistema é a complexidade associada ao comportamento da procura. Por vezes não é possível ajustar uma distribuição de probabilidade conhecida e, por outras, o comportamento dessa distribuição não torna exequível a implementação do modelo. Tome-se como exemplo a distribuição Gamma. É de notar que quando o número de observações aumenta, a distribuição Gamma tende para a normalidade. Assim, quando o número de observações da amostra é elevado, assume-se que o nível de reposição tem um comportamento Normal, como descreve Burgin [39].

O algoritmo de Hadley-Whitin proposto em 1963 pelos autores Hadley e Whitin, [40] soluciona esse problema, uma vez que produz soluções ótimas do custo e tempo de ciclo, assumindo que a procura tem um comportamento Normal. Alguns autores sugeriram diferentes métodos de gestão de inventário que têm como base o algoritmo de Hadley-Whitin, [41]. Este é recorrentemente aplicado a modelos de revisão periódica, [42].

3.1 Análise de séries temporais

3.1.1 Definição

Uma série temporal consiste numa amostra de observações extraída e organizada ao longo do tempo e pode ser decomposta nas suas componentes sazonais, cíclicas e restantes componentes. Assumindo uma decomposição aditiva, descreve-se uma série temporal da seguinte forma:

$$y_t = S_t + T_t + R_t \quad (3.1)$$

onde

y_t – observação t da amostra em estudo

S_t – componente sazonal

T_t – componente cíclica

R_t – restantes componentes

Caso haja uma relação de proporcionalidade entre a variável em estudo e alguma das suas componentes, pode também dar-se uma decomposição multiplicativa:

$$y_t = S_t \times T_t \times R_t \quad (3.2)$$

3.1.2 Processo de análise

O processo de análise de uma série temporal foi compilado em quatro passos.

1. Identificar a existência de possíveis padrões e ruído branco
2. Descrever o comportamento da série e contextualizá-lo
3. Identificar possíveis sazonalidades
4. Identificar possível estacionariedade

O padrão de uma série temporal corresponde a uma característica que se repete ao longo do tempo. Este pode ser sazonal, cíclico, entre outros. Se uma série temporal é afetada por fatores sazonais, como a estação do ano ou dia da semana, então existe um padrão sazonal, isto é o padrão dá-se com uma frequência fixa. Por outro lado, se esse padrão ocorre de acordo com uma frequência variável, então tem-se um padrão cíclico. Logo a grande diferença entre padrão sazonal e cíclico é a sua frequência.

Para além do padrão, existem também oscilações aleatórias, a que se dá o nome de ruído branco. Em situações que se torna difícil identificar padrões é por vezes necessário filtrar este ruído branco, para que os padrões emirjam, [43].

As sazonalidades presentes numa série temporal são descritas como dependências correlacionais de ordem k , entre os elementos i e $(i - k)$ da série. Assim, recorre-se à autocorrelação para identificar as sazonalidades presentes. Valores de autocorrelação mais próximos de um descrevem uma forte relação entre observações. Valores próximos de zero mostram a ausência de correlação ao longo do tempo. Dependências correlacionais de diferentes ordens representam diferentes tipos de sazonalidade. Ao analisar uma série diária, caso a correlação de ordem 7 seja elevada, pode concluir-se que esta série temporal tem uma sazonalidade semanal. A figura 3.1 ilustra uma série temporal com sazonalidade semanal com correlação de ordem 7 de 1.

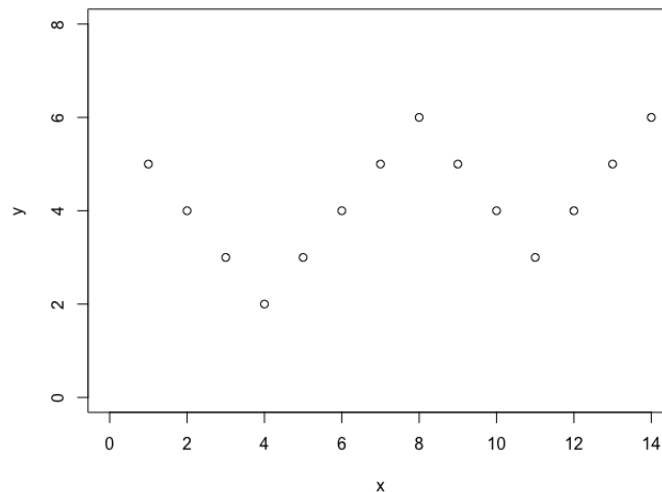


Figura 3.1: Série temporal com sazonalidade semanal

Considera-se que existe tendência numa série temporal quando se dá um crescimento ou decréscimo nos valores das observações ao longo do tempo, [6]. Esta tendência não é necessariamente linear. Ou seja, podem existir séries temporais com tendência e padrões que se repetem ao longo do tempo.

A existência de estacionariedade pode ser identificada através da utilização do teste estatístico de Dickey-Fuller. Este tem como hipótese nula a existência de uma raiz unitária. Seja $\{y_t\}_{t \in \mathbb{Z}}$ uma série temporal tal que:

$$y_t = \phi y_{t-1} + \epsilon_t \quad (3.3)$$

onde

y_t – observação t da amostra em estudo

ϵ_t – erro da observação t

Considere-se $H_0: |\phi| < 1$, então a série temporal é estacionária. Utilize-se a estatística

$$\hat{\tau} = \frac{\hat{\phi} - 1}{s(\hat{\phi})} \quad (3.4)$$

onde $\hat{\phi}$ é o estimador de mínimos quadrados, $s(\hat{\phi})$ o desvio padrão de $\hat{\phi}$ e S^2 o estimador de σ^2 .

$$\hat{\phi} = \frac{\sum_{t=2}^n y_{t-1} y_t}{\sum_{t=2}^n y_{t-1}^2} \quad (3.5)$$

$$s(\hat{\phi}) = \frac{S}{(\sum_{t=2}^n y_{t-1}^2)^{0.5}} \quad (3.6)$$

$$S = \frac{1}{n-2} \sum_{t=2}^n (y_t - \hat{\phi} y_{t-1})^2 \quad (3.7)$$

Caso o p-value obtido seja inferior a 0.05, é possível rejeitar a hipótese de que a série temporal apresenta estacionariedade, com um nível de confiança de 95%, [44].

3.2 Modelação através de regressão

3.2.1 Definição

Baseado em métodos estatísticos, um modelo de regressão estima a relação entre a variável resposta e o(s) seu(s) preditor(es). Identifiquem-se primeiramente as variáveis resposta e explicativa. A variável que cujos valores se pretendem estimar é a procura do material, sendo por isso designada por variável resposta ou dependente. A variável que possivelmente se relaciona com a variável resposta e cujo comportamento é conhecido é o número de PEAL, sendo por isso designada por variável explicativa ou independente.

Independente do tipo de regressão, é de notar que o objetivo é conhecer os parâmetros ou coeficientes do modelo que melhor descrevem a relação entre duas ou mais variáveis, [16]. Por isso, todo o método de regressão segue tipicamente esta estrutura:

1. Identificar a variável resposta, Y
2. Identificar o conjunto de possíveis variáveis explicativas que influenciam a variável resposta
3. Escolher uma amostra de Y e correspondentes valores das variáveis explicativas
4. Estimar os parâmetros da regressão
5. Testar as hipóteses efetuadas
6. Estimar valores da variável resposta

3.2.2 Avaliação de modelos de regressão

Existem vários métodos e métricas de avaliação de modelos de regressão. Uma das mais importantes é a medida R^2 . Esta medida calcula a percentagem de variabilidade da variável resposta que é explicada pelo modelo, sendo por isso uma medida relativa. Modelos cujo valor de R^2 é inferior a 0.7 não são considerados bons ajustes. Considerem-se bons ajustes os modelos cujo valor de R^2 varia entre 0.7 e 0.8. Os melhores modelos são aqueles que apresentam valores de R^2 próximos de 1. Logo, quanto maior o valor de R^2 , melhor é o modelo.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.8)$$

onde

y_i – i-ésima observação da variável resposta

\hat{y}_i – i-ésimo estimador da variável resposta

\bar{y} – valor médio da variável resposta

n – número de observações

A segunda medida de avaliação é o erro quadrático médio (MSE, *Mean squared error*). Este representa a média dos quadrados dos desvios entre os valores reais da amostra e os valores esperados, como indicado na equação 3.9:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.9)$$

Independentemente da métrica utilizada para o cálculo de erros, é de notar que modelos com média de erros próxima de zero tendem a ser os melhores modelos. De modo geral, a normalidade dos mesmos deve ser verificada. Contudo, caso a distribuição da variável resposta seja fortemente enviesada, a normalidade pode não verificar-se.

Outro método de avaliação de modelos é denominado por critério de informação de Akaike (AIC, *Akaike Information Criterion*). Este é considerado um método comparativo, uma vez que o AIC de um modelo em particular não é relevante. Os valores de AIC são utilizados para comparar vários modelos. De acordo com este critério, o melhor modelo é aquele que explica o máximo de informação usando o menor número possível de variáveis independentes, [45]. Assim, modelos que apresentem valores inferiores de AIC são considerados os melhores modelos.

$$AIC = 2K - 2\ln(L) \quad (3.10)$$

onde

K – número de variáveis independentes

L – estimador de log-verossimilhança

3.2.3 Regressão linear

O modelo de regressão mais simples é o método de regressão linear. Uma vez que existe apenas uma variável independente, o modelo é representado pela equação 3.11:

$$y_i = \beta_0 + \alpha x_i + \epsilon_i \quad , i = 1, \dots, n \quad (3.11)$$

onde

y_i – i-ésima observação da variável resposta

x_i – i-ésima observação da variável explicativa

β_0 – ordenada na origem do modelo

α – coeficiente angular ou declive do modelo

ϵ_i – erro residual da i-ésima observação

A figura 3.2 ilustra um ajuste de um modelo linear a uma amostra de dados.

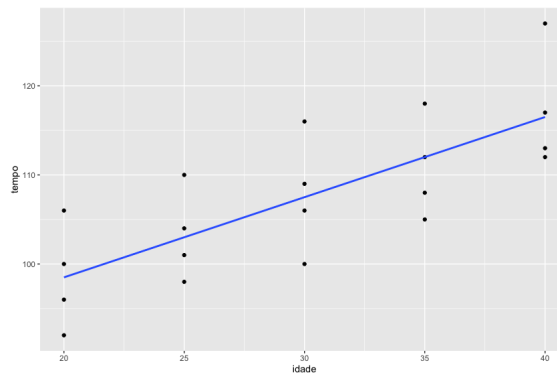


Figura 3.2: Ajuste a uma regressão linear

3.2.4 Regressão linear Piece-wise

O segundo modelo em estudo é o modelo de regressão linear piece-wise, [17].

$$y_i = \begin{cases} \beta_1 + \alpha_1 x_i + \epsilon_i & x_i < P.R. \\ \beta_2 + \alpha_2 x_i + \epsilon_i & x_i \geq P.R. \end{cases} \quad (3.12)$$

onde

$P.R.$ – ponto de rotura do modelo

O método consiste nos seguintes passos:

1. Construir um intervalo de confiança para o ponto de rotura (P.R.) dos dois grupos
2. Encontrar o ponto de rotura com menor MSE
3. Depois de escolhido o ponto, estimar os parâmetros necessários para a construção do modelo

3.2.5 Regressão linear generalizada

A construção de um modelo de regressão generalizada requer maior rigor. Tenham-se em conta três pressupostos para a modelação linear generalizada, [46]:

1. As observações são independentes entre si.
2. Os incrementos independentes. A cada Y_i corresponde um determinado intervalo de tempo.
3. Homogeneidade da distribuição de probabilidade – assumam-se que todas as observações têm a mesma distribuição de probabilidade.

Para além da metodologia geral indicada acima, é ainda necessário estudar a distribuição da variável resposta. Consequentemente, equacionar o preditor linear, $\eta_i = \sum_{j=1}^p \beta_j x_{ij}$, em função das variáveis explicativas, x_{ij} , parâmetros da regressão, β_j e escolher a melhor função de ligação, $h(\mu_i) = \eta_i$, tal que:

$$y_i = h^{-1}(\beta_j x_{ij}) + \epsilon_i \quad (3.13)$$

onde

$i = 1, \dots, n$ – número de observações

$j = 1, \dots, p$ – número de variáveis explicativas

Um método de avaliação do bom ajuste dos modelos lineares generalizados é a utilização de resíduos de Pearson. Estes são obtidos normalizando os resíduos pelo desvio padrão da variável resposta:

$$r_i = \frac{y_i - \hat{y}_i}{\sqrt{V(y_i)}} \quad (3.14)$$

onde

$V(y_i)$ – variância da variável resposta

3.2.6 Regressão quadrática

Quando existe uma relação parabólica entre as variáveis resposta e explicativa, utiliza-se um ajuste quadrático. Esta relação é uma particularização do modelo polinomial de grau n , onde $n = 2$. O modelo de regressão quadrática é definido através da equação 3.15.

$$y_i = \beta_0 + \alpha_1 x_i + \alpha_2 x_i^2 + \epsilon_i \quad (3.15)$$

A figura 3.3 demonstra um ajuste de um modelo quadrático a uma amostra de dados.

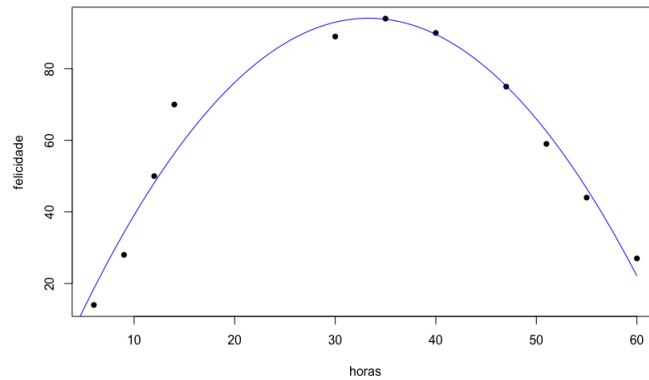


Figura 3.3: Ajuste a uma regressão quadrática

3.2.7 Regressão logarítmica

O modelo de regressão logarítmica é definido através da equação 3.16, adequado para amostras cuja relação entre variáveis é logarítmica.

$$y_i = \beta_0 + \alpha_1 \log(x_i) + \epsilon_i \quad (3.16)$$

A figura 3.4 é um exemplo a um ajuste de regressão logarítmica.

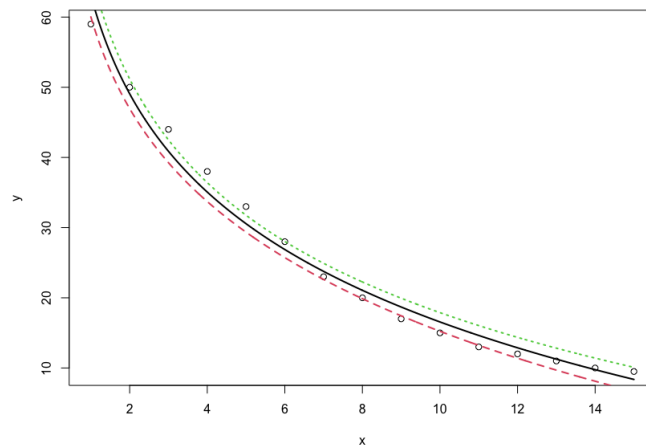


Figura 3.4: Ajuste a uma regressão logarítmica

3.3 Gestão de stocks

3.3.1 Definição

Uma política de gestão de stocks deve averiguar qual o melhor investimento de matéria prima a fazer em relação ao stock necessário para cobrir as ordens de produção. O seu objetivo é determinar como as encomendas devem ser processadas aos fornecedores – qual a quantidade ideal de material a encomendar e o período ideal entre processamento de encomendas. Para tal, é necessário conhecer todos os custos associados à cadeia de abastecimento, isto é, custos de processamento de encomendas, custos de aquisição de matéria prima e custos de armazenamento da mesma.

3.3.2 Processo de construção

Através da construção de gestão de stocks, é possível iniciar as previsões necessárias para as semanas seguintes. Construam-se duas políticas de gestão de stocks baseadas em princípios diferentes. A primeira é baseada no modelo determinístico, onde não é permitida rotura e a reposição é instantânea. A segunda é baseada no modelo de revisão periódica para procura variável, com recurso ao algoritmo de Hadley-Whitin. Abaixo segue-se a metodologia para as políticas de gestão de stocks:

1. Obter informações do número de passageiros das próximas semanas, através de fontes fidedignas como a IATA e o Instituto Nacional de Estatística (INE).
2. Utilizando o modelo de regressão escolhido, obter previsões da procura do material, baseados no passo 1.
3. Determinar os valores ótimos de lote, custo total e tempo de ciclo, baseados na procura obtida em 2.

3.3.3 Modelo determinístico

O modelo de gestão de stocks deve adequar-se ao tipo de procura do produto. Uma vez que se pretende fazer previsões semanais da procura do produto, recorra-se a modelos determinísticos. O modelo utilizado será um modelo onde a rotura de stock não é permitida, uma vez que um dos objetivos é não existir stock negativo, [47]. Assuma-se que a reposição de stock é instantânea, isto é, será calculada a frequência ideal por semana através da qual a reposição do stock será realizada. Definam-se as variáveis necessárias para a resolução do problema:

- $S(t)$ - evolução do stock em função do tempo, t
- C_T - custo total por ciclo
- K - custo total por unidade de tempo
- Q - lote económico de encomenda de material
- A - custo fixo de encomenda
- c_1 - custo unitário de aquisição
- c_2 - custo de armazenagem
- r - taxa de procura
- p - taxa de reposição de stock
- T - período do ciclo

Relacionando as diferentes variáveis, tem-se que:

$$C_T = A + c_1 Q + c_2 Q \frac{T}{2} \quad (3.17)$$

$$K = \frac{C_T}{T} = \frac{Ar}{Q} + c_1 r + c_2 \frac{Q}{2} \quad (3.18)$$

Para encontrar os valores ótimos do lote a encomendar, e o custo total ótimo por unidade de tempo, recorra-se às seguintes equações. Comece-se por derivar K em ordem a Q , de modo a obter o valor do lote económico:

$$Q^* = \sqrt{\frac{2Ar}{c_2}} \quad (3.19)$$

$$T^* = \frac{Q^*}{r} = \sqrt{\frac{2A}{rc_2}} \quad (3.20)$$

$$K^* = \sqrt{2Arc_2} + c_1 r \quad (3.21)$$

3.3.4 Modelo de revisão periódica para procura variável

O modelo de revisão periódica para procura variável tem como objetivo determinar o período ótimo de reposição para minimizar o custo total por ciclo. Este tem como hipótese que a procura tem um comportamento variável ao longo do tempo [48]. É de notar que a minimização do custo total por ciclo não implica a minimização do custo total por unidade de tempo. Assim, este é um modelo ideal quando o objetivo é minimizar os custos num determinado momento. No entanto, não é um modelo ideal a longo prazo. Definam-se as variáveis necessárias para a construção do modelo. Algumas das variáveis necessárias já foram definidas anteriormente em 3.3.3.

- R - nível de stock máximo
- O_p - custo de aquisição por ciclo
- O_s - custo de armazenagem por ciclo
- O_k - custo de rotura por ciclo
- s - nível de stock no final do ciclo
- D_L - valor esperado da procura para o período L
- z - variável aleatória da procura para o período $T + L$
- $f(z)$ - função de probabilidade de z
- D - valor esperado da procura z
- α - nível de serviço

O custo total por ciclo, C_T , é definido pela soma dos custos de aquisição, armazenagem e rotura por ciclo (O_p , O_s e O_k , respetivamente). O custo de aquisição corresponde à razão entre o custo fixo de encomenda, A , e o período do ciclo T . O custo de armazenagem por ciclo, O_s , é determinado multiplicando o valor esperado de stock por unidade de tempo, m pelo custo de armazenagem por unidade de tempo, c_2 .

Sabe-se que:

$$m = s + \frac{TD}{2} \quad (3.22)$$

$$s = R - D_L - TD \quad (3.23)$$

$$\alpha = \int_R^\infty f(z) dz = \frac{Tc_2}{c_1} \quad (3.24)$$

Assim, vem que:

$$C_T = D + \frac{A}{T} + c_2 \left(R - D_L + \frac{DT}{2} \right) + \frac{c_2}{T} \int_R^\infty (x - R) f(z) dz \quad (3.25)$$

Derivando C_T em ordem a R e igualando a 0, é possível calcular o tempo ótimo de ciclo:

$$T^* = \sqrt{\frac{2 \left(A + c_1 \int_R^\infty (x - R) f(z) dz \right)}{c_2 D}} \quad (3.26)$$

Por vezes, as equações 3.25 e 3.26 tornam-se impossíveis de calcular analiticamente, devido à complexidade do comportamento da procura. Por isso, tenha-se como hipótese que a procura segue uma distribuição Normal, de modo a implementar o método de Hadley-Whitin.

Algoritmo de Hadley-Whitin:

1. $T^* = \sqrt{\frac{2(A+c_1 \int_R^\infty (x-R)f(z) dz)}{c_2 D}}$
2. $\alpha = \frac{Tc_2}{c_1}$
3. Se $D \sim N(\mu, \sigma)$, então $R = D(T + L) + z_\alpha \sqrt{T_0 + L}$
4. $(C_T)^* = D_P + \frac{A}{T^*} + c_2 \left(R - D_L + \frac{DT^*}{2} \right) + \frac{c_2}{T^*} \int_R^\infty (x - R) f(z) dz$

Note-se que z_α é o valor obtido da tabela da distribuição Normal, de acordo com o valor calculado para α .

Para a construção dos modelos e tratamento de dados, serão utilizados os softwares R e Microsoft Excel. Os packages utilizados foram os seguintes: `gofit`, [49], `MASS`, [50], `fitdistrplus`, [51], `ggplot2`, [52], `weights`, [53], `aod`, [54], `vcd`, [55], `readxl`, [56], `fpp2`, [57], `corpcor`, [58], `Rfast`, [59], `glmDisc`, [60], `corrplot`, [61], `dse`, [62], `segmented`, [63] e `tseries`, [64].

CASO DE ESTUDO

4.1 Caracterização do problema

Sendo a ProcessWare uma empresa de software portuguesa, cujo produto, O2P, fornece soluções iterativas para problemas processuais de diversas indústrias. A sua solução para o setor de catering inclui a otimização de processos através da utilização de grandes quantidades de dados e a integração do software nos sistemas pré-existent dos clientes, como o aeroporto de Lisboa, [4]. A sua colaboração com a empresa Cateringpor está integrada em todos os seus departamentos, desde a criação de menus, processamento e gestão de encomendas, a sua produção, até ao empacotamento e sua entrega. Esta dissertação irá focar-se no processo de gestão de encomendas e toda a logística envolvente. Mais especificamente, na gestão de armazéns de produção.

O material em estudo é o material DBPA0001 utilizado na confeção de um produto alimentar da empresa Cateringpor, um dos clientes da ProcessWare. Este é um material congelado com validade de 6 meses. Atualmente a sua gestão da quantidade em stock e pontos de encomenda é realizada manualmente. Devido a esta gestão manual, é recorrente haver pedidos adicionais de material para reposição do stock durante a produção. Os lotes destas encomendas extraordinárias são mais pequenos que o normal e traduzem-se num aumento de custos do processo logístico.

Assim, com o objetivo de minimizar os pedidos de encomenda extra e todo o processo logístico, o problema em causa consiste em definir uma política de gestão de stocks adequada para um período escolhido, através do estudo do comportamento da procura do material DBPA0001 e sua previsão.

Foram cedidos dados sobre as quantidades de matéria prima consumidas durante a produção de produtos alimentares para fornecimento do aeroporto de Lisboa. Estes dados foram cedidos pela empresa ProcessWare, durante um estágio curricular. Uma vez que a Processware apenas armazena informação de inventário relativa aos últimos seis meses, estes dados são relativos ao período de julho de 2020 a junho de 2021. É de notar que este é um período demasiado curto para gerar conclusões baseadas apenas nesta série temporal, sendo assim necessário complementá-la com outra variável. Para analisar

a possível relação entre a procura do material e o tráfego aéreo, foram ainda obtidos dados do número de passageiros embarcados no aeroporto de Lisboa durante o período de setembro de 2020 a junho de 2021. Estes dados foram cedidos pela empresa ANA Aeroportos.

Inicialmente foi necessário conhecer o comportamento da série temporal do nível de stock para identificar os seus problemas. Conhecendo esta relação e com o objetivo de fazer previsões para os níveis de stock adequados, seria também necessário conhecer o comportamento da procura do material. Assim, seria possível prever o nível de stock mais adequado, conhecendo a procura do material durante as próximas semanas. Esta abordagem foi descartada, uma vez que atualmente não existe relação entre o nível de stock e a procura do material.

Ao analisar o nível de stock de material no armazém apresentado na figura 4.1, verificou-se que a partir do mês de abril, o nível de stock em armazém entra em rotura. Existem várias causas para esta aparente rotura tão prolongada. Atualmente, os valores de entrada e saída do material do armazém não estão a ser controlados de forma correta. Por vezes, os trabalhadores da linha de produção e armazém não registam as quantidades corretas de material utilizado. Desta forma, os valores não correspondem à realidade. Outra causa são os estragos do material antes de chegarem à produção – perda de stock. Isto pode ocorrer se o produto alimentar não estiver dentro dos parâmetros de qualidade estabelecidos pela Cateringpor e é necessário descartá-lo, o que reflete uma perda para a empresa.

É de notar ainda que o mês de março se insere na época baixa de turismo e da aviação. Este fator, combinado com a redução do tráfego aéreo por consequência da pandemia, contribui para as roturas de stock apresentadas na figura 4.1. Atualmente não existem políticas de stock na empresa e a linha de produção está a ser abastecida diretamente, sem recurso a este armazém. Conclui-se que atualmente não existe stock de segurança e é necessário uma reestruturação completa do modelo de stocks usado atualmente.

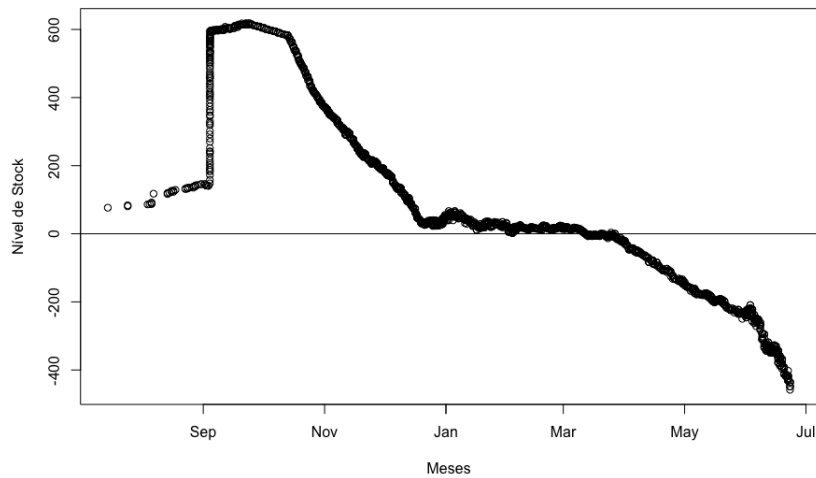


Figura 4.1: Nível de stock ao longo dos meses

4.2 Caracterização dos dados

4.2.1 Dados cedidos por ProcessWare

Os dados cedidos pela empresa ProcessWare são relativos ao material DBPA0001 e foram apresentados em três *datasets* diferentes:

1. Dados de entradas no armazém de produção
2. Dados de inventário
3. Dados de consumo de produção

Os dados relativos às entradas no armazém foram apresentados em quatro colunas: data e hora da entrada no armazém, identificação do material, unidade de medida (em Kg), quantidade de material. A tabela 4.1 contém cinco observações dos dados de entradas no armazém.

Tabela 4.1: Cinco observações dos dados de entradas no armazém

Dia e hora	Material	Unidade	Quantidade
04/09/2020 00:00	DBPA0001	KG	5
04/09/2020 02:03	DBPA0001	KG	6
04/09/2020 10:27	DBPA0001	KG	5
04/09/2020 12:00	DBPA0001	KG	5
04/09/2020 16:04	DBPA0001	KG	5

Os dados de inventário correspondem às entradas e transações extraordinárias no armazém. Estas encomendas não foram planeadas previamente e portanto representam custos mais elevados. Foram apresentados em quatro colunas: localização do armazém, data e hora da entrada no armazém, quantidade de material, tipo de transação (entrada extra no armazém ou perda de stock). A tabela 4.2 contém cinco observações destes dados.

Tabela 4.2: Cinco observações dos dados de inventário

Localização	Dias e hora	Quantidade	Tipo de transação
LI1	15/07/2020 17:30	0.173	INVENTORY
LI1	25/07/2020 15:04	4.145	INVENTORY
LI1	25/07/2020 16:23	2.998	INVENTORY
LI1	04/08/2020 14:10	2.622	INVENTORY
LI1	05/08/2020 14:39	0.615	INVENTORY

Os dados de consumo de produção correspondem às quantidades de material que foram utilizadas durante a produção. Estes foram apresentados em nove colunas: nome do produto final que foi produzido, identificação do cliente, data efetiva, localização da produção, identificação do material, quantidade de produto final produzida, quantidade de material consumida, status da produção e data e hora de produção, tal como mostra a tabela 4.3.

Tabela 4.3: Cinco observações dos dados de consumo de produção

Produto final	Cliente	Data efetiva	Localização	Material	Quantidade produto final	Material consumido	Status	Dia e hora
5D001261	43	22/10/2020	CF	DBPA0001	45	0.99225	CLOSED	27/10/2020 09:58
5N001634	1	30/03/2020	CF	DBPA0001	6	0.2016	CLOSED	28/10/2020 11:04
5D001261	43	22/10/2020	CF	DBPA0001	45	0.99225	CLOSED	29/10/2020 09:42
5D000636	0	01/01/2015	CF	DBPA0001	1	0.0441	CLOSED	29/10/2020 10:00
5D000636	0	01/01/2015	CF	DBPA0001	1	0,0441	CLOSED	30/10/2020 10:21

Finalmente, os dados foram compilados num *dataset* final. Uma vez que cada linha (observação) correspondia a uma transação de material, procedeu-se a organizar os dados diariamente. Assim, é possível estudar a relação entre a quantidade consumida de material, isto é, a sua procura e o número de PEAL.

Para além desta compilação, foi também calculado o stock a cada dia do processo. O stock no momento i é de:

$$S_i = S_{i-1} + E_i + EE_i - D_i \quad (4.1)$$

onde,

- S_i – nível de stock no momento i
- E_i – entrada de material no armazém no momento i
- EE_i – entrada extraordinária de material no armazém no momento i
- D_i – quantidade de procura de produto no momento i (Kg)

Construiu-se assim um *dataset* final relativo aos dados fornecidos pela ProcessWare que corresponde a uma cronologia de todo o processo. Existem 1901 observações e cada uma corresponde a uma movimentação de material. A tabela 4.4 contém algumas observações para mostrar a cronologia dos movimentos do armazém.

Tabela 4.4: Dez observações da cronologia do processo

Dia e hora	E_i	EE_i	D_i	S_i
03/02/2021 09:50		0.216		2.011045
03/02/2021 10:06			0.336	1.675045
03/02/2021 10:17	5			6.675045
03/02/2021 10:26			0.183	6.492045
03/02/2021 11:08			0.2646	6.227445
03/02/2021 11:08			1.323	4.904445
03/02/2021 11:12			0.03317	4.871275
03/02/2021 13:24	3			7.871275
04/02/2021 07:19			0.5292	7.342075
04/02/2021 07:19			0.3969	6.945175

4.2.2 Dados cedidos por ANA Aeroportos

De modo a averiguar se existe uma relação entre a procura de material e o número de passageiros do aeroporto, foram solicitados dados à empresa ANA Aeroportos. Os dados são relativos ao número de PEAL diário. Estes têm início no dia 1 de Setembro de 2020 a 30 de Junho de 2021 (10 meses). Este *dataset* é composto por 303 observações e 3 colunas: nome do aeroporto, data e número de passageiros nesse dia. A tabela 4.5 contém as dez primeiras observações.

Tabela 4.5: Primeiras observações dos dados ANA Aeroportos

Aeroporto ANA	Data	Passageiros Embarcados (Exceto em Transferência)
Aeroporto de Lisboa	01/09/2020	10 432
Aeroporto de Lisboa	02/09/2020	9 692
Aeroporto de Lisboa	03/09/2020	11 543
Aeroporto de Lisboa	04/09/2020	13 191
Aeroporto de Lisboa	05/09/2020	13 003
Aeroporto de Lisboa	06/09/2020	13 196
Aeroporto de Lisboa	07/09/2020	12 475
Aeroporto de Lisboa	08/09/2020	8 876
Aeroporto de Lisboa	09/09/2020	8 140
Aeroporto de Lisboa	10/09/2020	11 464

4.2.3 Dados finais

Compilando a informação fornecida por estas duas empresas, foi criado um terceiro *dataset*, que contém apenas dados relativos à procura do material e ao número de passageiros do aeroporto de Lisboa. Os dados foram agrupados diária, semanal e mensalmente. Por isso existem 9 colunas referentes à unidade de tempo, procura e número de passageiros para cada unidade de tempo. As variáveis de dados diários têm 303 observações. As variáveis semanais têm 40 observações e as variáveis mensais têm 10 observações. A tabela 4.6 mostra as dez primeiras observações do *dataset*.

Tabela 4.6: Dez observações da compilação final

Dia	D diária	PEAL diário	Semana	D semanal	PEAL semanal	Mês	D mensal	PEAL mensal
01/09/2020	0.3129	10432	36	48.67	38 902	09	52	316 801
02/09/2020	3.9081	9692	37	57.94	30 051	10	231	234 325
03/09/2020	0.6825	11543	38	68.59	27 130	11	242	132 780
04/09/2020	2.7657	13191	39	53.05	26 507	12	302	177 143
05/09/2020	0.672	13003	40	56.49	29 789	01	277	153 825
06/09/2020	3.36	13196	41	65.25	29 599	02	127	54 969
07/09/2020	0.3129	12475	42	74.78	46 561	03	138	78 238
08/09/2020	2.1945	8876	43	64.68	44 824	04	183	116 730
09/09/2020	0.7161	8140	44	54.19	36 560	05	211	195 190
10/09/2020	3.0345	11464	45	29.77	16 193	06	756	322 424

Após o tratamento e organização dos dados, proceda-se à sua análise e obtenção de resultados. Esta apresentação final dos dados sugere uma primeira análise através de séries temporais.

RESULTADOS E DISCUSSÃO

5.1 Análise da série temporal da procura

A figura 5.1 descreve a série temporal da procura de material diária, desde o dia 1 de setembro de 2020 até ao dia 30 de junho de 2021. Esta apresenta várias oscilações ao longo do tempo. Entre os meses de setembro e outubro de 2020, a quantidade de material consumida mantém-se muito baixa e com poucas variações. Existe apenas ruído branco devido às oscilações de produção nos fins de semana, que satisfazem o maior abastecimento dos aviões durante os fins de semana. A partir de novembro de 2020, dá-se um grande aumento na procura do produto. As oscilações durante as semanas tornam-se mais evidentes, uma vez que nesta altura do ano, a população tende a viajar mais. De janeiro a março de 2021, dá-se um decréscimo da quantidade produzida de material, coincidente com o novo período de confinamento. De março a maio de 2021, existe um ligeiro aumento de procura. Finalmente em junho dá-se um grande aumento na procura de material e ocorrem grandes oscilações de valores. Estes valores podem ser justificados pela entrada do verão e o início das férias dos passageiros.

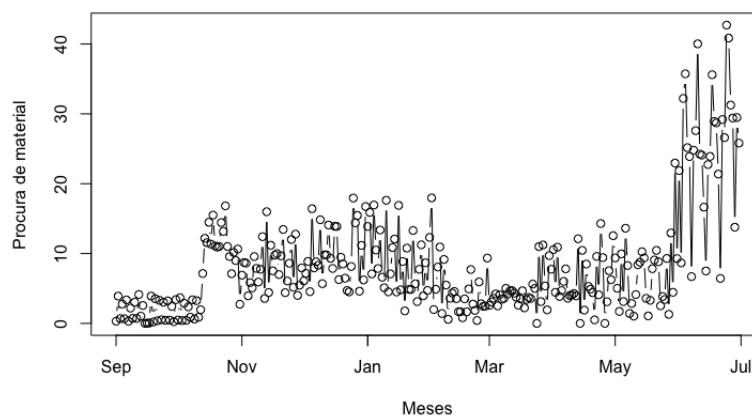


Figura 5.1: Série temporal da procura de material de setembro de 2020 a julho de 2021

Averiguou-se se existiria algum padrão, sazonalidade ou estacionariedade. É também necessário conhecer a sua autocorrelação para saber se as observações se relacionam ao longo do tempo. A autocorrelação da série é de 0.67, ou seja existe alguma relação, embora fraca, entre as observações i e $i - 1$. A figura 5.2 mostra a autocorrelação desta série temporal. É de notar que durante os primeiros meses, existe uma forte relação entre observações. Os valores de autocorrelação tendem a diminuir com o tempo.

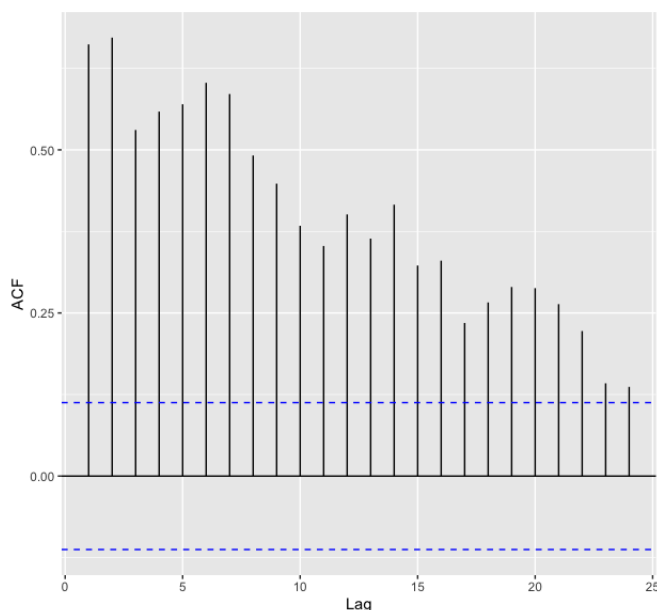


Figura 5.2: Autocorrelação da procura ao longo do tempo

Havendo uma autocorrelação bastante significativa, esta série temporal não tende a comportar-se de forma aleatória em torno de um ponto médio. Para confirmar esta afirmação, foi realizado o teste de estacionariedade de Dickey-Fuller. Uma vez que o p-value obtido foi de 0.99, então pode-se afirmar que esta série não é estacionária.

5.1.1 Agregação dos dados semanalmente

Com o objetivo de fazer um planeamento semanal de gestão de stocks, é necessário obter previsões semanais da procura do material. Assim, os dados foram agregados semanalmente. A série descrita pela figura 5.3 é muito semelhante à da figura 5.1. Tal como na série anterior, os meses de setembro e outubro são bastante calmos e dá-se um pico de procura no final de outubro e inícios de novembro. No final de dezembro dá-se outro pico de consumo de produto, que corresponde ao Natal. De março a maio, a procura mantém-se praticamente constante e volta a aumentar durante o mês de junho, devido ao início do verão. Conclui-se que os comportamentos observados na série diária tornaram-se mais evidentes nesta série semanal.

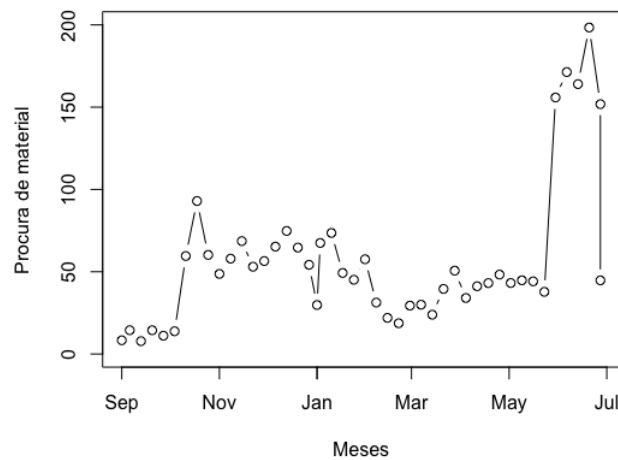


Figura 5.3: Série temporal da procura semanal de material

5.2 Análise da série temporal de passageiros

A figura 5.4 mostra algumas oscilações frequentes. Estas pequenas oscilações correspondem à variação do número de passageiros no aeroporto durante a semana. Os picos ligeiramente mais elevados caracterizam o comportamento do número de passageiros durante o final da semana e os picos ligeiramente mais baixos caracterizam o fluxo de passageiros durante o início e meio da semana.

Para além das oscilações semanais, o fluxo de passageiros também varia de acordo com o mês. Como é de esperar, os meses com maior número de passageiros no aeroporto são os meses de setembro, junho e julho. De outubro a dezembro, o fluxo de passageiros diminui significativamente, voltando a aumentar entre dezembro e janeiro. Este aumento é devido à época natalícia. Durante a primavera, é possível constatar um crescimento progressivo do número de passageiros até ao verão.

Para descrever uma série temporal, é importante averiguar se existe algum padrão, sazonalidade ou estacionariedade. É também necessário conhecer a sua autocorrelação para saber se as observações se relacionam ao longo do tempo.

A autocorrelação desta série temporal é de 0.857. Este valor significa que as observações do número de passageiros do aeroporto de Lisboa se relacionam ao longo do tempo.

A figura 5.5 mostra como se dão picos na correlação a cada 7 dias, o que comprova haver um padrão semanal: ao final de cada semana existe sempre um maior número de viajantes. Para comprovar esta afirmação, construiu-se o gráfico da figura 5.6, que descreve os picos de valores obtidos na série temporal. É de notar que entre os fins de semana de 15 e 22 de janeiro deu-se um grande desvio de valores. Este desvio pode dever-se a um decréscimo no número de passageiros no aeroporto, uma vez que coincide com o início do primeiro confinamento de 2021. Outro grande pico de desvios entre valores

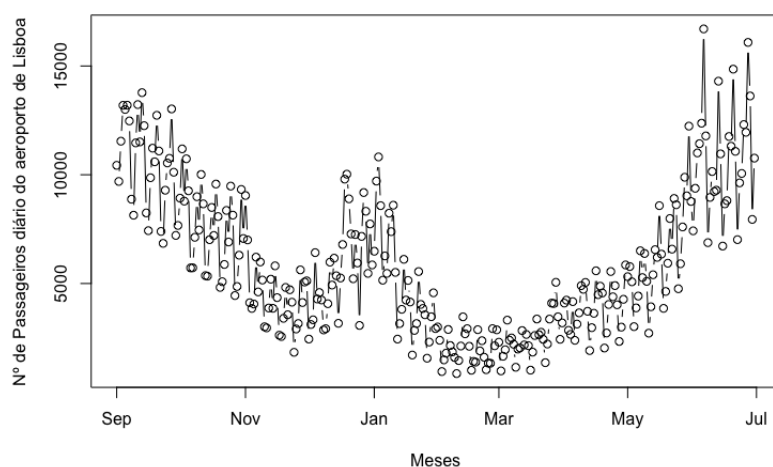


Figura 5.4: Série temporal do número de PEAL de setembro de 2020 a julho de 2021

registados nos fins de semana dá-se entre os fins de semana de 4 e 11 de junho. Este valor pode ser justificado com o início do verão, onde a população começa a usufruir das suas férias.

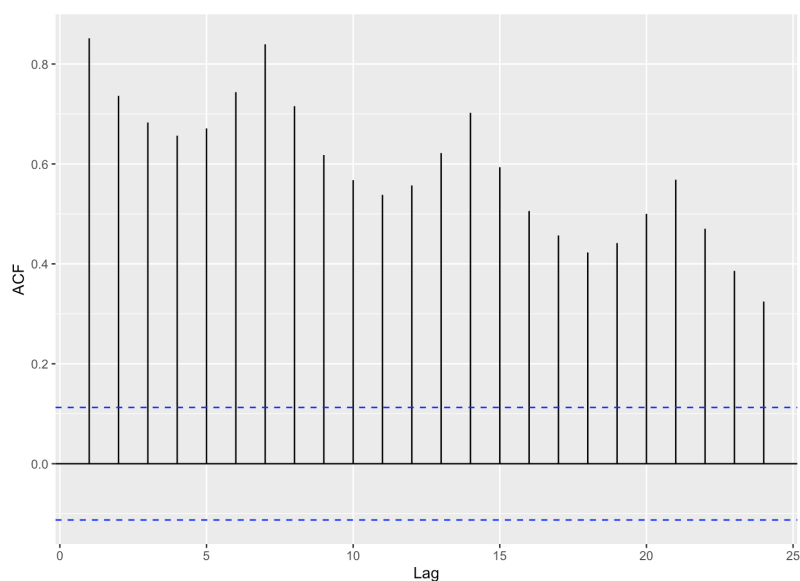


Figura 5.5: Autocorrelação do número de PEAL

Com o objetivo de confirmar que esta série tem realmente uma sazonalidade semanal, foi também calculada a autocorrelação da série temporal, removendo estes picos – 0.831. É de notar que este valor é inferior ao valor da autocorrelação da série temporal total (0.857), o que confirma a afirmação. A figura 5.7 mostra como ainda existe sazonalidade ao longo da série, com tendência decrescente e que a sua autocorrelação continua elevada.

Tendo esta série uma forte sazonalidade semanal, ela não é estacionária. Foi realizado

5.2. ANÁLISE DA SÉRIE TEMPORAL DE PASSAGEIROS

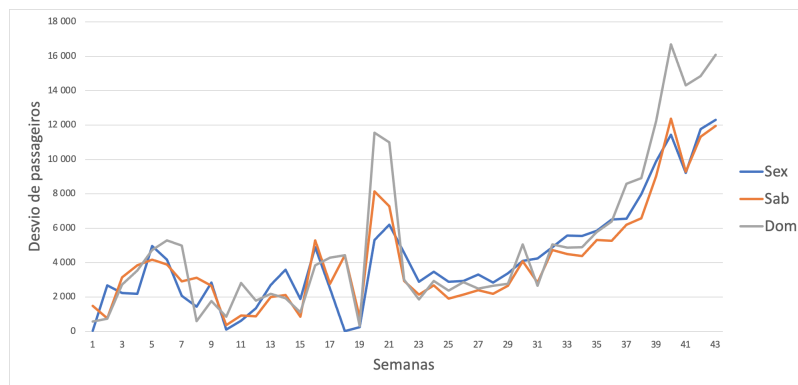


Figura 5.6: Desvio de valores obtidos entre sexta-feiras consecutivas, sábados e domingos, respectivamente

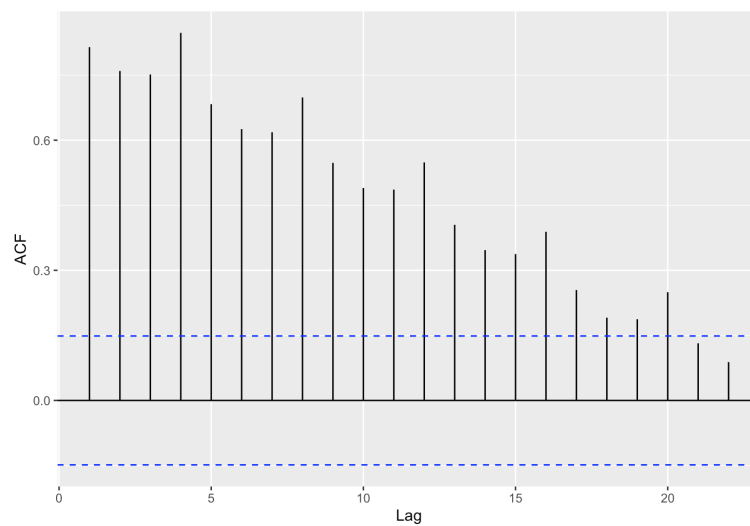


Figura 5.7: Autocorrelação do número de PEAL, excluindo os fins de semana

o teste Dickey-Fuller com p-value de 0.99, rejeitando a hipótese de estacionariedade.

5.3 Modelação da procura

Uma vez que os dados apresentados são relativos a um curto período de tempo, é necessário complementar a análise de séries temporais com o uso de regressões. Assim, é possível descrever com maior detalhe a relação entre a procura e o número de passageiros.

Primeiramente, analisando o gráfico da figura 5.8, repare-se que com exceção dos meses de setembro e outubro, a procura mensal do material tem o mesmo comportamento que o fluxo de passageiros do aeroporto de Lisboa.

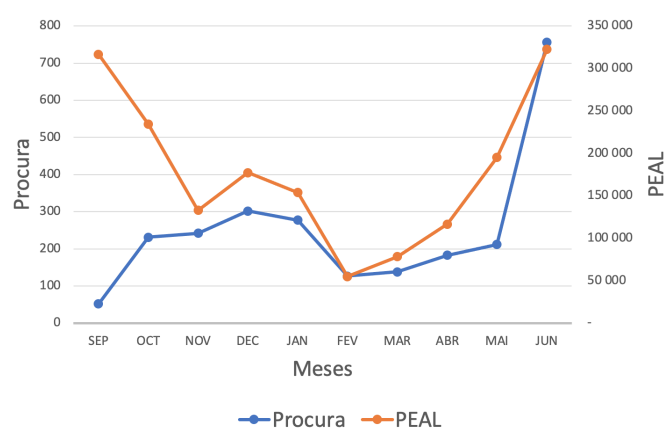


Figura 5.8: Relação entre a procura mensal e PEAL mensal

Esta diferença de comportamento durante os dois primeiros meses pode dever-se a vários fatores, como a sazonalidade do produto – se este for um produto pouco usado no final do verão e início do outono. Para além disso, pode ter-se dado uma paragem não explicada na sua utilização durante esses meses, ou da produção da própria fábrica da Cateringpor. Descartem-se assim os dados dos dois primeiros meses e proceda-se à avaliação da relação entre estas duas variáveis.

5.3.1 Agregação dos dados semanalmente

De acordo com a análise anterior, considere-se a agregação semanal de dados de 1 de novembro de 2020 a 30 de junho de 2021. A figura 5.9 representa a relação entre as duas variáveis, sendo que a procura é a variável resposta, D , e o número de PEAL é a sua variável explicativa, X . Os dados da variável explicativa foram também agregados semanalmente.

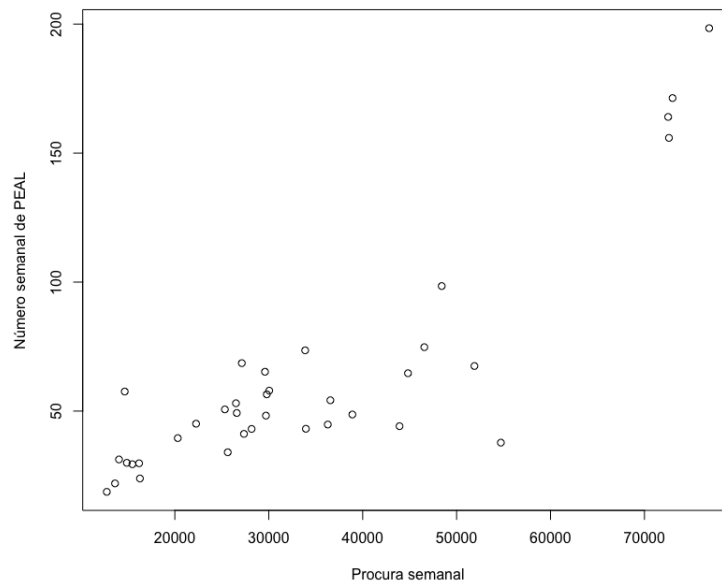


Figura 5.9: Gráfico de dispersão entre a procura semanal e número de PEAL semanal

Abordagem I - Regressão Linear

Analisando o gráfico da figura 5.9, testou-se a hipótese de haver uma relação de linearidade entre as duas variáveis. Para tal, recorrendo a regressão linear, construiu-se o modelo traduzido pela equação abaixo.

$$D = \beta_0 + \alpha X + \epsilon \quad (5.1)$$

onde

- $\beta_0 = -9.14$
- $\alpha = 0.002$
- $\epsilon \sim N(0, 1)$

A variável independente apresenta ser muito significativa para o modelo, o valor obtido para R^2 é de 0.7544, o que significa que este foi um bom ajuste. O valor médio dos seus resíduos é de $1.715e^{-15}$, que é muito próximo de 0 e a hipótese de estes seguirem uma distribuição Normal não é rejeitada. O seu valor de AIC é de 326.68.

Abordagem II - Regressão Linear Piece-wise

Para otimizar este modelo, foi utilizada regressão linear Piece-wise. A figura 5.9 mostra dois grupos de observações que podem originar duas regressões diferentes. Assim, calculou-se um ponto de rotura entre os dois grupos com menor erro quadrado absoluto. O intervalo de confiança escolhido foi [53000;72650]. O ponto de rotura obtido foi de 54712.

O modelo construído é o seguinte:

$$D = \begin{cases} \beta_1 + \alpha_1 X + \epsilon & X < 54712 \\ \beta_2 + \alpha_2 X + \epsilon & X \geq 54712 \end{cases} \quad (5.2)$$

Foram utilizados dois métodos de construção: um método iterativo e outro utilizando o package `Segmented` no software R.

Método iterativo

- $\beta_1 = 20.93$
- $\alpha_1 = 4.9.36 \times 10^{-4}$
- $\beta_2 = -449.55$
- $\alpha_2 = 7.50 \times 10^{-3}$
- $\epsilon \sim N(0, 1)$

O valor obtido para R^2 é de 0.9255, o que significa que este é um ótimo ajuste. O valor médio dos seus resíduos é de $6.23e^{-16}$, que é muito próximo de 0 e estes não rejeitam a hipótese do teste de normalidade de Shapiro-Wilk. O seu valor de AIC é de 286.42.

Método com recorrência ao package "Segmented"

Este método apresenta resultados ligeiramente menos interessantes, uma vez que o seu R^2 é de 0.9037. A figura 5.13 mostra o ajuste deste modelo para os dados reais. Pode-se verificar que é um bom ajuste.

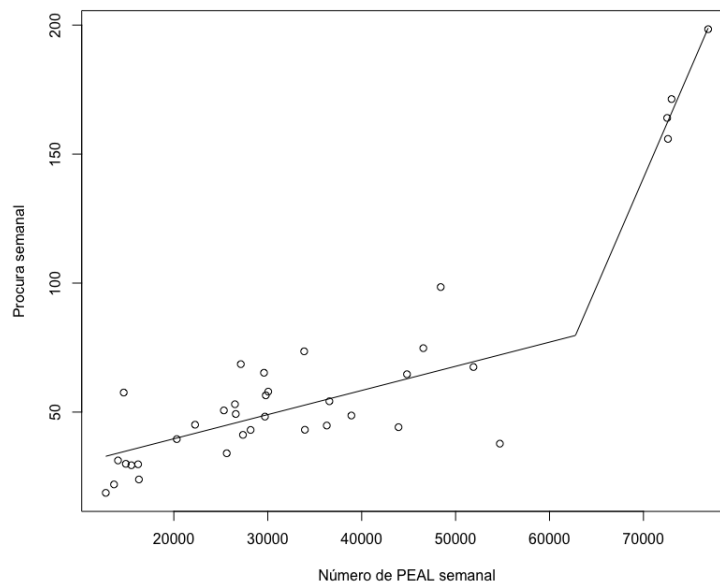


Figura 5.10: Ajuste do segundo método de regressão linear Piece-wise

Abordagem III - Regressão linear generalizada

Nesta tentativa, foram aplicados modelos lineares generalizados aos dois grupos encontrados anteriormente. Para tal, foi necessário estudar a distribuição da variável resposta.

Primeiro ramo

De modo a averiguar se existem condições para a utilização de modelos lineares generalizados, foi estudado o comportamento estatístico da variável resposta.

Analisando o histograma da figura 5.11, averiguou-se se D provinha de uma distribuição Gamma. O p-value obtido foi de 0.787, logo pode-se afirmar com confiança de 95% que a variável aleatória D segue uma distribuição Gamma de parâmetros $\alpha = 7.4$ e $\beta = \frac{1}{6.6}$.

Sabendo agora que a amostra da variável resposta pertence a uma distribuição Gamma, confirma-se que pertence à família exponencial.

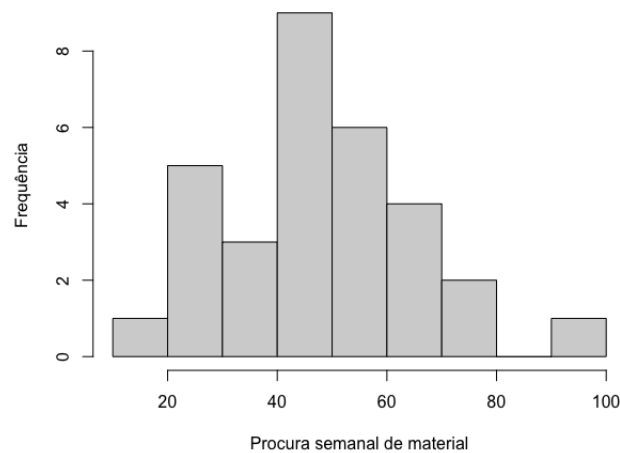


Figura 5.11: Histograma de D quando $x < 54712$

Segundo ramo

O segundo ramo contém apenas 5 observações. Ainda assim, averiguou-se se D provinha de alguma distribuição conhecida. Testando para a distribuição Gamma, o p-value obtido foi de 0.28, logo pode-se afirmar com confiança de 95% que a variável aleatória D segue uma distribuição Gamma de parâmetros $\alpha = 3.4.3$ e $\beta = \frac{1}{42.2}$.

Conclusão: Este modelo não cumpre o primeiro pressuposto descrito em 3.2, uma vez que as observações da variável da procura de material não são independentes entre si – estas relacionam-se ao longo do tempo. Uma vez que esta partição implica que o segundo ramo contenha apenas 5 observações, qualquer ajuste de distribuições não será favorável. Um bom ajuste de qualquer variável aleatória deve ter em conta que a amostra desta variável tenha pelo menos cerca de 30 observações. Assim, o modelo apresentado pela tentativa 3 não representa um bom ajuste.

Abordagem IV - Regressão quadrática

Para descrever um modelo de regressão quadrática, foi criada uma segunda variável explicativa (X^2), que representa o quadrado da variável explicativa inicial - número de PEAL semanal. Foi criado um modelo clássico de regressão quadrática e outro modelo de regressão quadrática por ramos.

$$D = \beta_0 + \alpha_1 X + \alpha_2 X^2 + \epsilon \quad (5.3)$$

tal que:

- $\beta_0 = 47.41$
- $\alpha_1 = -1.26 \times 10^{-3}$
- $\alpha_2 = 3.88 \times 10^{-8}$

e

$$D = \begin{cases} \beta_0^1 + \alpha_1^1 X + \alpha_2^1 X^2 + \epsilon & X < 54712 \\ \beta_0^2 + \alpha_1^2 X + \alpha_2^2 X^2 + \epsilon & X \geq 54712 \end{cases} \quad (5.4)$$

tal que:

- $\beta_0^1 = 7.56$
- $\alpha_1^1 = 1.75 \times 10^{-3}$
- $\alpha_2^1 = -9.39 \times 10^{-9}$
- $\beta_0^2 = -91.01$
- $\alpha_1^2 = -1.14 \times 10^{-3}$
- $\alpha_2^2 = 6.38 \times 10^{-8}$
- $\epsilon \sim N(0, 1)$

O primeiro modelo apresentou um R^2 de 0.8521. Os seus resíduos apresentam um valor médio próximo de 0, mas rejeitam a hipótese de normalidade. Isto deve-se ao facto do ajuste não ser linear. O seu valor de AIC é de 309.36. Os valores de R^2 do segundo modelo são de 0.5237 e 0.9896 respetivamente para cada ramo. Uma vez que estes valores são mais baixos, descarte-se o modelo de regressão quadrática por ramos e considere-se apenas o modelo clássico de regressão quadrática.

Conclusão: Estes ajustes, embora razoáveis não representam o comportamento da variável resposta. O ajuste do modelo clássico é menos complexo, uma vez que não existe

uma distinção entre grupos de observações. Contudo, o facto do primeiro ramo apresentar um valor de R^2 tão baixo é muito significativo, uma vez que este é o ramo com maior número de observações.

Abordagem V - Regressão logarítmica

Para criar um modelo de regressão logarítmica, foi também criada outra variável explicativa ($\log(X)$), que descreve o logaritmo da variável explicativa inicial X . Foi criado um modelo clássico de regressão quadrática e outro modelo de regressão quadrática por ramos.

$$D = \beta_0 + \alpha \ln(X) + \epsilon \quad (5.5)$$

tal que:

- $\beta_0 = -608.37$
- $\alpha = 65.02$
- $\epsilon \sim N(0, 1)$

$$D = \begin{cases} \beta_0^1 + \alpha^1 \ln(X) + \epsilon & X < 54712 \\ \beta_0^2 + \alpha^2 \ln(X) + \epsilon & X \geq 54712 \end{cases} \quad (5.6)$$

tal que:

- $\beta_0^1 = -276.72$
- $\alpha_1^1 = 31.98$
- $\beta_0^2 = -4949.9$
- $\alpha_1^2 = 457.1$
- $\epsilon \sim N(0, 1)$

O primeiro modelo apresentou um R^2 de 0.5886. Os seus resíduos apresentam um valor médio próximo de 0 e a um nível de significância de 95%, estes são Normais. O seu valor de AIC é de 345.26. Os valores de R^2 obtidos para cada ramo foram 0.5417 e 0.9869, respetivamente. Uma vez que estes valores são mais baixos, descarte-se o modelo de regressão logarítmica por ramos e considere-se apenas o modelo clássico de regressão logarítmica.

Conclusão: Mais uma vez, este modelo representa um ajuste razoável. Embora o valor de R^2 obtido no segundo ramo seja muito bom, este não reflete um bom ajuste devido ao número reduzido de observações que contempla.

5.3.2 Dados diários

Para além da modelação semanal, foram construídos modelos utilizando dados diários com o objetivo de averiguar se seria interessante fazer um planeamento diário de gestão de stocks.

De modo a averiguar se existe realmente uma relação entre elas, analise-se o gráfico de dispersão entre as duas variáveis. A figura 5.12 mostra que os valores entre as duas variáveis estão muito mais dispersos e sem aparente relação entre eles. Existe uma maior concentração de valores quando o número de PEAL é inferior a 5000 e a procura diária é inferior a 10 Kg de material. Foram realizadas várias tentativas para compreender melhor o tipo de relação.

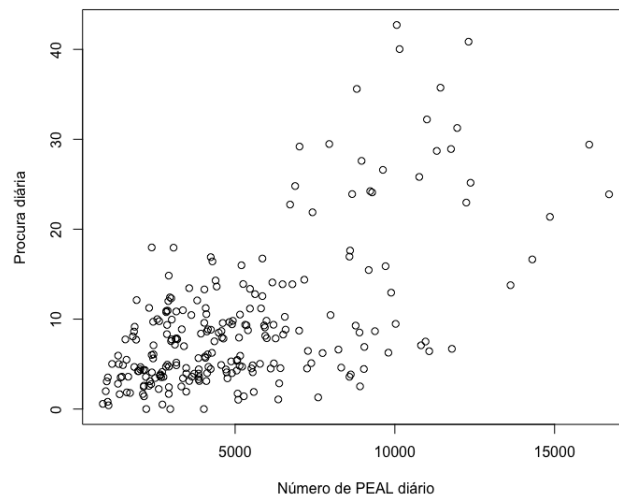


Figura 5.12: Gráfico de dispersão entre a procura diária e o número de PEAL diário

Abordagem I - Regressão Linear

Primeiramente, testou-se a hipótese de haver uma relação de linearidade entre as duas variáveis. Para tal, recorrendo a regressão linear, construiu-se um modelo tal que:

$$D = \beta_0 + \alpha X + \epsilon \quad (5.7)$$

onde

- $\beta_0 = 1.15007$
- $\alpha = 0.001597$
- $\epsilon \sim N(0,1)$

Apesar da variável independente ser bastante significativa para o modelo, o valor obtido para R^2 é de 0.3841, o que significa que este não foi um bom ajuste. O valor médio dos seus resíduos é muito próximo de 0, contudo estes falham o teste de normalidade.

Foi assim pensada noutra alternativa, que consiste em dividir as observações em dois grupos e aplicar uma regressão a cada um deles.

Abordagem II - Regressão Linear Piece-wise

Recorrendo agora a regressão linear Piece-wise, calculou-se o ponto de rotura de 8579. Assim, o modelo construído é:

$$D = \begin{cases} \beta_1 + \alpha_1 X + \epsilon & X < 8579 \\ \beta_2 + \alpha_2 X + \epsilon & X \geq 8579 \end{cases} \quad (5.8)$$

Foram utilizados dois métodos de construção: um método iterativo e outro utilizado o package `Segmented` no software R.

Método iterativo

Os resultados do método iterativo foram os seguintes:

- $\beta_1 = 3.34$
- $\alpha_1 = 9.46 \times 10^{-4}$
- $\beta_2 = 26.18$
- $\alpha_2 = 1.06 \times 10^{-3}$
- $\epsilon \sim N(0, 1)$

O valor obtido para R^2 é de 0.4161, o que significa que este ajuste não é ideal. O valor médio dos seus resíduos é de $-1.044e^{-16}$, que é muito próximo de 0, contudo estes falham o teste de normalidade de Shapiro-Wilk.

Método com recorrência ao package "Segmented"

Este método apresenta resultados ligeiramente menos interessantes, uma vez que o seu R^2 é de 0.3918. A figura 5.13 mostra o ajuste deste modelo para os dados reais. Pode-se verificar que o ajuste não é ideal.

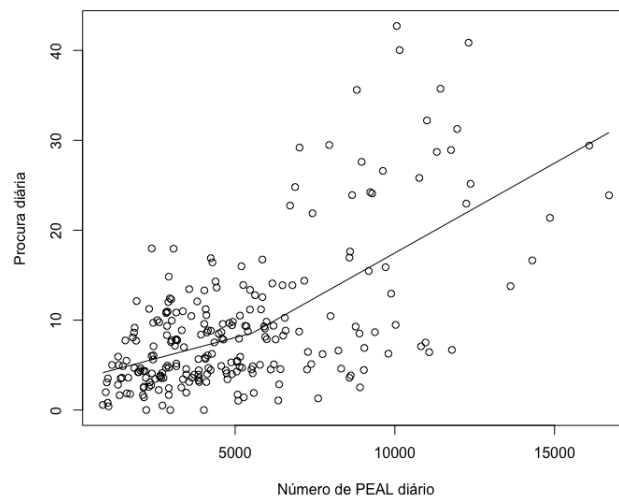


Figura 5.13: Ajuste do 2º método de regressão linear Piece-wise

Abordagem III - Regressão Linear generalizada por ramos

Nesta tentativa, foram aplicados modelos lineares generalizados aos dois grupos. Para tal, foi necessário estudar a distribuição da variável resposta.

Primeiro ramo

De modo a averiguar se existem condições para a utilização de modelos lineares generalizados, foi estudado o comportamento estatístico da variável resposta.

Analisando o histograma da figura 5.14, averiguou-se se D provinha de uma distribuição Gamma. O p-value obtido foi de 0.512, logo pode-se afirmar com confiança de 95% que a variável aleatória D segue uma distribuição Gamma de parâmetros $\alpha = 2.3$ e $\beta = \frac{1}{3.2}$.

Sabendo agora que a amostra da variável resposta pertence a uma distribuição Gamma, confirma-se que pertence à família exponencial.

Segundo ramo

Analisando o histograma da figura 5.15, averiguou-se se D provinha de alguma distribuição conhecida. Testando para a distribuição Gamma, o p-value obtido foi de 0.249, logo pode-se afirmar com confiança de 95% que a variável aleatória D segue uma distribuição Gamma de parâmetros $\alpha = 2.3$ e $\beta = \frac{1}{8.3}$.

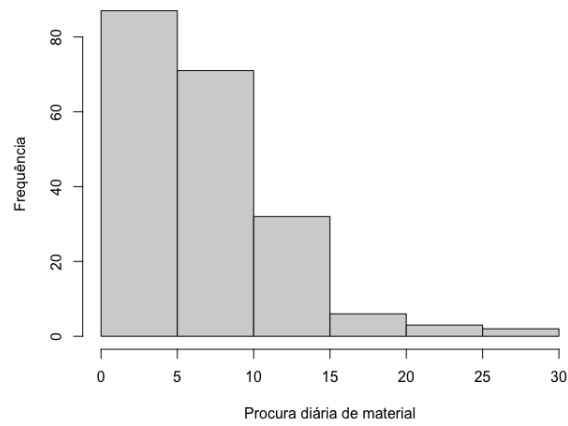


Figura 5.14: Histograma de D, quando $x < 8579$

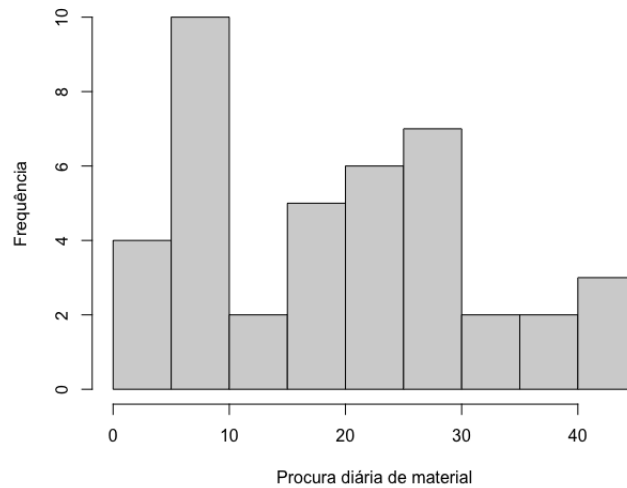


Figura 5.15: Histograma de D, quando $x \geq 8579$

Abordagem IV - Regressão quadrática por ramos

Para descrever um modelo de regressão quadrática, foi criada uma segunda variável explicativa (X^2), que representa o quadrado da variável explicativa inicial - número de PEAL diário. Assim, foi criado o seguinte modelo de regressão multilinear:

$$D = \begin{cases} \beta_0^1 + \alpha_1^1 X + \alpha_2^1 X^2 + \epsilon \\ \beta_0^2 + \alpha_1^2 X + \alpha_2^2 X^2 + \epsilon \end{cases} \quad (5.9)$$

tal que:

- $\beta_0^1 = 2.72$
- $\alpha_1^1 = 1.29 \times 10^{-3}$
- $\alpha_2^1 = -2.82 \times 10^{-8}$
- $\beta_0^2 = -52.97$
- $\alpha_1^2 = 1.13 \times 10^{-2}$
- $\alpha_2^2 = -4.06 \times 10^{-7}$
- $\epsilon \sim N(0,1)$

Abordagem V - Regressão logarítmica por ramos

Para criar um modelo de regressão logarítmica, foi também criada outra variável explicativa ($\ln(X)$), que descreve o logaritmo da variável explicativa inicial X . Foi criado o seguinte modelo de regressão logarítmica:

$$D = \begin{cases} \beta_0^1 + \alpha_1^1 \ln(X) + \epsilon \\ \beta_0^2 + \alpha_1^2 \ln(X) + \epsilon \end{cases} \quad (5.10)$$

tal que:

- $\beta_0^1 = -22.99$
- $\alpha_1^1 = 3.70$
- $\beta_0^2 = -158.48$
- $\alpha_1^2 = 19.20$
- $\epsilon \sim N(0,1)$

5.3.3 Resultados finais

Muito embora os modelos construídos através de dados diários pudessem trazer algumas vantagens como maior nível de precisão, os seus resultados não são tão bons quanto os resultados dos modelos construídos através de dados semanais. Contextualizando, os modelos semanais são mais vantajosos, uma vez que, por norma, é mais interessante e realista realizar um planeamento semanal de gestão de stocks. Assim, comparando os quatro melhores modelos semanais da tabela 5.1, verifica-se que o modelo com valor mais elevado de R^2 é o modelo de regressão linear piece-wise. Este modelo apresenta ainda o valor mais baixo de AIC, os seus resíduos têm valor médio aproximado de 0 e apresentam um comportamento Normal. Perante estes critérios de decisão, escolha-se o modelo semanal de regressão piece-wise como o modelo final que retrata o comportamento da procura de material dependente do número de PEAL.

Utilize-se este modelo para fazer previsões da procura do material e definir a melhor política de gestão de stocks.

Tabela 5.1: Compilação de resultados dos quatro melhores modelos

Modelo	R^2	Significância da variável explicativa	Resíduos	AIC
Regressão linear	0.7544	***	Valor médio: 0 Normais	326.68
Regressão linear piece-wise	0.9255	*	Valor médio: 0 Normais	286.42
Regressão quadrática	0.8521	***	Valor médio: 0 Não são Normais	309.36
Regressão logarítmica	0.5886	***	Valor médio: 0 Normais	345.26

onde

- *** – a significância da variável explicativa é superior a 99.9%
- ** – a significância da variável explicativa é superior a 99%
- * – a significância da variável explicativa é superior a 95%

5.4 Políticas de gestão de stocks

Realizem-se previsões para as semanas de julho de 2021. De acordo com o Instituto Nacional de Estatística, o número de PEAL foi de 588 159 passageiros. O mês de julho é um mês de grande movimentação constante de passageiros, com tendência a aumentar ao longo do mês. Escolheu-se este mês por ser a última atualização de dados do INE. Considere-se a seguinte distribuição de passageiros ao longo do mês, de acordo com a tabela 5.2:

Tabela 5.2: Distribuição do número de PEAL pelo mês de julho de 2021

Semana	Frequência Relativa	Frequência Absoluta
1	20%	117632
2	20%	117632
3	25%	147040
4	35%	205855
Total	100%	588159

Utilizando o modelo de regressão linear piece-wise semanal, realize-se uma previsão da procura do material de acordo com o número de PEAL destas quatro semanas (tabela 5.3).

Tabela 5.3: Previsão da procura durante o mês de julho de 2021

Semana	Número de PEAL	Previsão da procura
1	117632	542.44
2	117632	542.44
3	147040	790.44
4	205855	1286.44

5.4.1 Modelo determinístico

Considere-se que a procura é constante ao longo do mês. Muito embora tenha sido prevista a procura para cada semana do mês de julho, utilize-se agora o valor médio do número de passageiros semanais. Este valor é de 147 040 passageiros por semana. Utilizando o modelo de regressão linear piece-wise semanal, o valor da procura semanal média do mês de julho é de 790.44 Kg/ semana.

Considere-se a unidade temporal de uma semana. Foi necessário estimar os custos unitário de aquisição, c_1 , de armazenagem, c_2 , e fixo de encomenda, A , para calcular o lote económico, Q^* , a duração do ciclo, T^* , e o custo total por semana, K^* .

Custo unitário de aquisição

Foi realizada uma recolha de informação para averiguar o preço médio do produto praticado nos supermercados. Conclui-se que o seu valor é de 50.82 €/Kg. Tenha-se como hipótese que a margem de lucro praticada nos supermercados da Europa continental varia entre 2% e 3%, [65]. Contudo lojas do setor natural, biológico e gourmet tendem a praticar margens de lucro de 5% e 10%. Tendo em conta que este é um produto encontrado com maior frequência em supermercados gourmet, considere-se que a margem de lucro para este produto é de 6.5%. Assim, considere-se que o custo de aquisição do produto, c_1 , por kilograma é de 47.52 €.

Custo de armazenagem

Uma vez que este é um produto congelado, sabe-se que este é armazenado em arcas frigoríficas. Ao contactar gestores logísticos dentro da mesma área, concluiu-se que a tarifa de armazenagem de produtos congelados é de 0.551€ diários por palete de produto. Assumindo que cada paleta armazena 500Kg de material em estudo, considere-se que o custo unitário semanal do material é de 7.714×10^{-3} €, como mostra a equação 5.11.

$$c_2 = \frac{0.551}{500} \cdot 7 = 7.714 \times 10^{-3} \quad (5.11)$$

Custo fixo de encomenda

Sabendo que 55% do custo de aquisição do produto corresponde ao custo associado à cadeia de abastecimento, então o seu valor é de 26.13 €. Sabendo que o custo fixo de encomenda está englobado no custo associado à cadeia de abastecimento – e que este último engloba também os custos de transporte, inventário e associado à procura do produto – considere-se que o custo fixo de encomenda corresponde a um quarto do custo associado à cadeia de abastecimento. Assim, tenha-se como hipótese que o custo fixo de encomenda, A , é de 6.53 €, como mostra a equação 5.12.

$$A = \frac{55\% \cdot 47.52}{4} = 6.53 \quad (5.12)$$

Construção do modelo

Tendo estimado os custos anteriores, o modelo determinístico de gestão de stocks construído é o seguinte:

$$Q^* = \sqrt{\frac{2 \cdot 6.53 \cdot r}{7.714 \times 10^{-3}}} \quad (5.13)$$

$$T^* = \frac{Q^*}{r} = \sqrt{\frac{2 \cdot 6.53}{r \cdot 7.714 \times 10^{-3}}} \quad (5.14)$$

$$K^* = \sqrt{2 \cdot 6.53 \cdot r \cdot 7.714 \times 10^{-3}} + 47.52 \cdot r \quad (5.15)$$

Sabendo que a taxa média de procura do mês de julho, $r = 790.44$ Kg, é possível calcular a quantidade ótima de material a encomendar, Q^* . Deste modo, os valores ótimo do lote económico, da duração do ciclo e do custo total por semana estão representados pela tabela 5.4.

Tabela 5.4: Valores ótimos do modelo determinístico

Número de PEAL	Taxa da procura	Q^*	T^*	K^*
147 040	790.44 Kg	1156.82 Kg	1 semana e 3 dias	37 570.63 €

5.4.2 Modelo de revisão periódica

Abordagem I

Construa-se um modelo de revisão periódica, tendo como ponto de partida a primeira semana de julho – período L . Assim, o valor esperado da procura para no período L é de 542.44 Kg. Utilize-se o algoritmo descrito na metodologia. Tenham-se em conta os valores de c_1 , c_2 e A já calculados em 5.4.1. Ora, uma vez que a amostra da procura semanal tem mais de $36 > 30$ observações, tenha-se como hipótese que a procura semanal segue a distribuição Normal de parâmetros $\mu = 62.12$ e $\sigma = 42.59$. Assim, o valor esperado da procura que segue esta distribuição é $D = 62.12$ Kg semanais. Este é o valor esperado da amostra da variável da procura semanal. É de notar que esta amostra é referente ao período de 1 de novembro a 30 junho. Ou seja, a maioria dos valores inserem-se na época baixa da aviação e por isso este valor é muito baixo. Considere-se $L = 1$ semana. Dê-se início ao algoritmo.

Passo 1

$$T^* = \sqrt{\frac{2\left(A + c_1 \int_R^\infty (x - R)f(z) dz\right)}{c_2 D}} = \sqrt{\frac{2A}{c_2 D}} = 5.22 \quad (5.16)$$

Passo 2

$$\alpha = \frac{T^* c_2}{c_1} = 8.47 \times 10^{-4} \quad (5.17)$$

Passo 3

$$R = D(T^* + L) + z_\alpha \sqrt{T^* + L} = DT^* + D_L + z_\alpha \sqrt{T^* + L} = 866.71 \quad (5.18)$$

Passo 4

Assumindo que a previsão da procura inicial é a previsão feita para a primeira semana de julho de 2021. Este passo foi realizado com recurso ao software R.

$$(C_T)^* = D + \frac{A}{T^*} + c_2 \left(R - D_L + \frac{DT^*}{2} \right) + \frac{c_2}{T^*} \int_R^\infty (x - R)f(z) dz = 41195.09 \quad (5.19)$$

Este algoritmo assume uma procura semanal média de 62.12 Kg, uma vez que é o valor médio da amostra existente da variável. Assim, o valor ótimo de tempo de ciclo é de 5 semanas e 1 dia. Este modelo calcula um nível de stock máximo de 866.71 Kg, uma vez que faz um ajuste da procura média através da procura esperada para o período escolhido. Assim, o custo total deste ciclo é de 41 195.09€, dando um custo total por semana de 7 891.78€.

Este ajuste não é ideal uma vez que a procura média assumida é muito reduzida. Isto deve-se à hipótese de D ser um valor tão baixo.

Abordagem II

Construa-se agora um modelo de revisão periódica, considerando o valor esperado da procura semanal, $D = 790.44$ Kg, o valor da procura média do mês de julho.

Ou seja, considere-se uma diferenciação de D , consoante o período de previsão seja de época alta ou baixa. Considerem-se as restantes variáveis iguais aos valores usados para a tentativa 1. Dê-se início ao algoritmo.

Passo 1

$$T^* = \sqrt{\frac{2(A + c_1 \int_R^\infty (x - R)f(z) dz)}{c_2 D}} = \sqrt{\frac{2A}{c_2 D}} = 1.46 \quad (5.20)$$

Passo 2

$$\alpha = \frac{T^* c_2}{c_1} = 2.38 \times 10^{-4} \quad (5.21)$$

Passo 3

$$R = D(T^* + L) + z_\alpha \sqrt{T^* + L} = DT^* + D_L + z_\alpha \sqrt{T^* + L} = 1696.48 \quad (5.22)$$

Passo 4

Assumindo que a previsão da procura inicial a previsão feita para a primeira semana de julho de 2021. Este passo foi realizado com recurso ao software R.

$$(C_T)^* = D + \frac{A}{T^*} + c_2 \left(R - D_L + \frac{DT^*}{2} \right) + \frac{c_2}{T^*} \int_R^\infty (x - R)f(z) dz = 80632.25 \quad (5.23)$$

Assumindo agora uma procura média maior, o nível de stock máximo aumenta e o tempo ótimo de ciclo diminui. Assim, o valor ótimo de tempo de ciclo é de 1 semana e 3 dias. O nível de stock máximo de 1696.48 Kg. Assim, o custo total deste ciclo é de 80 632.25€, dando um custo total por semana de 55 227.57€. Este valor é muito superior que o valor obtido na tentativa 1, mas é mais realista.

CONCLUSÕES

6.1 Conclusões

Os resultados obtidos ao longo do estudo estão condicionados a um intervalo de tempo relativamente estreito. Muito embora houvesse informação relativa ao material e ao número de passageiros desde setembro de 2020 a junho de 2021, este elevado número de observações diárias não é suficiente para tirar conclusões sólidas acerca do padrão e sazonalidade das respectivas séries temporais. Por isso, o comportamento sazonal do material em estudo não é claro ao longo do ano. Para além disso, a diferença do comportamento do número de passageiros e da procura do material durante os meses de setembro e outubro de 2020 não é explícita e pode-se dever a um número elevado de fatores. Esses fatores podem incluir a sazonalidade do produto e seria interessante poder comprovar isso, tendo acesso a dados de setembro e outubro de futuros anos. Por isso, a informação relativa a esses dois meses não foi considerada, tornando o número de observações para modelação mais reduzido. Isto também apresenta uma limitação, uma vez que quanto maior é o número de observações das variáveis em estudo, mais precisos serão as modelações e previsões encontrados.

Contudo, é de notar que através da informação obtida durante 10 meses, a autocorrelação da série temporal da procura de material é de 0.67. Este valor reflete uma relação, embora fraca, entre as observações ao longo do tempo. A agregação dos dados em períodos semanais foi essencial, uma vez que eliminou o ruído branco existente na série diária. Esta agregação tornou mais perceptível a presença de picos da procura, tais como o final de outubro, a época do Natal e o início do verão de 2021, onde se registou a maior procura de material.

A série temporal do número de passageiros apresenta os mesmos picos que a série de procura de material, com exceção dos dois primeiros meses. A sua autocorrelação é mais elevada (0.857). A agregação dos dados mensalmente traz uma perspetiva interessante, uma vez que demonstra a semelhança do comportamento das duas variáveis, com exceção dos meses de setembro e outubro. Esta semelhança disputou a modelação da procura, tendo como variável explicativa o número de PEAL.

A modelação da procura agregada semanalmente apresentou resultados bastante fidedignos. Independentemente dos métodos de regressão utilizados, todos os modelos apresentaram um valor de R^2 superior a 0.5. O modelo linear generalizado (tentativa 3) não cumpre com os pressupostos necessários sendo por isso descartado. Foi realizada uma avaliação dos modelos de acordo com o seu valor de R^2 , a significância da variável explicativa, o comportamento estatístico dos seus resíduos e o valor de AIC. Conclui-se que o melhor é o modelo de regressão linear piece-wise, com R^2 de 0.9255, significância da variável explicativa superior a 95%, os resíduos têm um comportamento Normal com valor médio de 0 e com o menor valor de AIC.

A modelação da procura diária não apresenta resultados proveitosos uma vez que os valores de R^2 das várias regressões são inferiores a 0.5. A modelação diária não traz vantagens e contextualizando com o caso de estudo, é mais interessante fazer um planeamento semanal do abastecimento do armazém do que um planeamento diário. Logo, tomou-se como unidade de tempo a semana. No futuro, seria interessante verificar se, tendo acesso a dados de períodos de tempo mais alargados, os resultados das modelações diárias seriam melhorados.

Para a construção das políticas de stocks realizaram-se previsões da procura para as semanas de julho de 2021, tendo em conta o número de passageiros embarcados no aeroporto de Lisboa durante o mês de julho. Estes dados podem ser consultados pelo INE. Admitindo uma distribuição não uniforme do número de passageiros ao longo do mês, as previsões foram obtidas através do modelo de regressão linear piece-wise. Assim, foram construídas duas políticas de stocks. A primeira considera uma procura determinística, com reposição instantânea e rotura não permitida. A segunda considera uma procura variável com revisão periódica. Foi necessário estimar os custos unitário de aquisição, de armazenagem e fixo de encomenda. Estes custos são essenciais para a construção de ambos modelos. De modo a obter resultados mais precisos do que estimativas, seria ideal conhecer os valores exatos destes custos, contudo essa informação não foi disponibilizada.

O modelo determinístico foi construído com base numa procura constante durante o mês de julho. Assim, este é um modelo muito menos preciso uma vez que não assume as variâncias do número passageiros ao longo do tempo. O resultado tem menor qualidade que o resultado do modelo de revisão periódica. O comprimento do ciclo obtido é de 1 semana e 3 dias, ou seja, o modelo assume que a procura é constante durante este período. Para além disso o custo por unidade de tempo é de 37 570.63€.

O modelo de revisão periódica foi construído através do algoritmo de Hadley-Whitin. Foram realizadas duas tentativas. A primeira assume uma procura média baseada na amostra da variável procura. Esta é ideal para modelar os meses de inverno, uma vez que grande parte das observações são de semanas de época baixa. O seu custo é muito menor, uma vez que os níveis de stock máximos necessários são mais baixos. Os tempos de ciclo muito mais longos também. Contudo, este modelo não é realista, porque o nível máximo de stock é demasiado baixo para satisfazer procuras de época alta.

A segunda tentativa contempla um valor médio da procura baseado no mês de julho – mês de época alta na aviação. Os seus resultados são melhores, embora o seu custo tenha aumentado significativamente – 80 632.25€ por semana. O tempo de ciclo é menor.

Uma vez que o material em estudo é um produto alimentar congelado, com validade de 6 meses, é preferível assumir um nível máximo de stock mais elevado – 1 696.48 Kg, como sugerido na tentativa 2 do modelo de revisão periódica. A validade do produto é extensa o suficiente e assim a rotura de stock é evitada. Este nível de stock de segurança permite solucionar o problema em causa e reduzir os custos associados a compras extraordinárias, não planeadas, que eram feitas para satisfazer a procura.

6.2 Desenvolvimentos futuros

A continuação deste trabalho parte pela agregação de novos dados dos movimentos do material no armazém e juntá-los com os já existentes. Quanto maior a dimensão desta série temporal, mais claras se tornam as sazonalidades da mesma e a sua relação com a série temporal do número de passageiros. Deste modo, melhor serão os ajustes dos modelos de regressão.

Seria também interessante explorar a diversidade dos modelos de políticas de gestão de inventário e stocks. Analisando as vantagens e desvantagens de cada modelo, tendo em conta a natureza da procura, sendo estes determinísticos ou de procura variável – contínua ou discreta. Exemplos destes modelos são os modelos determinísticos com reposição não instantânea e rotura permitida e os modelos de revisão contínua para procura discreta ou contínua. Para além disso, o acesso ao valor real dos custos da cadeia de abastecimento traria uma maior precisão ao caso de estudo, de forma a reduzir o erro.

Seria ainda interessante expandir este estudo para um maior número de materiais utilizados na área produção da Cateringpor e averiguar se o modelo aqui construído pode ser generalizado para outros materiais e analisar se a procura dos vários materiais é independente entre si. Assim, dar-se-ia uma otimização global da gestão de armazéns da empresa.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] *The impact of COVID-19 on aviation*. <https://airlines.iata.org/news/the-impact-of-covid-19-on-aviation>. Consultado: 17-09-2021 (ver p. 1).
- [2] M. (1.) Jones et al. “In-flight test campaign to validate PIO detection and assessment tools.” Em: *Aerospace* 7.9 (2020). ISSN: 22264310 (ver p. 1).
- [3] S. Kot, K. Grondys e R. Szopa. “Theory of inventory management based on demand forecasting”. Em: *Polish Journal of Management Studies* 3 (out. de 2014), pp. 148–156 (ver p. 1).
- [4] *ProcessWare Group*. <http://www.processwaregroup.com>. Consultado: 18-10-2021 (ver pp. 1, 23).
- [5] S. Ravi. “Forecasting.” Em: *Salem Press Encyclopedia of Science* (2020) (ver p. 5).
- [6] R. Hyndman e G. Athanasopoulos. *Forecasting: Principles and Practice*. English. 2nd. Australia: OTexts, 2018 (ver pp. 5, 12).
- [7] G. Sun et al. “Short-term wind power forecasts by a synthetical similar time series data mining method.” Em: *Renewable Energy* 115 (2018), pp. 575–584. ISSN: 0960-1481 (ver p. 5).
- [8] G. Athanasopoulos et al. “Forecasting with temporal hierarchies.” Em: *European Journal of Operational Research* 262.1 (2017), pp. 60–74. ISSN: 0377-2217 (ver p. 5).
- [9] M. Umar e M. Denai. *Data Modeling and Forecasting Using Artificial Intelligence*. Lap Lambert Academic Publishing GmbH KG, 2014. ISBN: 9783659668593 (ver p. 5).
- [10] R. A. Wienclaw. “Time Series.” Em: *Salem Press Encyclopedia* (2021) (ver p. 6).
- [11] A. Bovas e L. Johannes. *Statistical Methods for Forecasting*. Wiley Series in Probability and Mathematical Statistics, Applied Probability and Statistics Applied probability and statistics. Wiley, 1983. ISBN: 9780471867647 (ver p. 6).

- [12] N. R. Draper e H. Smith. “Multiple Regression Applied to Analysis of Variance Problems.” Em: *Applied Regression Analysis* (1998), pp. 473–504. ISSN: 9780471170822. URL: <https://widgets.ebscohost.com/prod/customerspecific/ns000290/authentication/index.php?url=https%3a%2f%2fsearch.ebscohost.com%2fflogin.aspx%3fdirect%3dtrue%26db%3dedb%26AN%3d103046226%26lang%3dpt-pt%26site%3dedb-live%26scope%3dsite> (ver p. 6).
- [13] J. Ameet V. *Machine Learning and Artificial Intelligence*. Springer, 2020. ISBN: 9783030266219 (ver p. 6).
- [14] T. Dietterich. “Machine-learning research: Four current directions.” Em: *AI Magazine* 18.4 (1997), pp. 97–136. ISSN: 07384602 (ver p. 6).
- [15] M. Vidyasagar. “Recent Advances in Reinforcement Learning”. Em: *2020 American Control Conference (ACC)*. 2020, pp. 4751–4756. DOI: 10.23919/ACC45564.2020.9147512 (ver p. 6).
- [16] A. Sen e M. Srivastava. *Regression Analysis: Theory, Methods and Applications*. First. Springer - Verlag New York Inc., 1990. ISBN: 0-387-97211-0 (ver pp. 6, 14).
- [17] P. Julian, A. Desages e B. D’Amico. “Orthonormal high-level canonical PWL functions with applications to model reduction.” Em: *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on, IEEE Trans. Circuits Syst. I* 47.5 (2000), pp. 702–712. ISSN: 1057-7122 (ver pp. 7, 16).
- [18] L. Yang et al. “Mathematical programming for piecewise linear regression analysis”. Em: *Expert Systems with Applications* 44 (fev. de 2016), pp. 156–167. ISSN: 09574174. DOI: 10.1016/j.eswa.2015.08.034 (ver p. 7).
- [19] W. Tony. *Best Practice in Inventory Management*. Vol. 3 Edition. Routledge, 2018. ISBN: 9781138294424 (ver p. 7).
- [20] B.-C. Patricia Esperanza e S.-G. Gabriel De Las Nieves. *Problem Solving In Operation Management*. Vol. 1st ed. 2021. Springer, 2020. ISBN: 9783030500887 (ver p. 7).
- [21] E. D. Davison. “Logistics Management.” Em: *Salem Press Encyclopedia* (2021) (ver p. 7).
- [22] I. Lapinskaitė e J. Kuckailytė. “The Impact of Supply Chain Cost on the Price of the Final Product”. Em: *Business, Management and Education* 12 (mai. de 2014), pp. 109–126. DOI: 10.3846/bme.2014.08 (ver p. 7).
- [23] C. Gary, P. Terry e K. Tom. *Supply Chain Costing and Performance Management*. Vol. Second edition. Wiley, 2021. ISBN: 9781119793632 (ver p. 7).
- [24] J. M. Cândia. “Análise e otimização da gestão de stocks numa empresa industrial”. Fev. de 2019 (ver p. 8).
- [25] H. I. Alves. “Estudo e implementação de modelos de gestão de stocks em sistemas ERP para empresas retalhistas”. Universidade do Minho, 2015 (ver p. 8).

- [26] A. R. Gonzaga. “Gestão de Stocks e Armazém de Matérias-Primas”. Instituto Politécnico do Porto. Instituto Superior de Contabilidade e Administração do Porto, 2014 (ver p. 8).
- [27] T. Damron, W. T. Rupp e A. D. Smith. “Inventory control in the retail sector: case studies of best business practices”. Em: *International Journal of Procurement Management* 9.3 (2016), pp. 354–371. DOI: 10.1504/IJPM.2016.076309. eprint: <https://www.inderscienceonline.com/doi/pdf/10.1504/IJPM.2016.076309> (ver p. 8).
- [28] C. D. Lewis e M. Institute of Operations. *Demand Forecasting and Inventory Control*. Routledge, 1997. ISBN: 9781855732414 (ver p. 8).
- [29] D. J. Jakobczak. *Analyzing Risk Through Probabilistic Modeling in Operations Research*. Advances in Logistics, Operations, and Management Science (ALOMS) Book Series. IGI Global, 2016. ISBN: 9781466694583 (ver p. 8).
- [30] B. Gore. “Cases in Operations Management: Building Customer Value Through World-Class Operations.” Em: *Business Communication Quarterly* 69.2 (2006), 232–237. ISSN: 10805699 (ver p. 8).
- [31] G. E. Cantarella et al. *Stochastic equilibrium assignment with variable demand: Literature review, comparisons and research needs*. 2013. DOI: 10.2495/UT130271 (ver p. 8).
- [32] G. E. Cantarella, A. Cartenì e S. D. Luca. “Stochastic equilibrium assignment with variable demand: Theoretical and implementation issues”. Em: *European Journal of Operational Research* 241 (2 mar. de 2015), pp. 330–347. ISSN: 03772217. DOI: 10.1016/j.ejor.2014.09.010 (ver p. 8).
- [33] T. Düppe. “Equilibrium: History of the Concept”. Em: *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*. Ed. por J. D. Wright. Second Edition. Oxford: Elsevier, 2015, pp. 912–917. ISBN: 978-0-08-097087-5 (ver p. 8).
- [34] P. Berling e J. Marklund. “Multi-echelon inventory control: An adjusted normal demand model for implementation in practice”. Em: *International Journal of Production Research* 52 (11 jun. de 2014), pp. 3331–3347. ISSN: 1366588X. DOI: 10.1080/00207543.2013.873555 (ver p. 8).
- [35] J. Moors e L. Strijbosch. “Exact fill rates for (R; s; S) inventory control with gamma distributed demand”. Em: *Journal of the Operational Research Society* 53 (2002), pp. 1268–1274. DOI: 10.1057=palgrave.jors.2601441 (ver p. 8).
- [36] S. Axsäter. *Inventory Control*. Vol. 2nd ed. International Series in Operations Research and Management Science Vol. 90. Springer Science Business Media, 2006. ISBN: 9780387332505 (ver p. 9).

- [37] R. David Frederick. *Distribution Planning and Control : Managing in the Era of Supply Chain Management*. Vol. Third edition. Online Access with Purchase: Springer (t). Springer, 2015. ISBN: 9781489975775 (ver p. 9).
- [38] N. Steven. *Perishable Inventory Systems*. International Series in Operations Research and Management Science v. 160. Springer, 2011. ISBN: 9781441979988 (ver p. 9).
- [39] T. A. Burgin. “The Gamma Distribution and Inventory Control”. Em: *Operational Research Quarterly (1970-1977)* 26.3 (1975), pp. 507–525. ISSN: 00303623. URL: <http://www.jstor.org/stable/3008211> (ver p. 9).
- [40] G. Hadley e T. M. Whitin. *Analysis of inventory systems*. Rel. téc. 1963 (ver p. 9).
- [41] P. Cunha, F. Raupp e F. Oliveira. “A two-stage stochastic programming model for periodic replenishment control system under demand uncertainty.” Em: *Computers and Industrial Engineering* 107 (2017), pp. 313–326. ISSN: 0360-8352 (ver p. 9).
- [42] M. Permatasari, A. Y. Ridwan e B. Santosa. *PROPOSAL OF PERIODIC INVENTORY REVIEW POLICY FOR IRREGULAR DEMAND IN A CASE STUDY OF PT ABC DURI-RIAU 1*. 2017 (ver p. 9).
- [43] T. Hill e P. Lewicki. *Statistics: Methods and Applications, A Comprehensive Reference for Science, Industry and Data Mining*. First. StatSoft, Inc., 2006. ISBN: 1-884233-59-7 (ver p. 12).
- [44] D. A. Dickey e W. A. Fuller. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root”. Em: *Journal of the American Statistical Association* 74.366a (1979), pp. 427–431. DOI: 10.1080/01621459.1979.10482531 (ver p. 13).
- [45] A. Christoph. *Information Measures : Information and Its Description in Science and Engineering*. Vol. 2001. Signals and Communication Technology. Springer, 2012. ISBN: 9783540416333 (ver p. 15).
- [46] D. Annette J. e B. Adrian G. *An Introduction to Generalized Linear Models*. Vol. First edition. Texts in Statistical Science. Chapman e Hall/CRC, 2002. ISBN: 9781138741515 (ver p. 17).
- [47] F. Hillier. *Introduction to Operations Research*. McGraw-Hill Education, 2009. ISBN: 9780077298340 (ver p. 20).
- [48] Y.-J. Lin. “A stochastic periodic review integrated inventory model involving defective items, backorder price discount, and variable lead time.” Em: *4OR* 8.3 (2010), pp. 281–297. ISSN: 16142411 (ver p. 21).
- [49] E. Gonzalez-Estrada e J. A. Villasenor-Alva. *gof t: Tests of Fit for some Probability Distributions*. R package version 1.3.6. 2020. URL: <https://CRAN.R-project.org/package=gof t> (ver p. 22).
- [50] W. N. Venables e B. D. Ripley. *Modern Applied Statistics with S*. Fourth. ISBN 0-387-95457-0. New York: Springer, 2002 (ver p. 22).

-
- [51] M. L. Delignette-Muller e C. Dutang. “fitdistrplus: An R Package for Fitting Distributions”. Em: *Journal of Statistical Software* 64.4 (2015), pp. 1–34 (ver p. 22).
- [52] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN: 978-3-319-24277-4 (ver p. 22).
- [53] J. Pasek et al. *weights: Weighting and Weighted Statistics*. R package version 1.0.1. 2020 (ver p. 22).
- [54] Lesnoff et al. *aod: Analysis of Overdispersed Data*. R package version 1.3.1. 2012 (ver p. 22).
- [55] A. Zeileis, D. Meyer e K. Hornik. “Residual-based Shadings for Visualizing (Conditional) Independence”. Em: *Journal of Computational and Graphical Statistics* 16.3 (2007), pp. 507–525 (ver p. 22).
- [56] H. Wickham e J. Bryan. *readxl: Read Excel Files*. R package version 1.3.1. 2019 (ver p. 22).
- [57] R. Hyndman. *fpp2: Data for "Forecasting: Principles and Practice" (2nd Edition)*. R package version 2.4. 2020 (ver p. 22).
- [58] J. Schafer et al. *corpcor: Efficient Estimation of Covariance and (Partial) Correlation*. R package version 1.6.9. 2017 (ver p. 22).
- [59] M. Papadakis et al. *Rfast: A Collection of Efficient and Extremely Fast R Functions*. R package version 2.0.1. 2020 (ver p. 22).
- [60] A. Ehrhardt. *scoring: Credit Scoring tools*. 2018 (ver p. 22).
- [61] T. Wei e V. Simko. *R package 'corrplot': Visualization of a Correlation Matrix*. (Version 0.90). 2021 (ver p. 22).
- [62] Gilbert, P. D. *Brief User's Guide: Dynamic Systems Estimation*. 2006 or later (ver p. 22).
- [63] V. M. Muggeo. “Interval estimation for the breakpoint in segmented regression: a smoothed score-based approach.” Em: *Australian & New Zealand Journal of Statistics* 59 (2017), pp. 311–322 (ver p. 22).
- [64] A. Trapletti e K. Hornik. *tseries: Time Series Analysis and Computational Finance*. R package version 0.10-48. 2020 (ver p. 22).
- [65] H. Neil. *European Business*. Red Globe Press, 1999. ISBN: 9780333754078 (ver p. 48).

