

MARIA AMÉLIA ROCHA DA CUNHA

**Análises de regressão para determinação da viabilidade de estoque de segurança para
produtos lançamentos em uma empresa do ramo de cosméticos**

Maria Amélia Rocha da Cunha

Análises de regressão para determinação da viabilidade de estoque de segurança para produtos lançamentos em uma empresa do ramo de cosméticos

Trabalho de Graduação apresentado ao Conselho de Curso de Graduação em Engenharia Mecânica da Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista, como parte dos requisitos para obtenção do diploma de Graduação em Engenharia Mecânica.

Orientador: José Roberto Dale Luche

C972a	Cunha, Maria Amélia Rocha da Análises de regressão para determinação da viabilidade de estoque de segurança para produtos lançamentos em uma empresa do ramo de cosméticos / Maria Amélia Rocha da Cunha. – Guaratinguetá, 2018. 47 f : il. Bibliografia: f. 45-47 Trabalho de Graduação em Engenharia Mecânica – Universidade Estadual Paulista, Faculdade de Engenharia de Guaratinguetá, 2018. Orientador: Prof. Dr. José Roberto Dale Luche 1. Controle de estoque. 2. Análise de regressão. 3. Oferta e procura 4. Administração de produtos I. Título CDU 658.78
-------	---

Pâmella Benevides Gonçalves
Bibliotecária/CRB-8/9203

MARIA AMÉLIA ROCHA DA CUNHA

ESTE TRABALHO DE GRADUAÇÃO FOI JULGADO ADEQUADO COMO
PARTE DO REQUISITO PARA A OBTENÇÃO DO DIPLOMA DE
"GRADUADO EM ENGENHARIA MECÂNICA"

APROVADO EM SUA FORMA FINAL PELO CONSELHO DE CURSO DE
GRADUAÇÃO EM ENGENHARIA MECÂNICA

Prof. Dr. Marcelo Antônio Martins

Coordenador

BANCA EXAMINADORA:

Prof. Dr. José Roberto Dato Lucchese

Orientador UNESP-FEG

Prof. Dr. Anselmo Francisco da Silva

UNESP-FEG

Moisés Stefano Petros de Oliveira

Membro Externo

AGRADECIMENTOS

Agradeço à minha família por toda colaboração e apoio oferecidos durante a realização deste trabalho e durante todos os anos do curso de graduação. Em especial agradeço à minha mãe, que sempre batalhou muito para me proporcionar condições para que eu pudesse me desenvolver e que é motivo de muito orgulho para mim.

Agradeço aos amigos e ao meu parceiro. Sou imensamente grata pela cooperação, pela paciência e pelo incentivo. Nos momentos de dificuldades, sempre me incentivaram e reforçaram a capacidade que eu tinha para concluir todas as etapas.

Agradeço aos professores e à universidade por cada ensinamento, oportunidade e experiências que contribuíram não só com aprendizados técnicos, mas também com aprendizados pessoais que levarei para toda a vida.

Por fim, agradeço sobretudo a Deus, pela sabedoria e oportunidades concedidas, além da força para conquistar muito além do que eu imaginava ser capaz.

RESUMO

A previsão da demanda de um produto com características inovadoras pode não ser assertiva, principalmente no período de lançamento do produto, devido ao alto risco de variações ocasionado pela imprevisibilidade do comportamento de vendas e da aceitação do público. Empresas que possuem um portfólio que sofre constantes alterações, com produtos sendo lançados e descontinuados em um curto período de tempo, são as que mais sofrem com estimativa incorreta de demanda, pois variações para cima representam vendas perdidas e variações para baixo são excesso que pode nunca mais ser convertido em venda. Com isso, surge o dilema sobre qual quantidade de insumos deve ser comprada para produzir o mais próximo possível do volume requerido. Geralmente, essa decisão deve ser tomada com certa antecedência devido ao alto lead time dos insumos. O objetivo deste trabalho é estimar, através de análises de regressão, se existe relação entre a variação de demanda de um novo produto e determinados fatores prováveis de influência, para posteriormente identificar se é possível determinar com maior precisão para quais casos é viável a criação de um *Safety Stock* de materiais e produtos.

PALAVRAS-CHAVE: Lançamento de produtos. Variações de demanda. Modelos de Análise de regressão. Viabilidade de *Safety Stock*.

ABSTRACT

The demand estimation of a product with inovators characteristics might not be assertive, mostly in the product launch period, due to a high variation risk caused by the unpredictability sales behavior and by the public acceptance. Companies who have a portifolio that suffer constant changes, with products being launched and discontinued in a short period of time, are the most impacted by the uncorrect demand estimate, because upside and downside variations represent lost sales and excess that may never be converted in sales, respectively. Therefore, there is a question about the quantity that should be purchased to produce as close as possible to the amount requested. In general, this decision shall be taken in advance due to the long lead time of the materials. The aim of the present study is estimate, through regression analysis, if there is correlation between demand variation of new products and certain influence factors, in order to determine for which cases may be viable to create a safety stock.

KEYWORDS: Products launch. Demand variations. Regression analysis. Safety Stock.

SUMÁRIO

1	INTRODUÇÃO	07
1.1	OBJETIVOS	09
1.1.1	Gerais.....	09
1.1.2	Específicos	09
1.2	JUSTIFICATIVAS	09
1.3	MATERIAIS E MÉTODOS.....	10
2	FUNDAMENTAÇÃO TEÓRICA.....	11
3	ESTUDO DE CASO	15
3.1	LANÇAMENTO DE NOVOS PRODUTOS	15
3.1.1	Etapas de criação de um produto.....	15
3.1.2	Implementação dos projetos de novos produtos.....	17
3.2	CADEIA DE SUPRIMENTOS	17
3.2.1	Estimativa de demanda	18
3.2.2	Eventos de variações de demanda.....	19
3.2.3	Fluxo de planejamento da produção de novos produtos.....	21
3.2.4	Impacto das variações de demanda na produção	22
4	MODELOS DE ANÁLISE DE REGRESSÃO	23
4.1	REGRESSÃO LINEAR MULTIVARIADA	24
4.1.1	Modelo e técnica de estimação dos parâmetros da regressão multivariada.....	24
4.1.2	Aplicação: estimativa da variação de demanda de acordo com as características do novo produto	26
4.2	REGRESSÃO LOGÍSTICA	30
4.2.1	Modelo e técnica de estimação dos parâmetros da regressão logística.....	31
4.2.2	Aplicação: estimativa do nível de risco de variação de demanda de acordo com as características do novo produto.....	34
5	CONCLUSÃO.....	44
	REFERÊNCIAS	45
	BIBLIOGRAFIA CONSULTADA.....	47

1 INTRODUÇÃO

Equilibrar o risco de excesso em inventário com o risco de não atendimento é um dilema vivido por muitas empresas. Na maioria dos casos, os processos produtivos não são capazes de fornecer resposta instantânea à demanda. Por isso, a produção precisa ser acionada antes de se ter um conhecimento absoluto das quantidades e da variedade de produtos que serão solicitados, o que torna essencial a realização de algum tipo de previsão (PEINADO e GRAEML, 2007). Ainda segundo Peinado e Graeml (2007), existem quatro grandes modelos de previsão de demanda amplamente utilizados pelas empresas, que serão apresentados no capítulo 2.

No estudo de caso deste trabalho, a estimativa sobre quanto um novo produto irá vender é feita com base em produtos similares e resulta da combinação de parâmetros cujos pesos são dados considerando-se semelhança com comportamentos observados. São muitos os fatores que exercem influência sobre o número final e os métodos de estimativa utilizados ainda não conseguem prever com exatidão esse número, devido à aleatoriedade também presente no processo.

O grande problema da inexatidão do valor demandado são as consequências que essa variação acarreta para a companhia, devido ao seu modelo de negócio que não consegue reagir e absorver bem esses impactos. O portfólio de produtos ofertados varia constantemente, de acordo com o que estrategicamente pode trazer mais lucro. Se um novo produto não apresentar a performance esperada no período de lançamento, que é o período de pico de vendas da vida do produto, o excesso de material resultante pode ser difícil de ser exaurido e permanecer meses gerando impacto negativo. Mesmo um produto com vendas acima do esperado também pode trazer danos ao balanço da companhia, pois se não houver material para suprir a necessidade, o cliente ficará insatisfeito, além de contar como venda perdida. Como não há tempo hábil para adequar a produção à demanda final requerida, devido ao curto período de duração da campanha e ao alto lead time da maioria dos insumos, uma análise de riscos pode ser útil para auxiliar na tomada de decisão e minimizar os impactos gerados por essas imprecisões. Disso, surge a ideia de entender se é possível prever fatores que influenciam na maior probabilidade de variação de demanda para identificar para quais casos compensa investir em um estoque de segurança. (CHOPRA e MEINDL, 2003).

Existem métodos estatísticos capazes de determinar se existe relação entre duas ou mais variáveis, como as análises de correlação e regressão. Segundo Stevenson (1986), “A correlação mede a força, ou grau, de relacionamento entre duas variáveis; a regressão dá uma

equação que descreve o relacionamento em termos matemáticos”. De forma geral, a regressão compreende a análise de dados amostrais para saber como duas ou mais variáveis estão relacionadas uma com a outra e tem como resultado uma equação matemática que descreve o relacionamento. A equação pode ser usada para estimar, ou predizer, valores futuros de uma variável quando se conhecem ou se supõem conhecidos valores da outra variável.

Percebe-se que o método de análise de regressão pode trazer respostas a alguns questionamentos levantados neste estudo. Afinal, se a variação de demanda é o que preocupa (por causa da falta de capacidade para reagir a ela) e é o que gera a necessidade de estudo sobre a viabilidade de *Safety Stock*, nada melhor do que entender o que pode gerar essa variação e tentar identificar correlações para prever comportamentos futuros.

Ao longo deste trabalho, serão detalhadas as principais informações para entendimento do problema apresentado e serão mostradas as metodologias de estudo bem como os resultados obtidos.

1.1 OBJETIVOS

1.1.1 Gerais

Estimar, através de análises de regressão, se existe relação entre a variação de demanda de um produto e determinados fatores prováveis de influência para posterior determinação de viabilidade de estoque de segurança para lançamentos de novos produtos em uma empresa do ramo de Cosméticos.

1.1.2 Específicos

- Identificar os fatores que podem influenciar na variação de demanda;
- Identificar o método e a análise estatística mais apropriada para o estudo;
- Verificar se existe correlação entre as variáveis analisadas;
- Determinar se é possível propor a criação de *Safety Stock* com base nas informações obtidas no estudo.

1.2 JUSTIFICATIVAS

Empresas de bens de consumo geralmente enfrentam dificuldades para prever com exatidão o volume estimado de vendas e conseguir se planejar para atendê-lo. Entre muitas razões, isso ocorre devido ao tipo de negócio e ao fato de ser mais propenso a variações. Quando se trata de novos produtos, essas variações podem ser ainda mais evidentes e, por isso, um estudo para identificação de riscos pode auxiliar na tomada de decisão e, conseqüentemente, minimizar os impactos negativos. Dessa forma, uma boa alternativa seria entender se é possível prever fatores que influenciam na maior probabilidade de variação de demanda e, assim, identificar para quais casos compensa investir em um estoque de segurança.

1.3 MATERIAIS E MÉTODOS

Para determinar a possibilidade de prever uma variação de demanda mais susceptível a acontecer de acordo com certas características do produto a ser lançado, foi feito um levantamento com os dados dos lançamentos dos últimos 3 anos da empresa em estudo. As características analisadas foram: tipo do produto, tipo de inovação e grau de exposição. Os dados foram dispostos em uma tabela no excel e as análises foram feitas com o auxílio dos softwares Excel e SPSS Statistics. Para adequação de variáveis qualitativas aos modelos utilizados, foi inserido o conceito de variáveis dummies, que consiste em um recurso para mensurar variáveis qualitativas e possibilitar análises.

Inicialmente, foi utilizado o método de análise de regressão linear multivariada para determinar a possibilidade de estimar um valor X de variação de demanda. Como não houve aderência a esse modelo, foi verificado se pelo menos existia grau de relação confiável entre a variação de demanda e essas mesmas características dos produtos. Neste caso a análise não resultaria em uma estimativa numérica para a variação, mas poderia mostrar se as variáveis independentes predizem risco alto ou baixo de variação, por exemplo. Este segundo caso foi determinado através da regressão logística,

As técnicas utilizadas, bem como embasamento teórico, métodos de análises e resultados obtidos estão devidamente expostos no Capítulo 4 deste trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

A estratégia competitiva de uma empresa define o conjunto de necessidades do consumidor que ela pretende satisfazer por meio de seus produtos e serviços. A cadeia de valor de uma empresa começa com o desenvolvimento de novos produtos. Marketing e Vendas geram a demanda, divulgando as prioridades dos clientes ao qual os produtos e serviços deverão satisfazer. Utilizando as especificações do novo produto, as operações transformam os *inputs* de Marketing em *outputs* para a produção da mercadoria. A estratégia de cadeia de suprimentos determina a natureza da obtenção de matérias-primas, o transporte de materiais *Inbound* e *Outbound*, fabricação do produto e distribuição do produto ao consumidor (CHOPRA e MEINDL, 2003).

Uma cadeia de suprimentos engloba todos os estágios envolvidos no atendimento de um pedido de um cliente. É uma sequência de processos que acontecem entre diferentes estágios da cadeia, e que se combinam para atender à necessidade de um cliente. As atividades da cadeia de suprimento iniciam-se com o pedido de um cliente e terminam com o recebimento do produto pelo cliente. Os processos realizados na cadeia de suprimentos podem ser visualizados de duas maneiras (CHOPRA e MEINDL, 2003):

- Visão cíclica, onde os processos são divididos em uma série de ciclos. São eles: ciclo de pedido do cliente, ciclo de reabastecimento, ciclo de fabricação e ciclo de suprimentos;
- Visão *push/pull* (empurrados/puxados), onde os processos são divididos em duas categorias: acionados em resposta ao pedido dos clientes (*pull*) ou em antecipação aos pedidos dos clientes (*push*).

Todos os processos da cadeia de suprimentos recaem em uma das categorias (*push/pull*), dependendo do tempo de sua execução compatível com a demanda do cliente. Nos processos *pull*, a execução é iniciada em resposta aos pedidos dos clientes e assim, a demanda é conhecida com certeza. Os processos *push* são aqueles executados em antecipação aos pedidos dos clientes, assim, neste caso, a demanda não é conhecida e deve ser prevista (CHOPRA e MEINDL, 2003).

A previsão da demanda futura é a base de todas as decisões estratégicas e de planejamento em uma cadeia de suprimentos. Como citado anteriormente, todos os processos *push* são desempenhados em antecipação à demanda do cliente, demandando um planejamento do nível de produção. Os processos *pull* ocorrem em resposta à demanda do cliente, requerendo planejamento do nível de capacidade que será disponibilizada. Em ambos os casos, a primeira medida a ser tomada pelos responsáveis pela cadeia de suprimentos é

prever qual será a demanda dos clientes no futuro (CHOPRA e MEINDL, 2003). Segundo Peinado e Graeml (2007), existem três grandes modelos estáticos e um modelo dinâmico de previsão de demanda amplamente utilizados pelas empresas:

- Os modelos qualitativos, que são subjetivos e apropriados quando não existem dados históricos para serem analisados como base para a previsão. Os principais modelos qualitativos de previsão de demanda são: predição, opiniões de executivos, método Dephi, opiniões da equipe de vendas, pesquisas de mercado e analogia com produtos similares.

- Os modelos de decomposição de séries temporais se baseiam no estudo estatístico da demanda acontecida no passado para projetar a demanda futura. Dentre estes modelos tem-se: os modelos baseados na média (média móvel ponderada ou com suavização exponencial), que devem ser aplicados apenas a demandas que não apresentem tendência ou sazonalidade e os modelos de regressão linear, utilizados para demandas que apresentam tendência mas não apresentam sazonalidade.

- O modelo do ajustamento sazonal pode ser aplicado para séries temporais de demandas que apresentam nível, tendência e sazonalidade.

- O modelo dinâmico de previsão. Os três modelos acima são conhecidos como modelos estáticos de previsão, pois assumem que as características de nível, tendência e sazonalidade são constantes ao longo do tempo. Quando estes índices variam com o passar do tempo é necessário utilizar um modelo dinâmico de previsão. O modelo de Winter tem se destacado como bastante prático e de larga utilização, nestes casos.

Segundo Chopra e Meindl (2003), antes de escolher um modelo adequado de previsão, a empresa deve compreender alguns fatores, entre os quais pode-se citar:

- demanda passada;
- planejamento das campanhas publicitárias ou de marketing;
- localização em um catálogo;
- conjuntura econômica;
- planejamento de descontos nos preços;
- ações tomadas pelos concorrentes.

Definir qual modelo é o mais adequado para as previsões pode ser bastante complicado. Diversos estudos indicam que a utilização de múltiplos modelos de previsão e, posteriormente, a combinação de suas previsões em uma previsão final é mais eficaz do que a escolha de um modelo individual. De toda forma, um fato é certo: as previsões nunca

condizem exatamente com o real e devem, por isso, incluir o valor esperado (componente sistemático) e uma medida de erro de previsão (componente aleatório) (CHOPRA e MEINDL, 2003).

O erro de previsão mede a diferença entre a previsão de demanda e a demanda real. A previsão pode apresentar maiores ou menores distorções, dependendo, principalmente, do tipo de produto com que a empresa trabalha e do grau de profundidade da mudança dos fatores que foram usados como base de cálculo da previsão. Alguns tipos de produtos são mais suscetíveis à mudança dos fatores determinantes do comportamento, por isso é importante que a empresa conheça bem quais são os fatores e o seu grau de influência no comportamento da demanda. Algumas empresas podem desistir de realizar previsões por sentirem que não conseguem acertar, porém podem estar aceitando uma perigosa acomodação que acarretarão em um enorme estresse e prejuízo para as atividades de produção (PEINADO e GRAEML, 2007).

O comportamento aleatório da previsão pode ser captado, conforme dito anteriormente, por meio da medição dos erros de previsão. Um bom modelo de previsão capta o comportamento sistemático da demanda e indica o comportamento aleatório pela amplitude do erro. As características dos erros de previsão mais importantes de serem acompanhadas são: sua amplitude e a tendência de viés (PEINADO e GRAEML, 2007). Porém, quando esses erros não apresentam uma justificativa satisfatória e/ou não existem recursos disponíveis para eliminá-los, outras alternativas precisam ser estudadas para tentar minimizar os efeitos dessas discordâncias.

Uma forma de minimizar os efeitos das incertezas de demanda sobre a cadeia produtiva pode ser identificar uma possível relação entre as variações de demanda e as características dos produtos. Quando fala-se em relação entre variáveis e previsão de fenômenos futuros, um importante ramo da ciência vem à mente: a estatística. Segundo Normando, Tjäderhane e Quintão (2010), a escolha do teste estatístico apropriado requer conhecimentos básicos sobre: classificar o tipo de dado que está estudando (contínuo, categórico: ordinal ou nominal); determinar como esses dados estão distribuídos após o término da sua coleta (Distribuição Normal ou Distribuição Anormal), e quais são os tipos de amostras examinadas (Independentes ou Dependentes). Para o caso onde deseja-se conhecer os efeitos que algumas variáveis exercem, ou parecem exercer, sobre outras, a análise mais recomendada é a análise de regressão (HOFFMANN, 2016).

A análise de regressão estuda a relação entre uma variável dependente e outras variáveis independentes. Seu principal objetivo consiste em estimar o valor de uma variável com base

no(s) valor(es) de outra(s) e prever futuros valores. A regressão pode ser simples, quando apenas duas variáveis estão sendo relacionadas, ou múltipla, quando está sendo analisada a influência de mais de uma variável sob a variável de estudo. Pelo modelo de regressão linear, é possível identificar se existe relação linear entre as variáveis em estudo e o grau de compatibilidade com esse modelo. Em caso de não aderência ao modelo linear, ainda há a possibilidade de testar relação entre as variáveis através dos modelos não lineares, que determinam se existe alguma equação cuja curva descreve uma relação de influência entre as variáveis em questão. Tanto o modelo linear quanto o não linear resultam em uma equação que possibilita o cálculo do valor da variável dependente de acordo com os valores das variáveis independentes. Quando o problema a ser tratado possui uma variável dependente que não é contínua (ou seja, é discreta), o modelo mais adequado para prever relação é o modelo de regressão logística, pois ele permite a predição de valores tomados por uma variável categórica (MONTGOMERY, PECK e VINING, 2006).

3 ESTUDO DE CASO

A empresa é uma indústria do ramo de Cosméticos que possui a venda direta como sistema de comercialização. É uma multinacional norte-americana. Sua força de vendas é de cerca de 6 milhões de pessoas ao redor do mundo (maioria mulheres), sendo 1,5 milhão só no Brasil.

Os produtos são ofertados através de folhetos ou pela internet, sempre com o intermédio de uma revendedora. Estes folhetos são renovados aproximadamente a cada 18 dias, que é a duração de uma campanha. O ano possui 19 campanhas e para cada campanha existe um portfólio de produtos ofertados que varia consideravelmente de uma para a outra. Produtos são lançados e descontinuados a cada campanha, conforme estratégia adotada pela empresa. Para o segmento de cosméticos, os produtos são divididos em 6 categorias: *color*, *fragrance*, *personal care*, *skin care*, *hair care* e *children*. Cerca de 700 novos produtos são lançados por ano na planta do Brasil. Os produtos são fabricados em uma única planta (situada em São Paulo) e distribuídos para o resto do país.

3.1 LANÇAMENTO DE NOVOS PRODUTOS

Produtos considerados novos podem ser basicamente de três tipos: algo totalmente inovador, uma reformulação/redesenho de algo já existente ou uma variante de uma linha de certo produto existente (nova cor acrescentada, por exemplo). Para contextualização do objeto de estudo deste trabalho, uma breve explicação sobre as fases que abrangem o lançamento de um produto serão apresentadas a seguir.

3.1.1 Etapas de criação de um produto

A criação de um produto pode ser separada em duas principais fases, que são a fase de desenvolvimento e a fase de implantação do projeto do novo produto. Todo o processo dura em média 1 ano e meio. Ele é idealizado e desenvolvido no âmbito global e passa para controle das áreas locais após o “*Launch Gate*”, que é o início da fase de implementação. De forma resumida, as principais etapas dos projetos de criação de um novo produto são:

1. *Ideation* (Global):

- Definição de oportunidades;
- Geração de ideias;
- Seleção de ideias com potencial.

2. *Discovery* (Global):

- Exploração e refinamento de ideias;
- Revisão preliminar de tempos estimados;
- Compartilhamento dos produtos idealizados;
- Alinhamento de estratégia de marketing.

3. *Concept* (Global):

- Cotações e estimativas de volume de vendas, custos e preços;
- Estratégia de manufatura e análise de ferramental necessário, etc.

4. *Development* (Global/Local):

- Revisão das estimativas de vendas, custos e preços;
- Desenvolvimento de fórmula, embalagens, artes e especificações;
- Início da produção do ferramental;
- Aprovação final e aceitação dos projetos pelos mercados locais.

5. *Launch* (Local):

- Aprovação dos custos;
- Criação de códigos que caracterizam os produtos;
- Envio e aprovação de arte e especificação nos fornecedores;
- Compra de insumos e matérias-primas
- Produção

3.1.2 Implementação dos projetos de novos produtos

A fase de implementação é a fase “*Launch*” descrita no tópico anterior. Os projetos que passam para essa fase tiveram o seu desenvolvimento finalizado e foi dada aprovação para continuidade e início da implantação do projeto do novo produto. O projeto é considerado como finalizado depois que ocorre a primeira produção, ou seja, quando todas as etapas de

desenvolvimento e implementação foram finalizadas e o primeiro produto foi produzido e está apto a ser vendido. Um lançamento ocorre da seguinte forma:

- *Early needs*: primeiro o produto pode ser enviado para as gerentes de setor para que elas mostrem o que está por vir para o público;
- Campanha demo: aproximadamente um mês depois, o produto é ofertado no folheto que é destinado só para as revendedoras;
- Campanha intro: próxima campanha pós a demo, é a campanha oficial de lançamento do produto. É a primeira vez em que ele aparece no folheto que é disponibilizado para o público geral.
- Campanha *repeat* e *flow*: pós lançamento, o produto necessariamente é disponibilizado na campanha seguinte e, depois disso, é considerado encerrado o período de lançamento. O *flow*, que é a continuidade do produto nas campanhas, dependerá do seu resultado de vendas e da estratégia adotada por marketing.

3.2 CADEIA DE SUPRIMENTOS

A cadeia de suprimentos abrange todo o sistema envolvido na atividade de gerir produtos e serviços dos fornecedores aos seus clientes. Seu objetivo é alinhar as atividades de produção, armazenamento e transporte de forma sincronizada entre si e com as necessidades da companhia como um todo, visando redução de custos, qualidade em serviço, maximização da produtividade e, conseqüentemente, aumento nos lucros. Basicamente, envolve as atividades de:

- Compra de insumos e acompanhamento de pedidos/entregas;
- Planejamento da produção;
- Gerenciamento de estoques;
- Distribuição.

O bom funcionamento dessas atividades depende de atividades exercidas por outras áreas e da sincronia de estratégia entre elas. O planejamento de vendas e operações (S&OP) é um processo integrado de planejamento empresarial, adotado pela empresa em estudo. Seu objetivo é integrar funções administrativas às da produção, buscando coordenar esforços e atingir melhores resultados a partir de uma visão mais ampla e mais completa de informações. Este alinhamento entre as áreas envolvidas busca soluções em que se obtenha o máximo de lucratividade, combinando os recursos disponíveis com as restrições de fornecimento, produção e atendimento à demanda.

As cadeias de suprimentos estão cada vez mais complexas, em um contexto onde as expectativas são cada vez mais exigentes e as decisões mudam continuamente. Segundo Cooper (2000), a eficiência do *supply chain management* (gerenciamento da cadeia de suprimentos) não somente é indispensável para o sucesso como é uma condição primordial para a sobrevivência da empresa. Por isso, entender quais são as barreiras e fatores que limitam o alcance de melhores desempenhos é de extrema importância para alavancar os resultados da companhia.

3.2.1 Estimativa de demanda

O funcionamento da cadeia de suprimentos é baseado no número projetado pela área de planejamento de demanda. A produção e a compra de insumos é planejada de acordo com o volume projetado de vendas. A acuracidade desse número interfere em toda a cadeia produtiva e, assim, precisa apresentar alto grau de confiabilidade para evitar prejuízos.

Na empresa em questão, o método utilizado para estimar a demanda de um novo produto é baseado na comparação com vendas de produtos semelhantes e na junção de diversos fatores com pesos associados a cada um para obtenção de um valor final. De modo geral, o cálculo ocorre a partir de um número tomado como base (D_i), que é o valor final de vendas do período de lançamento de um produto com características semelhantes. A esse valor (D_i), são multiplicados fatores que podem aumentá-lo ou reduzi-lo, resultando no valor final estimado (D). O cálculo ocorre da seguinte forma:

$$D = D_p * E * V * P \quad (1)$$

Onde:

- D_p é a demanda projetada. É resultado da associação entre o número base D_i , a diferença de preço e a elasticidade do produto.
- Fator exposição (E): maior do que 1 em caso de exposição maior do que a do produto comparado, igual a 1 em caso de mesmo tipo de exposição e menor do que 1 em caso de exposição inferior.

- Fator veículos de venda (V): maior do que 1 caso a quantidade de veículos em que o produto será exposto seja maior do que o ocorrido com o produto comparado, igual a 1 em caso de mesma quantidade de veículos e menor do que 1 em caso de quantidade inferior.

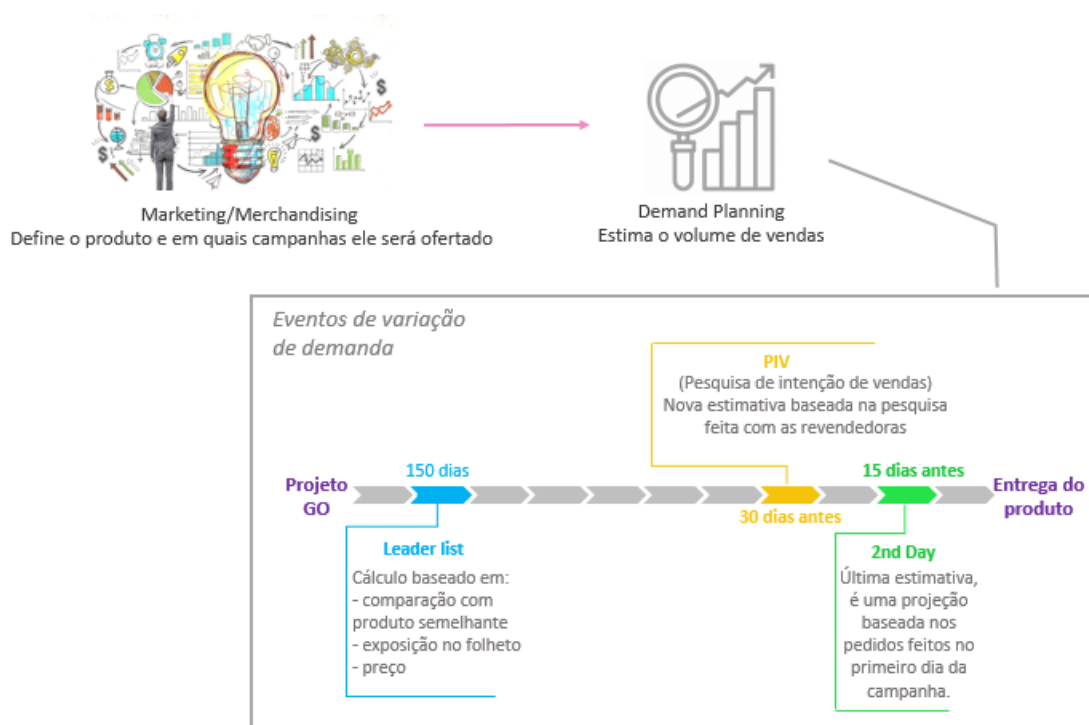
- Fator número de pedidos do período (P): compara o volume de vendas no período em que o produto comparado foi lançado com o volume de vendas atual da empresa. Caso a empresa esteja em uma situação melhor de venda, esse fator será maior do que 1, caso contrário (queda de vendas) será menor do que 1.

É importante ressaltar que o valor atribuído a cada fator parte do julgamento de cada pessoa responsável pela estimativa. Assim como a quantidade de fatores atribuídos ao cálculo, que pode variar de acordo com a pessoa, mas de modo geral esses são os fatores adotados.

3.2.2 Eventos de variações de demanda

O cálculo do volume de vendas estimado para o lançamento de um produto citado anteriormente resulta em um número chamado “*Leader List*”. Ele é a primeira estimativa feita e ocorre aproximadamente 150 dias antes do lançamento do produto (precisa ser determinado com essa antecedência para dar início à compra dos insumos, devido aos altos lead times). Mais próximo do lançamento, cerca de 30 dias antes, esse número começa a ser refinado. Isso é possível pois perto do lançamento estão disponíveis evidências mais concretas para embasamento do número. A Figura 1 mostra uma representação esquemática dos principais momentos em que a demanda é recalculada, chamado de eventos de demanda.

Figura 1 – Eventos de variação de demanda



Fonte: O autor (2018)

O primeiro refinamento é chamado de PIV, que significa pesquisa de intenção de vendas. Consiste em uma pesquisa de intenção de compra de acordo com a opinião das revendedoras e é feito cerca de 30 dias antes do lançamento. Os folhetos da campanha (que já estão impressos nesse momento) são enviados para um grupo de revendedoras previamente selecionado. Elas mostram quais produtos acreditam vender ou não e, com essas informações, é feito um novo cálculo para adequar o número à nova realidade, se for preciso.

O próximo importante refinamento é chamado de *2nd Day*. O cálculo é feito com base no volume de pedidos concretos já solicitados no primeiro dia da campanha. As revendedoras possuem cerca de 15 dias para inserir os pedidos de uma campanha no site. A estimativa do *2nd Day* é uma projeção do volume total baseada no rendimento do primeiro dia de colocação de pedidos.

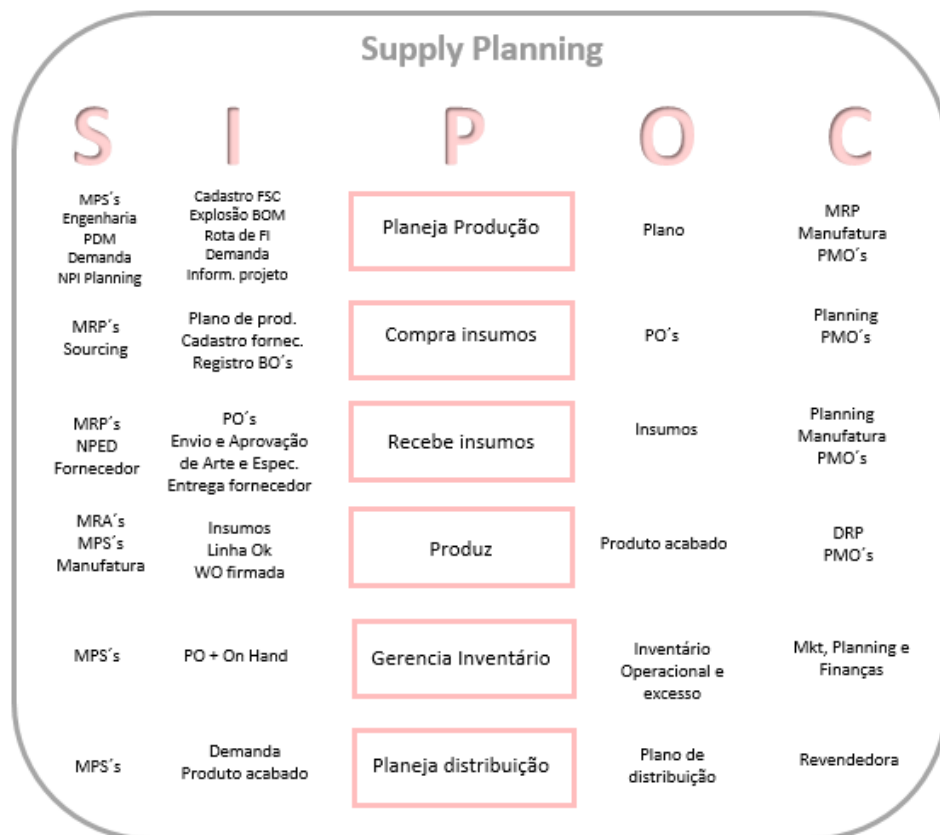
Esses refinamentos têm como objetivo minimizar os impactos de uma estimativa muito discrepante. A ideia é obter um número mais acurado para adequar a produção, caso seja preciso. Porém, como foi dito acima, o primeiro refinamento ocorre 30 dias antes do início da campanha. Considerando que os lead times dos insumos são de cerca de 90 dias (com casos que chegam a 120, 150, 200 dias), não há muitas chances de reação. O porquê disso ficará mais claro nos tópicos a seguir.

3.2.3 Fluxo de planejamento da produção de novos produtos

O planejamento da produção é uma das atividades sob responsabilidade de Supply Planning e consiste em planejar o que deve ser produzido, de acordo com a demanda solicitada, bem como tratar dos pedidos e entregas dos insumos necessários para que a produção ocorra.

Os novos produtos exigem um planejamento mais cauteloso, uma vez que algumas atividades de projeto ainda estão sendo concluídas e testadas, podendo aparecer imprevistos. Basicamente, para ocorrer a primeira produção, são necessários parâmetros como: cadastro do código do produto e dos insumos que o compõem (BOM); rotas de manufatura; permissão de compra de material (BO's), demanda em sistema e finalização das aprovações de projeto. A figura 2 esquematiza, de forma geral, as principais atividades exercidas por Supply Planning (coluna central) e as entradas/saídas desse processo.

Figura 2 – Representação, através da ferramenta SIPOC, de como funciona o processo seguido por *Supply Planning* no planejamento de novos produtos.



Fonte: O autor (2018)

O método de produção da empresa em estudo está mais voltado para o “*Make to Inventory*”. Uma certa demanda é estimada com antecedência afim de que o cliente final, pós pedido de compra, aguarde apenas o tempo de entrega do produto. A produção é baseada em uma estimativa e não na venda concreta, aproximando-se, assim, do conceito de produção empurrada. Porém, como os novos produtos apresentam todo o risco de imprevisibilidade já citado ao longo deste trabalho, tenta-se ao máximo controlar a produção e compra de insumos para alcançar o mais próximo possível de uma “produção puxada”, buscando um meio termo entre os dois métodos.

3.2.4 Impacto das variações de demanda na produção

Na Figura 3 está representada, de forma geral, a origem das informações que alimentam a cadeia de Supply Planning. Basicamente, as áreas de marketing definem os produtos com seus principais parâmetros de vendas e a área de planejamento de demanda calcula o volume demandado com base nessas informações. Esse número estimado é o que define grande parte das atividades de Supply Planning. É a referência utilizada para realização do planejamento de compra de insumos, programação da produção, programação da distribuição e gerenciamento dos estoques. Como a empresa trabalha com insumos de altos lead times, os insumos precisam ser comprados com certa antecedência com relação à data programada de produção. Isso gera uma baixa capacidade de reação tanto a *oversales* quanto a *undersales*. Por isso, é preciso entender o comportamento das variações de demanda pois, caso seja possível prever uma tendência de comportamento, o uso dessas informações pode ajudar na tomada de decisão sobre os casos em que a criação de *Safety Stock* é viável ou não, de acordo com o risco de variação do volume demandado.

Figura 3 – Visão macro do fluxo de informações que alimentam *Supply Planning*



Fonte: O autor (2018)

4 MODELOS DE ANÁLISE DE REGRESSÃO

Dada toda a problemática descrita até o momento, o principal questionamento a ser tratado é: como determinar a quantidade ideal que deve ser produzida sabendo-se dos riscos associados ao lançamento do produto? Para encontrar uma solução plausível, precisa-se entender qual é o comportamento da demanda de novos produtos, identificar tendências, analisar se existe correlação entre o risco de variação e algumas características do produto. Ou seja, se o foco é precaver-se ao risco de falta/excesso e o que o ocasiona é a estimativa equivocada da demanda, uma análise estatística de dados pode trazer algumas respostas e ajudar a prever melhor comportamentos futuros.

Neste estudo, inicialmente, será utilizado o método de análise de regressão linear multivariada para determinar se é possível estimar um valor de variação de demanda de acordo com algumas características dos produtos, como: tipo do produto, tipo de inovação e grau de exposição. Em caso de não aderência a esse modelo, é possível testar se pelo menos existe grau de relação confiável entre a variação de demanda e essas mesmas características dos produtos. Neste caso a análise não resultará em uma estimativa numérica para a variação, mas pode mostrar se as variáveis independentes predizem risco alto ou baixo de variação, por exemplo. Este segundo caso pode ser determinado através da regressão logística. Para adequação de variáveis qualitativas aos modelos de regressão utilizados, foi inserido o conceito de variáveis dummies.

Em um estudo estatístico completo, existem algumas fases que devem ser desenvolvidas para se chegar a resultados satisfatórios. As fases sugeridas por Correa (2003) e que foram seguidas para elaboração das análises são:

- Definição do problema;
- Planejamento;
- Coleta de dados;
- Apuração dos dados;
- Apresentação dos dados;
- Análise e Interpretação dos resultados.

Os modelos e as técnicas de regressão utilizadas neste estudo estão descritos nos tópicos 4.1 e 4.2 a seguir.

4.1 REGRESSÃO LINEAR MULTIVARIADA

No modelo de regressão linear multivariada, cada variável independente apresenta uma relação linear com a variável de resposta. A junção dessas individuais relações resulta na equação geral final. Conceitualmente, é equivalente a um gráfico de uma linha de regressão em um espaço n-dimensional. A seguir serão apresentadas as definições do modelo de regressão multivariada e as técnicas de estimação dos seus parâmetros.

4.1.1 Modelo e técnica de estimação dos parâmetros da regressão multivariada

Com base na demonstração feita por Hoffmann (2016), abaixo segue um resumo do embasamento teórico do modelo de regressão linear multivariada.

A equação básica de uma regressão linear múltipla consiste em uma resposta y que é função de k variáveis, conforme mostra a equação 2 a seguir.

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k + e \quad (2)$$

Dada uma amostra de n observações

$$\{(x_{11}, x_{12}, \dots, x_{1k}, y_1), \dots, (x_{n1}, x_{n2}, \dots, x_{nk}, y_n)\} \quad (3)$$

O modelo consiste em n equações:

$$\begin{aligned} y_1 &= b_0 + b_1 x_{11} + b_2 x_{12} + \dots + b_k x_{1k} + e_1 \\ y_2 &= b_0 + b_1 x_{21} + b_2 x_{22} + \dots + b_k x_{2k} + e_2 \\ &\vdots \\ y_n &= b_0 + b_1 x_{n1} + b_2 x_{n2} + \dots + b_k x_{nk} + e_n \end{aligned} \quad (4)$$

Sob a forma de aritmética matricial, temos:

$$y = Xb + e \quad (5)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad (6)$$

Conforme demonstrado por Hoffmann (2016) ou Nogueira (2007), o vetor b pode ser estimado através do método dos mínimos quadrados, resultando em:

$$b = (X^T X)^{-1} X^T y \quad (7)$$

Onde X é a matriz com os valores das variáveis independentes, X^T é a matriz X transposta e y é a matriz com os valores da variável resposta.

A precisão do modelo pode ser obtida calculando-se o coeficiente de determinação R^2 .

$$R^2 = \frac{SSR}{SST} \quad (8)$$

Onde:

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (9)$$

$$SSE = \sum e_i^2 = \{y^T y - b^T X^T y\} \quad (10)$$

$$SSR = SST - SSE \quad (11)$$

Sendo SSE a soma dos quadrados dos resíduos, SSR a soma dos quadrados de regressão e SST a soma total dos quadrados (mostra a variação em y em torno da própria média).

O R^2 indica a porcentagem de aderência dos dados ao modelo utilizado. Quanto maior for o R^2 obtido, maior é a confiabilidade e precisão do modelo.

4.1.2 Aplicação: estimativa da variação de demanda de acordo com as características do novo produto

Para determinar se é possível prever uma variação de demanda mais susceptível a acontecer de acordo com certas características do produto a ser lançado, foi feito um levantamento com os dados dos lançamentos dos últimos 3 anos da empresa em estudo. A ideia é entender se, por exemplo, um batom tem em média uma variação maior do que um perfume, ou se um produto totalmente inovador apresenta média de variação maior do que um que sofreu apenas uma reformulação, etc. Aplicando-se a análise de regressão multivariada a esses dados, pode-se determinar se existe relação entre cada variável independente (tipo de produto, tipo de inovação, tipo de exposição) e qual é a parcela de contribuição de cada uma delas no resultado final (variação da demanda estimada).

Uma análise prévia foi feita para identificar quais variáveis suspeita-se exercer influência na imprevisibilidade da demanda. Como já citado anteriormente, escolheu-se as variáveis: tipo do produto, tipo de inovação e grau de exposição.

- **Quantificação de variáveis qualitativas: Variável Dummy**

Para aplicar a análise de regressão multivariada, as variáveis precisam ser mensuradas. As variáveis quantitativas são facilmente mensuradas em alguma escala, já as variáveis qualitativas não. Como existem variáveis qualitativas no grupo de dados abordados neste trabalho, é preciso adotar um método confiável de quantificação de valores qualitativos em valores numéricos. Um método amplamente utilizado para “quantificar” esses atributos consiste em construir variáveis artificiais que assumam valores de 1 ou 0 (indicando presença e ausência de um atributo) que são conhecidas pela literatura existente como “variáveis dummy” (Missio, p.111-135, 2007). Uma representação da tabela de dados formada para realização das análises está representada na Tabela 1.

Tabela 1 – Representação do modo de agrupamento dos dados para realização da análise de regressão multivariada.

Y (Variação demanda)	Variáveis "Dummy"								
	X1 (Tipo de exposição)	X2 (Tipo de inovação)	X3 (É Perfume)	X4 (É Masc. de cílios)	X5 (É Batom)	X6 (É Esmalte)	X7 (É Anti-Idade)		Xn (É Hidratante)
-38%	1	1	1	0	0	0	0		0
10%	1	0	0	0	1	0	0		0
314%	3	1	0	0	1	0	0		0
-29%	1	1	1	0	0	0	0		0
-57%	1	1	0	0	0	0	1		0
-96%	2	0	0	0	0	1	0		0
16%	3	1	0	0	1	0	0		0
-20%	1	1	0	0	0	0	0		1
-45%	4	1	1	0	0	0	0		0
42%	4	1	0	0	0	0	1		0
43%	3	1	0	0	1	0	0		0
18%	1	1	0	0	0	0	1		0

Fonte: O autor (2018)

- **Aplicação do método e compreensão das respostas obtidas**

As variáveis qualitativas foram devidamente quantificadas e os dados de todas as variáveis foram agrupados em uma tabela, conforme mostra a Tabela 1. Como trata-se de um acumulado de cerca de 4 mil linhas de dados, o cálculo da regressão e de seus parâmetros pode ser feito com o auxílio de softwares, como Minitab, SPSS Statistics ou Excel.

O resultado que o software fornece é um resumo de fatores que devem ser interpretados para julgar se o modelo é útil para prever a informação desejada. Os resultados que precisam ser analisados para validação do modelo são:

- Teste F de significância global:

Mostra se o modelo, de forma geral, é útil para prever a informação y requerida. Se o valor-p do teste F for menor do que 0,05, há evidências estatísticas de que pelo menos uma variável no modelo está relacionada com a variável dependente y.

- Testes de significância individuais:

Mostra o valor-p de cada uma das variáveis independentes. Com isso, pode-se determinar quais variáveis estão relacionadas com a variável dependente e, assim, eliminar as que não tiverem evidência estatística de relação comprovadas (nesse caso, é necessário rodar a análise novamente sem essas variáveis para obtenção dos novos valores resultantes da regressão).

- Coeficientes de determinação R^2 e R^2 ajustado:

Mostra o grau de relação entre as variáveis independentes e a variável dependente. É expressa em porcentagem e determina quantos por cento da variabilidade de y é explicada pelo modelo com as variáveis x 's. O R^2 ajustado é utilizado quando deseja-se comparar modelos com diferentes quantidades de variáveis.

- Coeficientes da equação de regressão ($b_0, b_1, \dots, b_k$)

Por fim, a análise fornece os coeficientes que devem ser substituídos para construção da equação final de y .

- **Interpretação dos resultados**

Os resultados obtidos na realização da regressão linear multivariada dos dados analisados neste estudo estão dispostos nas Tabelas 2, 3 e 4 a seguir.

Tabela 2. Resumo dos resultados da análise de regressão linear multivariada.

RESUMO DOS RESULTADOS	
<i>Estatística de regressão</i>	
R múltiplo	0,146752728
R-Quadrado	0,021536363
R-quadrado ajustado	0,01710642
Erro padrão	2,079873451
Observações	3551

Fonte: O autor (2018)

Tabela 3. Tabela ANOVA da análise de regressão linear multivariada.

ANOVA					
	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	16	336,4868077	21,03043	4,861544	0,00
Resíduo	3534	15287,63721	4,325874		
Total	3550	15624,12402			

Fonte: O autor (2018)

Tabela 4. Resultados finais da análise de regressão linear multivariada.

	Coef.	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	0,076	1,476	0,051	0,959	-2,819	2,971	-2,819	2,971
X1 (Tipo exposição)	-0,101	0,037	-2,759	0,006	-0,173	-0,029	-0,173	-0,029
X2 (Tipo de inovação)	0,051	0,106	0,484	0,629	-0,157	0,260	-0,157	0,260
X3 (É Perfume)	0,114	1,474	0,077	0,938	-2,777	3,004	-2,777	3,004
X4 (É Creme/limpeza facial)	-0,161	1,480	-0,109	0,914	-3,063	2,741	-3,063	2,741
X5 (É Cabelos)	-0,325	1,489	-0,219	0,827	-3,244	2,594	-3,244	2,594
X6 (É Batom)	-0,060	1,473	-0,041	0,967	-2,949	2,828	-2,949	2,828
X7 (É Máscara de cílios)	-0,217	1,492	-0,146	0,884	-3,142	2,708	-3,142	2,708
X8 (É Esmaltes)	0,482	1,475	0,327	0,744	-2,410	3,373	-2,410	3,373
X9 (É Kids)	-0,176	1,484	-0,119	0,906	-3,086	2,733	-3,086	2,733
X10 (É Sabonete)	-0,178	1,485	-0,120	0,905	-3,089	2,733	-3,089	2,733
X11 (É Anti-Idade)	0,007	1,486	0,005	0,996	-2,905	2,920	-2,905	2,920
X12 (É Sombra/ ápis/delineador)	-0,240	1,478	-0,163	0,871	-3,138	2,657	-3,138	2,657
X13 (É Desodorante)	0,047	1,487	0,031	0,975	-2,868	2,962	-2,868	2,962
X14 (É Pó compacto/base/corretivo/blush)	0,570	1,473	0,387	0,699	-2,319	3,458	-2,319	3,458
X15 (É Hidratante corporal)	-0,091	1,476	-0,062	0,951	-2,984	2,802	-2,984	2,802
X16 (É Protetor solar)	0,176	1,551	0,113	0,910	-2,864	3,216	-2,864	3,216

Fonte: O autor (2018)

Os resultados mostram valores-p muito superiores a 0,05 e coeficiente de determinação R^2 extremamente pequeno (1,7%). Logo, não há evidência de relação linear entre as variáveis analisadas, ou seja, não pode-se determinar uma função matemática que relacione a variável dependente com as variáveis independentes. Assim, com a análise dos dados coletados, conclui-se que não é possível estimar, através de uma equação linear, um valor para a variação de demanda a partir de informações sobre o tipo de produto, tipo de inovação e grau de exposição.

4.2 REGRESSÃO LOGÍSTICA

De acordo com Figueira (2006), nos casos em que a variável dependente é qualitativa e expressa por duas categorias, o método dos mínimos quadrados, utilizado na regressão linear, não oferece estimadores plausíveis. Uma boa aproximação é obtida pela regressão logística, que é baseada no método da máxima verossimilhança e permite o uso de um modelo de regressão para se calcular ou prever a probabilidade de um evento específico.

A regressão logística é um recurso que permite estimar a probabilidade associada à ocorrência de determinado evento dado um conjunto de variáveis que suspeita-se interferir na

ocorrência desse evento. É uma técnica que busca estimar a probabilidade da variável dependente assumir um determinado valor em função de valores conhecidos de variáveis independentes. Esta é uma técnica recomendada para situações em que a variável dependente é de natureza dicotômica ou binária.

Como nesta parte do estudo deseja-se determinar se as variáveis independentes predizem certo risco de variação de demanda, percebe-se que o método de regressão logística adequa-se aos requisitos desta análise. A seguir serão apresentados os fundamentos teóricos deste método e, posteriormente, sua aplicação no estudo de caso deste trabalho.

4.2.1 Modelo e técnica de estimação dos parâmetros da regressão logística

Conforme apresentado por Hosmer, Lemeshow e Sturdivant (2013) ou ainda por Cabral (2013), abaixo está apresentado, de forma resumida, o embasamento teórico de um modelo de regressão logística.

O modelo logístico é dado por:

$$f(z) = \frac{e^z}{1 + e^z} \quad (12)$$

Ou ainda:

$$f(z) = \frac{1}{1 + e^{-z}} \quad (13)$$

Do mesmo modo como ocorre na regressão linear, este modelo faz uso de n variáveis independentes para avaliar o comportamento de uma variável dependente. Assim, tem-se que $z(x)$:

$$z = \alpha + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (14)$$

Substituindo no modelo logístico apresentado na equação 15, tem-se $f(x)$:

$$f(x) = \frac{1}{1 + e^{-(\alpha + b_1 x_1 + b_2 x_2 + \dots + b_n x_n)}} \quad (15)$$

Essa expressão é considerada como um conceito probabilístico. Considerando n eventos independentes, onde Y_n são variáveis dicotômicas e classificadas como n provas de Bernoulli, sua probabilidade de sucesso $P(Y)$ é dada por:

$$P(Y) = \begin{cases} 1, & \text{se sucesso} \\ 0, & \text{se fracasso} \end{cases} \quad (16)$$

Logo:

$$P(Y = 1 \setminus x_1, x_2, \dots, x_n) = \frac{1}{1 + e^{-(\alpha + b_1 x_1 + b_2 x_2 + \dots + b_n x_n)}} \quad (17)$$

Onde:

- $P(Y = 1 \setminus x_1, x_2, \dots, x_n)$ é a probabilidade de ocorrer o evento $Y=1$, onde Y é uma variável dicotômica dependente que pode assumir valor 0 ou 1;
- $x_1, x_2, \dots, x_n$ é o vetor X que contém as variáveis independentes;
- α e b são os parâmetros desconhecidos que devem ser estimados.

Conforme Menard (2002) apresenta em seu livro, se a probabilidade de sucesso for modelada como

$$P(Y = 1 \setminus x) = p = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n \quad (18)$$

os resultados poderiam apresentar probabilidades negativas ou superiores a 1, e sabe-se que a probabilidade dos valores de Y observados podem assumir apenas valores entre 0 e 1.

Para sanar essa incoerência, a probabilidade de $Y=1$ não é diretamente utilizada, mas sim a chance de $Y=1$, que pode ser definida como a razão entre a probabilidade de $Y=1$ e a probabilidade de $Y \neq 1$.

$$chance = \frac{P(sucesso)}{P(fracasso)} \quad (19)$$

Na regressão logística, probabilidade e chance (odds) são conceitos diferentes que descrevem o mesmo fenômeno. Probabilidade é o risco de determinado evento acontecer, já chance é a razão entre a probabilidade deste evento acontecer e a probabilidade de não acontecer.

A razão de chances (odds ratio) é a razão entre a chance de determinado evento ocorrer sob influência de um dado fator e a chance do evento ocorrer sem a influência deste fator. Ou seja, mede o efeito de um dado fator na probabilidade de certa condição ocorrer entre os dados analisados.

Todas as definições apresentadas têm como objetivo avaliar o impacto exercido pelas variáveis independentes na variável dependente. Transformando-se a chance obtida em uma razão de chances (odds ratio) e adequando-a em uma variável de base logarítmica, pode-se obter uma previsão de valores dentro do intervalo de 0 a 1. Logo, tem-se que a razão de chances pode ser determinada por:

$$\frac{P(sucesso)}{1 - P(sucesso)} = \frac{\frac{e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}}{1 + e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}}}{1 - \frac{e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}}{1 + e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}}} \quad (20)$$

Ajustando para simplificar a equação:

$$\frac{P(sucesso)}{1 - P(sucesso)} = \frac{\frac{e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}}{1 + e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}}}{\frac{1}{1 + e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}}} \quad (21)$$

Logo,

$$\frac{P(sucesso)}{1 - P(sucesso)} = e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n} \quad (22)$$

Aplicando-se o logaritmo natural à razão de chance, obtém-se o modelo de regressão logística, também conhecido como *logit*.

$$\text{logit} = \ln \left(\frac{P(\text{sucesso})}{1 - P(\text{sucesso})} \right) = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (23)$$

Os coeficientes do modelo de regressão logística mostram o impacto que cada variável independente exerce sobre a razão de chance. Segundo Bissacot (2015), o método que pode ser utilizado para estimar esses coeficientes é o método da máxima verossimilhança, cujo princípio é demonstrado por Hoffmann (2016).

Graficamente, a regressão logística apresenta comportamento no formato de um “S”, caracterizada por áreas onde a mudança é mais acentuada e áreas onde ela é praticamente imperceptível ou nem ocorre. As regiões onde pequenas variações de x ocasionam altas variações de Y são regiões de maior probabilidade de mudança de estado da variável Y em função de x .

Análises com alto grau de confiabilidade exigem grande quantidade de dados na amostra. Quanto maior o número de dados observados, maior é a representatividade da análise e os resultados tendem a descrever melhor a realidade. Essas análises necessitam do auxílio de softwares para serem feitas, devido ao elevado esforço que demandaria e, dependendo da quantidade de informações, é humanamente inviável ser feito à mão com os recursos atuais, da mesma forma que acontece com a regressão linear citada no tópico anterior. Um software comumente utilizado para análises estatísticas é o SPSS *Statistics* e ele foi utilizado na análise de regressão logística aplicada neste trabalho e descrita no tópico a seguir.

4.2.2 Aplicação: estimativa do nível de risco de variação de demanda de acordo com as características do novo produto

A análise de regressão logística foi aplicada ao mesmo conjunto de dados da análise anterior. Apenas uma alteração foi realizada: a variável dependente Y (porcentagem de variação de demanda) foi categorizada. Variações acima de 30% foram consideradas como 1 (alto risco) e variações abaixo de 30% foram consideradas como 0 (baixo risco). Percebe-se que, desta forma, o objetivo é avaliar se as variáveis independentes predizem se o risco de vendas acima do esperado é alto ou baixo. Nesta etapa, não está sendo avaliado risco de variação geral, mas sim risco de incremento de demanda acima de 30%, como parte da

determinação da viabilidade do estoque de segurança. Para realização desta análise, foi utilizado o software *SPSS Statistics*.

A priori, a análise foi feita considerando-se:

- Y = Variação de demanda (1 para variações acima de 30% e 0 para variações abaixo desse percentual);
- X1 = Tipo de exposição (Valores possíveis: 1, 2, 3 e 4);
- X2 = Tipo de inovação (1 para produtos inovadores e 0 para reformulações/ redesenhos/ incremento a uma linha já existente);
- X3 / X4 / X5 / ... / X10 = Cada um desses X's representam um tipo de produto (os valores possíveis são 1 ou 0, representando se refere ao tipo de produto abordado ou não).

Para dar continuidade à análise, 2 verificações precisaram ser feitas. São elas:

- Presença de multicolinearidade entre as variáveis: foi realizado o teste de multicolinearidade para validação das variáveis independentes. Na estatística de colinearidade, valores de tolerância acima de 0,1 e de VIF (fator de inflação da variância) abaixo de 10 indicam ausência de multicolinearidade entre as variáveis (MONTGOMERY, PECK e VINING, 2006). Como o teste forneceu resultado de ausência de multicolinearidade para as variáveis em estudo (vide Tabela 5), pode-se dar continuidade à análise com as variáveis escolhidas;

Tabela 5. Verificação da multicolinearidade entre as variáveis.

		Coeficientes ^a					Estatísticas de colinearidade	
Modelo		Coeficientes não padronizados	Erro	Coeficientes padronizados	t	Sig.	Tolerância	VIF
B	Erro	Beta						
1	(Constante)	,155	,042		3,719	,000		
	X1 (Tipo exposição)	-,008	,007	-,020	-1,193	,233	,950	1,053
	X2 (Tipo de inovação)	-,006	,020	-,006	-,300	,764	,733	1,364
	X3 (É Perfume)	,065	,040	,052	1,627	,104	,268	3,732
	X4 (É Batom)	,044	,038	,044	1,155	,248	,191	5,240
	X5 (É Máscara de cílios)	-,058	,060	-,020	-,977	,329	,666	1,502
	X6 (É Esmaltes)	,248	,041	,215	6,121	,000	,219	4,561
	X7 (É Anti-Idade)	-,016	,053	-,006	-,302	,763	,588	1,702
	X8 (É Sombra/lapis/delineador)	-,001	,045	,000	-,013	,989	,409	2,443
	X9 (É Pó compacto/base/corretivo/blush)	,110	,039	,103	2,850	,004	,207	4,830
	X10 (higiene/cuidado pessoal)	,019	,038	,019	,514	,607	,190	5,265

a. Variável Dependente: Y (Variação >30%)

Fonte: O autor (2018)

- Presença de *Outliers*: a segunda verificação é a determinação de presença de *Outliers*. Eles são valores atípicos, ou seja, valores que apresentam um grande afastamento dos demais da série ou que são inconsistentes. *Outliers* podem inviabilizar todo o resultado de uma análise, por isso eles precisam ser identificados e tratados para não direcionar negativamente o modelo (MARONA, MARTIN e YOHAI, 2006). Para os dados em questão, não foram identificados *outliers*.

Assim, os dados foram devidamente dispostos em uma tabela e foi aplicada a análise de regressão logística com auxílio do software *SPSS Statistics*. A análise foi segmentada em 2 etapas:

- 1ª etapa da análise: todas as variáveis independentes citadas foram abordadas na regressão.

Nessa primeira etapa, a ideia foi aplicar a regressão com todas as variáveis sendo consideradas e, caso algumas delas apresentassem valor p maior do que 0,05, seriam desconsideradas para obtenção de um modelo contendo apenas as variáveis com significância comprovada. Os resultados obtidos estão dispostos nas Figuras 4 e 5.

Figura 4. *Print Screen* da tela de resultados do software SPSS: resumo do caso e Bloco 0 da primeira análise de regressão logística.

Regressão logística

[Conjunto de Dados3]

Resumo de processamento do caso

Casos não ponderados ^a		N	Porcentagem
Casos selecionados	Incluído na análise	3551	100,0
	Casos omissos	0	,0
	Total	3551	100,0
Casos não selecionados		0	,0
Total		3551	100,0

a. Se a ponderação estiver em vigor, veja a tabela de classificação para o número total de casos.

Codificação de variável dependente

Valor original	Valor interno
0	0
1	1

Bloco 0: Bloco Inicial

Tabela de Classificação^{a,b}

Observado		Previsto			
		Y (Variação >30 %)	0	1	Porcentagem correta
Etapa 0	Y (Variação >30 %)	0	2801	0	100,0
		1	750	0	,0
	Porcentagem global				78,9

a. A constante está incluída no modelo.

b. O valor de recorte é ,500

Variáveis na equação

	B	S.E.	Wald	df	Sig.	Exp(B)
Etapa 0 Constante	-1,318	,041	1027,140	1	,000	,268

Variáveis não presentes na equação

		Score	df	Sig.
Etapa 0	Variáveis	X1 (Tipo exposição)	1,998	1
		X2 (Tipo de inovação)	30,123	1
		X3 (É Perfume)	,308	1
		X4 (É Batom)	5,569	1
		X5 (É Máscara de cílios)	8,131	1
		X6 (É Esmaltes)	114,183	1
		X7 (É Anti-Idade)	4,951	1
		X8 (É Sombra/lapis/delineador)	8,125	1
		X9 (É Pó compacto/base/corretivo/b lush)	5,128	1
		X10 (higiene/cuidado pessoal)	17,076	1
	Estatísticas globais	146,155	10	,000

Fonte: O autor (2018)

Figura 5. *Print Screen* da tela de resultados do software SPSS: bloco 1 da primeira análise de regressão logística.

Bloco 1: Método = Enter

Testes de Omnibus do Modelo de Coeficientes

		Qui-quadrado	df	Sig.
Etapa 1	Etapa	137,693	10	,000
	Bloco	137,693	10	,000
	Modelo	137,693	10	,000

Resumo do modelo

Etapa	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
1	3523,764 ^a	,038	,059

a. Estimação finalizada no número de iteração 5 porque as estimativas de parâmetro mudaram foram alteradas para menos de ,001.

Tabela de Classificação^a

Observado		Previsto		Porcentagem correta
		Y (Variação >30%)		
Etapa 1	Y (Variação >30%)	0	1	
	0	2801	0	100,0
	1	750	0	,0
	Porcentagem global			78,9

a. O valor de recorte é ,500

Variáveis na equação

		B	S.E.	Wald	df	Sig.	Exp(B)
Etapa 1 ^a	X1 (Tipo exposição)	-,054	,045	1,430	1	,232	,948
	X2 (Tipo de inovação)	-,033	,118	,076	1	,783	,968
	X3 (É Perfume)	,467	,281	2,762	1	,097	1,595
	X4 (É Batom)	,335	,274	1,491	1	,222	1,398
	X5 (É Máscara de cílios)	-,669	,530	1,593	1	,207	,512
	X6 (É Esmaltes)	1,373	,277	24,620	1	,000	3,946
	X7 (É Anti-Idade)	-,141	,399	,125	1	,724	,868
	X8 (É Sombra/lapis/delineador)	-,011	,325	,001	1	,973	,989
	X9 (É Pó compacto/base/corretivo/blush)	,730	,271	7,239	1	,007	2,075
	X10 (higiene/cuidado pessoal)	,157	,273	,329	1	,566	1,169
	Constante	-1,736	,287	36,662	1	,000	,176

a. Variável(is) inserida(s) na etapa 1: X1 (Tipo exposição), X2 (Tipo de inovação), X3 (É Perfume), X4 (É Batom), X5 (É Máscara de cílios), X6 (É Esmaltes), X7 (É Anti-Idade), X8 (É Sombra/lapis/delineador), X9 (É Pó compacto/base/corretivo/blush), X10 (higiene/cuidado pessoal).

O Bloco 0 da regressão mostra os resultados de uma análise que não abrange as variáveis independentes, ou seja, refere-se a um modelo de regressão baseado apenas na variável Y. Esse bloco presume que o risco de variação de demanda pode ser modelado sem a influência das variáveis independentes e mostra a aderência da adoção desta premissa. Observando a tabela de classificação da Figura 4, que mostra a porcentagem de acerto caso seja adotado risco 0 para todos os casos, pode-se interpretar que, dos 3551 casos analisados, 2801 casos seriam previstos como baixo risco e condiriam com a realidade, porém 750 casos seriam também previstos como baixo risco porém seriam de alto risco, não condizendo com a realidade. Assim, o modelo sem considerar a influência das variáveis independentes e com premissa de que todos os casos têm baixo risco teria uma porcentagem de 78,9% de acerto. Parece uma porcentagem de acerto alta porém isso só ocorre porque a grande maioria dos casos observados são de baixo risco de variação. Como a ideia é entender mais a fundo quais são os casos mais propensos a elevadas variações de demanda, não é satisfatório assumir este resultado. Além disso, a tabela “Variáveis na equação” do Figura 4 mostra valor de significância menor do que 0,05 para a influência de pelo menos uma das variáveis independentes na equação, o que significa que ao menos uma variável X exerce influência considerável na variável Y e pode enriquecer de forma significativa o modelo obtido. Ainda no Bloco 0, na tabela “Variáveis não presentes na equação”, é possível observar o valor p (significância) para cada variável X. Um primeiro refinamento do modelo já pode ser feito com base nessa informação, uma vez que valores de p acima de 0,05 demonstram significância insuficiente e, por isso, essas variáveis podem ser retiradas pois estão prejudicando a confiabilidade do modelo. Sabendo-se disso, o Bloco 1, que mostra a análise com todos os X's levados em consideração, já pode ser descartado pois não fornece o resultado desejado. A análise dos valores de p comprovam isso (muitos valores acima de 0,05, mostrando que o modelo com todas essas variáveis consideradas não apresenta significância relevante).

- 2ª etapa da análise: refinamento onde apenas as variáveis independentes com significância comprovada na análise anterior foram inseridas no modelo.

Na segunda etapa da análise, as variáveis cuja significância não foi satisfatória foram retiradas para não interferir negativamente no modelo. O objetivo foi tentar identificar se pelo menos as variáveis restantes podem auxiliar na estimação de alto risco de variação ou não. Assim, foram consideradas como variáveis independentes apenas os tipos de produtos X5

(máscara de cílios), X6 (esmaltes), X8 (sombra/lápis/delineador) e X10 (higiene/cuidado pessoal). Como as variáveis correspondentes aos outros tipos de produtos foram excluídas, um novo conjunto de dados foi determinado, onde só as informações referentes aos 4 tipos de produtos resultantes foram analisadas. Com isso, dos 3551 lançamentos analisados, restaram 1680, que eram apenas os dados relativos aos tipos de produtos representados por X5, X6, X8 e X10.

Os resultados obtidos estão dispostos nas Figuras 6 e 7 nas páginas seguintes.

Figura 6. *Print Screen* da tela de resultados do software SPSS: resumo do caso e Bloco 0 da segunda análise de regressão logística

Regressão logística

Resumo de processamento do caso

Casos não ponderados ^a		N	Porcentagem
Casos selecionados	Incluído na análise	1680	100,0
	Casos omissos	0	,0
	Total	1680	100,0
Casos não selecionados		0	,0
Total		1680	100,0

a. Se a ponderação estiver em vigor, veja a tabela de classificação para o número total de casos.

Codificação de variável dependente

Valor original	Valor interno
0	0
1	1

Bloco 0: Bloco Inicial

Tabela de Classificação^{a,b}

Observado			Previsto		Porcentagem correta
			Y (Variação >30%)		
Etapa 0	Y (Variação >30%)	0	1310	0	100,0
		1	370	0	,0
	Porcentagem global				78,0

a. A constante está incluída no modelo.

b. O valor de recorte é ,500

Variáveis na equação

		B	S.E.	Wald	df	Sig.	Exp(B)
Etapa 0	Constante	-1,264	,059	461,158	1	,000	,282

Variáveis não presentes na equação

			Escore	df	Sig.
Etapa 0	Variáveis	X5 (É Máscara de cílios)	9,150	1	,002
		X6 (É Esmaltes)	123,346	1	,000
		X8 (É Sombra/lapis/delineador)	10,555	1	,001
		X10 (higiene/cuidado pessoal)	31,979	1	,000
	Estatísticas globais		126,251	4	,000

Fonte: O autor (2018)

Figura 7. *Print Screen* da tela de resultados do software SPSS: bloco 1 da segunda análise de regressão logística.

Bloco 1: Método = Enter

Testes de Omnibus do Modelo de Coeficientes

		Qui-quadrado	df	Sig.
Etapa 1	Etapa	120,925	4	,000
	Bloco	120,925	4	,000
	Modelo	120,925	4	,000

Resumo do modelo

Etapa	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
1	1650,497 ^a	,069	,107

a. Estimativa finalizada no número de iteração 5 porque as estimativas de parâmetro mudaram foram alteradas para menos de ,001.

Teste de Hosmer e Lemeshow

Etapa	Qui-quadrado	df	Sig.
1	,000	3	1,000

Tabela de Classificação^a

		Previsto		Porcentagem correta
		Y (Variação >30%)	0	1
Etapa 1	Observado			
	Y (Variação >30%)	0	1310	0
		1	370	0
Porcentagem global				78,0

a. O valor de recorte é ,500

Variáveis na equação

		B	S.E.	Wald	df	Sig.	Exp(B)
Etapa 1 ^a	X5 (É Máscara de cílios)	-,704	,529	1,768	1	,184	,495
	X6 (É Esmaltes)	1,389	,269	26,640	1	,000	4,010
	X8 (É Sombra/lapis/delineador)	-,026	,325	,006	1	,936	,974
	X10 (higiene/cuidado pessoal)	,161	,273	,347	1	,556	1,174
	Constante	-1,846	,254	52,965	1	,000	,158

a. Variável(is) inserida(s) na etapa 1: X5 (É Máscara de cílios), X6 (É Esmaltes), X8 (É Sombra/lapis/delineador), X10 (higiene/cuidado pessoal).

Algumas conclusões podem ser feitas a partir da análise do Bloco 1 mostrado no Figura 7, são elas:

- A primeira importante interpretação é sobre a significância disposta na tabela “Teste de Omnibus do Modelo de Coeficientes”. Como os valores de significância são menores do que 0,05, pode-se aferir que o modelo com as variáveis adotadas é mais significativo do que o modelo que não as considera (modelo do Bloco 0);

- Uma outra informação importante pode ser retirada da tabela “Resumo do Modelo”. O valor de R quadrado Nagelkerke informa a porcentagem do conjunto de dados que pode ser explicada pelo modelo obtido. Como pode-se observar, o valor mostra que o modelo é capaz de explicar apenas 10% das variações registradas na variável dependente.

- Por último, uma interpretação dos valores de significância mostra que apenas a variável que indica produto do tipo esmalte possui valor p menor do que 0,05.

De modo geral, pode-se perceber que o conjunto de dados analisados não apresentou modelagem satisfatória. As interpretações identificadas são suficientes para interromper a análise e comprovar que os dados não favorecem uma modelagem através das técnicas de regressão, uma vez que não se pode predizer uma correlação satisfatória entre as variáveis desejadas.

5 CONCLUSÃO

Este estudo foi motivado por um problema vivenciado atualmente em uma empresa do ramo de cosméticos cujo tipo de comercialização de produtos é a venda direta. O problema está relacionado às frequentes variações de demanda observadas para os novos produtos, mais especificamente no período de lançamento destes produtos. Como toda a cadeia produtiva é baseada no volume de vendas estimado, essas variações precisam ser evitadas e, assim, é preciso entender o comportamento da demanda para tomar ações preventivas e se precaver ao risco de suas variações.

Para tentar identificar o comportamento dessas variações, foram realizadas análises de regressão. O objetivo principal foi determinar se existe correlação entre as variações de demanda e algumas características do produto, como: tipo de produto, tipo de inovação, grau de exposição no folheto. Primeiramente foi realizada a análise de regressão linear multivariada, para determinar se era possível estimar um valor provável de variação de acordo com as características inseridas. A análise foi feita com base em dados de vendas coletados dos lançamentos dos últimos 3 anos. Como os resultados foram insatisfatórios para comprovação de adequação ao modelo de regressão linear, pode-se concluir, assim, que os dados coletados não demonstram relação linear entre as variáveis analisadas. Um segundo recurso foi utilizado para tentar encontrar algum grau de relação: já que não pode-se estimar um valor provável de variação, foi analisado se era possível ao menos prever alto ou baixo risco de variação. Pensando nisso e na variação que causa mais dano para a companhia, o estudo foi direcionado para risco alto igual a variações acima de 30% e risco baixo para variações abaixo disso. As variáveis foram categorizadas, transformadas em binárias e a análise de regressão logística foi aplicada. Os resultados mostraram que as variáveis também não se ajustam ao modelo de regressão logística, logo não pode-se estimar, de acordo com as características dos produtos, se ele é mais propenso a altas variações ou não.

Assim, pode-se aferir que o conjunto de dados estudados não apresenta correlação capaz de ser comprovada pelas análises de regressão e, conseqüentemente, não é possível prever a informação desejada sobre o risco de variação de demanda através das variáveis analisadas.

REFERÊNCIAS

- BISSACOT, A. C. G. **Estudo comparativo entre regressão logística binária e redes neurais artificiais na avaliação dos resultados clássicos de Hosmer, Lemeshow e Sturdivant**. 2015. 106 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Itajubá, Itajubá, 2015.
- CABRAL, I. S. C. **Aplicação do modelo de regressão logística em um estudo de mercado**. 2013. 58 f. Dissertação (Mestrado em Matemática Aplicada à Economia e à Gestão) – Faculdade de Ciências da Universidade de Lisboa, Lisboa, 2013.
- CHOPRA, S.; MEINDL, P. **Gerenciamento da cadeia de suprimentos: estratégia, planejamento e operação**. 3ª ed. New Jersey: Prentice Hall, 2003.
- CORREA, S. M. B. B. **Probabilidade e estatística**. 2ª ed. 2003. Minas Gerais: Pontifícia Universidade Católica de Minas Gerais, 2003.
- FIGUEIRA, C. V. **Modelos de regressão logística**. 2006. 138f. Dissertação (Mestrado em Matemática) – Instituto de Matemática da Universidade Federal do Rio Grande do Sul, Rio Grande do Sul, 2006.
- HOFFMANN, R. **Análise de regressão: uma introdução à econometria**. 5ª ed. São Paulo: Editora Hucitec, 2016.
- HOSMER, D. W.; LEMESHOW, S.; STURDIVANT, R. X. **Applied logistic regression**. 3rd ed. Hoboken: John Wiley & Sons, 2013.
- LAMBERT, D. M.; COOPER, M. C. **Issues in supply chain management**. Flórida: The International Journal of Logistics Management, v. 29, p. 65-83, 2000.
- MARONA, R. A.; MARTIN, R. D.; YOHA, V. J. **Robust statistics: theory and Methods**. Hoboken: John Wiley & Sons, 2006.
- MISSIO, F.; JACOBI, L. F. Variáveis dummy: especificações de modelos com parâmetros variáveis. **Ciência e Natura**, Santa Maria, Departamento de Estatística da UFSM, Universidade Federal de Santa Maria, 2007.
- MONTGOMERY, D. C.; PECK, E. A.; Vining, G. G. **Introduction to linear regression analysis**. 4th ed. Hoboken: John Wiley & Sons, 2006.
- NOGUEIRA, F. E. **Modelos de regressão multivariada**. 2007. 89 f. Dissertação (Mestrado em Ciências) – Instituto de Matemática e Estatística da Universidade de São Paulo, Universidade de São Paulo, São Paulo, 2007.
- NORMANDO, D.; TJÄDERHANE, L.; QUINTÃO, C. C. A. A escolha do teste estatístico – um tutorial em forma de apresentação em PowerPoint. **Dental Press Journal of Orthodontics**, Maringá, v. 15, n. 1, p. 101-106, 2010.

PEINADO, J.; GRAEML, A. R. **Administração da produção:** operações industriais e de serviços. Curitiba: UnicenP, 2007.

STEVENSON, W. J. **Estatística aplicada à administração.** São Paulo: Harbra, 1986.

BIBLIOGRAFIA CONSULTADA

FLEURY, P. F. **Supply chain management: conceitos, oportunidades e desafios da implementação**. Rio de Janeiro: Editora Atlas, 2000.

FOGLIATTO, F. S. Estratégias para modelagem de dados multivariados na presença de correlação. **Gestão & Produção**, Porto Alegre, v.7, n.1, 2000. Disponível em: <http://dx.doi.org/10.1590/S0104-530X2000000100002>. Acesso em: 10 nov. 2018.

HAIR JÚNIOR., J.F. *et al.* **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009.

LAPIDE, L. Sales and operations planning (S&OP) mindsets. **The journal of Business Forecasting**, Flushing, v. 26, n. 1. 2007.

MESQUITA, P. S. B. **Um modelo de regressão logística para avaliação dos programas de pós-graduação no Brasil**. 2014. 90 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Estadual do Norte Fluminense, Rio de Janeiro, 2014.

MUNDSTOCK, E. *et al.* **Introdução à análise estatística utilizando o SPSS 13.0**. Porto Alegre: Instituto de matemática da Universidade Federal do Rio Grande do Sul, 2006. (trabalho de apoio didático).

VICINI, L. **Análise multivariada da teoria à prática**. Santa Maria: UFSM, CCNE. 2005.