

Aplicação da Regressão Linear Simples para análise do preço de um produto

Leonardo Zammar (Universidade Tecnológica Federal do Paraná) leonardozammar@alunos.utfpr.edu.br

Raniela Samanta da Silva (Universidade Tecnológica Federal do Paraná) rani.sam@hotmail.com

Yslene Rocha Kachba (Universidade Tecnológica Federal do Paraná) yslener@utfpr.edu.br

João Luiz Kovaleski (Universidade Tecnológica Federal do Paraná) kovaleski@utfpr.edu.br

Resumo:

O presente artigo trata-se de uma análise do preço da gasolina, bem como, da composição do preço, utilizando o etanol como base de cálculo e variável independente, vide correlação linear entre ambos os produtos. Utilizou-se dados reais do ano de 2016 para obtenção de valores de intercepção e inclinação da reta, para aplicação no modelo de Regressão Linear Simples e posterior previsão do valor da gasolina para o ano de 2017. Os valores previstos para o ano de 2017 mostraram convergência no modelo estatístico escolhido para a análise do tema, e, constata-se, ainda, que há uma dependência de fatores externos, que fogem da regra explicitamente padronizada. Assim, na tentativa de encontrar uma explicação para tais comportamentos, diversos fatores são levantados, evidenciando que existem relações de extrema importância no cunho econômico e político para composição dos preços.

Palavras chave: Previsão, Gasolina, Análise.

Application of Simple Linear Regression to analyze the price of a product

Abstract

The present article deals with an analysis of the price of gasoline, as well as of the composition of the price, using ethanol as the basis of calculation and independent variable, see linear correlation between both products. Real data from the year 2016 were used to obtain intercept and slope values for the application of the Simple Linear Regression model and subsequent prediction of the value of gasoline for the year 2017. The predicted values for the year 2017 showed convergence in the statistical model chosen for the analysis of the theme, and, it is also observed that there is a dependence on external factors, which deviate from the explicitly standardized rule. Thus, in an attempt to find an explanation for such behavior, several factors are raised, evidencing that there are extremely important relations in the economic and political nature for price composition.

Key-words: Prediction, Gasoline, Analysis.

1. Introdução

O cenário econômico brasileiro, quando analisado historicamente, possui a característica mercadológica de instabilidade. Essas mudanças, provenientes deste contexto de oscilações, atingem diretamente o cotidiano do consumidor porque interferem no preço, qualidade e outros aspectos determinísticos em bens de consumo essenciais.

Exemplificando, um bem de consumo essencial é a gasolina, combustível utilizado no funcionamento de motores à explosão. Através dela, é possível que pessoas possam se deslocar até seus respectivos trabalhos, levar os filhos à escola, ir ao supermercado comprar alimentos, que é outro bem de consumo essencial, entre outras atividades.

Partindo deste princípio, o objetivo do presente trabalho é realizar uma previsão do preço deste produto, visto que possui ampla interferência na vida da população brasileira. Analisando a série histórica de preços de combustíveis no Brasil fornecidas pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, realizou-se a previsão de valores da gasolina, utilizando como base o ano de 2016.

Cada mercado possui determinadas necessidades que irão intervir na previsão, como por exemplo, a economia do país no momento estudado, taxas de juros e até mesmo escândalos de corrupção.

2. Referencial Teórico

Com as crises econômicas recorrentes não só no Brasil, mas no mundo, os empresários atuantes sentiram necessidade de desenvolver novos planos estratégicos para alavancarem suas organizações.

Um possível plano, crucial para a tomada de decisão, é a previsão de demanda. Segundo Lemos (2015), a formulação de planejamentos e o direcionamento estratégico das empresas dependem da identificação e previsão correta das mudanças emergentes[...], identificando riscos.

Previsões de demanda podem ser elaboradas com uma gama de diferentes métodos de obtenção e análise de dados, como quantitativos, qualitativos e até mesmo a combinação dos dois anteriores.

Para a realização de métodos quantitativos de demanda, faz-se necessária a obtenção de dados históricos, que permitam uma análise e aplicação para uma demanda futura. Já a demanda futura, tem como requisito básico, uma construção de modelo, seja matemático ou estatístico, que melhor se adeque aos dados que estão disponíveis. A disponibilidade dos dados que melhor descrevem a evolução da demanda ao longo do tempo, é chamada de série temporal. Partindo desse princípio, as diversas técnicas reconhecidas para resolução de tais problemas, são as chamadas técnicas de *forecasting*.

Na maioria das vezes, os métodos de origem qualitativa baseiam-se em opiniões de pessoas especializadas, seja por experiência pessoal, expectativas por parte de consumidores ou por conceitos adquiridos. Porém, como essa técnica pode ter interpretações distintas, isso compromete a confiabilidade das amostras analisadas, seja por tendências ou variabilidades.

As previsões possuem níveis estratégicos, táticos e operacionais, podendo variar entre longo prazo (5 anos), médio prazo (1 ano) e curto prazo, respectivamente. Para tanto, necessitamos

de métodos quantitativos matemáticos. A escolha do modelo matemático será em decorrência dos resultados das análises das séries temporais.

Modelos de suavização exponencial têm, também, uma utilização de maneira ampla para as diversas análises de previsão de demandas, devido a sua simplicidade, ligada à facilidade de ajustes. Tal método utiliza um sistema de ponderações, de forma diferente para cada valor que seja observado, dos mais antigos aos mais recentes. De tal forma, assim, a junção de todas as ponderações forma um conjunto, que, por sua característica, tem um decaimento exponencial, partindo dos valores que estão mais recentes à análise.

Para a obtenção do método de Suavização Exponencial, faz-se necessária a obtenção de uma constante de suavização, denominada de α ; tal valor é arbitrário, ou, de forma que utilize-se programação computacional, encontra-se um valor ótimo para a constante.

No caso do método arbitrário de obtenção da constante, pode-se fazer o uso do método iterativo, utilizando alguma forma de comparação. Ao selecionar um valor aleatório inicial para tal constante, as previsões serão geradas partindo desse ponto. Comparam-se os valores de previsão com os valores reais. Para fins de análise de erros, calcula-se a média do quadrado das diferenças entre os mesmos. Como o objetivo da utilização é para a obtenção do menor erro projetado, escolhido será aquele modelo que melhor adequar-se a essa premissa.

$$B_t = \alpha \times D_t + (1 - \alpha) \times B_{t-1}$$

Equação 1 – Equação da Suavização Exponencial Simples

Em que B_t é a base de suavização para o período em que se deseja conhecer, D_t a demanda no período t e α a constante de suavização, como mencionado anteriormente

O modelo de Regressão Linear Simples visa identificar a relação entre as variáveis dependentes e independentes. No caso da simples, apenas uma variável é independente.

A função regressão é delimitada por:

$$y = a + b \times x$$

Equação 2 – Equação linear simples

Em que y é a variável dependente, x a variável independente, a , a constante de intercepção de y e b é a inclinação da linha de tendências.

Segundo Rodrigues, Medeiros e Gomes (p. 610, 2013), é necessário desenvolver um trabalho preliminar para determinar se um modelo linear é adequado. O processo mais simples consiste em fazer o diagrama de dispersão[...]. Esses diagramas de dispersão podem apresentar correlação fraca, forte ou perfeita, podendo ser ainda inversamente ou diretamente proporcional.

A análise de correlação é necessária em diversos tipos de pesquisas, pois, frequentemente, procura-se verificar se existe relação entre duas ou mais variáveis, isto é, saber se as alterações sofridas por uma das variáveis são acompanhadas por alterações nas outras, que se deseja realizar o estudo e validar tal pesquisa.

$$r = \frac{n \times \sum_{t=1}^n x_t \times d_t - \sum_{t=1}^n x_t \times \sum_{t=1}^n d_t}{\sqrt{n \times \sum_{t=1}^n x_t^2 - (\sum_{t=1}^n x_t)^2 \times [n \times \sum_{t=1}^n d_t^2 - (\sum_{t=1}^n d_t)^2]}}$$

Equação 3 – Equação da correlação linear

De tal forma, o coeficiente de correlação r é um número que varia de -1 a $+1$, e sua interpretação dependerá do valor numérico e do sinal.

Para $r = -1$ diz-se que a correlação é perfeita negativa; para $r = 1$ diz-se que a correlação é perfeita positiva; para $\pm 0,2 < r < \pm 0,4$ diz-se que a correlação é fraca; para $\pm 0,4 < r < \pm 0,7$ diz-se que a correlação é moderada e para $\pm 0,7 < r < \pm 0,9$ diz-se que a correlação é forte.

3. Metodologia

Para a realização das análises presentes neste artigo, inicialmente buscou-se um levantamento histórico de preços dos combustíveis, mais precisamente Gasolina e Etanol. Tais dados estão disponíveis para acesso público, pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis.

Após levantamento dos dados, realizou-se a separação dos anos 2016 e 2017 para análise, com uma amostra dos 12 meses para cada ano.

Com os dados contidos em uma planilha, verificou-se graficamente o comportamento apresentado durante os períodos mensais, para posterior adequação aos modelos estatísticos. Em seguida, verificou-se a correlação linear entre os dados de Gasolina e Etanol.

Dois modelos estatísticos para a projeção de demanda foram eleitos para análise, sendo eles: Regressão Linear Simples e Suavização Exponencial.

Assim, por meio de software (Microsoft Excel®), obteve-se os coeficientes de intercepção e inclinação com os dados de Gasolina e Etanol do ano de 2016, para posterior utilização no método de Regressão Linear Simples.

Em posse dos mesmos dados do ano de 2016 para os combustíveis, realizou-se a aplicação do método de Suavização Exponencial, em que a Base Inicial foi a média dos valores de demanda.

A fim de comprovar a viabilidade da pesquisa, aplicou-se ambos os métodos. Após a obtenção das projeções, realizou-se a comparação do valor real, com o obtido, por meio da equação:

$$|\text{Valor Real} - \text{Valor Obtido}| \times 100$$

Equação 4 – Erro de previsão percentual

Dessa forma, tornou-se possível realizar a média dos erros de previsão percentuais, os quais integralizaram o erro de previsão médio percentual.

Em posse de tais informações, projetou-se o valor de Gasolina para o ano de 2017, utilizando o Etanol do ano de 2017 como variável independente, e os valores de intercepção e inclinação já obtidos na aplicação do ano de 2016.

4. Resultados e Discussão

A escolha do método de Regressão Linear Simples justifica-se pelo fato de que apresentou um erro de previsão médio percentual de 1% (vide Tabela 1), se comparado ao valor real da amostra.

O resultado da correlação linear entre as variáveis de Gasolina e Etanol foi de 0,83; em que, conforme referencial teórico, apresenta uma correlação forte.

Período	Preço Gasolina	Preço Etanol	F(t)	Erro
Janeiro	3,68	2,72	3,69	2%
Fevereiro	3,71	2,81	3,71	0%
Março	3,73	2,87	3,72	1%
Abril	3,72	2,74	3,70	2%
Maio	3,67	2,48	3,65	2%
Junho	3,65	2,47	3,65	0%
Julho	3,65	2,45	3,65	0%
Agosto	3,65	2,49	3,65	0%
Setembro	3,65	2,51	3,66	1%
Outubro	3,66	2,68	3,68	2%
Novembro	3,67	2,81	3,71	4%
Dezembro	3,73	2,83	3,71	2%
				EM 1%

Fonte: os autores (2017)

Tabela 1 – Projeção do ano de 2016 pelo método de Regressão Linear Simples

Já a Suavização Exponencial, apresentou 3% de erro de previsão médio (vide Tabela 2), se comparado ao valor real da amostra.

Período	Preço Gasolina	B(t)	F(t)	Erro
		3,68		
Janeiro	3,68	3,68	3,68	0%
Fevereiro	3,71	3,68	3,68	3%
Março	3,73	3,68	3,68	4%
Abril	3,72	3,69	3,69	3%
Maio	3,67	3,68	3,68	1%
Junho	3,65	3,68	3,68	4%
Julho	3,65	3,68	3,68	4%
Agosto	3,65	3,67	3,67	3%
Setembro	3,65	3,67	3,67	3%
Outubro	3,66	3,67	3,63	1%
Novembro	3,67	3,67	3,67	0%
Dezembro	3,73	3,67	3,67	5%
				EM 3%

Fonte: os autores (2017)

Tabela 2 – Projeção do ano de 2016 pelo método de Suavização Exponencial

Em posse dos dados de intercepção (3,229) e inclinação (0,170) e da projeção do ano de 2016, realizou-se, então, a projeção para o ano de 2017.

Os resultados de demanda projetados para a Gasolina no ano de 2017 encontram-se explícitos na tabela a seguir:

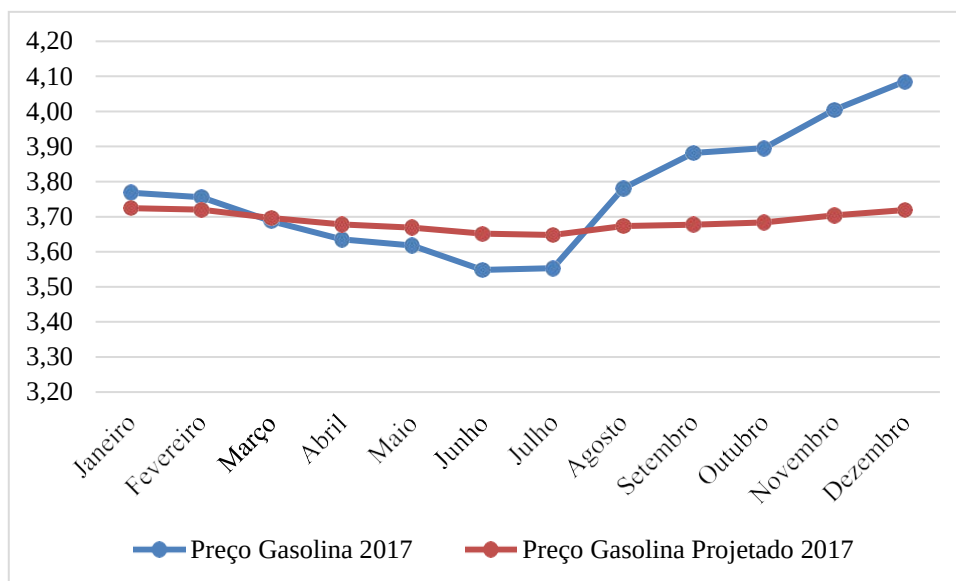
Período	Preço Gasolina	Preço Etanol	F(t)	Erro
Janeiro	3,77	2,91	3,72	4,41%
Fevereiro	3,76	2,88	3,72	3,55%
Março	3,69	2,75	3,70	0,88%
Abril	3,64	2,64	3,68	4,28%
Maio	3,62	2,58	3,67	5,14%
Junho	3,55	2,49	3,65	10,36%
Julho	3,55	2,46	3,65	9,43%
Agosto	3,78	2,61	3,67	10,80%
Setembro	3,88	2,63	3,68	20,42%
Outubro	3,90	2,67	3,68	21,23%
Novembro	4,00	2,79	3,70	30,05%
Dezembro	4,09	2,88	3,72	36,64%
				EM 13,10%

Fonte: os autores (2017)

Tabela 3 – Projeção do ano de 2017 pelo método de Regressão Linear Simples

Dessa forma, pode-se analisar que os modelos de Regressão Linear Simples realmente foram convergentes para ambas as projeções, apresentando um valor próximo do real. Porém, a informação que chama atenção é a crescente no erro de previsão apresentado a partir de junho do ano de 2017, com uma leve queda em julho, para posterior manutenção e dobra no mês de setembro. Elevando-se o erro aos 36,64% no último mês do ano. Fato que refletiu de forma direta no erro médio.

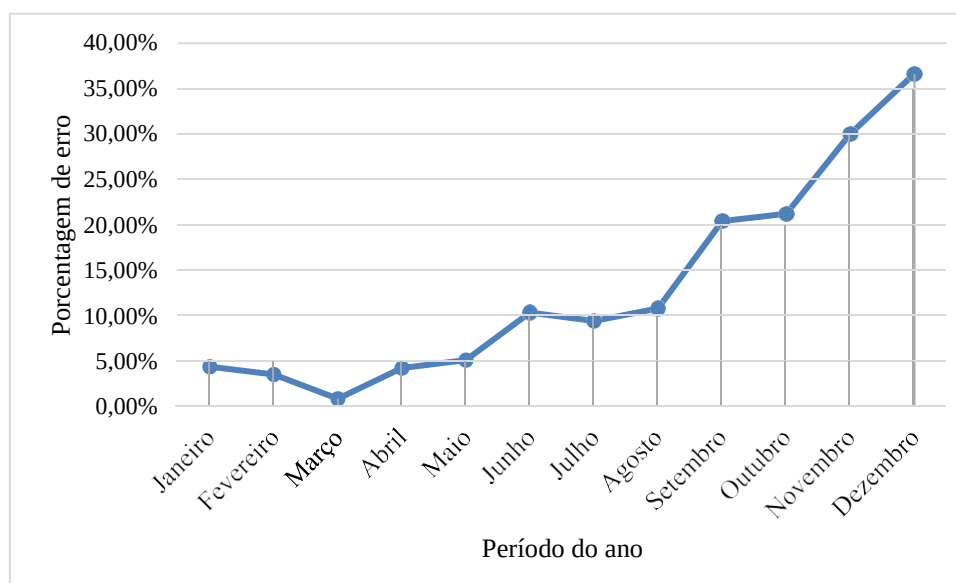
Nos dois gráficos a seguir, pode-se notar explicitamente tais variações. O gráfico 1 trata da comparação entre o preço real praticado pela Gasolina no ano de 2017 e o preço projetado para o ano de 2017.



Fonte: os autores (2017)

Gráfico 1 – Comparação entre o preço real e projetado para a gasolina em 2017

O gráfico 2 demonstra o erro de projeção percentual ao longo do ano de 2017.



Fonte: os autores (2017)

Gráfico 2 – Progressão do erro de projeção percentual para o ano de 2017

Para explicar essa crescente na variação, existem diversos fatores que podem ser analisados.

Um dos principais componentes relevantes para a formação de preço dos combustíveis é o repasse de impostos, que reflete em todos os tipos de energéticos e derivados. Assim, a partir do momento em que a dependência é externa ao que se prevê estatisticamente, eleva-se o nível do erro de previsão, visto que são fatores que não estão intrínsecos ao modelo estatístico, dificultando análise de progressão. As alíquotas de PIS/Cofins são responsáveis por boa parte da formação do preço final.

Segundo Gadotti (2017), outro fator relevante para a alta foi o preço do etanol anidro, que subiu 32% desde que julho de 2017.

Um ponto extremamente relevante, é o reajuste praticamente diário e seguido, realizado pela refinaria, evidenciado a partir de julho do ano de 2017. Sobre a nova política de repasses, (Colombo 2017) identifica que o barril de petróleo experimentou uma elevação de 33,4% nos mercados internacionais. Dessa forma, passou de US\$ 48,69, no mês de julho, para US\$ 62,57, no mês de novembro.

A refinaria alega que todo o reajuste repassado aos distribuidores, e, por conseguinte, consumidores, foi feito para se adequar ao elevado reajuste do preço dos barris de petróleo no mercado externo. Tal atitude fere a todos os stakeholders, visto que para as distribuidoras, se torna complicada a gestão de estoque, porque o pagamento é realizado para a refinaria no dia da entrega do combustível. Então, visto que a tendência é que haja um novo aumento num momento próximo, dessa forma, na chegada da produto, pode ser que o preço praticado seja maior do que o esperado no momento da realização da encomenda.

Diversas são as consequências da adotada política de reajuste de preços, que, inegavelmente, é responsável pelo impacto direto na inflação e no custo de vida dos brasileiros. As seguidas altas nos preços dos combustíveis, eleva o preço, também, de qualquer tipo de frete realizado dentro das instalações do país, seja no transporte de alimentos, vestuário, entre outros. Todos fazem parte de uma cadeia logística produtiva, que depende, diretamente, dos combustíveis, e, por consequência, precisam também, repassar seus custos para equilíbrio do balanço.

5. Considerações finais

De acordo com o que pode ser visto nesse estudo, a projeção do preço da Gasolina, baseando-se no preço do Etanol, obedeceu a lógica estatística com um baixo erro relativo em grande parte dos períodos analisados, se comparados aos valores realmente praticados no ano de 2017.

Porém, tornou-se extremamente difícil a análise a partir do momento em que a alta dos preços deixou de seguir especificamente a relação com o preço do Etanol, como visto anteriormente.

Ainda, além dos fatores que tentam explicar os aumentos consecutivos, já tratados durante e ao longo do presente trabalho, há o fator político e o estado atual em que o Brasil se encontra. São inúmeras incertezas em um período de ajustes e recuperação da confiança de investimento.

A incerteza e desconfiança política fazem com que o comércio interno, bem como, os produtos que são exportados, sejam extremamente dependentes da demanda monetária de países com maior poder aquisitivo, como os Estados Unidos e o Dólar.

A alta praticada no preço cobrado pelos barris de petróleo no exterior reflete diretamente nas refinarias e no bolso do consumidor, visto que, boa parte do petróleo utilizado para o combustível brasileiro é de origem de importação, mesmo com o Brasil sendo um dos maiores produtores mundiais.

Segundo Iramina (2016), tal ponto se explica pelo fato de que se olharmos pela quantidade, o Brasil é autossuficiente. Agora, pela qualidade, ainda somos dependentes de petróleo importado.

Portanto, para uma análise mais aprofundada sobre a composição e projeção dos preços de combustíveis, faz-se necessário o conhecimento aprofundado da parte econômica, tributária e, mais especificamente, do comércio exterior; bem como, o conhecimento dos fatores que acompanham a balança econômica do petróleo no Brasil.

Referências

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS (ANP). *Levantamento de preços: Série histórica do levantamento de preços e margens de comercialização de combustíveis.* Disponível em: <<http://www.anp.gov.br/wwwanp/precos-e-defesa/234-precos/levantamento-de-precos/868-serie-historica-do-levantamento-de-precos-e-de-margens-de-comercializacao-de-combustiveis>> Acesso em: 05/04/2018.

DE SOUZA MELO, A.; SAMPAIO, Y. de S. B. *Impactos dos preços da gasolina e do etanol sobre a demanda de etanol no Brasil.* *Revista de economia Contemporânea*, v. 18, n. 1, p. 56-83, 2014.

HONAISSER, E. H. R.; SAUAIA, A. C. A. *Desenvolvimento e esenvolvimento aplicacao de um modelo para previsao de demanda em jogos de empresas.* *RAC-Electronica*, v. 2, n. 3, p. 470-486, 2008.

LEMONS, F. de O. *A importância da previsão de demanda no processo de tomada de decisão.* Disponível em: <<http://www.crea-sc.org.br/portal/index.php?cmd=artigos-detalle&id=3370#.Wssep4jwbIU>> Acesso em 05/04/2018.

MICROSOFT EXCEL. Ver. 1802 (2016). Microsoft Co., Redmond, WA.

MONTGOMERY, D. C., RUNGER, G. C. *Applied statistics and probability for engineers*, John Wiley, New York, 1994.

SEBER, G. A. F. *Linear regression analysis*, John Wiley, New York, 1977.

SOUZA, J. G. de M.; POMPERMAYER, F. M. *Variações no preço do etanol em comparação ao preço da gasolina: uma análise da resposta do consumidor.* 2015.

RESENDE, C. B.; SCARPEL, R. A. *Análise dos Fatores Determinantes na Precificação de Bens de Consumo—Um Caso na Industria Automobilística.* *X Simpósio de Pesquisa Operacional e Logística da Marinha*, 2007.

RIOS, D. M.; DA SILVA, J. G.; SILVEIRA, K. B. V. *A modelagem matemática e o preço da gasolina c no Brasil.* *Anais do Salão Internacional de Ensino, Pesquisa e Extensão*, v. 7, n. 1, 2016.

RODRIGUES, R. L.; MEDEIROS, F. P. A.; GOMES, A. S. *Modelo de Regressão Linear aplicado à previsão de desempenho de estudantes em ambiente de aprendizagem.* *CBIE*, 2013. Disponível em: <<http://br-ie.org/pub/index.php/sbie/article/view/2539>> Acesso em 08/04/2018.