

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

Resumo: As organizações deparam-se constantemente com questões críticas relativas a sua competitividade futura e crescimento organizacional. Dentro desse contexto as decisões gerenciais desempenham um papel crucial sendo responsável pela competitividade e os resultados frente a um ambiente dinâmico. Por isso os gestores necessitam de orientação tanto quanto a direcionamentos futuros quanto a variáveis que influenciam nos seus planejamentos. Uma orientação correta e precoce aumenta a probabilidade de chances de sucesso nos processos de tomada de decisão. O presente trabalho tem o objetivo de construir um modelo econométrico que auxilie no processo de tomada de decisão gerencial. Esse modelo deverá ser capaz de identificar o comportamento da demanda do produto bem como as principais variáveis externas que afetam sua demanda, assim contribuindo para que a decisão seja tomada com maior confiabilidade. Para isso observou-se o comportamento das vendas de uma grande instituição utilizando a técnica de análise de regressão múltipla. O modelo proposto obteve um resultado satisfatório.

Palavras-chave: econometria, previsão, estatística, finanças.

Bruno Bourgard Magalhães Garcia

Universidade Federal Fluminense

bbourgard@bol.com.br

Ricardo Bordeaux-Rego

Universidade Federal Fluminense

ribordeaux@gmail.com

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

1. Introdução

Castiglioni em seu estudo em 2008 definiu a previsão da demanda como base para o planejamento estratégico da produção, vendas e finanças de qualquer empresa, que assim, pode desenvolver seus planos de capacidade, fluxo de caixa, vendas, produção, estoque, mão de obra, compras, entre outras atividades empresarias.

De uma maneira geral, sistemas de previsão de demanda têm o propósito de oferecer aos gestores um apoio para a compreensão do comportamento da demanda de seus clientes. Podem ser usados na tomada de decisão sobre o estabelecimento de preços para produtos, no planejamento da capacidade e na inserção de um novo produto no mercado. (Destro 2011).

Os métodos propostos para a estimação da demanda são muitos, e dependem das características específicas do mercado no qual o bem cuja demanda será estimada está inserido.

Em nosso estudo os métodos causais são utilizados e, de acordo com Bowersox e Closs (2010), as previsões comprovadamente mais confiáveis são baseadas em relações causais, embora não seja necessário haver uma relação de causa e efeito entre a venda de um produto e as variáveis explicativas utilizadas.

Ballou (2001) comenta que modelos causais podem aparecer nas seguintes formas:

- Estatístico: Modelos Econométricos (Regressão por exemplo)
- Descritivo: Modelos de Ciclo de vida, simulação e modelos de entrada e saída.

Ainda de acordo com Bowersox e Closs (2010) uma previsão por regressão estima as quantidades vendidas para cada produto com base em variáveis independentes ou explicativas. Conforme os estudos de Slack, Chambers & Johnston (2008) esses modelos econométricos são utilizados em ambientes onde o nível de previsão é altamente exigente e complexo, exigindo normalmente altos custos de desenvolvimento, porém, permitindo representações mais realistas.

2. O Modelo Clássico de Regressão

Quando se imagina um par de variáveis aleatórias ligadas por uma relação linear da forma:

$$Y = \alpha + \beta X + \varepsilon \quad (1)$$

Este modelo é chamado uma regressão de Y em X, onde:

Y é a variável dependente

X é a variável independente

ε é o erro aleatório

α e β são os coeficientes da regressão

Os parâmetros α e β podem ser estimados utilizando um método adequado, por exemplo, o Método de Mínimos Quadrados.

Estendendo (1) ao caso em que há mais de uma variável independente, temos a seguinte equação:

$$Y = \beta X + \varepsilon \quad (2)$$

Onde:

β é um vetor $k \times 1$ de parâmetros a serem estimados

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

X é uma matriz nxk composta pelas variáveis explicativas

Y é o vetor nx1 composto pelos valores observados da variável dependente.

Por extenso temos:

$$\begin{aligned} Y_1 &= \beta_1 X_{11} + \beta_2 X_{12} + \dots + \beta_k X_{1k} + \varepsilon_1 \\ Y_2 &= \beta_1 X_{21} + \beta_2 X_{22} + \dots + \beta_k X_{2k} + \varepsilon_2 \\ &\dots \\ Y_n &= \beta_1 X_{n1} + \beta_2 X_{n2} + \dots + \beta_k X_{nk} + \varepsilon_n \end{aligned}$$

A relação (2) é chamada regressão linear múltipla e, para este caso, o método de Mínimos Quadrados pressupõe as seguintes hipóteses:

- A média dos erros é zero: $E(\varepsilon) = 0$
- $Cov(\varepsilon) = \sigma^2 I$ onde I é a matriz identidade
- Homocedasticidade: $E(\varepsilon_i, \varepsilon_j) = \sigma^2$
- X é não estocástica

O MQO (mínimos quadrados ordinários) é equivalente à procura de parâmetros estimados que minimizem a soma dos quadrados dos erros de predição (ou resíduos), ou seja, devemos minimizar

$$\sum (Y_i - \beta X_i - \alpha)^2$$

Os valores minimizadores são:

$$\begin{aligned} \alpha_1 &= \frac{\sum Y}{n} \\ \alpha &= \bar{Y} - \bar{X}\beta \\ \beta &= \frac{n \sum (XY) - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} \end{aligned}$$

2.1 O Modelo de Regressão Múltipla em termos Matriciais

A expressão do modelo linear geral de regressão é dada por:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{ip-1} + \varepsilon_i$$

Em termos matriciais podemos definir:

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

$$\underset{nx1}{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} \quad \underset{nxp}{X} = \begin{bmatrix} 1 & X_{11} & \cdot & \cdot & X_{1p-1} \\ 1 & X_{21} & \cdot & \cdot & X_{2p-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & X_{n1} & \cdot & \cdot & X_{np-1} \end{bmatrix} \quad \underset{px1}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \beta_{p-1} \end{bmatrix} \quad \underset{nx1}{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix}$$

X é o vetor de variáveis aleatórias independentes e normalmente distribuídas com esperança $E(X) = 0$ e matriz de variância-covariância dada por:

$$\sigma^2(\varepsilon) = \begin{bmatrix} \sigma^2 & 0 & \cdot & 0 \\ 0 & \sigma^2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \sigma^2 \end{bmatrix} = \sigma^2 I$$

Assim, o vetor das observações Y tem esperança e variância dadas por:

$$\underset{nx1}{E(Y)} = X\beta \quad \underset{nxn}{\sigma^2(Y)} = \sigma^2 I$$

O sistema de equações normais para o modelo em questão é dado por:

$$X' X \beta = X' Y$$

E os estimadores de mínimos quadrados são:

$$\beta = (X' X)^{-1} X' Y$$

que é linear, não tendencioso e de variância mínima (BLUE) cuja variância é estimada por:

$$S^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{(n - k)}$$

2.2 Resíduos

O i-ésimo valor do resíduo é a diferença entre o valor observado Y_i e o correspondente valor ajustado \hat{Y}_i . Denota-se esse valor por ε_i , e pode ser escrito como:

$$\varepsilon_i = Y_i - \hat{Y}_i = Y_i - \beta_0 - \beta_1 X_i$$

É preciso distinguir entre o erro do modelo e o resíduo ε_i . O erro do modelo é o desvio de Y_i da linha real de regressão desconhecida. E o resíduo é o desvio de Y_i do valor

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

estimado da regressão, por isso é conhecido. Os resíduos são úteis para responder se uma regressão é apropriada para os dados que estão sendo trabalhados. Mais à frente, na fase de validação do modelo será visto como os resíduos podem auxiliar.

2.3 Propriedades de Regressão Linear Ajustada

A regressão ajustada pelo método de mínimos quadrados tem algumas propriedades:

- 1 - A soma dos resíduos é igual à zero
- 2 - A soma dos quadrados dos resíduos é mínima.
- 3 - A soma dos valores observados é igual à soma dos valores ajustados
- 4 - A soma do produto dos valores ajustados e resíduos é igual à zero
- 5 - A soma do produto dos valores de x e os resíduos é igual à zero

3. Descrição dos dados

Para a construção dessa análise observamos os dados relacionados às vendas de uma grande companhia no período de janeiro de 2011 até dezembro de 2014.

Com objetivo de estimar a demanda diária para o bem em questão construímos um modelo econométrico.

As variáveis utilizadas para a construção do modelo são descritas seguir:

I-Venda Quantidade do produto vendida.	II- Vendas em Períodos Anteriores Quantidade do produto vendida em períodos passados.
III- Condição Climática É a condição diária do tempo, podendo ser tempo bom, tempo nublado ou tempo chuvoso. A condição do tempo é medida preferencialmente pela manhã, dado que a compra do produto em estudo tem caráter matinal.	IV- Dia da Semana Dia da semana que se refere à venda.
V-Mês Mês que se refere à venda.	VI- Feriado Referente aos dias em que a venda foi realizada em algum feriado.
VII- Campanha Promocional Dias em que o produto foi vendido acompanhado de alguma promoção do tipo "casadinha", que se caracteriza pela venda de algum produto vinculado.	VIII-Preço dos Concorrentes Valor dos dois principais concorrentes vendidos na cidade do Rio de Janeiro.
IX-Salário Mínimo Real Valor referente ao salário mínimo real vigente em cada mês.	

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

4. Objetivos

O objetivo do estudo é a construção de um modelo econométrico baseado em regressão múltipla utilizando como alternativa ao método dos Mínimos Quadrados Ordinários (MQO) o de Mínimos Quadrados Ponderados para realização do cálculo das estimativas dos parâmetros do modelo, uma vez que problemas de heterocedasticidade são muito frequentes em estimativas de demanda.

Dessa maneira busca-se a construção de uma ferramenta de apoio a decisão e planejamento da produção visando a minimizar o erro médio entre a demanda real e a demanda estimada aumentando a receita total.

5. Formulação do modelo

Nosso objetivo nessa etapa é encontrar o modelo que melhor se ajuste às vendas do produto. Foi necessário o uso de variáveis dummy, pois em algumas datas as vendas são diferenciadas, são exemplos dessas datas: carnaval e natal.

Segue abaixo (Tabela 1) o comportamento de acordo com o mês das vendas.

Tabela 1 - Médias e desvios padrões por mês e ano na região local

Ano	jan	fev	mar	abr	mai	jun	jul	Ago	set	out	nov	dez
2011	31.520	34.902	30.775	31.030	33.155	29.820	28.940	32.838	29.796	34.200	31.026	29.577
	28.443	31.872	31.261	29.113	32.025	29.142	28.011	34.421	29.800	31.823	29.942	27.546
2012	34.367	33.597	31.460	32.699	33.945	30.932	34.048	30.540	29.395	29.503	27.314	25.060
	30.826	29.678	29.756	30.030	31.987	29.237	29.866	28.368	27.665	28.094	26.384	21.895
2013	28.905	29.032	28.421	31.060	28.100	27.063	29.732	29.805	31.161	33.681	26.728	26.384
	24.787	26.176	26.162	27.364	25.798	23.713	26.240	29.870	30.339	31.686	24.384	22.762
2014	26.997	27.393	26.515	28.967	25.364	26.454	30.666	26.802	28.632	24.755	25.474	24.624
	23.454	22.877	24.486	25.178	23.433	24.829	28.066	26.494	27.098	23.322	23.545	22.923

As seguintes variáveis foram consideradas para o estudo: clima, preço, preço dos concorrentes, salário mínimo, venda no mesmo dia do mês anterior, venda no mesmo dia do ano anterior e as seguintes variáveis *dummy*: para os anos de 2011, 2012 e 2013, carnaval, feriado, promoção do produto, para os meses e para os dias da semana.

5.1 Modelo Inicial

Percebe-se na Tabela 2 que se consegue um bom ajuste, já que o R^2 (coeficiente de determinação), medida de proporção da variância da variável dependente em torno de sua média que é explicada pelas variáveis independentes, do modelo é alto, cerca de 99,3% da variabilidade das vendas conseguiu ser explicada. Além dessa estatística observa-se também na mesma tabela um alto valor para o R^2 ajustado (ajusta o R^2 em função do número de variáveis independentes) Pela estatística F concluímos que o modelo é adequado, porém pelo teste T o qual analisa separadamente cada variável concluímos que algumas variáveis não são significativas ao nível de significância de 5 % são estatisticamente iguais a zero. Esse problema nos sugere que poderíamos estar com problemas de multicolineariedade.

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

Tabela 2

R-2	0.993832	Média da Var Dependente	29647.96
R-2 Ajustado	0.993676	D.P da Var Dependente	27423.55
Estatística F	1.780.122	Prob(F-statistic)	0.000000

Pela matriz de correlação observa-se que algumas variáveis são altamente correlacionadas e após uma análise de suas correlações algumas variáveis foram excluídas do modelo.

Foram eliminadas as variáveis: CLIMA1, CLIMA2, DABR, DFERIADO, DJUN, DMAI, DMAR, DNOV.

5.2 Modelo Inicial

Obtiveram-se para o modelo resultante as estatísticas na Tabela 3. Percebe-se que apesar da variável preço ser não significativa ela permanece no modelo, pois ela é fundamental para a explicação do nosso modelo, pois segundo a teoria econômica a quantidade demandada por um bem está diretamente relacionada com o seu preço. O r^2 encontrado é alto o que significa que grande parte (cerca de 99,36%) da variabilidade das vendas foi explicada. O modelo foi representado na Saída 03.

Tabela 3

R-2	0.993694	Média da Var Dependente	29644.36
R-2 Ajustado	0.993624	D.P da Var Dependente	27414.50
F-statistic	14220.41	Prob(F-statistic)	0.000000

Saída 03

Equação Estimada

=====

=====

VENDAS = C(1) + C(2)*D2005 + C(3)*D2006 + C(4)*DCARNAVAL + C(5)*DDOM + C(6)*DFEV + C(7)*DJAN + C(8)*DJUL + C(10)*DOUT + C(11)*DSAB + C(12)*DSEG + C(13)*DSEX + C(14)*DTER + C(16)*SALARIO + C(17)*PRECO

Substituindo os
coeficientes:

=====

=====

VENDAS = - 2545.763 + 499.3438*D2005 + 274.9816*D2006 - 3517.301*DCARNAVAL + 10438.41*DDOM + 722.2103*DFEV + 519.1466*DJAN + 513.7440*DJUL + 667.6918*DOUT + 1667.550*DSAB - 560.9762*DSEG + 611.3000*DSEX - 669.4199*DTER + 5.791902*SALARIO - 470.1676*PRECO

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

6. Validação do modelo

Nesta etapa testaremos algumas hipóteses do modelo.

6.1 Heterocedasticidade

Pelo teste White, visto na Tabela 4, rejeita-se a hipótese de homocedasticidade do modelo, que é um pressuposto básico do MQO. Tentaremos a correção usando o estimador de MQP.

Tabela 4

Teste de Heterocedasticidade de White :			
Estatística F	3.064.956	Probabilidade	0.000000
Obs*R-2	1.054.504	Probabilidade	0.000000

Procedimento para a estimação por MQP:

- Estimamos o modelo por MQO e obtemos os resíduos;
- Fazer regressão do resíduo ao quadrado em constante;
- Usar como ponderador o valor ajustado de (b).

Percebemos, observando a tabela 5, que ao usar o estimador de Mínimos Quadrados Ponderados, tanto o R^2 quanto o $R^2_{ajustado}$ aumentaram, ou seja, conseguimos um ajuste melhor. Quanto ao problema de heterocedasticidade, se observamos novamente o teste de White que está na Tabela 6, percebemos que a hipótese de homocedasticidade foi aceita.

Tabela 5

R-2 Modelo Ponderado	0.993694	Média da Var Dependente	29644.36
R-2 Ajustado	0.993624	D.P da Var Dependente	27414.50
Estatística F	14220.41	Prob(F-statistic)	0.000000

Tabela 6

Teste de Heterocedasticidade de White :			
Estatística F	Estatística	Probabilidade	0.01178
Obs*R-2	Obs*R-2	Probabilidade	0.013978

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

Modelo Final

Saída 07

Equação Estimada =====

VENDAS = C(1) + C(2)*D2006 + C(3)*DDOM + C(4)*DFEV + C(5)*DJAN + C(6)*DJUL + C(8)*DOUT + C(9)*DSAB + +
C(11)*VENDA(-1) + C(12)*VENDA(-7) + C(13)*VENDA(-35)

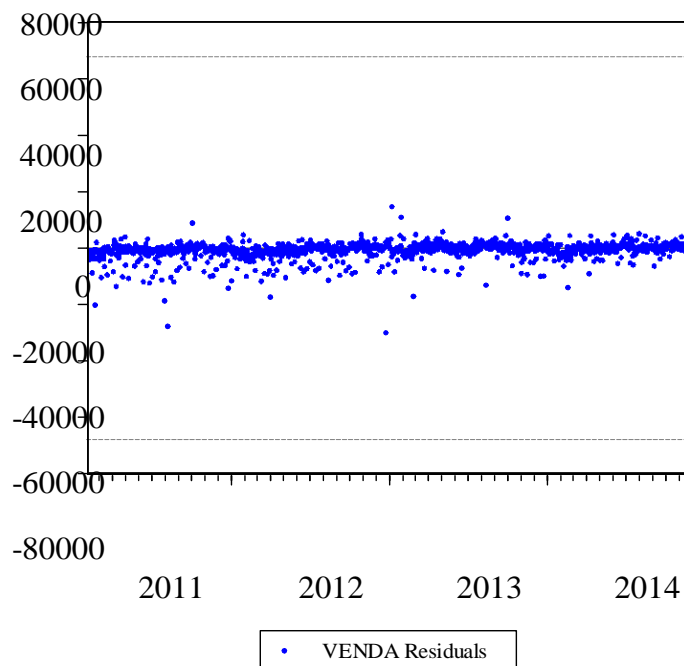
Substituindo os coeficientes: =====

VENDAS = - 5430.131 - 452.1334*D2006 - 11823.83*DDOM + 2583.727*DFEV + 934.9996*DJAN + 499.3431*DJUL -
664.5925*DOUT - 667.6293*DSAB + 0.012095*VENDA(-1) +
0.293899*VENDA(-7) + 0.152716*VENDA(-35)

6.2 Autocorrelação dos Resíduos

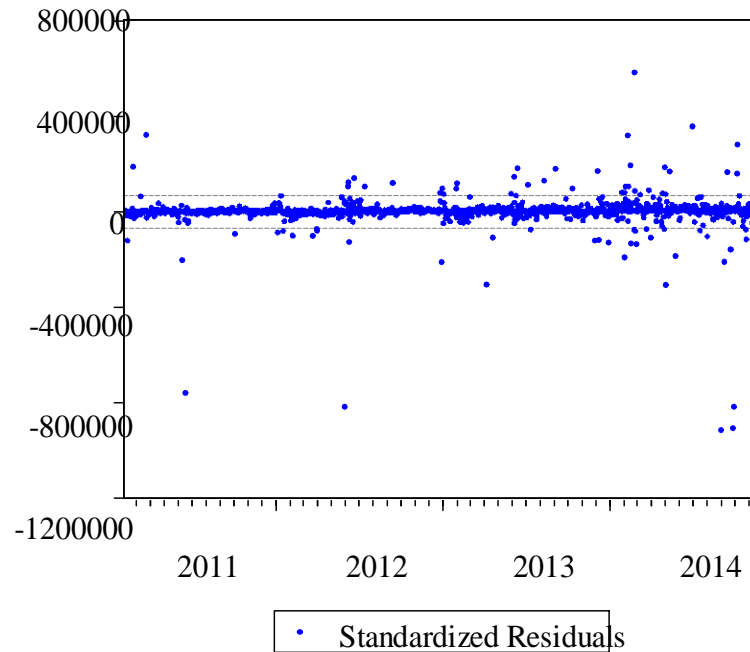
Pelo teste de Durbin-Watson, aceitamos a hipótese de não autocorrelação de primeira ordem.

Gráfico 3 – Dispersão dos resíduos



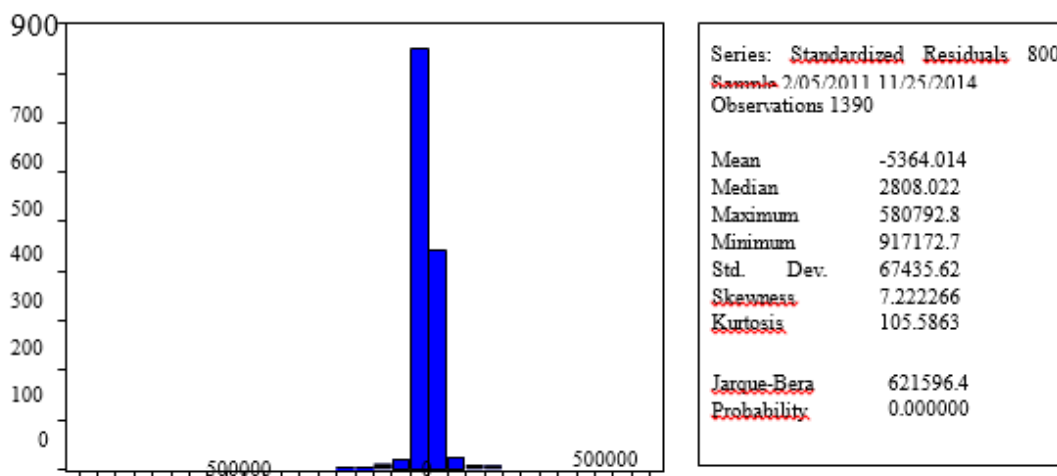
MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

Gráfico 4 – Dispersão dos resíduos padronizados



Através da análise dos resíduos padronizados (corresponde ao resíduo bruto dividido pelo erro padrão dos resíduos) observado no gráfico 4 percebemos que a maioria dos resíduos está disposta dentro do seu intervalo de confiança. Apesar do aparente bom ajuste não se consegue obter a normalidade dos resíduos, como pode ser visto no gráfico 5. A violação dessa hipótese é considerada leve e pode tornar os intervalos de confiança não confiáveis.

Gráfico 5 – Histograma dos resíduos



MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

7. Conclusão

O modelo proposto obteve um resultado satisfatório. O ajustamento de modelo como este contribui para a tomada das decisões gerenciais que desempenham um papel crucial na busca pela competitividade e os ganhos de resultados frente a um ambiente dinâmico. O modelo serve como orientação tanto quanto a direcionamentos futuros quanto a variáveis que influenciam nos seus planejamentos. Uma orientação correta e precoce aumenta a probabilidade de chances de sucesso nos processos de tomada de decisão. Neste estudo, os estimadores dos parâmetros de regressão estimados por Mínimos Quadrados Ponderados se mostraram eficientes ao corrigir o problema da heterocedasticidade e contribuíram para um melhor ajuste do modelo. Para estudos futuros podemos buscar novas variáveis explicativas para o modelo com o objetivo de aperfeiçoar sua capacidade preditiva.

8. Referências

- (1) Azevedo, R. C., Bremer, C.F. (2006) O uso do ERP e CRM no suporte à gestão da demanda em ambientes de produção MAKE-TO-STOCK. Gest. Prod., v. 13, n. 2, p. 179-190.
- (2) Ballou, R. H. (2010) Logística empresarial: transportes, administração de materiais, distribuição física. Rio de Janeiro: Atlas.
- (3) Bowersox, D. J., Closs, D. J. and Stank, T. P. (2000) Ten mega-trends that will revolutionize supply chain logistics, Journal of Business Logistics, 21, 2, 1-15.
- (4) Castiglioni, J.A.M. Logística Operacional: guia prático. São Paulo: Érica
- (5) Cox, J. F.; Blackstone, J. H., Jr. (1998) APICS Dictionary - Ninth Edition. Alexandria, VA: APICS, 1998.
- (6) Destro, I. R. (2011) A variação da Demanda em Cadeia de Suprimentos. Artigo Mundo Logístico. Disponível em: www.otimis.com
- (7) Melech, E. C. (2012) "Morte do jornal de papel e novas tecnologias digitais: desafios e perspectivas" Rev. Comunicação & Mercado. Vol.01, p.56-64
- (8) Melo, D. C., Alcântara, R.L.C. (2012) Proposição de um modelo para a gestão da demanda: um estudo entre os elos atacadista e fornecedores de produtos de mercearia básica". Rev. Gest. Prod., São Carlos, v. 19, n. 4, p. 759-777.
- (9) Proud, J. F. (1991) Master Scheduling. New York: John Wiley & Sons.
- (10) Pires, S. R. I., Bremer, C. F., Azevedo, R. C.; Nasciutti, A. C., Santa Eulália, L. A. (2011) A reference model and a case study in demand management. In: international conference of the production and operations management society.
- (11) Rodrigues, L. H., Antunes, J.A.V, Pellegrin, I., Vaccaro, G. (2001) Proposta para desenvolvimento de um protótipo de um sistema integrado de gestão de demanda. ENEGEP, 2001, Salvador. Anais do XXI ENEGEP.
- (12) Slack, Nigel; Chambers, Stuart; Johnston, Robert. (2007) Administração da produção. 2. ed. São Paulo: Atlas.
- (13) Souza, F. B., Pires, S.R.I. (2014) Produzindo para disponibilidade: uma aplicação a teoria das restrições em ambientes de produção para estoque. Gest. Prod., v. 21, n. 1, p. 65-76.
- (14) Veríssimo, A. J. & Alves, C. C. & da Cruz, A. C. (2012) Métodos Estatísticos De Suavização Exponencial Holt-Winters Para Previsão De Demanda Em Uma Empresa Do Setor Metal Mecânico, Gestão Industrial, v8, n4, p154-71

MODELO DE REGRESSÃO NA PREVISÃO DA DEMANDA

- (15) Willcox, B. (1998) Study notes for master planning of resources: Action MRPII. São Paulo: SAP, 2000. Apostila para treinamento interno. Cox, J. F.; Blackstone, J. H., Jr. APICS Dictionary - Ninth Edition. Alexandria, VA: APICS.