

:Área de Influência: Um Estudo Exploratório de um Modelo Quantitativo para Previsão

Autoria: Juracy Gomes Parente e Heitor Takashi Kato

RESUMO

Neste trabalho procurou-se investigar o fenômeno da Área de Influência, conceito de fundamental importância nos estudos de localização varejistas, com base em pesquisa empírica com clientes de supermercados brasileiros. A distribuição dos clientes em torno das lojas foi retratada por meio de curvas de Área de Influência. Procurou-se identificar diferentes modelos de funções curvilíneas que pudessem representar as curvas de Área de Influência. A curva cúbica mostrou o melhor ajuste entre as curvas com dados reais e as curvas estimadas. Utilizando como variáveis explicativas a área de vendas das lojas e a densidade populacional ao redor de 1 km da loja, conseguiu-se estimar os parâmetros das curvas cúbicas da Área de Influência de cada loja estudada. Verificou-se que os parâmetros das curvas cúbicas podem ser explicadas pela área de vendas e pela densidade populacional. Obteve-se um ajuste satisfatório entre as curvas com dados reais e as curvas estimadas. O pequeno número de dados pesquisados não permite uma generalização dos dados, mas os resultados indicam um caminho promissor para a previsão quantitativa da Área de Influência de lojas.

INTRODUÇÃO

Diferente de outros setores, onde a dimensão geográfica da demanda de mercado é uma variável controlável que pode ser previamente delimitada, os varejistas, entretanto, não conseguem exercer controle para definir os contornos da região geográfica onde está contida sua demanda de mercado. Embora as empresas varejistas possam escolher o local onde localizar uma loja, a região geográfica de seu mercado pode ser mais ou menos extensa, dependendo não só do mix varejista da loja, mas também da concorrência e das características de sua localização. A **Área de Influência** é um conceito fundamental nos estudos de localização pois reflete a dimensão geográfica da demanda de mercado do varejo. A maior parte das vendas de uma loja vem de clientes que moram dentro de uma razoavelmente restrita área geográfica em torno da loja. Ao investigar onde os clientes estão localizados, o conceito de **Área de Influência** permite identificar a demanda de mercado disponível para certa loja. É por esta razão que esse conceito recebe tanta importância nos estudos varejistas, pois sua determinação torna-se um pré-requisito para se desenvolver avaliações de desempenho (ex., fatia de mercado) de lojas atuais e estimativas de vendas para novas lojas.

Segundo a AMA – *American Marketing Association*, a Área de Influência é uma área geográfica contendo os consumidores de uma empresa particular ou grupo de empresas para bens ou serviços específicos¹. O conceito de Área de Influência tem sido pesquisado há bastante tempo, com os estudos de Reilly², Converse e de McGill³, com os trabalhos de Applebaum⁴ e de Huff⁵. A análise da dispersão geográfica dos clientes em torno de uma loja permitiu identificar 3 segmentos de uma Área de Influência⁶: PRIMÁRIA (cerca de 60 a 75% dos clientes da loja), SECUNDÁRIA (15 a 25% dos clientes) e Terciária (parcela restante, cerca de 10% dos clientes que moram mais afastados da loja).

Os Modelos Gravitacionais, desenvolvidos inicialmente por Reilly⁷ em 1929 e aperfeiçoados por Converse⁸ e Huff⁹, estão baseados nas leis de física de gravitação dos corpos celestes. O trabalho de Huff foi importante pois foi o primeiro que sugeriu que as áreas de influência eram complexas, contínuas e probabilísticas¹⁰, definidas através de contornos

que apresentavam uma mesma probabilidade (ex. 90% , 60%, 5%, etc.) de serem atraídos para a loja. Perterson¹¹ introduziu o mapeamento da tendência da superfície, para descrever o fenômeno da distribuição dos clientes em torno das lojas, utilizando modelo estimado por meio da técnica da regressão múltipla cúbica. Boots e South¹² utilizaram os Diagramas de Voronoi de Alta-Ordem Ponderados Multiplicativamente (*High-Order, Multiplicatively Weighted Voronoi Diagrams*) para delimitar Áreas de Influência teóricas para um conjunto similar (incluindo competidores) de instalações no espaço¹³.

Esse trabalho, apoiado nos resultados de vasta pesquisa empírica realizada em 8 supermercados da cidade de São Paulo, procura explorar uma maneira de explicar e prever a distribuição geográfica dos clientes em torno de uma loja. Como a Área de Influência tem um efeito fundamental no desempenho de lojas, construir um modelo que permita uma explicação e uma previsão deste fenômeno seria de grande utilidade para o planejamento da localização de unidades varejistas.

OBJETIVOS

O objetivo desse artigo consiste em buscar melhor entendimento e explicações sobre os padrões e dimensões da Área de Influência de supermercados brasileiros e desenvolver um modelo quantitativo que permita representar melhor os padrões e dimensões destas Áreas de Influência. Esse artigo está ordenado nas seguintes três partes: 1 – Descrição da área de influência; 2 – Explicação da área de influência; 3 – Previsão da área de influência. A metodologia adotada e os resultados obtidos estão detalhados em cada uma dessas partes.

PARTE 1 – DESCRIÇÃO DA ÁREA DE INFLUÊNCIA

METODOLOGIA

A pesquisa empírica foi concentrada na investigação da Área de Influência em 8 supermercados de São Paulo. Adotou-se a clássica técnica do mapeamento da residência dos clientes (*customer spotting*) em torno de cada loja, utilizando-se o software MAP-INFO.

AMOSTRA DE LOJAS

Para diminuir o número de variáveis pesquisadas, foram escolhidas lojas pertencentes a uma única empresa. Assim, conseguiu-se concentrar a análise nos fatores que variam entre as diferentes lojas de uma mesma rede, como o tamanho da loja (refletindo sua variedade e nível de conforto) e características da região onde a loja está localizada, como a densidade populacional. As lojas pesquisadas foram selecionadas, procurando-se representar diferentes características do universo de supermercados de São Paulo.

AMOSTRA DE CONSUMIDORES

Para cada loja foi selecionada, de forma probabilística, uma amostra de 800 consumidores, portadores e usuários ativos do cartão fidelidade da empresa.

ÁREA DE INFLUÊNCIA POR LOJA

O mapeamento dos clientes foi efetuado pela localização do endereço da residência de cada cliente, através da técnica de “geo-referenciamento”. Mapas computadorizados (versão “Map-Info”) da Grande São Paulo foram utilizados para esse processo, permitindo também o cálculo das distâncias entre a loja e a residência de cada cliente pesquisado. Após o cálculo dessas distâncias para todos os elementos da amostra, desenvolveu-se uma matriz, para cada loja, relacionando distâncias e o percentual de clientes contido em cada região..

CURVAS DE ÁREA DE INFLUÊNCIA

Com base na matriz, relacionando distâncias e o percentual de clientes contido em cada região, os autores desenvolveram uma representação gráfica, para proporcionar um sumário mais completo dos fenômenos da distribuição dos clientes em torno de lojas. Conforme ilustrado na Figura 1, o eixo horizontal indica a distância em linha reta entre a loja e a residência do cliente, e o eixo vertical representa a percentagem acumulada de clientes dentro de cada intervalo de distância. Ao capturar a maior parte das informações contidas nos mapas de *customer spotting*, as curvas de área de influência representam um excelente e completo resumo do fenômeno da dispersão geográfica dos clientes, pois a qualquer ponto da curva uma relação específica pode ser estabelecida entre o raio da distância (entre loja e residência de cliente) e a percentagem de clientes contida nesses raios. A natureza cumulativa do eixo vertical ajuda a oferecer respostas a questões práticas perguntadas por varejistas, tais como: qual raio contém uma certa percentagem de clientes (ex.: 60%) da loja? Ou, qual percentagem de clientes vem de uma certa distância (ex.: raio de 2 km)? Essas curvas oferecem uma visão precisa da relação espacial entre a loja e seus clientes, e refletem a capacidade da loja de atrair clientes.

RESULTADOS DA PESQUISA EMPÍRICA

Iniciou-se a análise dos dados empíricos, procurando-se investigar como os consumidores estão distribuídos em torno dos 8 supermercados pesquisados. Mapas de *customer spotting* foram então construídos através das técnicas de geoprocessamento. Buscou-se sumarizar esses dados através de tabelas, e posteriormente por meio das curvas de Área de Influência..

Tabela 1 - Porcentagem acumulada de clientes em raios concêntricos

Segmentos da área de Influência	% do número de clientes							
	Loja 4	Loja 2	Loja 3	Loja 5	Loja 6	Loja 7	Loja 8	Loja 1
Até 0,5 KM	33	5	8	6	62	45	8	15
Até 1,0 KM	59	19	24	17	85	61	23	29
Até 1,5 KM	66	43	37	34	89	64	35	45
Até 2,0 KM	69	61	49	50	90	67	47	55
Até 2,5 KM	72	68	60	61	91	69	57	65
Até 3,0 KM	73	79	69	65	92	72	66	71
Até 3,5 KM	75	80	77	68	92	75	74	76
Alem de 3,5 km	25	20	23	32	8	25	26	24
TOTAL	100	100	100	100	100	100	100	100

TABELA INDICANDO VALORES ACUMULADOS POR DISTÂNCIA

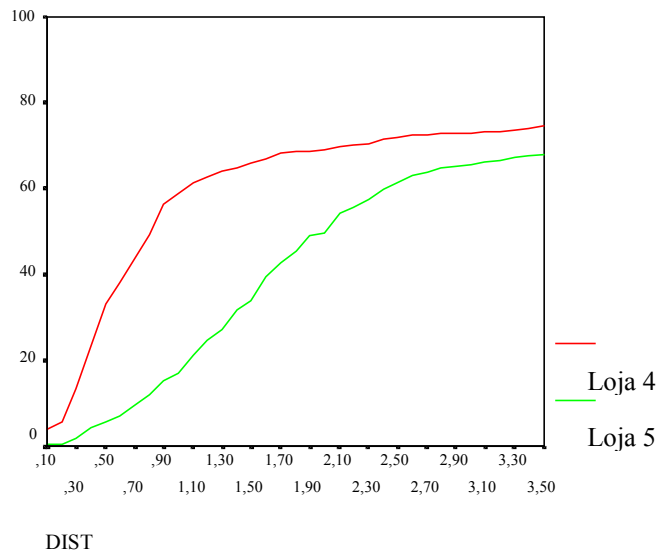
A tabela 1 mostra a percentagem acumulada de clientes dentro de diferentes raios traçados em torno da loja. Esse formato cumulativo permite uma melhor comparação entre as diferentes lojas, pois proporciona uma informação agregada, refletindo o padrão geral da distribuição geográfica em torno da loja. Para o estudo quantitativo em si, foi usado outro grau de agregação, com um espaçamento de 100 m para os segmentos da Área de Influência.

CURVAS DE ÁREA DE INFLUÊNCIA

A figura 1 ilustra as curvas de Área de Influência da Loja 4 e da Loja 5. Pode-se observar que a curva da Área de Influência da Loja 4 apresenta uma inclinação vertical, quase retilínea até 60%, atingido perto do 1 km de distância da loja. A partir daí, a taxa de inclinação cai bastante. É uma loja pequena, localizada perto de um centro de comércio, bem

servido de transporte público, em uma confluência de avenidas, em região de alta densidade populacional. A curva da loja 5 apresenta inclinação mais horizontal, indicando uma Área de Influência muito extensa. Trata-se de uma loja grande localizada ao longo de importante via de acesso. Algumas das regiões além dos 3 km apresentam carência de bons supermercados, justificando assim a alta percentagem de clientes vindo de regiões mais distantes (35% além dos 3 km).

Figura 1 - Curvas de Área de Influência



PARTE 2 – EXPLICAÇÃO DA ÁREA DE INFLUÊNCIA.

Analisando-se os resultados de Área de Influência na Tabela 2 e as curvas da loja 4 e da loja 5, verifica-se que existe enorme variação no formato das curvas e no padrão de dispersão geográfica dos clientes de diferentes lojas. Essa grande variação remete-nos à necessidade de investigarmos mais profundamente dois aspectos para melhor compreensão do fenômeno da Área de Influência:

- Quais fatores explicam as diferentes dimensões de Área de Influência das diversas lojas?
- Quais tipos de funções curvilíneas melhor refletem o comportamento das curvas de Área de Influência?

METODOLOGIA

VARIÁVEIS EXPLICATIVAS DA ÁREA DE INFLUÊNCIA

Os vários autores que estudaram e desenvolveram modelos sobre Área de Influência identificaram muitos fatores que influenciam o padrão e a dimensão da Área de Influência. Em se tratando de um estudo exploratório, resolveu-se utilizar duas variáveis, reconhecidas como sendo fatores fundamentais na determinação da capacidade de diferentes lojas atraírem clientes.

TAMANHO DA LOJA – A maior parte dos modelos¹⁴ comentados anteriormente utilizam o tamanho da loja como variável *proxi* para refletir a atratividade da unidade varejista. Efetivamente lojas maiores apresentam em geral uma maior variedade, mais conforto e serviços, e preços muito competitivos, conseguindo, assim, uma maior capacidade de atrair clientes. A extensão da Área de Influência é afetada pelo próprio tamanho da loja, mas o tamanho desta Área de Influência não cresce proporcionalmente com o aumento do tamanho da loja. Ou seja, trata-se de uma relação não linear¹⁵.

Nesse estudo, o tamanho da loja foi medido através de sua área de vendas (metros²). A informação sobre o tamanho da loja foi obtida junto a rede de supermercados que participou do projeto, e confirmada nas lojas, através de pesquisa de observação realizada pelos pesquisadores desse trabalho.

DENSIDADE POPULACIONAL - Uma maior densidade populacional vem, em geral, acompanhada de maior número de concorrentes e de maior lentidão no tráfego, fatores que restringem a extensão da Área de Influência.. Padrões de distribuição das moradias também afetam a Área de Influência de uma loja. Em áreas urbanas centrais, as pessoas vivem agrupadas em apartamentos perto de centros de comércio, e as Áreas de Influência tendem a ser menores. Já em áreas de menor densidade, as pessoas tendem a viver em unidades individuais, casas, e a Área de Influência tende a ser maior.

Nesse estudo, a densidade populacional foi determinada no raio de 1 km em torno da loja, sendo estimada com dados extraídos do último censo (1991) do IBGE, e reflete o número de habitantes por hectare, localizado dentro do raio de 1 km em torno da loja.

FUNÇÕES CURVILÍNEAS

Pela análise das curvas pode-se verificar que existem diferentes tipos de funções curvilíneas que parecem refletir o comportamento da Área de Influência. Existem procedimentos tradicionais na área da estatística utilizados para o estudo da adaptação de distribuições teóricas para modelagem de dados empíricos, como o de Pearson¹⁶, no qual o estudo das distribuições é baseado na análise das estatísticas das distribuições dos dados observados¹⁷. Para uma primeira aproximação no estudo da distribuição das lojas, será feito uma análise comparativa com 3 tipos de curvas: a Logística, a Logarítmica Recíproca (Curva S) e a Curva Cúbica.

Curva Logística –

É uma curva que pertence a família de curvas em S. Como as curvas de Área de Influência, a curva Logística apresenta uma natureza assintótica tanto em seu limite superior como no inferior em direção a sua origem. A maior parte das curvas analisadas teve um

comportamento que se parece com uma curva logística. O modelo estatístico da curva logística é definido por¹⁸:

$$Y_t = \frac{1}{k + ab^t} \text{ com } b > 0$$

Para a estimação do modelo as variáveis representadas foram as seguintes:

Y=proporção acumulada de clientes

k= inverso do maior limite superior, igualado a 100

T=comprimento do raio a partir da loja

a e b=parâmetros a serem estimados pela estimação

Curva S (logarítmica recíproca) –

Trata-se de "uma função com formato de S muito simples, que pode ser usada para modelar o padrão de saturação de vendas..."¹⁹, mas que pelo seu formato poderia também ser adequada para modelar os dados de algumas lojas. A Curva S (logarítmica recíproca) é definida da seguinte forma²⁰:

$$Y = e^{(b_0 + (\frac{b_1}{t}))}$$

Para a estimação do modelo as variáveis representadas foram as seguintes:

Y=proporção acumulada de clientes

e= 2,718282 (base do logaritmo natural)

T=comprimento do raio a partir da loja

b₀ e b₁=parâmetros a serem estimados pela estimação

Este modelo não seria o mais apropriado para explicar o fenômeno concentração dos clientes ao redor de uma loja por não ser uma curva de densidade de probabilidade (com máximo igual a 1), pois a variável Y é uma variável contínua qualquer, sem limitação do valor mais alto. Entretanto, como trata-se de um estudo exploratório, utilizado para verificar os resultados proporcionador por um estudo com este tipo de metodologia, estaremos investigando também este modelo devido à facilidade de sua aplicação para uma primeira análise do fenômeno observado nos dados.

Curva Cúbica –

De acordo com Peterson, outro modelo que poderia modelar este tipo de fenômeno seria a curva cúbica²¹. Segundo Peterson, o componente linear representaria uma tendência geral, o componente quadrático representaria a tendência regional e o componente cúbico a tendência local. O modelo de regressão cúbica seria definido da seguinte forma:

$$Y = b_0 + b_1t + b_2t^2 + b_3t^3$$

Para a estimação do modelo as variáveis representadas foram as seguintes:

Y=proporção acumulada de clientes

T=comprimento do raio a partir da loja

b₀, b₁ b₂ e b₃=parâmetros a serem estimados pela estimação

RESULTADOS

FATORES QUE INFLUENCIAM A ÁREA DE INFLUÊNCIA

A Tabela 2 apresenta informações sobre Área de Influência, tamanho da loja e densidade populacional, para os oito supermercados pesquisados. Apesar da amostra de lojas ser muito pequena (apenas oito unidades) para possibilitar o desenvolvimento de relações estatísticas de causa e efeito, a análise dos dados sugere que:

- Área de influência está positivamente relacionada com o tamanho da loja
- Área de influência está negativamente relacionada com a densidade populacional

Tabela 2 - Características das lojas estudadas

	Loja 4	Loja 2	Loja 3	Loja 5	Loja 6	Loja 7	Loja 8	Loja 1
Área de Vendas	780	1.100	3.400	4.800	450	440	2.030	3.500
Densidade Populacional Raio de 1 km (hab./ha)	195	50	80	108	323	218	43	111
Amostra de consumidores	800	800	800	800	800	800	800	800

Tanto estudos anteriores conduzidos pelos autores, como a lógica dedutiva parecem confirmar essas hipóteses. A maior parte dos modelos²² comentados anteriormente utilizam o tamanho da loja como variável *proxi* para refletir a atratividade da unidade varejista. A extensão da Área de Influência é afetada pelo tamanho da loja, mas a área de influência não cresce proporcionalmente com o aumento do tamanho da loja, ou seja a relação é não-linear (Bell, Ho e Tang, 1998, p.355). Tem-se também observado que uma maior densidade populacional vem, em geral, acompanhada de maior oferta de supermercados concorrentes, e de maior lentidão no tráfego, fatores que restringem a extensão da Área de Influência de cada loja.

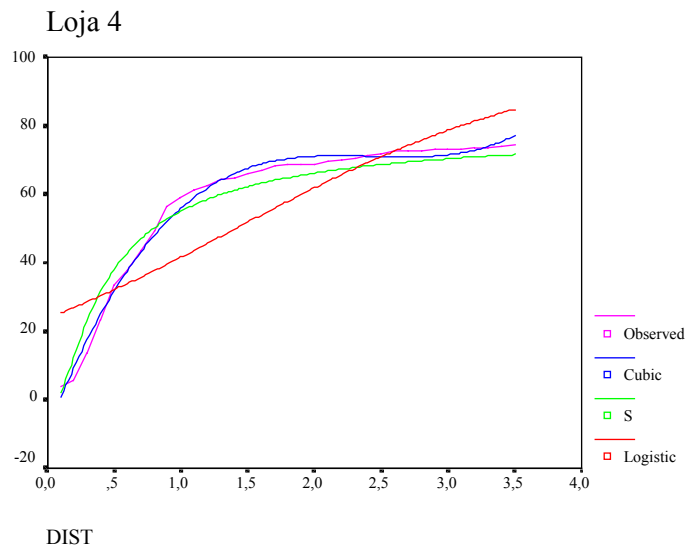
Devido à natureza exploratória deste trabalho, estas constatações não podem ser generalizadas, mas sugerem hipóteses para investigações posteriores.

IDENTIFICAÇÃO DAS MELHORES FUNÇÕES CURVILÍNEAS

Para identificar as funções curvilíneas mais adequadas, procedeu-se tanto uma análise visual das curvas e como também uma análise estatística do grau de ajuste entre os modelos e os dados reais. O processo de análise visual comparou o ajuste entre a curva dos dados reais e as curvas geradas pelos três modelos de curvas, para cada uma das 8 lojas pesquisadas. Para ilustrar o processo de análise visual, as figuras 2 e 3 mostram os resultados da Loja 4 e da Loja 5.

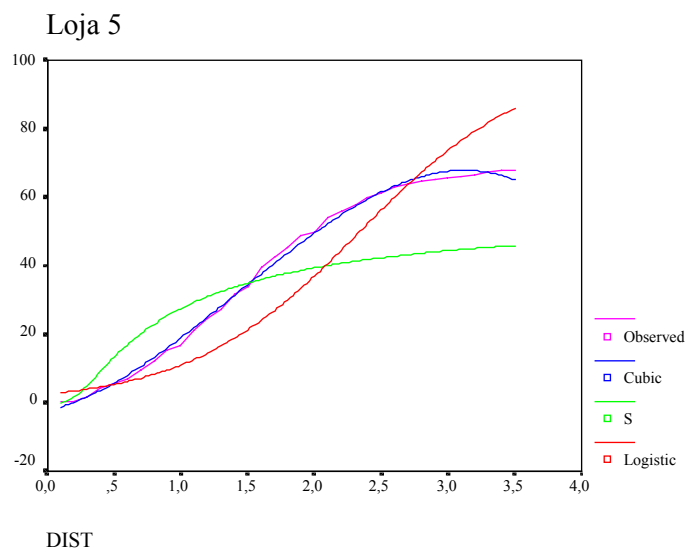
A curva S mostra um ajuste melhor para as lojas onde há um crescimento da proporção dos clientes de forma abrupta, e depois uma parcela muito pequena se estendendo por uma distância maior, como exemplificado na Figura 2.

Figura 2 - Estimação das curvas para loja 4



A curva cúbica também mostra um bom ajuste para estas lojas. Já para as lojas onde há um crescimento mais lento em relação à distância, como exemplificado pela Figura 3, a curva cúbica mostra-se a de melhor ajuste.

Figura 3 Estimação das curvas para loja 5



A tabela 3 resume a análise feita com os modelos estimados e uma avaliação subjetiva do melhor modelo para cada loja, e do segundo e do terceiro melhor modelos no ajuste. Podemos ver por esta tabela que a curva cúbica foi o melhor modelo para um maior número de lojas, seguida da curva S. Embora a curva S mostre melhor ajuste para 3 lojas, este não é um número adequado para fazer um estudo de estimação.

Entretanto, todas as 3 lojas onde a curva S foi a melhor tiveram o modelo da curva cúbica como segundo melhor modelo. Desta forma, foi escolhido o modelo cúbico para se fazer um estudo preliminar sobre a possibilidade de se proceder a um modelo preditivo para as curvas de Área de Influência das lojas.

Tabela 3 - Avaliação do ajuste das curvas aos dados originais das lojas

Tipo de curva	Loja 4	Loja 2	Loja 3	Loja 5	Loja 6	Loja 7	Loja 8	Loja 1
Curva S	Melhor	3°	3°	3°	Melhor	Melhor	3°	3°
Curva cúbica	2°	Melhor	Melhor	Melhor	2°	2°	Melhor	Melhor
Logística	3°	2°	2°	2°	3°	3°	2°	2°

PARTE 3 – PREVISÃO DA ÁREA DE INFLUÊNCIA

METODOLOGIA

O primeiro passo no estudo de um modelo preditivo foi fazer a análise dos dados de maneira a encontrar a melhor função que representasse uma distribuição mais aproximada com os dados originais das lojas. Verificou-se que o modelo da curvas cúbica foi o que proporcionou o melhor ajuste.

Uma vez escolhido este modelo, passou-se para a fase da estimação dos parâmetros do modelo para cada loja. Uma vez estimado esse conjunto de parâmetros, o terceiro passo consistiu em utilizar as variáveis explicativas, densidade populacional e área de vendas, para tentar estimar, por meio de uma regressão linear múltipla, os parâmetros do modelo preditivo (via curva cúbica) de área de influência. O último passo foi utilizar o modelo de curva cúbica para estimar a curva de área de influência de cada loja, e, então, analisar o ajuste existente entre os dados reais e as estimativas proporcionadas pelo modelo.

ESTIMAÇÃO DOS PARÂMETROS

Após realizar um teste preliminar, verificou-se que o coeficiente b_0 seria dispensável, já que os modelos cúbicos para as lojas foram bem ajustados sem a constante. Desta forma, preferiu-se usar estes modelos sem o coeficiente b_0 , o que oferece a vantagem de ter todos os parâmetros diretamente dependentes da distância, isto é o raio dos diferentes segmentos da Área de Influência estudada. O resumo dos resultados da estimação das curvas cúbicas está expresso na tabela 4, onde vê-se que a maioria dos parâmetros mostrou-se significativo, a 95% de confiança..

Tabela 4 - Resultado da estimação dos parâmetros do modelo da Curva Cúbica

	Loja 4	Loja 2	Loja 3	Loja 5	Loja 6	Loja 7	Loja 8	Loja 1
Adjusted R Square	,997	,996	,999	,999	,997	,996	,999	,999
B1	82,242	4,518	20,631	4,368	142,192	96,379	16,929	30,207
Desvio padr	2,426	2,591	1,284	1,215	3,686	3,308	1,054	1,393
T	33,897	1,743	16,067	3,596	38,579	29,130	16,067	21,681
Significância	,0000	,0909	,0000	,0011	,0000	,0000	,0000	,0000
B2	-31,492	21,390	7,354	18,380	-68,402	-44,318	6,034	,783*
Desvio padr	2,118	2,263	1,121	1,060	3,218	2,889	,919	1,216
T	-14,866	9,452	6,559	17,332	-21,255	-15,342	6,559	,644
Significância	,0000	,0000	,0000	,0000	,0000	,0000	,0000	,5242
B3	4,038	-4,733	-1,726	-4,103	10,256	6,638	-1,417	-,960
Desvio padr	,4414	,471	,2336	,2209	,670	,601	,1917	,253
T	9,149	-10,038	-7,390	-18,567	15,294	11,028	-7,390	-3,789
Significância	,0000	,0000	,0000	,0000	,0000	,0000	,0000	,0006

* substituído por zero na regressão linear

ESTIMAÇÃO DOS PARÂMETROS DO MODELO PREDITIVO

Uma vez definidos um conjunto de coeficientes b_0 , b_1 , b_2 e b_3 para as curvas cúbicas das lojas estudadas, precisou-se verificar a existência de um modelo capaz de estimar estes mesmos parâmetros a partir de variáveis independentes explicativas (área de vendas e densidade populacional) observadas para cada loja.

Para esta estimação utilizou-se o modelo de regressão linear, onde a formulação geral toma a seguinte forma²³:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + \varepsilon_i, \text{ onde}$$

Y = variável dependente

β_1 = constante

β_k = coeficientes das variáveis independentes

X_k = variáveis independentes

ε = termo de erro

As variáveis que usadas nesta estimação foram as variáveis Densidade Populacional e Área de Vendas, discriminadas na tabela 1.

Tabela 5 - Resumo da estimação dos parâmetros das Curvas Cúbicas

	b_1	b_2	b_3
Adjusted R Square	,953	,894	,853
LDens	106,071	-62,944	10,027
Desvio padrão	10,959	7,824	1,461
T	9,679	-8,045	6,864
Significância	,000	,000	,000
LArea	-53,305	37,335	-6,205
Desvio padrão	10,959	5,102	0,953
T	-7,459	7,318	-6,513
Significância	,000	,000	0,001

ESTIMAÇÃO DO MODELO DE CURVA CÚBICA

Após realizar um teste adotando diferentes formulações, verificou-se que o coeficiente b_0 seria dispensável, já que os modelos lineares sem a constante foram bastante adequados.

Apresentando um satisfatório poder de explicação (R quadrado ajustado de 0,953), o modelo que melhor explica a variável b_1 inclui as variáveis logaritmo de Área de Vendas (Larea) e logaritmo da Densidade Populacional (LDens). A análise dos coeficientes da regressão (Tabela 5) mostra que o coeficiente b_1 sofre uma influência negativa do logaritmo da Área de Vendas e positiva do logaritmo da Densidade Populacional. Ambas variáveis são significativas a um grau de 95% de confiança. Pode-se observar, que o modelo para o coeficiente b_2 tem um bom poder de explicação (R quadrado ajustado de 0,894). Embora tenha um bom ajuste, o modelo tem um poder de explicação menor do que o modelo para o coeficiente b_1 e, ao contrário de b_1 , sofre uma influência negativa do logaritmo da Densidade Populacional e positiva do logaritmo da Área de Vendas. Já o modelo que melhor explica a variável b_3 revela também um bom ajuste (R quadrado ajustado de 0,853) sendo também

influenciado positivamente pela variável Logaritmo da Densidade Populacional(LDens) e negativamente pela variável logaritmo da Área de Vendas (Larea).

ESTIMAÇÃO DA CURVA DE ÁREA DE INFLUÊNCIA DE CADA LOJA

O próximo passo no nosso estudo foi estimar os coeficientes da Curva Cúbica para as oito lojas, a partir dos dados sobre Área de Vendas e Densidade Populacional. A partir destes dados, estimou-se os coeficientes b_1 , b_2 e b_3 com os parâmetros obtidos nos estudos de regressão feitos na etapa anterior (Tabela 5). Os dados podem ser vistos na Tabela 6.

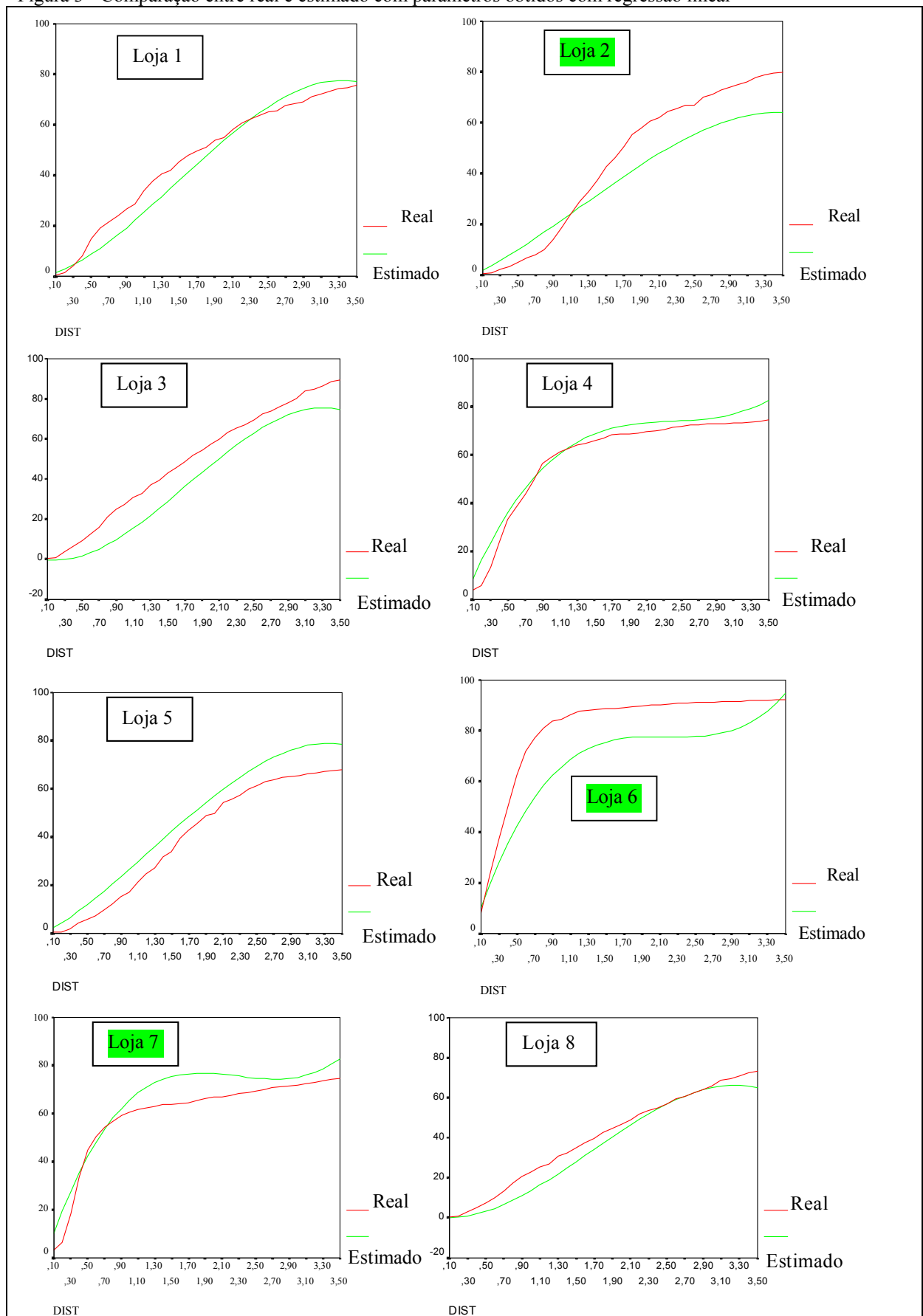
Tabela 6 - Estimação dos parâmetros b_1 , b_2 e b_3 para Curvas Cúbicas das lojas

	Loja 4	Loja 2	Loja 3	Loja 5	Loja 6	Loja 7	Loja 8	Loja 1
Área de Vendas	780	1.100	3.400	4.800	450	440	2030	3500
Densidade Populacional Raio de 1 km (hab./ha)	195	50	80	108	323	218	43	111
Larea	2,892	3,041	3,5319	3,681	2,653	2,643	3,307	3,872
LDens	2,290	1,698	1,903	2,033	2,509	2,338	1,633	2,045
Coef Larea p/ b_1	-53,305	-53,305	-53,305	-53,305	-53,305	-53,305	-53,305	-53,305
Coef LDens p/ b_1	106,071	106,071	106,071	106,071	106,071	106,071	106,071	106,071
Coef Larea p/ b_2	37,335	37,335	37,335	37,335	37,335	37,335	37,335	37,335
Coef LDens p/ b_2	-62,944	-62,944	-62,944	-62,944	-62,944	-62,944	-62,944	-62,944
Coef Larea p/ b_3	-6,205	-6,205	-6,205	-6,205	-6,205	-6,205	-6,205	-6,205
Coef LDens p/ b_3	10,027	10,027	10,027	10,027	10,027	10,027	10,027	10,027
Estimativa b_1	88,745	18,008	13,585	19,427	124,714	107,109	-3,066	10,518
Estimativa b_2	-36,169	6,657	12,081	9,465	-58,877	-48,487	20,679	15,841
Estimativa b_3	5,017	-1,844	-2,834	-2,456	8,696	7,043	-4,146	-3,521

A partir dos dados da Tabela 6, utilizando as estimativas para b_1 , b_2 e b_3 e estimou-se a curva cúbica para cada loja. A Figura 3 apresenta o confronto entre os dados reais, obtidos do mapeamento dos clientes das lojas (Tabela 2) e dos obtido a partir das curvas cúbicas estimadas. Como se trata de um estudo exploratório preliminar, desenvolvido com um número limitado de lojas (oito) e portanto deixando poucos graus de liberdade para a estimação da regressão cúbica, os resultados devem ser analisados com cautela.

Apesar da natureza exploratória que caracteriza esse trabalho, e do limitado número de variáveis independentes consideradas, os resultados são extremamente promissores. Na maioria das lojas, a curva de Área de Influência estimada apresentou resultados extremamente próximos da realidade das lojas pesquisadas. A discrepância maior (entre estimado e real) ocorreu nas Lojas 2, 6 e 7. Verificou-se que a densidade populacional na loja 2 apresenta um padrão atípico – sua densidade populacional aumenta consideravelmente, além do raio de 1 km. Caso tivéssemos utilizado, na estimação, a densidade que caracteriza a área de influência além de 1 km, a previsão da curva de sua Área de Influência seria menor, e mais próxima da realidade. Para a Loja 6, observou-se que é uma loja sem estacionamento, com pouca visibilidade, e localizada em região de difícil acesso. A Loja 7 apresenta pequeno estacionamento e embora tenha fácil acesso, fica localizada numa região mais comercial, com uma ampla rede de transporte público, que faz com que uma proporção maior de clientes venha de regiões mais distantes, mesmo a loja sendo pequena.

Figura 3 - Comparação entre real e estimado com parâmetros obtidos com regressão linear



CONCLUSÕES

Embora este seja um estudo exploratório, pudemos verificar que existe uma boa possibilidade de usar modelos estatísticos para tentar ajustar os dados empíricos obtidos da distribuição dos clientes de uma loja ao redor da mesma, na tentativa de prever a Área de Influência destas lojas. Sendo um primeiro estudo neste sentido, algumas restrições foram relaxadas, de maneira a poder explorar mais livremente a linha de raciocínio utilizada. Pudemos ver que as curvas acumuladas de clientes mostram-se propícias a serem ajustadas a dois tipos de modelos testados, a curva logarítmica recíproca (curva S) e a curva cúbica, onde a distância foi a variável explicativa. A partir dos parâmetros obtidos, buscou-se modelar estes conjuntos de parâmetros através de regressões utilizando como variáveis explicativas a área de vendas das lojas e a densidade populacional ao redor de 1 km da loja. Verificou-se que os parâmetros das curvas cúbicas podem ser explicadas por estas variáveis, permitindo uma projeção das mesmas.

Através das análises comparativas dos gráficos dos dados reais e dos dados obtidos com as regressões cúbicas utilizando parâmetros estimados com dados de área de vendas e densidade populacional das lojas, pode-se verificar que há um ajuste muito promissor entre as curvas. Devido à natureza exploratória deste trabalho, estas constatações não podem ser generalizadas. Sugerimos entretanto aprofundar essa investigação, utilizando-se uma maior amostra de lojas e um tratamento estatístico mais sistemático, para possibilitar resultados com maior validade conceitual e capacidade preditiva.

BIBLIOGRAFIA

- APPLEBAUM, W.. “Methods for determining store trade areas, market penetration and potential sales”. *Journal of Marketing Research*, Vol. III, May, 1966, pág. 127-141.
- ARANHA, Francisco. “Atlas dos Setores Postais: Uma Nova Geografia a Serviço da Empresa”. *RAE – Revista de Administração de Empresas*. V.37, n.3, Jul./Set. 1997, pg. 20-27.
- BELL, David R; HO, Teck-Hua; TANG, Christopher S. “Determining where to shop: Fixed and variable costs of shopping”, *Journal of Marketing Research*, Aug 1998 p: 352-369.
- BERMAN, Barry; EVANS, Joel R.. *Retail Management: A strategic Approach*. Upper Saddle River, Prentice Hall, 1998.
- BLACK, W.. “Choice-Set Definition in Patronage Modeling”. *Journal of Retailing*, vol 60, pg. 63-85.
- BOOTS, Barry; SOUTH, Robert. “Modeling Retail Trade Areas Using High-Order, Multiplicatively Weighted Voronoi Diagrams”. *Journal of Retailing*, Vol 73, no.4,1997, pg. 519-536.
- CRAIG, C. Samuel; GHOSH, Avijit;McLAFFERTY, Sara. “Models of the Retail Location Process: A Review”. *Journal of Retailing*, Vol 60, no.1, Spring 1984, pg. 5-36.
- GHOSH, Avijit; CRAIG, C.Samuel. “FRANSYS: A Franchise Distribution System Location Model”. *Journal of Retailing*, Vol 67, no. 4, Winter 1991, pg. 466-495.
- HUFF, David L.. “Defining and Estimating a Trade Area”. *Journal of Marketing*, July 1964,pág.37.

JOHNSTON, J.. *Econometric Methods*. 3rd ed.. New York, McGraw-Hill Book Company, 1984.

KENDALL, Maurice G.; STUART, Alan. *The Advanced Theory of Statistics*. Vol.1.2nd ed.. London, Charles Griffin & Company Limited, 1963

KOTLER, Philip. *Administração de Marketing*. 10a. ed.. São Paulo, Prentice-Hall, 2000.

McGILL, Kenneth H.. "A Method for Delineating Retail Trade Area". *Journal of Retailing*, April 1933, pg. 10-14.

PETERSON, Robert A.. "Trade Area Analysis Using Trend Surface Mapping". *Journal of Marketing Research*. Vol. XI, August 1974, pg. 338-342.

PINDYCK, Robert S.; RUBINFELD, Daniel L.. *Econometric Models & Economic Forecasts*. 3rd Ed.. New York, McGraw-Hill, Inc., 1991.

TAYMAN, Jeff; POL, Louis. "Retail Site Selection and Geographic Informations Systems". *Journal of Applied Business Research*, Vol 11, no.2, pg.46-54.

NOTAS:

¹ BENNETT, Peter D.. (org.). *Dictionary of Marketing Terms*, 2nd ed.. Chicago, American Marketing Association, 1995, p. 287. In: BERMAN, Barry; EVANS, Joel R.. *Retail Management: A Strategic Approach*. Upper Saddle River, Prentice Hall, 1998, p. 276.

² REILLY, William J., *Method for the Study of Retail Relationships*, Reserch Monograph n° 4 – Austin: University of Texas Press, 1929. University of Texas Bulletin n°2944)

³ McGILL, Kenneth H.. "A Method for Delineating Retail Trade Area". *Journal of Retailing*, April 1933, p. 10-14.

⁴ APPLEBAUM, W.. "Methods for determining store trade areas, market penetration and potential sales". *Journal of Marketing Research*, Vol. III, May, 1966, p. 127-141.

⁵ HUFF, David L.. "Defining and Estimating a Trade Area". *Journal of Marketing*, July 1964, p.37.

⁶ APPLEBAUM, W. "Methods for determining store trade areas, market penetration and potential sales". *Journal of Marketing Research*, Vol. III, May, 1966, p. 127-141.

⁷ REILLY, William J., *Method for the Study of Retail Relationships*, Reserch Monograph n° 4 – Austin: University of Texas Press, 1929. University of Texas Bulletin n°2944)

⁸ CONVERSE, Paul D. *New Laws of Retail Gravitation*. *Journal of Marketing*, January 1949: p. 379-384.

⁹ HUFF, David L. Op. Cit.

¹⁰ CRAIG, C. Samuel; GHOSH, Avijit;McLAFFERTY, Sara. "Models of the Retail Location Process: A Review". *Journal of Retailing*, Vol 60, no.1, Spring 1984, p. 15.

¹¹ PETERSON, Robert A.. "Trade Area Analysis Using Trend Surface Mapping". *Journal of Marketing Research*. Vol. XI, August 1974, p. 338-342.

¹² BOOTS, Barry; SOUTH, Robert. "Modeling Retail Trade Areas Using High-Order, Multiplicatively Weighted Voronoi Diagrams". *Journal of Retailing*, Vol 73, no.4,1997, p. 519-536.

¹³ BOOTS, Barry; SOUTH, Robert. "Modeling Retail Trade Areas Using High-Order, Multiplicatively Weighted Voronoi Diagrams". *Journal of Retailing*, Vol 73, no.4,1997, p. 519.

¹⁴ Modelos gravitacionais de Reilly, Converse, e o de Voronoi.

-
- ¹⁵ De acordo com Bell et al, o modelo de Huff mediria o custo fixo de se fazer as compras, ou seja, mediria o custo do deslocamento ao local de compra, que seria fixo qualquer que fosse a compra. Eles chamam os custos variáveis aqueles relacionados às quantidades de produtos compradas em cada compra e às expectativas de seus preços respectivos. BELL, David R; HO, Teck-Hua; TANG, Christopher S. “Determining where to shop: Fixed and variable costs of shopping”, *Journal of Marketing Research*, Aug 1998 p: 355.
- ¹⁶ KENDALL, Maurice G.; STUART, Alan. *The Advanced Theory of Statistics*. Vol.1.2md. ed.. London, Charles Griffin & Comapny Limited, 1963, p. 148.
- ¹⁷ STATSOFT, INC.. Process Analysis. [online]
<http://www.statsoft.com/textbook/stprocan.html>. Capturado em 26/4/2001.
- ¹⁸ PINDYCK, Robert S.; RUBINFELD, Daniel L.. *Econometric Models & Economic Forecasts*. 3rd. Ed.. New York, McGraw-Hill, Inc., 1991, p. 421.
- ¹⁹ PINDYCK, Robert S.; RUBINFELD, Daniel L.. *Econometric Models & Economic Forecasts*. 3rd. Ed.. New York, McGraw-Hill, Inc., 1991, p. 422. Há também uma curva descrita em Johnston chamada de modelo recíproco logaritmico (logarithmic reciprocal model) que tem definição semelhante. Ela seria definida da seguinte forma:
- $$Y = e^{\alpha - \frac{\beta}{x}};$$
- onde se muda apenas a notação e um dos sinais dos parâmetros, subtração no lugar de adição, que na prática não representaria mudanças no resultado da estimação. JOHNSTON, J.. *Econometric Methods*. 3rd ed.. New York, McGraw-Hill Book Company, 1984.
- ²⁰ Esta notação é usada pelo SPSS para a estimação da curva denominada de "S Shape" nos modelos de estimação para regressão. Na versão de Pindyck e Rubinfeld os parâmetros b_0 e b_1 são substituídos por k_1 e k_2 . PINDYCK, Robert S.; RUBINFELD, Daniel L.. *Econometric Models & Economic Forecasts*. 3rd. Ed.. New York, McGraw-Hill, Inc., 1991, p. 422.
- ²¹ PETERSON, Robert A.. “Trade Area Analysis Using Trend Surface Mapping”. *Journal of Marketing Research*. Vol. XI, August 1974, p. 338-342.
- ²² Modelos gravitacionais de Reilly, Converse, e o de Voronoi.
- ²³ PINDYCK, Robert S.; RUBINFELD, Daniel L.. *Econometric Models & Economic Forecasts*. 3rd. Ed.. New York, McGraw-Hill, Inc., 1991, p. 73.