O modelo Logito Multinomial
Aplicado à Análise do Comportamento
de Compra do Consumidor

Andréa Lemos Nery

Dissertação Apresentada

ao

Instituto de Matemática e Estatística

da

Universidade de São Paulo

para a Obtenção do Grau de

Mestre em Estatística

Área de Concentração : **Estatística**

Orientador : **Adolpho Walter Pimazoni Canton**

- São Paulo, Fevereiro de 1994 -

AGRADECIMENTOS

São muitos os agradecimentos à várias pessoas que estiveram comigo durante todo o tempo em que levei para concluir mais uma fase em uma grande caminhada que tracei para minha vida.

Um agradecimento especial gostaria de dedicar aos professores Clóvis de Araújo Peres, Wilton de Oliveira Bussab, Lisbeth K. Cordani , Abraham Laredo Secsu e Vanderlei da Costa Bueno por uma imensa colaboração e estímulo que deram a minha formação estatística , à eles devo um pouco do gosto que possuo pela profissão à que me dedico.

Ao professor Adolpho W. P. Canton agradeço por todos os minutos dispendidos para que chegassemos a um final feliz.

Ao professor Celso C. de H. e Grisi por sua imensa paciência ao ler e colaborar com sugestões em meio a uma série de outras obrigações , e por ter estimulado a minha iniciação na área de marketing.

Aos meus amigos de trabalho uma atenção especial pelo empréstimo dos ouvidos nos meus momentos de cansaço. Aos verdadeiros amigos que fiz no Instituto de Matemática da USP . E a muitos outros que souberam sempre ser Grandes Amigos.

Aos **meus pais** , que em nenhum momento deixaram de acreditar em mim dando todo o apoio que sempre precisei.

Ao **Luciano** por tantas coisas que seriam impossíveis de serem descritas, e principalmente por todo amor que tem me dado durante todos esses anos.

Aos meus avôs

ÍNDICE

Capítulo 1 - Introdução

1.1 - Introdução	1
------------------	---

Capítulo 2 - O Modelo Logito Multinomial

2.1 - Introdução	16
2.2 - O desenvolvimento do modelo logito multinomial	16
2.3 - A função de utilidade linear	23
2.4 - Tratamento da heterogeneidade	26
2.5 - A função de verossimilhança	34
2.6 - Conclusões	41

Capítulo 3 - Algumas Aplicações

3.1 - Introdução	42
3.2 - Aplicação 1	42
3.3 - Aplicação 2	67
3.4 - Aplicação 3	80
3.5 - Conclusões	95

Capítulo 4 - Considerações Finais

Referências Bibliográficas	i
----------------------------	---

Resumo

Os pesquisadores de marketing tem grande interesse em saber como as variáveis do composto de marketing afetam as decisões de compra do consumidor, e conseqüentemente as vendas das marcas. Essa dissertação apresenta o desenvolvimento do modelo logito multinomial que tem sido usado para compreender melhor a decisão de compra do consumidor. O modelo logito multinomial fornece as probabilidades de escolha de uma alternativa levando em consideração os atributos de todas alternativas disponíveis. O modelo é estocástico e ainda admite variáveis de decisão. A forma de incorporar a heterogeneidade existente na preferência dos consumidores ao modelo logito também é apresentada. Três aplicações do modelo para área de marketing são descritas, todas elas usando produtos de baixo envolvimento de compra. Um fator de grande importância para o bom resultado do modelo é o banco de dados rico em detalhes sobre o histórico de compras dos consumidores e as marcas competitivas no mercado. O modelo apresentado é parcimonioso, uma vez que, os coeficientes das variáveis são modelados de modo a serem os mesmos para todas alternativas.

Abstract

A question of continuing interest to marketing researchers and practitioners is how marketing mix variables affect consumers' purchase decisions and thus the sales of a brand. In this dissertation we developed the multinomial logit model frequently used to solve these questions. The multinomial logit model computes the probability of choosing an alternative as a function of the attributes of all the alternatives available. The model has the appeal of being stochastic and yet admitting decision variables. We will present the heterogeneity in preferences across households being incorporated in a logit model of brand choice, too. The description of three applications of the model in this topic is part of this dissertation. All of them use low-commitment products. The success of the model can be attributed in part to the level of detail and competitive completeness of the scanner panel data. The model is parsimonious in that the coefficients of the variables are the same for all alternatives.

Capítulo 1 - Introdução

1.1 - Introdução

Há muito tempo o comportamento de escolha individual tem sido objeto de estudos em áreas como psicologia , economia e administração. Compreender melhor esse comportamento pode ser uma "grande arma" para profissionais dessas áreas.

O trabalho dessa dissertação está relacionado a um caso especial do comportamento de escolha individual : a decisão de compra do consumidor. Esse assunto vem sendo estudado por profissionais da área de administração de marketing devido a sua importância. A seguir vamos descrever mais detalhadamente o desenvolvimento desse estudo.

O mercado consumidor em questão é formado por todos os indivíduos e famílias que compram ou adquirem **produtos e serviços** para consumo pessoal. Os consumidores variam muito em idade, renda , nível educacional , padrões de migração e gosto. Por isso , o estudo do mercado consumidor vem sendo de grande importância para empresas que desejam obter sucesso em suas vendas.

Cada vez mais os gerentes de marketing tem necessidade de conhecer o mercado com que estão lidando e para isso , devem conhecer os chamados **Sete O's do Mercado** : **Ocupantes** : o público que constitui o mercado , **Objetos** : o que o mercado está comprando , **Objetivos** : o motivo que o mercado compra , **Organização** : quem participa nas compras , **Operações** : como o mercado compra , **Ocasões** : em que momento o mercado compra, **Onde** : local em que o mercado compra.

Um interesse central é saber como os consumidores irão responder aos vários estímulos de marketing pois , entendendo como o mercado vai responder às diferentes características dos produtos (ou serviços) , aos diferentes níveis de preços, aos diversos estímulos promocionais e as variadas formas de distribuição , a empresa terá grande vantagem sobre seus concorrentes. Por isso os pesquisadores de marketing tem dispensado

muitos esforços para estudar a relação entre os estímulos de marketing e a resposta do consumidor.

Kotler (1992) [19] propôs um modelo de comportamento de compra do consumidor como vemos na Figura 1.1.

Figura 1.1 - Modelo de Comportamento de Compra

O ponto inicial é o modelo de estímulos representado na Figura 1.1. Essa figura mostra os esforços de marketing e as variáveis ambientais estimulando positiva ou negativamente o comprador e , provocando em consequência uma resposta de compra.

Os estímulos externos são de dois tipos : os de marketing , formado pelos **quatro P's** : **P**roduto , **P**reço , **P**romoção e **P**ontos de distribuição e , os ambientais , que são os elementos relevantes às decisões de compras , presentes no macroambiente - eventos econômicos , tecnológicos , políticos e culturais. Esses estímulos atingem o comprador e produzem decisões de aquisição mostradas à direita da Figura 1.1 como : escolha do produto, da marca , do revendedor , da oportunidade de compra e da quantidade de compra.

É importante procurar compreender o que acontece entre essas duas fases ; ou seja , o que ocorre na mente do consumidor que o leva a uma determinada decisão de compra.

Nesse estudo quando se fala em compreender o comportamento de compra do consumidor deseja-se estudar como os estímulos externos e as características do comprador influenciam no momento da escolha de determinado produto (ou serviço).

Para cada produto (ou serviço) comprado o consumidor esteve diante de uma tomada de decisão que se inicia no momento em que ele reconhece a necessidade de compra e termina quando o produto está sendo consumido.

Os níveis de envolvimento do consumidor para produtos e marcas decorrem de vários fatores que se apresentam no processo de decisão de compra. O maior ou menor envolvimento é caracterizado por esses fatores e , a intensidade com que eles se apresentam ao consumidor. x3)

O grau de envolvimento relaciona-se diretamente com os benefícios esperados e com a parcela que esse valor significa na renda do comprador , e também , com o grau de interesse que o consumidor dedica a compra do produto. Os riscos percebidos e as possíveis conseqüências negativas também são fatores determinantes para o alto ou baixo envolvimento de compra.

Os produtos adquiridos para consumo e em situações rotineiras parecem não representar situações de muito risco por isso , são associados ao baixo nível de envolvimento.

O envolvimento parece aumentar , quando o produto avaliado reflete uma imagem social. Ou seja , quando o produto está relacionado a uma imagem a qual o consumidor deseja se identificar , ele se envolverá mais com a compra do produto , se preocupando mais com as conseqüências de sua escolha , do que se o mesmo , não fosse refletir sua posição social.

O alto e baixo envolvimento , como uma forma alternativa de entender e explicar o relacionamento entre os consumidores e os produtos , possibilita o planejamento e desenvolvimento de ações diferenciadas dirigidas à conciliação dos interesses de fabricantes , intermediários e seus mercados.

Robertson, Zielenski e Ward (1984) [29] especificaram um conjunto de elementos comportamentais que sintetizam dimensões cognitivas nas quais o processo de decisão com alto e baixo envolvimento podem ser distinguidos. Essa distinção pode ser observada na Figura 1.2 a seguir.

Figura 1.2 - Ótica de Alto e Baixo Envolvimento de Compra

Dimensão Coportamental	Alto Envolvimento	Baixo Envolvimento
Busca de Informações	Busca informações ativamente sobre produtos/marcas	Busca informações limitadas sobre produtos/marcas
Resposta Cognitiva	Resistem à informações discrepantes e intensificam suas contrargumentações	Recebem passivamente informações discrepantes com baixa contrargumentação
Processamento de Informações	Processa informações numa sequência decisória em vários estágios	Processa informações em esquema simplificado , decidindo de forma abreviada
Mudança de Atitude	É difícil e rara	É frequente , mas transitória
Repetição	Persuasão pelo conteúdo da mensagem e não pela repetição	A repetição pode provocar a persuasão
Preferência a Marca	Lealdade a marca é comum	Podem comprar a mesma marca por rotina
Dissonância Cognitiva	Dissonância pós compra é comum	Dissonância pós compra é rara
Influência Pessoal	Obtem informações com outras pessoas ou as usa como fonte para imitação social	Outras pessoas exercem baixa influência

Existem ainda , quatro tipos de comportamento de compra.

•
•Comportamento complexo de compra onde existe uma diferença significativa entre as marcas oferecidas no mercado e um alto envolvimento do consumidor para tomada de decisão , como por exemplo , a compra de um carro novo;

•Comportamento de compra com dissonância reduzida onde existem poucas diferenças entre as marcas e um alto envolvimento como por exemplo , a compra de carpete ;

•Comportamento de compra à procura de variedade onde existem diferenças nas marcas mas um baixo envolvimento como por exemplo , a compra cadernos escolares e ;

•Comportamento habitual de compra onde quase não há diferença entre as marcas e existe um baixo envolvimento do consumidor como por exemplo , a compra de café.

Na Figura 1.3 abaixo são representados os quatro tipos de comportamento de compra .

Figura 1.3. - Tipos de Comportamento de Compra

	Alto Envolvimento	Baixo Envolvimento
Diferença Significativa de Marca	Comportamento Complexo de Compra	Comportamento à Procura de Variedade
Pouca Diferença de Marca	Comportamento com Dissonância Reduzida	Comportamento Habitual de Compra

Um modelo de processo de compra com cinco estágios foi proposto em Kotler (1992) [19] como apresentado na Figura 1.4 . Esse modelo sugere o

(x2)

que foi ressaltado anteriormente : que o processo de compra começa muito antes da compra se efetivar e tem conseqüências logo após a compra. É importante verificar que esse modelo pode não ser verdadeiro quando se trata de compras com baixo envolvimento pois , nesse caso os consumidores podem saltar ou reverter alguns desses estágios.

Figura 1.4 - Modelo de Cinco Estágios do Processo de Compra (X4)

Para os pesquisadores de marketing essas informações são o material básico para suas decisões pois, irão permitir melhor conhecimento do mercado bem como , possibilitarão a melhor avaliação do desempenho de seu produto.

Nesse trabalho serão tratados produtos (ou serviços) que tenham um **baixo envolvimento de compra**.

Entre as características do comprador que parecem influenciar fortemente as decisões de compra está a lealdade à marca ou seja, a compra de um produto (ou serviço) de uma marca que o consumidor está acostumado a adquirir não importando as outras marcas que são apresentadas pelo mercado.

O fato de algumas dessas variáveis influenciarem a escolha do consumidor vem sendo estudado por diversos pesquisadores.

No artigo de Dodds, Monroe e Grewal (1991) [12] , foi feito um estudo para avaliar os efeitos de preço , marca e local de compra na percepção da qualidade e do valor do produto e também, a disposição do consumidor de realizar a compra. Os autores criaram um modelo para avaliação do produto

e mostraram através de uma análise empírica que esse modelo tem validade. O modelo proposto é apresentado na Figura 1.5.

Figura 1.5 - Modelo Conceitual do Efeito de Preço, Marca, e Loja na Avaliação do Produto

Nesse modelo foi mostrada a existência de uma influência de marca, de local de compra e de preço na disposição de compra de um produto.

Para cada uma dessas variáveis foi associada a percepção do consumidor que é, a maneira como cada um nota e avalia o item em questão. A marca, o local de compra e o preço geram uma qualidade percebida do produto. O consumidor irá adquirir o produto medindo a sua disposição de compra que está relacionada com a percepção do valor do produto e do sacrifício. Com isso, tem-se cada consumidor associando um valor percebido diferente ao

mesmo produto provocando assim , uma disposição de compra diferente em cada indivíduo.

No artigo de Davis, Inman e McAlister (1992) [11] , foi feito um estudo do efeito da promoção na avaliação da marca de produto. Os autores testaram a hipótese de que a promoção de um produto pode fazer com que o consumidor crie uma imagem negativa do mesmo e reduza a sua probabilidade de recompra. Vários produtos de baixo envolvimento de compra foram testados e a conclusão dos autores foi que a promoção não provoca uma má avaliação do produto ou seja, há uma rejeição da hipótese de que realizar promoções denigre a imagem da marca.

Um estudo sobre a influência da quantidade comprada e da marca em produtos de baixo envolvimento de compra foi apresentado por Simonson e Winer (1992) [30] .

Até mesmo o efeito que o "zapping" causa no comportamento de compra de uma determinada marca vem sendo estudado como pode ser encontrado no artigo de Zufryden , Pedrick e Sankaralingam (1993) [34] .

Esses são apenas alguns exemplos de que avaliar as influências dos estímulos externos na decisão de compra do consumidor é um assunto muito atual e importante.

Nessa dissertação vamos estudar um modelo estatístico que possibilite compreender o comportamento de compra do consumidor de produtos com baixo envolvimento.

Alguns modelos foram propostos para avaliar as decisões de escolha de marca de consumidores entre eles , um dos mais utilizados é o modelo logito multinomial desenvolvido por McFadden (1974) [24] .

O modelo logito multinomial calcula a probabilidade de escolher uma alternativa como uma função dos atributos de todas as outras alternativas

disponíveis. Ele tem como atração o fato de ser estocástico e ainda admitir variáveis de decisão.

O modelo é bastante utilizado porque é baseado na teoria comportamental da utilidade, permitindo variáveis explicativas e considerando competições entre marcas do mesmo produto e entre produtos similares.

Um exemplo da teoria comportamental da utilidade é um indivíduo que está diante de várias marcas de uma mesma categoria de produtos e ele associa a cada uma delas uma utilidade baseado em seus estímulos externos (variáveis de marketing e do ambiente), a marca para qual ele associar a maior utilidade será sua marca escolhida.

Esse modelo tem uma extensa história de aplicação em marketing, como no artigo de Punj e Staelin (1978) [28] onde foi empregado o modelo para descrever a escolha das escolas feitas por estudantes de administração. Nesse artigo eles desejavam avaliar quais características um estudante considera na escolha de uma escola e qual a influência dessas variáveis (atributos).

O modelo foi usado também, para investigar a escolha dos compradores quanto a loja em que farão suas compras. Esse assunto foi descrito no artigo de Gensch e Recker (1979) [14].

No artigo de Domencich e McFadden (1975) [13], o modelo foi usado no campo de planejamento de transporte para prever a escolha individual do tipo de transporte que será utilizado para viajar: carro, ônibus ou avião.

Nessa dissertação esse modelo vai ser usado para tratar da escolha de marca de produtos que tenham baixo envolvimento de compra.

Alguns pesquisadores tem trabalhado nesse assunto: Guadagni e Little (1983) [15] fizeram um estudo da compra de café moído e utilizaram

informações de compras regulares do produto por 100 donas de casa , durante 32 semanas. Atributos como lealdade à marca , lealdade ao tamanho , presença ou ausência de promoção na loja , preço real do produto e preço promocional são colocados no modelo a fim de produzir uma boa descrição do comportamento de compra.

Krishnamurthi e Raj (1988) [20] avaliaram conjuntamente o modelo de escolha de marca e a quantidade comprada . O maior interesse está no papel do atributo preço na marca e na quantidade compradas. Argumentaram que quando o consumidor está escolhendo entre várias alternativas ele considera o preço de todas as marcas disponíveis mas, quando vai tomar a decisão da quantidade a ser adquirida ele considera o preço só da marca que decidiu adquirir. Para avaliar suas hipóteses utilizaram um painel diário das compras de uma classe de produtos adquiridos freqüentemente. 12

Para estudar como as promoções afetam o momento da compra , a decisão do que comprar e quanto comprar , Gupta (1988) [16] propôs um modelo onde as vendas são decompostas em vendas devido a troca de marca , antecipação do momento de compra e estocagem , e o modelo utilizado para definir a escolha da marca foi o modelo logito multinomial. Para realizar essa análise também foram utilizadas compras regulares de café moído de um painel composto por 100 donas de casa por um período de 100 semanas. Atributos como lealdade à marca , lealdade ao tamanho , preço real , preço promocional , apresentação e "display" foram utilizados no modelo para descrever a decisão da marca a ser adquirida.

No artigo de Chintaguta , Jain e Vilcassim (1991) [9] um painel de 135 donas de casa que realizaram compra de bolachas salgadas de cinco marcas diferentes foi utilizado para analisar a escolha de marca. Eles propuseram uma maneira diferente das existentes até o momento para considerar a presença da heterogeneidade entre os participantes do painel e utilizaram atributos como "display" , apresentação , preço e lealdade para considerar o comportamento do comprador.

Todos esses estudos tem procurado mostrar através do modelo logito multinomial que , os estímulos externos tem um grande impacto seja na

escolha da marca , no momento em que se realizará a compra ou ainda , na quantidade que será comprada.

Uma mercadoria possui atributos que por si só contribuem para sua venda e desempenho (Kotler (1992) [19]) mas, para os pesquisadores de marketing interessa saber como essa mercadoria interage dentro da sua categoria .

Para se entender essas interações é preciso modelar toda categoria de produtos e também ter disponíveis variáveis que sejam capazes de quantificar o ambiente competitivo a que o consumidor é exposto no momento da compra.

Diante do modelo os pesquisadores de marketing seriam capazes de aumentar suas vendas , realizando simulações e , através delas , manipulando convenientemente os fatores de marketing ou ainda, se necessário , realizando um reposicionamento da marca de modo a provocar um aumento nas vendas , através da influência de fatores que são inerentes ao consumidor (fatores que constituem o perfil do consumidor).

A obtenção de dados capazes de "varrer" uma extensa gama de variáveis e atributos bem como, um grande número de consumidores ao longo do tempo foi a primeira preocupação ao se pensar em modelar o comportamento de compra do consumidor.

O avanço de estudos como esses se deve em grande parte a uma técnica de coleta de dados denominada "dados de varredura" (*scanner data*). Eles são assim denominados por possuírem informações completas de um grande número de indivíduos e de itens por um longo período de tempo. Esse tipo de dado permite conhecer a dinâmica do comportamento de escolha.

Os "dados de varredura" (*scanner data*) normalmente, são apresentados de duas formas: "dados de painel" (*panel data*) e "dados estocados" (*store data*).

Os "dados de painel" (*panel data*) fornecem históricos por exemplo, toma-se um indivíduo e faz-se o acompanhamento das suas compras ou seja, toda vez que ele for ao supermercado são armazenadas informações sobre os produtos por ele adquirido , quantidade , marca , preço , data e uma série de outras informações.

Os "dados estocados" (*store data*) são informações de vários itens obtidas em diversos momentos por exemplo , toda semana são coletados os preços de mercadorias que fazem parte da cesta básica além do preço , outras informações podem ser coletadas como promoção, display e propaganda.

Os "dados de varredura" (*scanner data*) tem se apresentado como uma maneira muito eficaz de obter informações.

Os "dados de painel" (*panel data*) ou, dados longitudinais como também são conhecidos, fornecem o acompanhamento ao longo do tempo de uma amostra fixa sendo dessa maneira, o mais conveniente para o tipo de estudo que é realizado quando trata-se do comportamento de compra do consumidor.

Em geral esse acompanhamento está associado a um custo elevado principalmente porque , é preciso um bom tempo para que hajam observações suficientes que sejam capazes de fornecer boas estimativas e , é preciso dispendir esforços para que não haja perdas de unidades amostrais durante o decorrer do tempo.

Os dados longitudinais constituem um caso especial dos chamados dados com medidas repetidas, o que os distingue é que neles existe uma forma sistemática com que as informações de cada unidade participante são obtidas.

Várias áreas de estudo desenvolveram painéis. Podemos citar aqui no Brasil painéis realizados pelo IBOPE em que domicílios são observados em relação ao seu comportamento de compra, ou então , painéis econômicos que acompanham o indivíduo para verificar as mudanças no seu padrão de

vida, e ainda , painéis financeiros onde acompanha-se empresas a fim de verificar seu comportamento em relação aos bancos que utilizam.

Esses dados são coletados por empresas especializadas que desenvolvem técnicas especiais para obtenção dessas informações.

Nos Estados Unidos essas técnicas se desenvolveram bastante em especial, para os que trabalham com comportamento de compra , as empresas estão oferecendo três categorias desse serviço muito interessantes:

- O primeiro são dados coletados em um grupo de lojas localizadas em um único mercado (cidade ou região) oferecendo oportunidades para experimentos variados dentro dessas lojas; a esses dados pode-se ainda acrescentar os compradores que são leais a uma loja aumentando-se as possibilidades de experimentação. - "in-store behavior"

- O segundo são dados coletados em uma amostra nacional de lojas permitindo fazer experimentação em todo o país na tentativa de uma generalização de resultados. O problema ainda apresentado por essa categoria é que o número de lojas disponíveis para coleta desse tipo de serviço ainda é pequeno não sendo assim uma amostra muito representativa porém, é uma alternativa melhor do que tomar um só mercado para generalizar os resultados obtidos.

- Por último , temos o mercado controlado onde é tomada uma região de tamanho médio ou pequeno e todas as suas lojas são equipadas para receber as informações, juntamente com isso trabalha-se de maneira a veicular propagandas por TV, jornal, revistas em locais diferentes e detectar as reações dos consumidores.

Sem as informações obtidas por esse tipo de coleta de dados estudar o comportamento de compra do consumidor através do modelo logito multinomial não seria viável. Pode-se notar que em todas as pesquisas realizadas está associada uma empresa que trabalha com dados de painel : nos artigos de Gupta (1988) [16] , de Chintagunta , Jain e Vilcassim (1991)

[9] os dados são da Information Resources, Inc. (IRI) no de Guadagni e Little (1983) [15] os dados são da Selling Areas Marketing Inc. (SAMI) , e no artigo de Krishnamurthi e Raj (1988) [20] os dados são da ADTEL.

De uma maneira ideal os parâmetros do modelo logito multinomial deveriam ser estimados a nível individual. No entanto, o número de observações para cada indivíduo freqüentemente não é suficiente para obter estimativas consistentes e eficientes dos parâmetros específicos do indivíduo seria preciso um painel muito extenso para obtenção de resultados razoáveis. Além disso, do ponto de vista dos decisores de marketing os parâmetros estimados só tem algum significado a nível agregado ou de mercado.

Porém , ao se trabalhar com as informações agregadas é preciso lembrar que cada unidade que compõe um painel possui características próprias formando um grupo heterogêneo.

Portanto, uma questão pertinente que foi abordada fortemente no artigo de Chintagunta , Jain e Vilcassim (1991) [9] é a do efeito da heterogeneidade na estimativa de parâmetros do modelo logito multinomial. Em geral, ignorar a heterogeneidade quando ela está presente fornecerá estimativas viesadas e inconsistentes dos efeitos dos atributos. (Hsiao (1986) [18])

Vários métodos foram usados na literatura para tratar a heterogeneidade em preferência. Alguns pesquisadores incorporaram a heterogeneidade no contexto do modelo logito multinomial usando o comportamento de escolha passado observado por exemplo, Guadagni e Little (1983) [15] usaram os atributos lealdade à marca e lealdade ao tamanho para considerar a heterogeneidade , Krishnamurthi e Raj (1988) [20] usaram uma variável específica do indivíduo baseada na participação de uma marca em particular no mercado em relação a todas as marcas. Nesses casos , o valor da variável que pretende incorporar a heterogeneidade muda a cada ocasião de compra.

Uma outra abordagem proposta por Hsiao (1986) [18] foi , estimar os parâmetros do modelo assumindo que os termos fixos variam entre os

indivíduos de acordo com alguma distribuição de probabilidade, esse método é conhecido como especificação de efeitos aleatórios. O problema principal dessa abordagem é que se uma distribuição de probabilidade incorreta for imposta para heterogeneidade entre os participantes do painel isso irá resultar em estimativas viesadas dos efeitos dos atributos (Heckman e Singer (1984) [17]).

Uma solução proposta foi utilizar a especificação de efeitos aleatórios semi-paramétrico (Chintagunta, Jain e Vilcassim (1991) [9]). Nessa especificação não se impõe nenhuma forma paramétrica para heterogeneidade existente entre os participantes do painel mas, se estima uma distribuição empírica usando os dados. A estratégia consiste em aproximar uma distribuição de probabilidade para cada marca por um número finito de pontos suporte. A estimação vai envolver a determinação dos pontos de massa e as probabilidades associadas a cada um deles. Esse método foi apresentado por Vilcassim e Jain (1991) [32] e por Heckman e Singer (1984) [17] .

Nessa dissertação vai-se mostrar o tratamento da heterogeneidade desenvolvido no modelo logito multinomial.

Serão apresentados ainda os exemplos de aplicações desses modelos desenvolvidos nos Estados Unidos e através deles a grande utilidade desse estudo.

No Capítulo 2 será apresentado o desenvolvimento do modelo logito multinomial através de uma abordagem axiomática , será mostrado o tratamento da heterogeneidade usando especificação de efeitos aleatórios e a estimação do modelo. No Capítulo 3 serão apresentados três exemplos de aplicações do modelo para produtos de baixo envolvimento de compra desenvolvidos nos Estados Unidos através de painéis de consumidores e , no Capítulo 4 , serão feitas algumas considerações e sugestões para pesquisas futuras.

Capítulo 2- O Modelo Logito Multinomial

2.1 - Introdução

Como foi visto no capítulo anterior o modelo logito multinomial tem sido muito utilizado para compreender a decisão de compra do consumidor. E , com a utilização de dados de painel esse estudo pôde avançar bastante.

Nesse capítulo , na seção 2.2 será mostrada a derivação axiomática desse modelo ou seja, como através de axiomas pôde ser construído o modelo logito multinomial e também , será feita a apresentação da função utilidade ; na seção 2.3 a função utilidade será desenvolvida mostrando como os estímulos externos e as características do consumidor vão ser incorporados ao modelo ; na seção 2.4 serão mostradas as maneiras de incorporar a heterogeneidade ao modelo ; na seção 2.5 será mostrada a função de verossimilhança e por fim , na seção 2.6 a estimação do modelo será tratada.

2.2 - O Desenvolvimento do Modelo Logito Multinomial

Como um modelo de escolha o logito multinomial permite uma derivação axiomática. Considere um indivíduo i confrontando uma escolha de um conjunto $S_i \subset S$ onde S , é o universo de objetos de escolha. —

Seja P_x a probabilidade do indivíduo i escolher a alternativa x , $x \in S_i$. A observação de uma escolha desse indivíduo pode ser vista como a ocorrência de uma distribuição multinomial com probabilidades de seleção P_x para $x \in S_i$.

Suponha que o indivíduo i tenha uma função utilidade que pode ser descrita na forma:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad (1)$$

com i - indicador do indivíduo e ,

j : indicador do objeto ;

onde, V é a componente determinística da utilidade de i a ser calculada de variáveis observáveis e, reflete o gosto "representativo" da população e, ε é a componente aleatória da utilidade de i que varia a cada momento de escolha possivelmente , como um resultado de variáveis não observáveis e, reflete as particularidades do indivíduo no gosto por uma alternativa.

O indivíduo irá escolher a alternativa que maximiza a sua função de utilidade. A probabilidade que um indivíduo selecionado aleatoriamente de uma população escolha a alternativa $x_k \in S_i$, $S_i = \{x_1, \dots, x_J\}$ é dada por:

$$P_{x_k} = P(U_{x_k} > U_{x_j} \text{ para todo } j \neq k) \quad (2)$$

$$= P(V_{x_k} + \varepsilon_{x_k} > V_{x_j} + \varepsilon_{x_j} \text{ para todo } j \neq k)$$

$$= P(\varepsilon_{x_j} - \varepsilon_{x_k} < V_{x_k} - V_{x_j} \text{ para todo } j \neq k)$$

Essa probabilidade induz a uma função de distribuição acumulada conjunta $F(\varepsilon_{x_1}, \dots, \varepsilon_{x_J})$ sobre os valores ε_{x_j} para todo $j=1, \dots, J$.

$$F(\varepsilon_{x_1}, \dots, \varepsilon_{x_J}) = P(\varepsilon_{x_1} \leq \varepsilon_1, \dots, \varepsilon_{x_J} \leq \varepsilon_J)$$

Faça F_l ser a derivada parcial de F em relação ao seu l -ésimo argumento então,

$$P_{x_i} = \int_{\varepsilon=-\infty}^{\infty} F_i(\varepsilon + V_i - V_1, \dots, \varepsilon + V_i - V_J) d\varepsilon \quad (3)$$

Um procedimento para determinar as probabilidades de seleção P_X seria especificar uma distribuição conjunta.

A distribuição conjunta por exemplo, poderia ser a Normal Conjunta que conduziria a uma família de probabilidades dependente de parâmetros desconhecidos da distribuição. Em geral, seria necessário impor hipóteses mais restritas nos parâmetros desconhecidos para torná-los identificados em um experimento de escolha.

Na prática porém, é difícil definir as distribuições conjuntas F que permitam um cálculo das fórmulas de P_i . Uma outra abordagem seria especificar as fórmulas para as probabilidades de seleção e então, verificar se elas podem ser obtidas através de alguma função de distribuição acumulada conjunta ou de uma função utilidade.

McFadden e Richter (1971) [25] caracterizaram as condições necessárias e suficientes das probabilidades de seleção para satisfazer a equação (3). Esse método será utilizado a seguir.

Considere um axioma poderoso para probabilidade de seleção que foi introduzido por Luce (1959) [21], o axioma mostra que a razão relativa de uma alternativa ser escolhida sobre uma segunda seria independente da presença ou ausência de terceiras alternativas consideradas então, irrelevantes.

AXIOMA 1 - Independência das Alternativas Irrelevantes - Para todo possível conjunto de alternativas S_i e $x, y \in S$ tem-se:

$$P(x/\{x, y\}) \cdot P(y/S_i) = P(y/\{x, y\}) \cdot P(x/S_i) \quad (4)$$

A seguir , será mostrado que esse axioma é consistente com a equação (3) e que conduz a especificação simples das seleções de probabilidades.

Quando $P(x/S_i)$ é positiva, a equação (4) implica que $P(x/\{x,y\})$ também é positiva e pode-se escrever:

$$\frac{P(y/\{x,y\})}{P(x/\{x,y\})} = \frac{P(y/S_i)}{P(x/S_i)} \quad (5)$$

Essa condição mostra que a razão de y ser escolhido sobre x em uma situação de múltipla escolha, quando ambas são avaliadas, é igual a razão de uma escolha binária entre x e y .

Como empiricamente uma probabilidade zero é indistinguível em relação a uma probabilidade muito pequena, há pouca perda de generalidade em assumir que as probabilidades de seleção são todas positivas para todo conjunto de alternativas possíveis no experimento.

AXIOMA 2 - Positividade - $P(x/S_i) > 0$ para todo possível conjunto de alternativas S_i e $S_i \in S$.

Considere um conjunto de escolha S_i contendo as alternativas x, y, z e seja $P_{xy} = P(x/\{x,y\})$, então:

$$P(x/\{x,y\}) \cdot P(y/S_i) = P(y/\{x,y\}) \cdot P(x/S_i)$$

$$P(y/S_i) = \frac{P_{yx}}{P_{xy}} \cdot P(x/S_i) \quad (6)$$

e

$$1 = \sum_{y \in S_i} P(y/S_i) = P(x/S_i) \sum_{y \in S_i} \frac{P_{yx}}{P_{xy}} \quad (7)$$

e assim as probabilidade de seleção de escolha múltipla podem ser escritas em termos das razões binárias:

$$P(x/S_i) = \frac{1}{\sum_{y \in S_i} \frac{P_{yx}}{P_{xy}}} \quad (8)$$

Permutando os índices x, y, z na equação (6) e multiplicando, tem-se a condição :

$$P(y/S_i) = \frac{P_{yx}}{P_{xy}} \cdot P(x/S_i)$$

$$P(y/S_i) = \frac{P_{yz}}{P_{zy}} \cdot P(z/S_i)$$

$$P(z/S_i) = \frac{P_{zx}}{P_{xz}} \cdot P(x/S_i)$$

$$\frac{P_{yz}}{P_{zy}} \cdot \frac{P_{zx}}{P_{xz}} \cdot P(x/S_i) = \frac{P_{yx}}{P_{xy}} \cdot P(x/S_i)$$

$$\frac{P_{yx}}{P_{xy}} = \frac{P_{yz}}{P_{zy}} \cdot \frac{P_{zx}}{P_{xz}} \quad (9)$$

Seja z uma alternativa de referência ao conjunto de alternativas S_i e define-se:

$$V_{xz} = \log \left(\frac{P_{xz}}{P_{zx}} \right) \text{ logo ,}$$

$$\frac{P_{xz}}{P_{zx}} = e^{I_{xz}}$$

e , a equação (8) poderá ser escrita :

$$P(x/S_i) = \frac{1}{\sum_{y \in S_i} \frac{P_{yz}}{P_{zy}} \cdot \frac{P_{zx}}{P_{xz}}}$$

$$P(x/S_i) = \frac{1}{\frac{P_{zx}}{P_{xz}} \cdot \sum_{y \in S_i} \frac{P_{yz}}{P_{zy}}}$$

$$P(x/S_i) = \frac{P_{xz}/P_{zx}}{\sum_{y \in S_i} P_{yz}/P_{zy}}$$

$$P(x/S_i) = \frac{e^{I_{xz}}}{\sum_{y \in S_i} e^{I_{yz}}} \quad (10)$$

Na função V_{xz} pode-se pensar no argumento x como dando um "efeito da alternativa escolhida" e o argumento z como um "efeito do conjunto de alternativas" . Em um experimento com vários conjuntos de escolha S_i e várias repetições das ocorrências seria possível identificar cada um desses efeitos. Na ausência de réplicas é impossível identificar o "efeito do conjunto de alternativas" e uma restrição é necessária para isolar o "efeito da alternativa escolhida".

AXIOMA 3 - Irrelevância do efeito do conjunto de alternativas -A função V_{xz} determinando as probabilidades de seleção na equação (10) tem a forma aditiva separada :

$$V_{xz} = V_x - V_z$$

De acordo com esse axioma pode-se escrever a equação (10) como :

$$\begin{aligned} P(x/S_i) &= \frac{e^{I'_{xz}}}{\sum_{y \in S_i} e^{I'_{yz}}} \\ P(x/S_i) &= \frac{e^{I'_x - I'_z}}{\sum_{y \in S_i} e^{I'_y - I'_z}} \\ P(x/S_i) &= \frac{e^{I'_x}}{\sum_{y \in S_i} e^{I'_y}} \end{aligned} \quad (11)$$

e a função V pode ser interpretada como um indicador da utilidade dos postos "representativos".

O resultado seguinte justifica a terminologia em termos do comportamento de uma população de consumidores.

LEMA 1

Suponha que cada indivíduo da população de consumidores tem uma função utilidade $U_{x_j} = V_{x_j} + \varepsilon_{x_j}$ onde V_{x_j} é uma função não estocástica refletindo o gosto representativo e ε_{x_j} é uma função que varia aleatoriamente na população com a propriedade que em cada conjunto de alternativas possível $S = \{x_1, \dots, x_J\}$, os valores de ε_{x_j} são independentes, identicamente distribuídos com uma distribuição de Weibull.

$$P(\varepsilon_{x_j} \leq \varepsilon) = e^{-e^{-\varepsilon}} \quad (12)$$

Então , as probabilidades de seleção dadas pela equação (3) satisfazem a equação (11).

A demonstração desses resultados pode ser encontrada no artigo de McFadden (1974) [24].

A forma final da função de distribuição é dada por :

$$\begin{aligned} F_l(\varepsilon + V_l - V_1, \dots, \varepsilon + V_l - V_J) = \\ = \exp(-\varepsilon) \cdot \exp\left(-\exp(-\varepsilon) \left[\sum_{j=1}^J \exp(V_j - V_l) \right]\right) \end{aligned}$$

Uma prova não construtiva desse resultado foi dada pela primeira vez por Marschak (1959) [23] , o argumento acima aparece em Luce e Suppes (1965) [22] e é atribuído a E.Holman e A.Marley.

E assim é obtida a derivação axiomática do modelo logito multinomial.

2.3 - Função de Utilidade Linear

A fórmula (11) para probabilidades de seleção obtidas dos axiomas 1 a 3 podem ser adaptadas para análise empírica especificando a forma da função utilidade , mais especificamente, da componente determinística V_{x_j} . Uma suposição conveniente é que V seja linear com parâmetros desconhecidos.

Então, a componente determinística da utilidade do consumidor para alternativa x_j será expressa como uma função linear de variáveis observadas chamadas atributos de x_j . Alguns desses , serão atributos da

alternativa e outros , poderão ser atributos do consumidor ou de seu meio ambiente que irão favorecer uma alternativa a outra por alguma razão. Em geral, tem-se:

$$V_{ix_j} = \sum_{k \in T} \beta_{kx_j} \cdot z_{ikx_j} \quad (13) \quad \text{onde,}$$

z_{ikx_j} - valor observado do atributo K da alternativa x_j para o consumidor i

β_{kx_j} - peso da utilidade do atributo K da alternativa x_j

Quando não houver dúvidas pode-se omitir o subscrito i que indica o consumidor.

Do ponto de vista da modelagem é conveniente separar os atributos em duas classes :

(1) T_j - atributos exclusivos a alternativa x_j - um produto pode possuir itens que outros produtos não tem e que contribuem para sua utilidade. Para cada atributo $k \in T_j$ os coeficientes podem ser denotados por β_{kj} e as variáveis por z_{ikx_j} .

(2) T_C - atributos comuns a todas as alternativas - embora um atributo tal como preço possa ser associado unicamente a cada produto, um modelo mais parcimonioso no número de parâmetros usaria preço como atributo único para todos os produtos. O preço terá o mesmo coeficiente para todas alternativas. Para cada atributo $k \in T_C$ os coeficientes podem ser denotados por β_k e as variáveis por z_{ikx_j} .

Assim a fórmula (13) pode ser escrita :

$$V_{ix_j} = \sum_{k \in T_j} \beta_{ki} \cdot z_{ikx_j} + \sum_{k \in T_c} \beta_k \cdot z_{ikx_j} \quad (14)$$

Embora uma forma linear para utilidade seja uma colocação natural para começar a construção do modelo, nota-se que a linearidade de V_{x_j} conduz a probabilidades de seleção não lineares nas variáveis observáveis z_{kx_j} .

$$e^{V_{x_j}} = e^{\sum \beta_{kx_j} \cdot z_{kx_j}} = \prod e^{\beta_{kx_j} \cdot z_{kx_j}} \quad (15)$$

então, o modelo é mais multiplicativo do que aditivo.

$$P(x / S_i) = \frac{e^{V_x}}{\sum_{y \in S_i} e^{V_y}} = \frac{e^{\beta_{kx_j} \cdot z_{kx_j}}}{\sum_{y \in S_i} e^{\beta_{ky} \cdot z_{ky}}} \quad (16)$$

Esse é o **modelo logito multinomial**.

Na prática, não são observadas nem as utilidades nem as probabilidades diretamente. As escolhas e os valores dos atributos é que são observados. Os dados consistem de um conjunto de informações sobre as escolhas de cada indivíduo e, os valores dos atributos das alternativas em cada ocasião de escolha.

$y_{x_i}^i(n)$ - 1 - se o consumidor i escolher a alternativa x_i na n -ésima ocasião de escolha e, 0 - caso contrário

$x_{kx_j}^i(n)$ - valor do atributo k para alternativa x_j na n -ésima ocasião de escolha para consumidor i .

Em um caso mais específico, a escolha ou observação é a compra de um produto por um consumidor em determinada ocasião. Os β_{ki} e β_k são

constantes desconhecidas que devem ser determinadas (estimadas) usando-se parte dos dados para calibração.

A calibração é feita através da máxima verossimilhança usando-se (10) e (14) para calcular a função de verossimilhança. Um programa para esse cálculo foi desenvolvido por Manski e Ben Akiva em Ben-Akiva (1973) [3]. Os parâmetros estimados de máxima verossimilhança são consistentes e assintoticamente eficientes e com distribuição normal sob condições gerais.

2.4 - Tratamento da Heterogeneidade

Nos itens anteriores foi mostrado que o desenvolvimento teórico do modelo logito é baseado na maximização da utilidade do comportamento de compra de um indivíduo.

Então , o ideal seria que os parâmetros do modelo logito fossem estimados a nível individual. Em dados de painel, no entanto, o número de observações por indivíduo frequentemente não é o suficiente para fornecer estimativas consistentes e eficientes dos parâmetros específicos do indivíduo. E mais, do ponto de vista das decisões de marketing, os parâmetros estimados tem mais significado a nível agregado.

Por isso , os pesquisadores tem feito uso da junção dos dados e estimado um conjunto de parâmetros a nível agregado.

Quando da união dessas informações, deve-se lembrar que cada indivíduo possui um comportamento diferente e o efeito da heterogeneidade na estimação dos parâmetros do modelo logito para escolha da marca deve ser levada em consideração.

Em geral, ignorar a heterogeneidade quando ela está presente levará a estimativas inconsistentes e viesadas.

Alguns pesquisadores incorporaram a heterogeneidade na preferência da marca no contexto do modelo logito , usando o comportamento passado

de escolha de marca observado. Isso pode ser feito usando variáveis no modelo do tipo "lealdade à marca" e "lealdade ao tamanho" (Guadagni e Little (1983) [15]), ou usando uma variável baseada na participação de compras de uma determinada marca em relação a todas outras (Krishnamurthi e Raj (1990) [20]).

Uma operacionalização diferente foi proposta por Allenby e Rossi (1990) [2] onde , a heterogeneidade foi captada por um termo fixo que é medido como a diferença entre as probabilidades de escolha previstas e as probabilidades estimadas através de uma abordagem de frequências relativas.

Uma justificativa para o uso dessas variáveis é que , elas procuram captar as diferenças na preferência das marcas entre indivíduos.

A utilização dessas variáveis que usam o comportamento de escolha de marca passada observado para captar a heterogeneidade pode afetar as estimativas das variáveis do composto de marketing tais como , preço e promoção.

O que ocorre é que , o efeito das escolhas de marca passada na atual escolha terá um peso maior se não forem ajustados para os possíveis efeitos de preço e promoção. Da mesma maneira, o efeito de algumas das outras variáveis no modelo terão peso menor.

Uma abordagem alternativa possível para captar a heterogeneidade , seria incluir um termo constante ou específico de um indivíduo em adição ao conjunto de parâmetros comuns para variáveis do composto de marketing.

A idéia é que a preferência de um indivíduo para cada marca é estável e pode ser captada adequadamente por um termo fixo.

Essa abordagem não depende do comportamento passado e assim , não contamina os efeitos estimados das variáveis do composto de marketing.

Existem duas maneiras de estimar os parâmetros do modelo com essa especificação. Uma , é estimar o termo constante para cada indivíduo e para cada marca , que é conhecido como modelo de efeito-fixo.

No entanto, essa especificação depende da estimação de um grande número de observações por indivíduo. Quando existem poucas observações por indivíduo, a estimação de máxima verossimilhança convencional conduz a estimativas inconsistentes não só dos termos fixos mas também, dos efeitos das variáveis do composto de marketing.

Uma segunda abordagem para estimar os parâmetros do modelo , é assumir que o termo fixo varia entre os indivíduos de acordo com alguma distribuição de probabilidade. Esse método de considerar a heterogeneidade é conhecido como especificação de efeitos aleatórios.

Os parâmetros do modelo sob tal especificação podem ser estimados de duas maneiras. A primeira , é assumir uma forma paramétrica particular para distribuição de heterogeneidade entre indivíduos.

Um grande problema nessa abordagem é que se uma distribuição de probabilidade incorreta for imposta para heterogeneidade entre indivíduos, isso irá resultar em estimativas viesadas dos efeitos das variáveis de marketing na escolha da marca.

Uma solução seria usar uma especificação de efeitos aleatórios semi-paramétrica na qual não se impõe nenhuma forma paramétrica para heterogeneidade entre indivíduos mas , se estima a distribuição empiricamente através dos dados.

A seguir será mostrado como fica o modelo logito quando se incorpora o termo que vai captar a heterogeneidade entre os indivíduos participantes do painel.

Em (1) tem-se :

$$U_{ij} = V_{ij} + \varepsilon_{ij}$$

Como visto anteriormente V é a componente determinística que foi escrita da forma:

$$V_{ix_j} = \sum_{k=1}^N \beta_{kx_j} \cdot z_{kx_j}^i = \underline{Z}_{ix_j}' \cdot \underline{B}$$

onde, $\underline{Z}_{ix_j}$ é o vetor de valores observados e , $\underline{B}$ é o vetor dos parâmetros desconhecidos.

A essa expressão vamos acrescentar um termo fixo α_{ix_j} que servirá para captar a preferência intrínseca do indivíduo i à "marca" x_j .

$$V_{ix_j} = \underline{Z}_{ix_j}' \cdot \underline{B} + \alpha_{ix_j} \quad (17)$$

De acordo com o que foi visto anteriormente , sob as suposições de maximização da utilidade e , ε_{ix_j} independente , identicamente distribuídos com distribuição de Weibull , a probabilidade que o indivíduo i escolha a marca x_j na ocasião de escolha t , condicionadas aos valores $\alpha_{ix_1} , \dots , \alpha_{ix_n}$ é dado por :

$$P_{it}(x_j / \alpha_{i_1}, \dots, \alpha_{i_n}) = \frac{\exp(\underline{Z}_{ix_j}' \cdot \underline{B} + \alpha_{ix_j})}{\sum_{k=1}^N \exp(\underline{Z}_{ix_k}' \cdot \underline{B} + \alpha_{ix_k})} \quad (18)$$

Para um indivíduo i , os valores α_{ix_k} assume-se que sejam N escalares fixos que não variam ao longo do tempo.

Mas, entre os indivíduos, os valores α_{ix_k} podem variar. Essa variação (heterogeneidade) vai ser captada quando se assume uma distribuição de probabilidade para cada marca.

Seja α_{x_k} uma variável aleatória associada com a distribuição de probabilidade para marca x_k , $k = 1, 2, \dots, N$. Assim, α_{ix_k} representa uma realização particular de α_{x_k} para o indivíduo i . As distribuições de preferência entre indivíduos para uma marca particular são independentes das distribuições de preferência para outras marcas.

Para identificar os parâmetros do modelo pode-se normalizar a equação (18) tomando como base, por exemplo, a marca x_n .

$$\begin{aligned} P_{it}(x_j / \alpha_{i_1}, \dots, \alpha_{i_n}) &= \\ &= \frac{\exp\left\{(\underline{Z}_{ix_j} - \underline{Z}_{ix_n})' \cdot \underline{B} + (\alpha_{ix_j} - \alpha_{ix_n})\right\}}{1 + \sum_{k=1}^{N-1} \exp\left\{(\underline{Z}_{ix_k} - \underline{Z}_{ix_n})' \cdot \underline{B} + (\alpha_{ix_k} - \alpha_{ix_n})\right\}} \\ &= \frac{\exp(\underline{W}_{ix_j}' \cdot \underline{B} + v_{ix_j})}{1 + \sum_{k=1}^{N-1} \exp(\underline{W}_{ix_k}' \cdot \underline{B} + v_{ix_k})} \end{aligned}$$

onde , $\underline{W}_{ix_j} = (\underline{Z}_{ix_j} - \underline{Z}_{ix_n})'$ e, $v_{ix_j} = (\alpha_{ix_j} - \alpha_{ix_n})$ para $j=1,2,\dots,n$

Em geral, como o número de observações é insuficiente para um indivíduo i , não será viável estimar o parâmetro específico do indivíduo V_{ix_k} $k=1,2,\dots,n-1$. Para contornar esse problema pode-se derivar a probabilidade não condicional que um indivíduo selecionado aleatoriamente irá escolher a marca x_j .

Isso poderia ser obtido integrando a probabilidade condicional na equação (19) com relação a distribuição de V_{ix_k} $k=1,2,\dots,n-1$ onde $v_{xk} = (\alpha_k - \alpha_n)$.

Pode-se notar porém , que embora as variáveis aleatórias α_k $k=1,2,\dots,n$ sejam por suposição independentes, os V_k não tem distribuição independente porque:

$$E[(v_k, v_i)] = E[(\alpha_k - \alpha_n), (\alpha_i - \alpha_n)] = E[\alpha_n^2] \neq 0$$

Por isso, quando se trabalha com as variáveis aleatórias V_k $k=1,2,\dots,n-1$ é preciso usar uma distribuição de probabilidade conjunta variante $(n-1)$, o que torna o procedimento de estimação intratável.

Para poder realizar a estimação foi proposta uma abordagem onde é realizada uma decomposição que explora a suposição de independência das variáveis aleatórias α_k $k=1,2,\dots,n$. A variável α_k é decomposta como segue:

$$\alpha_k = \mu_k + \gamma_k \quad k=1,2,\dots,n \quad (20)$$

onde , μ_k é variável aleatória com média zero e , γ_k é escalar específico da marca de valor invariante entre os indivíduos e igual ao valor

médio de α_k . Como as variáveis aleatórias α_k são independentes a equação (20) implica que as variáveis aleatórias μ_k $k=1,2,\dots,n$ também são independentes. Assim, para o i -ésimo indivíduo tem-se :

$$\alpha_{ik} = \mu_{ik} + \gamma_k \quad k=1,2,\dots,n \quad (21)$$

A equação (18) pode então ser escrita como:

$$\begin{aligned} P_{it}(x_j / \mu_{i1}, \dots, \mu_{in}) &= \\ &= \frac{\exp(\underline{Z}_{ix_j}' \cdot \underline{B} + \mu_{ix_j} + \gamma_{x_j})}{\sum_{k=1}^N \exp(\underline{Z}_{ix_k}' \cdot \underline{B} + \mu_{ix_k} + \gamma_{x_k})} \quad (22) \end{aligned}$$

Para fazer a normalização uma consideração importante é ser capaz de preservar a suposição de independência das variáveis aleatórias μ_k $k=1,2,\dots,n$. Assim, divide-se a equação (22) por

$\exp(\underline{Z}_{ix_n}' \cdot \underline{B} + \gamma_{x_n})$ e obtem-se :

$$\begin{aligned} P_{it}(x_j / \mu_{i1}, \dots, \mu_{in}) &= \\ &= \frac{\exp\left\{(\underline{Z}_{ix_j} - \underline{Z}_{ix_n})' \cdot \underline{B} + \mu_{ix_j} + (\gamma_{x_j} - \gamma_{x_n})\right\}}{\exp(\mu_{ix_n}) + \sum_{k=1}^{n-1} \exp\left\{(\underline{Z}_{ix_k} - \underline{Z}_{ix_n})' \cdot \underline{B} + \mu_{ix_k} + (\gamma_{x_k} - \gamma_{x_n})\right\}} \end{aligned}$$

fazendo

$$\underline{W}_{ix_j}' = (\underline{Z}_{ix_j} - \underline{Z}_{ix_n})' \text{ e } \theta_{x_j} = \gamma_{x_j} - \gamma_{x_n}$$

$j=1,2,\dots,n$ obtem-se :

$$\begin{aligned}
P_{it}(x_j / \mu_{i1}, \dots, \mu_{in}) &= \\
&= \frac{\exp(W'_{ix_j} \cdot \underline{B} + \mu_{ix_j} + \theta_{x_j})}{\exp(\mu_{ix_n}) + \sum_{k=1}^{n-1} \exp(W'_{ix_k} \cdot \underline{B} + \mu_{ix_k} + \theta_{x_k})} \quad (24)
\end{aligned}$$

Seja $\xi_{ix_n} = \exp(\mu_{ix_n})$ $k=1,2,\dots,n$. A probabilidade condicional em (24) pode então ser escrita como:

$$\begin{aligned}
P_{it}(x_j / \xi_{i1}, \dots, \xi_{in}) &= \\
&= \frac{\xi_{ix_j} \exp(W'_{ix_j} \cdot \underline{B} + \theta_{x_j})}{\xi_{ix_n} + \sum_{k=1}^{n-1} \xi_{ix_k} \exp(W'_{ix_k} \cdot \underline{B} + \theta_{x_k})} \quad (25)
\end{aligned}$$

A heterogeneidade entre os indivíduos será agora captada através do termo ξ_k $k=1,2,\dots,n$. Quando se exclui μ_{ix_n} na normalização preserva-se a propriedade de independência de ξ_k $k=1,2,\dots,n$. Isso torna a estimação tratável. E mais, os parâmetros θ_k com $k=1,2,\dots,n-1$ representam a média da preferência intrínseca para cada marca em relação a marca n e assim, são análogas as constantes usadas no modelo logito convencional.

Para obter a probabilidade não condicional de escolher a marca x_j na ocasião de compra t toma-se a integral da equação (25) com relação a todas n variáveis aleatórias independentes $\xi_{x_1}, \dots, \xi_{x_n}$.

$$(26) \quad P_{it}(x_j) = \int_0^\infty \int_0^\infty \dots \int_0^\infty P_{it}(x_j / \xi_{ix_1}, \dots, \xi_{ix_n}) f(\xi_{x_1}) \dots f(\xi_{x_n}) d(\xi_{x_1}) \dots d(\xi_{x_n})$$

onde $f(\xi_{x_k})$ é a função densidade de ξ_{x_k} $k=1,2,\dots,n$.

E assim , são obtidas as probabilidades de um indivíduo i , na ocasião de compra t selecionar uma determinada alternativa considerando a heterogeneidade existente entre os indivíduos.

2.5 - A Função de Verossimilhança

Para estimar os parâmetros do modelo será utilizado o método da máxima verossimilhança.

Considerando o modelo que leva em conta a heterogeneidade através da especificação de efeitos aleatórios. Os parâmetros são: B efeitos das variáveis que fazem parte do composto da marketing e que estão na função utilidade; θ_k $k=1,2,\dots,N-1$ média da preferência intrínseca para cada marca, e os parâmetros da distribuição de probabilidade de ξ_k $k=1,2,\dots,N$.

Seja T_i o número de compras realizadas pelo i -ésimo indivíduo participante do painel.

A função de verossimilhança associada condicionada em $\xi_1, \xi_2, \dots, \xi_N$ é dada por:

$$(27) \quad L_i / \xi_1, \dots, \xi_N = \prod_{t=1}^{T_i} \left\{ \prod_{j=1}^N P_t(j / \xi_1, \dots, \xi_N)^{\delta_{ijt}} \right\}$$

onde $\delta_{ijt} = 1$ se o i -ésimo indivíduo escolher a marca j na ocasião de compra t e , 0 caso contrário.

A função de verossimilhança não condicional é obtida integrando a equação (27) nas distribuições de probabilidades de $\xi_1, \xi_2, \dots, \xi_N$.

$$(28) L_i = \int_0^\infty \dots \int_0^\infty (L_i / \xi_1, \dots, \xi_N) f(\xi_1) \dots f(\xi_N) d(\xi_1) \dots d(\xi_N)$$

A função de verossimilhança amostral para todos os participantes do painel (M) é dada por:

(29)

$$L = \prod_{i=1}^M L_i$$

$$= \prod_{i=1}^M \int_0^\infty \dots \int_0^\infty \left\{ \prod_{t=1}^{T_i} \left\{ \prod_{j=1}^N \left[\frac{\xi_j \exp(\underline{Z}_{jt}' \cdot \underline{B} + \theta_j)}{\xi_N + \sum_{k=1}^{N-1} \xi_k \exp(\underline{Z}_{kt}' \cdot \underline{B} + \theta_k)} \right]^{\delta_{jt}} \right\} \right\} f(\xi_1) \dots f(\xi_N) d(\xi_1) \dots d(\xi_N)$$

Através dessa função de verossimilhança amostral , é que se vai obter as estimativas dos parâmetros do modelo.

Para maximizar a função de verossimilhança amostral acima é preciso especificar a função de distribuição de probabilidade $f(\xi_1), \dots, f(\xi_N)$. Para especificação paramétrica de efeitos aleatórios podem ser consideradas duas funções de distribuições como candidatas : a função Gama e a função Normal que já foram utilizadas anteriormente na literatura para modelar heterogeneidade.

No caso da distribuição Gama é preciso definir a variável aleatória no intervalo de $(0, \infty)$. Nota-se que $\xi_k = \exp(\mu_k)$ satisfaz essa

condição. Então, assume-se que as variáveis aleatórias ξ_k $k=1,2,\dots,N$ tem distribuição Gama com parâmetros de forma e de escala ω_k e δ_k respectivamente:

$$(30) f(\xi_k) = \frac{\delta_k^{\omega_k}}{\Gamma(\omega_k)} \exp(-\delta_k \xi_k) \xi_k^{\omega_k-1}$$

onde $\Gamma(.)$ é a função Gama.

Recorda-se que μ_k é definida como uma variável com média zero devido a presença de uma constante específica de marca γ_k . Como $\log(\xi_k) = \mu_k$ a variável aleatória $\log(\xi_k)$ também tem média zero o que impõe uma relação funcional entre ω_k e δ_k da distribuição de ξ_k . Então, sem perda de generalidade o parâmetro δ_k será fixado igual a 1 e somente o parâmetro de forma será estimado.

$$(31) f(\xi_k) = \frac{1}{\Gamma(\omega_k)} \exp(-\xi_k) \xi_k^{\omega_k-1}$$

Pode-se ainda considerar a distribuição Normal para heterogeneidade, a variável aleatória em questão deve ser livre para assumir qualquer valor nos reais, positivo ou negativo. A variável aleatória ξ_k não é apropriada pois, ela só assume valores não negativos. Então, considera-se a variável aleatória μ_k que pode assumir valores positivos e negativos. Note que μ_k tem média zero por causa da presença da constante específica de marca γ_k .

A função densidade para μ_k é

$$g(\mu_k) = \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left(-\frac{\mu_k^2}{2\sigma_k^2}\right) \quad (32)$$

Dessa forma, a probabilidade que um indivíduo selecionado aleatoriamente compre a marca j na ocasião de compra t é dada por:

$$(33) \quad P_t = \int_0^\infty \dots \int_0^\infty P_t(j/\mu_1, \dots, \mu_N) g(\mu_1), \dots, g(\mu_N) d\mu_1 \dots d\mu_N$$

Quando a expressão da função de probabilidade da distribuição Gama ou da distribuição Normal são substituídos, a função de verossimilhança amostral não possui uma expressão de forma fechada. Usar técnicas de integração numérica é complicado e freqüentemente leva a estimativas instáveis.

Para contornar esse problema é utilizada a técnica da Quadratura Gaussiana para aproximar as integrais da função de verossimilhança em somas de um número finito de pontos.

A principal vantagem desse método sobre a técnica de integração numérica é que ele requer um número pequeno de pontos para obter uma boa aproximação das integrais. A seguir são apresentados mais detalhes sobre o método da Quadratura Gaussiana.

"Quadratura Gaussiana"

É preciso fazer a suposição que a distribuição de heterogeneidade para as N marcas são independentes. Então, as integrais em questão podem ser avaliadas com relação a distribuição de uma marca de cada vez. A técnica vai ser ilustrada para uma marca.

A forma da integral que precisa ser aproximada é a seguinte:

$$(Q1) \quad \int_0^\infty W(\xi) f(\xi) d\xi \quad , \text{ onde:}$$

$$W(\xi) = \left\{ \prod_{t=1}^{T_i} \left\{ \prod_{j=1}^N \left[\frac{\xi_j \exp(Z'_{ijt} \cdot \underline{B} + \theta_j)}{\xi_N + \sum_{k=1}^{N-1} \xi_k \exp(Z'_{ikt} \cdot \underline{B} + \theta_k)} \right]^{\delta_{ijt}} \right\} \right\}$$

Usando a quadratura gaussiana pode-se transformar a integral (Q1) na soma:

$$(Q2) \quad \sum_{h=1}^H W(\xi_h) \lambda_h$$

onde , a função densidade $f(\xi)$ é aproximada por um número finito H, dos pontos de suporte ξ_h com probabilidade de massa correspondentes λ_h $h=1,2,\dots,H$. A função $W(\xi)$ é conseqüentemente , avaliada somente em cada um desses valores ξ_h finitos.

Butler e Moffit (1982) [6] mostraram que a quadratura permite bons resultados mesmo com a aproximação de dois pontos (H=2). Dessa forma é necessário obter ξ_1 e ξ_2 e os correspondentes λ_1 e λ_2 para a aproximação.

Quando a variável aleatória ξ em (Q1) tem distribuição Gama com parâmetros α e 1, foi mostrado que os pontos de suporte ξ_h , $h=1,2,\dots,H$ são as raízes do polinômio de Laguerre $l_H(\xi)$ de ordem H (Press e outros ,(1986) [27]). O polinômio é dado pela seguinte relação de recorrência:

$$(Q3) \quad H l_H(\xi) = (\alpha - 1 + 2H - \xi) l_{H-1}(\xi) - (\alpha + H - 1) l_{H-2}(\xi)$$

para H=2 tem-se:

$$(Q4) \quad 2 l_2(\xi) = (\alpha + 3 - \xi) l_1(\xi) - (\alpha + 1) l_0(\xi)$$

Da tabela 22.4 de Abramowitz e Stegun (1972) [1]:

$$(Q5) \quad l_0(\xi) = 1 ; \quad l_1(\xi) = (\alpha + 1 - \xi)$$

Substituindo (Q5) em (Q4) e simplificando obtém-se:

$$(Q6) \quad l_2(\xi) = 1/2 \{ \xi^2 - (2\alpha + 4)\xi + (\alpha^2 + 3\alpha + 2) \}$$

As raízes do polinômio em (Q6) são dadas por:

$$(Q7) \quad \begin{aligned} \xi_1 &= (\alpha + 2) + \sqrt{\alpha + 2} \\ \xi_2 &= (\alpha + 2) - \sqrt{\alpha + 2} \end{aligned}$$

Para obter os valores de λ_h correspondentes a ξ_h é usado o seguinte resultado:

$$(Q8) \quad \lambda_h = \frac{\phi_2(\xi_h)}{l_2'(\xi_h)} \quad h = 1, 2$$

onde l_2' é a derivada de l_2 com relação a ξ e ϕ_2 segue a relação de recorrência descrita em (Q4), exceto que:

$$(Q9) \quad \phi_0(\xi) = 0, \quad \phi_1(\xi) = l_1'$$

Substituindo (Q9) em (Q4) :

$$(Q10) \quad \phi_2(\xi) = 1/2 \{ (\alpha + 3 - \xi) l_1'(\xi) \}$$

Usando a expressão (Q10) em (Q8) :

$$\lambda_1 = 1/2[1 - 1/\sqrt{\alpha + 2}]$$

(Q11)

$$\lambda_2 = 1/2[1 + 1/\sqrt{\alpha + 2}]$$

Dessa forma , a solução da integral dada por Q1 é obtida. Repetindo de maneira análoga para todas as integrais na função de verossimilhança elas são transformadas em somas que são de uso mais ameno em rotinas de estimação de máxima verossimilhança padrão. E , conseqüentemente , obtem-se as soluções dos estimadores dos vários parâmetros da função de verossimilhança. Esses estimadores são utilizados no modelo logito multinomial que fornece as probabilidades de preferência para cada marca considerada.

Para uma especificação de efeitos aleatórios semi-paramétrica não é imposta nenhuma forma paramétrica para função densidade de probabilidade $f(\cdot)$ e $g(\cdot)$. ao invés disso, se aproxima cada uma das distribuições de probabilidade por um número finito de pontos de suporte e determina-se a localização e probabilidade da massa associada a cada ponto. O número de pontos suporte necessários para aproximar cada uma das distribuições é determinado empiricamente por um procedimento iterativo onde vai se acrescentando pontos até que 2 deles se tornem tão próximos um do outro que não seja possível distinguí-los empiricamente. A função de verossimilhança amostral para essa especificação é dada por :

$$(34) \quad L = \prod_{i=1}^M \left\{ \sum_{h=1}^{H_N} \sum_{h=1}^{H_{N-1}} \dots \left\{ \sum_{h=1}^{H_i} \left\{ \prod_{t=1}^{T_i} \left\{ \frac{\xi_j \exp(\underline{Z}'_{jt} \cdot \underline{B} + \theta_j)}{\xi_N^h + \sum_{k=1}^{N-1} \xi_k^h \exp(\underline{Z}'_{kt} \cdot \underline{B} + \theta_k)} \right\}^{\delta_{jt}} \right\} \lambda_1^h \right\} \dots \lambda_N^h \right\}$$

onde ξ_j^h e λ_j^h são, respectivamente, a locação e massa associada com a h-ésimo ($h=1,2,\dots,H_j$) ponto de suporte para marca j. Nessa expressão as variáveis aleatórias ξ_k são tais que, $E(\log \xi_k) = 0$

$k=1,2,\dots,N$. É necessário impor essa restrição na estimação paramétrica e semi-paramétrica.

2.6 - Conclusões

Dessa maneira , são apresentadas as várias funções de verossimilhança obtidas de acordo com a abordagem considerada.

Uma vez obtida a função de verossimilhança amostral , adota-se uma técnica para maximização dessa função de forma a obter as estimativas dos parâmetros do modelo.

Assim , com as estimativas dos parâmetros do modelo logito multinomial , é possível realizar uma série de outros cálculos para obter medidas agregadas que são de grande utilidade em marketing , ou ainda , fazer simulações com o modelo obtido.

Fica apresentado nesse capítulo a derivação axiomática do modelo logito multinomial para comportamento de escolha individual , bem como, todo procedimento para tratamento da heterogeneidade e estimação dos parâmetros.

No capítulo 3 serão apresentadas aplicações do modelo apresentado nesse capítulo.

Capítulo 3 -Algumas Aplicações

3.1 - Introdução

Nesse capítulo , serão apresentados três exemplos de aplicações do modelo logito multinomial encontrados na literatura. O objetivo é mostrar a utilidade desse modelo para a área de marketing.

A primeira aplicação a ser apresentada será do artigo de Guadagni e Little (1983) [15] "*A Logit Model of Brand Choice Calibrated on Scanner Data*" onde é analisado o comportamento de compra de café moido cafeinado ; a segunda aplicação está no artigo de Chintagunta , Jain e Vilcassim (1992) [9] "*Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data*" onde é estudado o comportamento de compra de bolachas e por fim , a terceira aplicação de Gupta (1988) [16] "*Impact of Sales Promotions on When , What , and How much to Buy*" onde , novamente se estuda o comportamento de compra de café moído.

3.2 - Aplicação 1

3.2.1 - Formação do Banco de Dados

No artigo de Guadagni e Little (1983) [15] o modelo logito multinomial para escolha de marca será utilizado para analisar o comportamento de compra de café.

O objetivo é entender como preço , promoção e outras variáveis de marketing afetam as vendas e as participações desse produto no mercado. Para compreender esse comportamento é preciso um modelo capaz de avaliar toda categoria de produto considerada.

A participação de mercado é uma medida agregada das escolhas individuais do consumidor. Se for possível entender como e porquê os consumidores escolhem um produto a outro , será obtido um maior entendimento para o sucesso ou falha de um produto no mercado.

Com os dados de painel e o modelo logito surge a oportunidade de aumentar o conhecimento sobre o comportamento de compra em relação aos bens de consumo.

Ao ser aplicado o modelo , as alternativas serão os produtos mas, como considerar a agregação do produto e quais incluir no conjunto relevante no momento da análise não é uma decisão necessariamente óbvia para o pesquisador. Cada produto vai possuir uma particularidade que deve ser estudada detalhadamente para que se atinja o objetivo da pesquisa. Diferentes sabores, cores ou tamanhos podem algumas vezes ser tratados como diferentes produtos e algumas vezes , devem ser mantidos juntos.

Pode-se exemplificar com a categoria dos vegetais. Dentro dessa categoria a análise pode ser feita entre os vegetais naturais, congelados ou enlatados , ou então, pode-se focar a atenção na competição mais específica de marcas de ervilhas enlatadas.

Um outro aspecto importante é o grau de homogeneidade dos consumidores de uma população. É importante que sejam considerados consumidores idênticos no sentido de que os pesos das variáveis possam ser utilizados indiscriminadamente para todos eles. Isso não implica que todos tenham que se comportar da mesma maneira mas sim, que as diferenças devem ser levadas em consideração e , incorporadas ao modelo.

Nesse artigo o café foi escolhido por ser um produto com grande frequência de compra e também , por ser um produto que constantemente sofre variação de preço devido ao mercado de "commodities" tornando-se assim , um excelente objeto para estudo. As lojas também , costumam fazer muitas promoções de café e , como muitos produtos nos Estados Unidos , quando esse produto está em promoção em determinada loja os consumidores se deslocam a fim de comprá-lo com o melhor preço, fazendo muitas vezes, com que as lojas percam sua participação de mercado para seus concorrentes.

O mercado considerado nesse estudo é a cidade de Kansas nos Estados Unidos .

No momento da escolha do produto a ser avaliado foi feita uma seleção das alternativas baseado na pesquisa de Urban, Johnson, Brudnick e Hauser (1982) [31] sobre a estrutura do mercado de café. Eles dividem o mercado de café em 6 grupos de produtos definidos na Figura 3.1.

Figura 3.1 - Tipos de Café

Para obter um conjunto homogêneo de alternativas restringiu-se o estudo ao café moído cafeinado (1). Apenas essa linha em toda à categoria deu ensejo ao estudo de preferência por marca.

Um assunto importante a ser tratado para formação do banco de dados é a idéia de marca versus tamanho.

No mercado de Kansas City os tamanhos populares são de 1 libra e 3 libras que serão chamados respectivamente , de pequeno e grande. Foram

incluídos nos pequenos , tamanhos de 13 a 14 onças anunciados como capazes de produzir o mesmo número de xícaras que 1 libra de outro café.

Do ponto de vista do consumidor e do vendedor tem sido notado que produtos da mesma marca mas , com tamanhos diferentes são tratados como produtos diferentes. Em geral , o consumidor mostra lealdade distinta ao tamanho e , o vendedor promove os tamanhos separadamente. Isso naturalmente ocorre porque o comprador costuma associar os tamanhos e as quantidades às ocasiões e finalidades de consumo.

Por isso , nesse estudo marca significa a marca-tamanho de um produto ou seja, a cada marca será também associado um tamanho e, marcas iguais mas com tamanhos diferentes serão consideradas diferentes.

No mercado de Kansas City existem 10 marca-tamanho de café moído cafeinado mas, duas delas foram eliminadas da análise pois, representavam menos de 1% das compras no período considerado. Sobraram então, 8 marca-tamanho que possuem os maiores percentuais de vendas no mercado.

Tamanho	Marca				
pequeno	A ₁	B ₁	C ₁	D ₁	E ₁
grande	A ₂	B ₂	C ₂		

As observações foram realizadas em quatro supermercados selecionados na cidade de Kansas nos Estados Unidos.

Dois bancos de dados que se complementam foram utilizados para análise. Um deles é o banco de dados dos supermercados , onde estão contidas as informações sobre as promoções das mercadorias , posição na gôndola , se houve propaganda em algum meio de comunicação e o preço

da mercadoria no dia da compra. Essa informação é fornecida pelas lojas que foram consideradas e é o que se chama de dados estocados (*store data*).

O outro banco de dados utilizado é do painel de consumidores (*panel data*) que foi coletado pela empresa Selling Areas Marketing Inc. (SAMI) e , que foi cedido para pesquisa. Esse painel contém o histórico de compra de cada consumidor com a data que ele foi fazer a compra e as mercadorias adquiridas com todas as suas especificações como: marca , tamanho , preço e o código do consumidor.

Para coleta dos dados de painel a empresa possui um grupo de consumidores selecionados aleatoriamente da população e que se dispõe a participar do painel , cada indivíduo desse grupo vai possuir um código que permite sua identificação no momento da compra ; ao realizar uma compra o indivíduo deve apresentar o seu código ao caixa que vai registrá-lo e , automaticamente , toda informação sobre os produtos que estão sendo comprados passa a ser armazenada com isso, é formado um histórico longitudinal do consumidor.

Esses dados são coletados através de leitura óptica. Nos Estados Unidos os produtos possuem uma identificação denominada Universal Product Code (UPC) onde, estão contidas todas as informações sobre o produto, aqui no Brasil , este processo apenas está começando com a implementação dos códigos de barras em supermercados. O UPC funciona como uma impressão digital do produto , de maneira que um produto fica completamente identificado através de seu UPC e nenhum outro produto poderá confundir-se com ele. Esse código facilitou bastante o estudo uma vez que, esse produto foi identificado de forma única , independentemente do local onde as compras foram realizadas.

O painel de consumidores desse mercado possui aproximadamente 2000 indivíduos . Desses, 90% ou mais realiza suas compras em algum dos quatro supermercados que foram selecionados para essa pesquisa.

O estudo desenvolveu-se durante um período de 78 semanas , entre 14 de setembro de 1978 e 12 de março de 1980. Esse período foi dividido

em três sub-períodos distintos: o primeiro , de 26 semanas , que foi usado para inicializar variáveis que irão considerar a lealdade do consumidor à marca e ao tamanho ; o segundo , de 32 semanas , que foi usado para calibração do modelo e o terceiro , de 20 semanas , que serviu para validação do modelo.

O volume de dados seria muito grande se todos os indivíduos do painel fossem utilizados para modelagem. Por essa razão , foram adotados critérios para eliminar alguns indivíduos.

Foram eliminados primeiramente aqueles que não faziam compras com frequência em nenhum dos quatro supermercados considerados , depois foram eliminados os indivíduos que entraram no painel no meio do período utilizado para análise e, por fim , foram determinados aqueles indivíduos que realizaram menos de cinco (5) compras do produto durante o período , sendo considerados consumidores leves de café e portanto de pouco interesse para análise. Também foram eliminados os que não consumiam café.

Dos indivíduos restantes, foram selecionados 200 aleatoriamente. Um grupo de 100 fez parte da amostra de calibração e serviu para inicializar as variáveis de lealdade e outros 100 fizeram parte da amostra de validação reservada para testar o modelo final.

O grupo de calibração fez 718 compras no período que foi usado para inicialização das variáveis de lealdade e , 1037 compras , das marca-tamanho escolhidas , durante o período de 32 semanas usado para calibração. Dezeses (16) compras realizadas no período de calibração apresentam falta de informação nos dados ("missing data") da loja restando 1021 compras utilizáveis para modelagem.

O grupo de validação tem 921 compras nas 32 semanas correspondentes ao período de calibração e outras 666 no período de 20 semanas seguintes. Cinco (5) dessas compras tem dados incompletos, restando 1582 compras para o propósito de validação.

3.2.2 - As Variáveis de Modelo

A função utilidade do consumidor foi expressa como função linear das variáveis (atributos) como apresentado no capítulo 2.

$$U_{ij} = V_{ij} + \varepsilon_{ij}$$

onde, i representa o indivíduo, j representa a marca-tamanho do produto; V é a parte determinística da função utilidade e, ε é a parte aleatória.

E a parte determinística da função utilidade foi escrita como:

$$V_{ij} = \sum_k \beta_k x_{ikj}$$

Os atributos utilizados nesse modelo se dividem em dois grupos: um de atributos específicos para cada alternativa (T_j) e, um de atributos comuns para todas alternativas (T_C).

Assim, a parte determinística é escrita da seguinte forma:

$$V_{ij} = \sum_{k \in T_j} \beta_k x_{ikj} + \sum_{k \in T_C} \beta_k x_{ikj}$$

(T_j) Atributo específico para cada alternativa

O único atributo específico para cada alternativa prestou-se a captar qualquer particularidade que uma alternativa possuisse e que não fosse possível captar com outros atributos usados para descrever o comportamento de escolha médio das observações.

Portanto, a função utilidade para marca-tamanho incluiu uma constante única aditiva para cada uma das alternativas. Isso foi feito através de um conjunto de variáveis "dummy".

$$x'_{0j}(n) = \begin{cases} 1 & \text{para alternativa } j \\ 0 & \text{caso contrario} \end{cases}$$

Cada alternativa j tem seu próprio atributo exceto uma delas, digamos a n -ésima, que foi omitida para evitar a singularidade na estimação de máxima verossimilhança.

O valor β_{0n} foi tomado como 0 (zero). As constantes específicas marca-tamanho resultantes na função utilidade serão β_{0j} , $j=1,2,\dots,n-1$.

(TC) Atributos comuns a todas alternativas

Os principais atributos considerados foram: o preço real do produto (sem promoção), a presença ou ausência de promoção no momento da compra para marca-tamanho e, em caso de promoção, o valor do corte no preço (pode ser 0) durante a promoção. O primeiro desses atributos é o preço usual na ausência da promoção.

$x'_{1j}(n)$ - preço real (sem promoção) da marca-tamanho j no momento da n -ésima compra de café do consumidor i .

Se a marca-tamanho estava em promoção no momento de compra o seu preço real (sem promoção) é o preço antes de começar a promoção.

O segundo atributo é o que determina a existência de promoção:

$$x'_{2j}(n) = \begin{cases} 0 & \text{se } j \text{ está em promoção} \\ 1 & \text{caso contrário} \end{cases}$$

Nos dados da amostra que estavam disponíveis não havia qualquer informação de "display" e cartaz. Então, para obter a informação da existência de promoção foi usado um critério. Sabe-se que 3 eventos podem ser esperados acompanhando uma promoção: (1) redução do preço por curto período (de 1 a 3 semanas), (2) menção do item no jornal semanal da loja, (3) um movimento de vendas alto por pouco tempo, incomum para esse produto com vendas maiores que dois desvios acima da média do produto sem promoção.

Como qualquer um desses eventos sozinho pode ser um indicador falso de promoção, assumiu-se que uma promoção seria caracterizada somente se 2 ou mais desses eventos ocorressem em uma dada semana. Embora um indicador de movimento incomum pareça ser potencialmente uma explicação para compras notou-se que: (1) um movimento incomum é determinado pelos dados da loja e não pelos dados do painel. (2) os movimentos são grandes, tipicamente de 3 a 10 vezes o normal de vendas e são claramente um resultado de um fenômeno da loja e não do acaso. (3) um movimento incomum nunca define uma promoção por causa da regra de pelo menos 2 dos 3 eventos mencionados anteriormente ocorrerem.

A dificuldade em identificar uma promoção , nesse caso , ocorre devido a uma falha no banco de dados. Talvez , fosse possível através de um trabalho de campo determinar quando um produto esteve em um cartaz e quanto de "display" ele tinha disponível. Esse trabalho poderia dispende muito tempo por isso , a necessidade de procurar uma maneira para inferir a existência de promoção.

A variável utilizada para promoção (0-1) não foi capaz de captar toda variação ocorrida tais como , a diferença na resposta entre o "display" no fim do corredor e uma chamada no meio da prateleira.

No entanto, pareceu possível explicar um pouco mais da variação pelo acréscimo ao modelo de uma variável de desconto ou de redução do preço.

•

$x^i_{3j}(n)$ - redução do preço na marca-tamanho j no momento da n -ésima compra do consumidor i

Essa variável assume o valor zero quando não houver promoção de j . Então, o preço a ser pago pela marca-tamanho j é sempre $(x^i_{1j} + x^i_{3j})$.

Com o objetivo de verificar se uma compra de promoção é de alguma forma tão boa quanto uma compra comum, ou seja, se o indivíduo que trocou a marca quando ela estava em promoção irá voltar a comprar mais tarde no caso de uma compra sem promoção, foram introduzidas duas outras variáveis ao modelo:

$x^i_{4j}(n) = 1$ se a compra anterior do consumidor i foi uma compra promocional de uma alternativa com a mesma marca que a marca-tamanho j e, 0 caso contrário.

$x^i_{5j}(n) = 1$ se a segunda compra anterior do consumidor i foi uma compra promocional de uma alternativa com a mesma marca que a marca-tamanho j e, 0 caso contrário.

Esse conjunto de atributos está relacionado com a loja. Agora, é apresentado um conjunto de atributos relacionados com as características do consumidor chamadas de preferência, hábito ou lealdade que podem ser captadas através do comportamento passado do indivíduo.

Como observado antes, a heterogeneidade nesse caso, vai ser captada por essas variáveis.

A lealdade é utilizada para descrever a tendência do consumidor de recomprar a mesma marca-tamanho.

$x^i_{6j}(n)$ - lealdade à marca para marca-tamanho j para n -ésima compra do consumidor i .

Ela será definida por:

$$x'_{6j}(n) = \alpha_m x'_{6j}(n-1) + (1 - \alpha_m) \cdot \begin{cases} 1 \text{ se comprou marca } j \text{ na } (n-1) \text{ ocasião} \\ 0 \text{ caso contrário} \end{cases}$$

Operacionalmente, então, lealdade é tomada como sendo média ponderada exponencialmente das compras passadas da marca, tratadas como variáveis 0 - 1. A constante de persistência é α_m . Para inicializar a lealdade à marca faz-se $x'_{6j}(1) = \alpha_m$ se a marca j foi a primeira compra nos dados do consumidor i , caso contrário $(1 - \alpha_m) / (\text{número de marcas} - 1)$, no caso, igual a 7, assegurando, assim, que a soma das lealdades das marcas sempre seja igual a 1 para um consumidor.

A variável lealdade ao tamanho é análoga:

$x'_{7j}(n)$ - lealdade ao tamanho para a marca-tamanho j para a n -ésima compra de café do consumidor i .

$$x'_{7j}(n) = \alpha_t x'_{7j}(n-1) + (1 - \alpha_t) \cdot \begin{cases} 1 \text{ se comprou tamanho } j \text{ na } (n-1) \text{ ocasião} \\ 0 \text{ caso contrário} \end{cases}$$

α_t é a constante de persistência para tamanho onde, a inicialização é análoga a da lealdade à marca e a soma das lealdades para todos tamanhos é, da mesma forma, igual a 1.

Para todas variáveis, exceto a variável específica da alternativa, os coeficientes β_k são os mesmos para todas marca-tamanho. Então, ao calcular a utilidade do consumidor i os coeficientes são os mesmos para toda marca-tamanho com relação aos atributos preço, promoção e lealdade. Isso nos dá um modelo com poucos parâmetros, considerando o número de produtos listado. Usar o mesmo coeficiente entre marca-tamanho não implica as mesmas respostas para variáveis de controle. A resposta irá

dependem de todo o ambiente como visto no Capítulo 1 e das variáveis de lealdade do consumidor.

Usando os atributos listados a parte determinística da função utilidade pode ser escrita como:

$$V_{ij} = \beta_0 x_{i0j} + \beta_1 x_{i1j} + \beta_2 x_{i2j} + \beta_3 x_{i3j} + \beta_4 x_{i4j} + \beta_5 x_{i5j} + \beta_6 x_{i6j} + \beta_7 x_{i7j}$$

3.2.3 - Inicialização das Constantes para Variáveis de Lealdade

Setecentos e dezoito (718) compras realizadas nas 26 semanas iniciais foram utilizadas para inicializar as variáveis de lealdade. Cada compra foi tratada como uma observação de maneira que combinou-se dados longitudinais e dados de perfil (transversais).

Por isso pode-se considerar as variáveis de lealdade particularmente importantes, pois irão incorporar muito da heterogeneidade e uma boa parte da dinâmica longitudinal das compras.

Para obter as variáveis de lealdade foi preciso primeiro obter os valores de α_m e α_t que precisam ser uniformes. Uma maneira para se obter isso foi primeiro tomar valores experimentais para α_m e α_t e depois, refiná-los.

Valores aproximados de α (no caso, $\alpha_m = \alpha_t = 0.75$) foram usados e o modelo foi desenvolvido com todos outros atributos. Feito isso, tomou-se 10 ocasiões de compra e a variável lealdade à marca foi substituída por 10 variáveis "dummies" cada uma delas denotando se a marca foi ou não comprada na n -ésima ocasião de compra ($n=1,2,\dots,10$). Os coeficientes obtidos no modelo vão indicar o impacto da n -ésima ocasião de compra.

O modelo foi ajustado e os valores do coeficiente foram obtidos. Os coeficientes foram obtidos para cada ocasião de compra e depois foram normalizados de maneira a tornar o primeiro coeficiente igual a 1 (um). Para obter o valor de α_m , foi traçada uma função exponencial que melhor aproximava o comportamento dos coeficientes nas 10 ocasiões de compra tomadas e, α_m foi obtido de maneira que a diferença entre a exponencial traçada e os coeficientes tivessem média igual a zero. O mesmo procedimento foi utilizado para o caso do atributo lealdade ao tamanho.

Assim, no caso dos dados em questão os valores das constantes obtidos foram: $\alpha_m = 0.875$ e $\alpha_l = 0.812$.

3.2.4 - Resultados Obtidos

O modelo logito ajustado aos dados forneceu informações sobre uma série de questões em marketing. Na Figura 3.2 abaixo têm-se as estimativas dos parâmetros e as estatísticas t correspondentes entre parentêses. O modelo final é obtido fazendo a inclusão dos parâmetros para obter a contribuição de cada uma das especificações (M1 à M7)..

Figura 3.2 -Modelo Ajustado

	M1	M2	M3	M4	M5	M6	M7
Lealdade à Marca		2.78 (22.1)	3.47 (21.9)	3.47 (21.5)	3.79 (21.3)	3.92 (21.6)	
Lealdade ao Tamanho		2.12 (14.6)	2.74 (16.1)	2.72 (15.9)	2.74 (15.9)	2.97 (15.9)	
Promoção			2.22 (15.5)	2.00 (13.6)	2.07 (13.9)	2.11 (14.1)	1.40 (11.5)
Redução Preço			18.12 (5.5)	29.66 (7.2)	29.20 (7.1)	29.21 (7.1)	26.98 (8.0)
Preço Real				-26.36 (-6.0)	-26.49 (-5.9)	-29.94 (-6.6)	-28.02 (-7.8)
Compra ant. Promoção					-0.60 (-3.5)	-0.22 (-1.3)	0.62 (4.5)
2a Compra ant. Promoção					-0.72 (-3.9)	-0.46 (-2.47)	0.49 (3.3)
Cte. Marcas							
A1	0.28 (2.2)	-0.12 (-0.9)	-0.06 (-0.4)	0.03 (0.2)	0.01 (0.1)	-0.09 (-0.6)	0.48 (3.5)
A2	0	0	0	0	0	0	0
B1	0.48 (3.9)	0.28 (2.0)	0.11 (0.7)	0.19 (1.2)	0.20 (1.2)	0.10 (0.6)	0.41 (3.1)
B2	-0.06 (-0.4)	0.24 (1.6)	0.05 (0.3)	-0.06 (-0.4)	-0.05 (-0.3)	-0.08 (-0.5)	-0.35 (-2.3)
C1	0.99 (8,8)	0.14 (1.0)	0.31 (2.0)	0.57 (3.6)	0.62 (3.8)	0.41 (2.5)	1.49 (11.6)
C2	0.38 (3.1)	0.01 (0.1)	0.03 (0.2)	0.22 (1.3)	0.27 (1.6)	0.12 (0.7)	0.59 (4.2)
D1	-1.25 (-6.1)	-1.23 (-5.3)	-0.66 (-2.6)	0.15 (0.5)	0.10 (0.3)	-0.05 (-0.2)	0.10 (0.4)
E1	-1.83 (-6.8)	-1.52 (-5.2)	-0.76 (-2.3)	-1.5 (-4.1)	-1.47 (-4.1)	-1.72 (-4.7)	-1.88 (-6.3)
Log(n= 1021)	-1896	-1440	-1025	-1007	-1002	-977	-1488
U ^{2*}	0	0.24	0.46	0.47	0.47	0.48	0.22

Com o objetivo de observar a contribuição dos vários atributos e investigar a estabilidade dos coeficientes quando ocorre a mudança na especificação do modelo, o modelo final foi construído com a introdução dos atributos pouco a pouco e a avaliação da sua contribuição para o modelo é analisada toda vez. A quantidade de incerteza descrita pelo modelo pode ser avaliada através dos valores de U^2 que, é calculado como segue:

$$* \quad U^2 = 1 - \frac{L - \theta}{L(0)} \quad \text{onde, } L \text{ é o log da verossimilhança, } \theta \text{ é o}$$

número de parâmetros do modelo que está sendo estimado e $L(0)$ é o log da verossimilhança do modelo nulo.

A primeira especificação contém somente as variáveis "dummy" de marca-tamanho. O efeito dessa especificação foi fazer com que cada probabilidade de compra do indivíduo para uma marca-tamanho fosse a mesma e igual a participação dessa marca-tamanho no total das compras. É um modelo das participações das marcas no mercado. Esse foi escolhido como modelo nulo portanto, $U^2 = 0$.

Os primeiros atributos introduzidos no modelo foram os que representam a lealdade à marca e ao tamanho, como pode ser observado existe um grande acréscimo no valor de U^2 que passa a ser igual a 0.24. Isso mostra que, como esperado, as variáveis lealdade explicam muito do comportamento entre os indivíduos.

Na terceira especificação os atributos de promoção foram introduzidos e um outro grande salto é observado em U^2 que passa a ser igual a 0.46. Ao ser introduzido o atributo preço real do produto, o valor de $U^2=0.47$ apresenta um aumento bem pequeno. Isso mostra claramente que o preço real do produto não é um fator que caracteriza a mudança das participações. Embora o acréscimo de U^2 seja pequeno, foi realizado um teste de Qui-quadrado e mostrou-se que a diferença entre M3 e M4 é significativa a nível 0.01.

Na especificação M5 os atributos relacionados a compra promocional anterior foram introduzidos e nesse caso o valor de U^2 permaneceu o mesmo.

O último passo a ser realizado foi a atualização das variáveis α_m e α_r como mencionado anteriormente . O resultado foi a especificação M6 que foi tomada como a especificação final com valor de $U^2 = 0.48$ indicando uma melhora nas funções de lealdade como resultado da interação.

Pode-se notar ainda na especificação final que o atributo compra anterior promocional não tem valor da estatística t significativa mas, mesmo assim ele foi mantido pois , o atributo segunda compra anterior promocional foi significativa.

As constantes de marca-tamanho seriam zero se fossem obtidas variáveis que explicassem todas diferenças entre marca e tamanho. Obviamente que isso seria difícil de obter e nesse modelo isso não foi alcançado mas , se forem observadas as especificações de M1 a M6 verifica-se que os valores das constantes foram diminuindo , bem como as estatísticas t.

Todos os outros coeficientes uma vez introduzidos no modelo parecem ter se mantidos estáveis excepcionalmente , o coeficiente do preço promocional que foi mais afetado depois da introdução do preço regular. Mesmo assim , isso parece razoável uma vez que o preço promocional é baseado no preço regular .

Uma questão importante é se o modelo de nível individual realmente oferece alguma vantagem substancial em relação ao modelo de nível agregado . Então , na especificação M7 os atributos relacionados ao comportamento específico (lealdade à marca e ao tamanho) são retiradas do modelo e permanecem todos os outros. Nota-se com isso uma queda no valor de U^2 que passa a ser igual a 0.22.

É interessante analisar esse modelo para se avaliar a importância das variáveis de lealdade. A simples mudança no valor de U^2 demonstra claramente a força dessas variáveis. Pode-se notar que os coeficientes dos atributos preço, promoção e redução no preço permaneceram como antes. Caso isso fosse diferente, as variáveis desagregadas explicariam mais incerteza e assim, seriam preferidas a outras. O atributo promoção anterior tem o sinal de seu coeficiente modificado e se torna estatisticamente significativo. Esse fato não é surpreendente caso se observe que esse é agora o atributo mais próximo no modelo ao que se chamava variável de lealdade.

Pode-se notar também um aumento nas constantes relacionadas a marca-tamanho. Isso pode ser justificado talvez por estarem agora com uma parte do peso anteriormente colocado nas variáveis lealdade.

Em resumo, pode-se dizer que a saída das variáveis de lealdade claramente deterioram o modelo, justificando então o uso de um modelo desagregado.

3.2.5.- Discussão dos Coeficientes

Todos os coeficientes que aparecerem na função utilidade estavam com o sinal que seria esperado.

A importância relativa dos atributos para explicar o comportamento do mercado é de grande interesse. A magnitude dos coeficientes não é uma boa medida. Cada um aparece com uma unidade diferente e, então os melhores indicadores a serem utilizados são o valor de U^2 e o valor da estatística t .

Usando então, esses indicadores verifica-se que os atributos mais importantes são lealdade à marca, seguido de lealdade ao tamanho. Lealdade à marca possui maior coeficiente e maior valor de t que lealdade ao tamanho. No entanto, pode-se dizer que lealdade ao tamanho também é de grande importância e, provavelmente de importância muito maior que os fabricantes e revendedores imaginam.

O outro atributo que segue é a promoção , indicando que a atenção dedicada pelos fabricantes e revendedores às promoções não é inútil. Depois , em grau de importância surgem os dois atributos de preço e finalmente , o atributo de compra promocional anterior.

As constantes de marca-tamanho formam um grupo especial. Elas podem representar uma qualidade única do produto e/ou erro de especificação. Como mencionado anteriormente , se os atributos utilizados estivessem explicando "bastante coisa" os valores desses coeficientes seriam muito próximos de zero.

O valor de A_2 foi assumido como sendo igual a zero por definição. A maioria dos outros coeficientes obtidos tem valor menor que zero. Pode-se comparar as constantes marca-tamanho com a participação de mercado de cada uma das marca-tamanho que representam.

	C_1	B_1	C_2	A_1	A_2	B_2	D_1	E_1
Mercado	28.4	17.1	15.5	14.0	10.6	9.8	3.0	1.6
Constante	0.41	0.10	0.12	-0.09	0	-0.08	-0.05	-1.72

Talvez exista uma relação fraca embora , somente C_1 e E_1 tenham estatísticas t significantes. Poder-se-ia ainda pensar que tenha faltado um atributo nesse modelo e , um candidato natural seria a qualidade do "display" durante a promoção. Mas , por falta de informações no banco de dados , não é possível acrescentar esse atributo ao modelo.

E_1 sobressai bastante e , de fato ao analisar-se o produto verifica-se que se trata de um produto diferente dos outros e com características únicas.

Promoção é uma variável de controle bem interessante. Não somente seu efeito é grande , mas normalmente , vem acompanhada por uma redução de preço da mercadoria o que aumenta seu impacto total.

Alguns pesquisadores acreditam que o efeito das promoções pode ser inteiramente devido ao preço mas , essa pesquisa não deu suporte a essa hipótese. Mesmo com uma redução no preço igual a zero, uma grande probabilidade de ganho é calculada. Esse resultado foi consistente com outras pesquisas que dizem que o "display" desempenha um forte papel na resposta de compra do consumidor.

É muito importante saber se um produto que foi colocado em promoção vai ser comprado novamente ou se a compra se realizou unicamente devido a promoção. O modelo que foi proposto permite verificar diretamente esse efeito. Uma compra promocional vai afetar compras subsequentes através dos atributos lealdade à marca , lealdade ao tamanho e compra anterior promocional. Então pode-se calcular a contribuição da compra para a função utilidade no próximo período (serão usados os coeficientes de M6 os valores de α calculados anteriormente e as definições de x_6 e x_7) :

$$(3.92)(0.125)(1) + (2.97)(0.188)(1) = 1.05$$

$$\beta_6(1 - \alpha_m)(1) + \beta_7(1 - \alpha_t)(1)$$

Essa contribuição para função utilidade diminui com o atributo compra anterior promocional:

$$(-0.22)(1) = -0.22$$

$$\beta_4(1)$$

Isso leva a uma contribuição relativa de 0.83 em uma situação onde não houve a compra da marca-tamanho . Continuando com o processo para a próxima compra , a contribuição da compra original para utilidade através da variável lealdade é:

$$(3.92)(0.875)(0.125)(1) + (2.97)(0.812)(0.188)(1) = 0.88$$

$$\beta_6(\alpha_m)(1 - \alpha_m)(1) + \beta_7(\alpha_t)(1 - \alpha_t)(1)$$

Ela foi diminuída pelo atributo de segunda compra anterior promocional :

$$(-0.46)(1) = -0.46$$

$$\beta_5(1)$$

A contribuição relativa final foi igual a 0.42. O efeito negativo da compra anterior promocional vem em concordância com pesquisas anteriores mas , a contribuição nos dois períodos permanece positiva.

Em outras palavras , o cálculo diz que o consumidor está mais propenso a comprar uma marca-tamanho promovida se ela não tiver sido comprada uma primeira vez . Mas , está menos propenso a comprar uma marca-tamanho promovida se tiver tido uma compra não promocional .

3.2.6. - Validação do Modelo

Para fazer a validação do modelo foi usada uma amostra de 100 consumidores durante um período de 32 semanas correspondente ao período de calibração e mais , 20 semanas que corresponde ao período de validação.

Foram usados dois períodos com objetivo de verificar como o modelo se comporta ao serem avaliados outros consumidores no mesmo período em que foi feita a sua calibração e o que ocorre no período subsequente ao da calibração.

Uma maneira de se avaliar é calcular as participações previstas das marcas no mercado e verificar as participações reais. Ou seja , avaliar a capacidade preditiva do modelo.

Com esse objetivo , os pesquisadores criaram medidas para calcular a participação de mercado de cada uma das marca-tamanho em questão e também criaram uma medida de dispersão.

Através desses valores foram construídos gráficos que fazem um rastreamento das participações de mercado e , que mostram como seria o comportamento previsto pelo modelo e como foi o comportamento real. Os resultados obtidos foram bastante satisfatórios , mostrando que o modelo segue as mudanças e tendências das participações de mercado razoavelmente bem.

Para o período de validação , as variáveis de lealdade à marca e lealdade ao tamanho , que dependem do comportamento passado , foram atualizadas utilizando o método de Monte Carlo. Foi construído um histórico de compra para cada consumidor.

No período usado para validação foram detectados alguns valores fora das margens do intervalo de confiança de 90%. Isso pode sugerir que o modelo ainda não seja capaz de captar todos os fenômenos e que , como já mencionado anteriormente a ausência de determinadas variáveis supostamente relevantes à explicação do comportamento do consumidor , tenham reduzido o poder explicativo do modelo.

Enfim , pode-se concluir que o modelo obtido foi capaz de descrever o comportamento de compra do consumidor de café cafeinado. Deve-se também ressaltar , que esse modelo obteve bons resultados principalmente pela riqueza de detalhes no banco de dados que foi utilizado , muito embora o mesmo ainda apresente-se com informações não completas e/ou faltantes.

3.27 - Resposta do Mercado para Variáveis de Controle

Um uso prático do modelo requer medidas que mostrem aos pesquisadores como o mercado responde aos esforços de marketing dos revendedores.

O modelo obtido é um modelo de nível individual mas, as respostas agregadas do mercado são de maior interesse para os decisores de marketing e para serem obtidas necessitam de um maior esforço computacional.

Antes de mais nada , é importante ressaltar que , embora os coeficientes sejam os mesmos para todos os consumidores , as respostas individuais irão mudar bastante , dependendo do ambiente da loja e das lealdades a "priori".

Isso pode ser ilustrado , mostrando um mercado hipotético composto de dois produtos onde, um consumidor com lealdade igual para os dois produtos possui uma elasticidade de preço de participação igual a -2.2 enquanto, um consumidor com lealdade 0.8 para um produto e 0.2 para o outro tem uma elasticidade de -0.06 para o primeiro.(ver figura 3.3)

Figura 3.3. -Mercado Hipotético

Mercado Hipotético com dois produtos A e B com coeficiente de lealdade à marca iguais a 7.0 e coeficiente de preço igual a -30 oz/dol . Não são inclusas outros atributos no modelo.

	Situação 1	Situação 2
Lealdade Produto A	0.5	0.8
Lealdade Produto B	0.5	0.2
Preço Produto A	0.15 dol/oz	0.15 dol/oz
Preço Produto B	0.15 dol/oz	0.15 dol/oz
Probabilidade de A *	0.5	0.985
Probabilidade de A **	0.61	0.991
Elasticidade de Preço	-2.2	-0.06

** cálculo da probabilidade com 10% redução no preço de A = 0.135 dol/oz

Analogamente, embora as variáveis de controle no modelo de oito produtos tenham o mesmo coeficiente para todas as marca-tamanho as

respostas do mercado , para um produto particular irão depender das ações das outras marca-tamanho e distribuição de lealdade entre os consumidores.

Analiticamente , a determinação da resposta agregada do consumidor é uma conjugação da função resposta do consumidor e da distribuição conjunta da lealdade do consumidor, preço e promoção para todas as marcas.

Uma abordagem padrão para esse tipo de procedimento é a utilização da simulação pelo método de Monte Carlo. Uma outra maneira mais simples é usar os consumidores reais, períodos de tempo e atributos das variáveis dos dados. Mudando-se ,digamos, o preço real de uma marca-tamanho de 1% durante todo o período e observando as correspondentes mudanças na participação de mercado, é possível obter uma resposta de participação agregada para preço real.

Muitos cálculos diferentes de respostas do mercado são possíveis. Alguns exemplos são : (1) elasticidade de preço regular ; (2) Elasticidade cruzada de participação de uma marca-tamanho e o preço de outra ; (3) Resposta a promoção através de participação de mercado ; (4) Canibalização de um tamanho de uma marca pela promoção de outra e ainda , (5) Elasticidade de preço durante a promoção.

Cada um deles pode ser avaliado como um efeito de curto prazo; isto é, como uma alteração das probabilidades de escolha na ocasião de compra na qual a variável é mudada ou , como um efeito de longo prazo, levando também em consideração o impacto futuro da mudança de lealdade. Um exemplo será apresentado restringindo-se ao efeito de curto prazo para elasticidade de preço regular. Os cálculos foram feitos da seguinte maneira:

(a) Elasticidade do preço regular : Tomou-se o período de calibração de 32 semanas e os consumidores que foram avaliados neste período. Foi feita a mudança no preço regular de uma das marcas-tamanho , acrescentando 1% enquanto as outras foram mantidas

constantes. Com isso uma nova probabilidade de escolha foi obtida e o cálculo da participação de mercado também foi feito. A porcentagem de mudança na participação de compra dividida pela atual mudança no preço foi tomada como a elasticidade do preço regular a curto prazo.

Os resultados são apresentados na Figura 3.4.

Figura 3.4 - Respostas de Mercado a Curto Prazo

	% mercado	(a)
C ₁	28.4	-1.9
B ₁	17.1	-2.3
C ₂	15.5	-2.3
A ₁	14.0	-2.5
A ₂	10.6	-2.7
B ₂	9.8	-2.6
D ₁	3.0	-3.4
E ₁	1.6	-2.4

3.2.8 - Discussão dos Resultados

A variável resposta oscila bastante para cada marca tamanho. Isso era esperado embora todas marca-tamanho usem os mesmos coeficientes. Pode-se atribuir essa variação ao fato de que no cálculo da variável resposta foi levado em consideração todo ambiente competitivo e as lealdades ao produto. Essas variações na resposta tem grande significado para fabricantes e vendedores. Por exemplo , a diferença nas elasticidades de preço sugerem diferença nas políticas de preço entre as marca-tamanho.

Na Figura 3.4 as marca-tamanho foram listadas em ordem decrescente de participação de mercado. Pode-se observar que as magnitudes das respostam tendem a crescer a medida que decresce a participação.

Marcas com maior participação são menos sensíveis à variável de controle que as marcas de menor participação. Se ainda , for multiplicada cada participação de mercado pela variável resposta , obtem-se o incremento absoluto da participação para mudança na variável controle e pode-se observar que as marcas com maior participação obtém um incremento maior que as marcas de menor participação no mercado.

Uma observação final sobre a variável resposta diz respeito à abrangência do modelo. Ele dá informações não somente sobre uma marca-tamanho mas , sobre todo o mercado. Isso permite um estudo comparativo das estratégias assumidas pelos concorrentes , em retaliação aos esforços de marketing realizados por um produtor/revendedor específico.

3.2.9 - Conclusões

O modelo logito multinomial permitiu uma excelente representação do mercado de café moido cafeinado para o consumidor e a nível de revendedores. Os coeficientes do modelo foram estatisticamente significantes. O modelo foi capaz de prever razoavelmente bem , no período usado para validação.

É interessante verificar como a manipulação dos dados permite avaliar as variáveis de resposta agregadas para o mercado.

Essa aplicação foi baseada em um banco muito rico em informações. À medida que a coleta de dados fica mais abrangente , torna-se mais fácil melhorar os modelos para avaliação do comportamento de compra do consumidor.

Vamos apresentar agora a utilização do modelo logito multinomial para o consumo de bolachas salgadas do artigo de Chintagunta , Jain e Vilcassim (1992) [9] "*Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data*".

3.3 - Aplicação 2

3.3.1 - Objetivos

Nesse estudo os autores propõe uma abordagem semi-paramétrica para tratar a heterogeneidade entre indivíduos. A teoria desse tratamento foi mostrada no capítulo 2.

Então , o objetivo principal desse artigo foi mostrar que a abordagem proposta produz resultados melhores que as abordagens até então conhecidas. Entre as abordagens analisadas está a apresentada no item 3.2 dessa dissertação , referente ao artigo de Guadagni e Little (1983) [15] , onde , a heterogeneidade foi captada através dos atributos de lealdade.

É apresentada ainda, uma maneira de fazer comparação entre os modelos e, após essa comparação entre os modelos propostos é feita uma análise das variáveis de mercado.

O procedimento para estimação desse modelo é semelhante ao modelo anterior , com exceção das variáveis de lealdade pois , nesse caso elas não existem.

3.3.2 - Formação do Banco de Dados

O banco de dados final nesse caso , também é a composição de dois bancos de dados. Um deles contém a informação fornecida pela loja sobre promoções , "display" e propaganda (*store data*) e o outro , é um painel de consumidores que foi fornecido para pesquisa pela empresa IRI (Information Resources , Inc.).

A categoria de produtos que foi analisada é a de bolachas salgadas tipo "cream cracker", dentro dessa categoria selecionaram-se as quatro marcas de maior participação no mercado : Sunshine , Keebler , Nabisco e a Marca Genérica. A Marca Genérica é uma marca oferecida pelos estabelecimentos , no Brasil , a marca Carrefour é um exemplo dessa

marca Genérica. Entre as marcas escolhidas tomou-se o tamanho de 16 onças que é o comprado com mais frequência.

Dos consumidores que participam do painel da IRI foram selecionados aleatoriamente 135 participantes.

Essa amostra foi dividida para ser utilizada da seguinte forma: setenta (70) participantes fizeram parte da amostra usada para calibração , onde foram definidos os coeficientes do modelo e , sessenta e cinco (65) participantes fizeram parte da amostra de validação que foi usada para testar o modelo obtido anteriormente. O número de compras realizadas pela amostra de calibração é 1736 e pela amostra de validação é 1542.

Os consumidores selecionados não realizaram nenhuma compra múltipla no período ou seja , não compraram mais de uma marca em uma mesma ocasião de compra.

3.3.3 - As Variáveis do Modelo

Os atributos do composto de marketing : preço , "display" e propaganda foram selecionados para fazer parte da função utilidade e formar o modelo de compra.

O atributo preço foi medido como o preço no momento em que o consumidor realizou a compra.

Para representar os atributos "display" especial e propaganda foram usadas variáveis "dummy" .

Esses são os atributos comuns a todas marcas.

Temos ainda , as constantes específicas de cada marca que , como no artigo anterior , procuram captar qualquer diferença que os atributos comuns não tenham captado. São quatro constantes específicas de marca

sendo que , a constante que corresponde a Marca Genérica é igual a zero para evitar a singularidade.

3.3.4 -Resultados Obtidos

Como o objetivo do artigo foi comparar resultados de diversos tratamentos de heterogeneidade foram feitas as estimações dos parâmetros para cada um dos modelos que se desejava comparar e os resultados são apresentados na Figura 3.5.

A Figura vai apresentar as estimativas dos parâmetros e seus desvios entre parentêses nos cinco (5) casos .

Os cinco (5) modelos considerados são os seguintes: (a) modelo de Guadagni e Little (1983) [15] onde a heterogeneidade foi captada através das variáveis de lealdade (no caso , só lealdade à marca) ; (b) modelo proposto por Krishnamurthi e Raj (1988) [20] , esse modelo difere do anterior por considerar que a marca possui seu próprio efeito de preço e estima separadamente coeficiente de preço para cada marca considerada e , ainda, a variável lealdade à marca foi calculada de maneira diferente : em cada ocasião de compra o percentual de vezes que o participante do painel escolhe cada marca nas ocasiões de compra anteriores foi calculada. Isso fornece o histórico de participação de compras de cada marca e um índice foi calculado baseado nessa participação ; (c) o terceiro modelo considerou a distribuição Gamma como sendo a distribuição de probabilidade para tratar a heterogeneidade entre os participantes do painel ; (d) o quarto modelo considerou a distribuição Normal e por fim , (e) o modelo onde a distribuição da heterogeneidade entre os participantes do painel não foi definida e trabalhava-se com uma abordagem semi-paramétrica.

Figura 3.5 - Parâmetros Estimados e Erros Padrão

Variável	(a)	(b)	(c)	(d)	(e)
"display"	0.366 (0.128)	0.376 (0.132)	0.384 (0.177)	0.497 (0.193)	0.546 (0.183)
Propaganda	1.071 (0.190)	1.048 (0.191)	0.883 (0.127)	0.974 (0.140)	0.903 (0.159)
Preço	-3.964 (0.414)	-4.085 (0.409)	-4.276 (0.094)	-5.131 (0.088)	-5.630 (0.183)
Lealdade*	3.570 (0.115)	3.532 (0.115)			
Cte Sunshine	0.424 (0.165)	0.475 (0.167)	-0.902 (0.152)	-0.746 (0.177)	-0.328 (0.160)
Cte Keebler	1.148 (0.217)	1.223 (0.215)	1.059 (0.142)	0.816 (0.142)	0.933 (0.192)
Cte Nabisco	1.640 (0.193)	1.772 (0.192)	1.012 (0.178)	2.232 (0.201)	4.527 (0.211)
Log	-878.42	-892.30	-960.01	-876.68	-838.49
g.l.	1729	1729	1726	1726	1710
U ² *	0.533	0.524	0.488	0.532	0.546

*Foram usadas 5 observações para calcular a variável lealdade

*A normalização na estimação é feita em relação a cte Marca Genérica

*U² calculado como no item 3.2.4

A primeira vista, os resultados da estimação parecem ser similares. Por exemplo, todos apresentaram resultados que tem validade (i.e., os valores dos parâmetros estimados para variáveis de marketing foram todas significantes a 5% e os sinais foram como esperado). Existem no entanto, algumas diferenças importantes.

A primeira diferença pode ser percebida em termos de ajuste geral. Observem que na Figura 3.5 a especificação de efeitos aleatórios semi-paramétrica teve a maior medida do log da verossimilhança (-838.49) e o maior valor de U^2 (0.546). Também as magnitudes dos parâmetros estimados variaram através das diferentes especificações. A maior variação apareceu no efeito estimado da variável preço.

Essas diferenças sugerem a necessidade de um teste estatístico formal para escolha da especificação que seja mais consistente com os dados.

Ben-Akiva e Lerman (1985) [4] sugerem o uso de um teste baseado nos valores calculados de U^2 . Nesse teste um modelo é tratado como "nulo" e os outros modelos são testados contra o modelo nulo. O teste estatístico é dado por:

$$P(U^2_2 - U^2_1 > \tau) \leq \Phi\{-[-2\tau L(0) + (\eta_2 - \eta_1)]^2\} \quad \tau \geq 0$$

onde, η_1 e η_2 são o número de parâmetros estimados para os dois modelos que serão testados e Φ denota a função de distribuição normal acumulada. A expressão a direita, é o limite superior para probabilidade que a diferença entre o U^2 do modelo nulo e do modelo alternativo seja maior que τ .

Nota-se na Figura 3.5 que das especificações apresentadas as que possuem maior valor de U^2 foram, respectivamente, a especificação semi-paramétrica e a de Guadagni e Little (1983) [15]. Então, verificou-se quais dessas duas especificações era a mais adequada aos dados.

O objetivo é saber se a diferença de 0.013 nos valores de U^2 dos dois modelos foi significativa estatisticamente. Substituindo os valores obtem-se:

$$P(U^2_2 - U^2_1 > 0.013) \leq \Phi\{-8.02\}$$

A probabilidade que essa diferença tenha ocorrido por acaso é menor que 0.001. Então pode-se concluir que a especificação semi-paramétrica forneceu o melhor ajuste dos dados.

A solução apresentada anteriormente foi obtida através de um teste estatístico. Pode-se ainda utilizar uma outra maneira para avaliar os modelos e verificar sua capacidade de previsão através da amostra de validação. Com esse objetivo foi realizada a validação preditiva:

Primeiro, foram comparadas as previsões para toda amostra de validação. Depois, foi feita a comparação dividindo a amostra de validação em 2 partes: uma caracterizada pelo auto grau de compras repetidas e outra pelo auto grau de troca de marca.

Como observado anteriormente a especificação de efeitos aleatórios semi-paramétrica foi o modelo de melhor ajuste das 3 abordagens de efeitos aleatórios para considerar heterogeneidade. Então, esse modelo foi comparado as outras 2 especificações que usam o comportamento de escolha passada para considerar heterogeneidade. Os resultados da validação preditiva são apresentados na Figura 3.6.

Para realizar a comparação foi utilizada a participação de cada marca no mercado. Primeiro, foi verificada a participação real no mercado de cada marca e então, foi feito o cálculo da participação obtida utilizando os modelos propostos.

Figura 3.6 - Resultado da validação do modelo

Marca	Real	(e)	(a)	(b)
Leais				
Sunshine	0.02	0.04	0.03	0.03
Keebler	0.04	0.05	0.03	0.03
Nabisco	0.65	0.64	0.63	0.63
Particular	0.29	0.27	0.31	0.31
R.E.Q.M.	-	0.03	0.03	0.03
%acerto		80%	81%	81%
Não Leais				
Sunshine	0.20	0.18	0.14	0.15
Keebler	0.10	0.12	0.11	0.11
Nabisco	0.38	0.37	0.42	0.41
Particular	0.32	0.33	0.33	0.33
R.E.Q.M.	-	0.03	0.07	0.06
%acerto	-	74%	48%	50%
Amostra Tot				
Sunshine	0.08	0.08	0.06	0.07
Keebler	0.06	0.07	0.06	0.05
Nabisco	0.56	0.56	0.56	0.56
Particular	0.30	0.29	0.32	0.32
R.E.Q.M.	-	0.01	0.03	0.03
%acerto	-	78%	71%	71%

*R.E.Q.M. - Raiz do Erro Quadrático Médio

*(a) modelo de Guadagni e Little , (b) modelo de Krishnamurth e Raj , (e) modelo semi-paramétrico

Através da Figura 3.6 verifica-se que o modelo semi-paramétrico prevê melhor que o modelo proposto por Guadagni e Little (1983) [15] pois , as porcentagens de acerto foram de 78% e 71% respectivamente , no caso da amostra total.

Para a amostra dividida como mencionado anteriormente observa-se que no caso de consumidores leais à marca (ou seja, que realizam pouca troca de marca) os dois modelos se comportam de maneira muito similar, 80% e 81% de acerto, respectivamente. Uma diferença significativa foi observada no caso em que a amostra é de consumidores não leais à marcas (ou seja, que realizam diversas trocas de marcas durante o período em estudo), nesse caso, a especificação semi-paramétrica previu muito melhor do que o modelo de Guadagni e Little, 74% e 48% de acerto, respectivamente.

Uma explicação para esse resultado é que em Guadagni e Little os parâmetros de lealdade são uma função do preço e da atividade promocional das marcas compradas pelo consumidor. Os consumidores sensíveis a promoção e, com lealdade moderada frequentemente mudam a marca quando uma delas está em promoção. Então, a medida de lealdade baseada nas compras passadas pode se confundir com a atividade promocional para esses consumidores.

Dadas as previsões dos modelos na Figura 3.6 pode-se concluir que de maneira geral a abordagem de efeitos aleatórios semi-paramétrica seria a preferida.

3.3.5 - Resultados Agregados

Definido um modelo, deseja-se mostrar a implicação dos resultados para o marketing. A discussão será centrada nas elasticidades de escolha de marca estimadas e na distribuição de heterogeneidade das preferências de marca.

As elasticidades são medidas de grande importância em marketing. Elas estão relacionadas com o quanto a oscilação de uma variável vai afetar o comportamento de outra. Uma das medidas de elasticidade mais utilizadas é a de preço.

Pode-se exemplificar o cálculo com a elasticidade de preço-demanda:

$$\text{Elasticidade} = \frac{\% \text{variação de quant. demandada}}{\% \text{variação do preço}}$$

$$E = \frac{\Delta Q}{Q} \div \frac{\Delta P}{P}$$

Para variáveis "dummy" a elasticidade representa a mudança relativa na probabilidade devido a variável analisada.

As elasticidades cruzadas e próprias são apresentadas na Figura 3.7 a seguir, para o modelo de efeitos aleatórios semi-paramétrico.

Figura 3.7 - Elasticidades de escolha para modelo de efeitos aleatórios semi-paramétrico

Mudança na	Mudança na probabilidade de escolha			
	Sunshine	Keebler	Nabisco	Particular
Elasticidade Preço				
Sunshine	-3.07	0.53	0.29	0.19
Keebler	0.46	-2.55	0.23	0.15
Nabisco	2.08	1.83	-1.02	0.97
Particular	0.60	0.55	0.45	-1.33
Elasticidade Propaganda				
Sunshine	0.64	-0.10	-0.06	-0.04
Keebler	-0.10	0.58	-0.05	-0.03
Nabisco	-0.31	-0.28	0.17	-0.17
Particular	-0.11	-0.10	-0.08	0.24
Elasticidade "Display"				
Sunshine	0.37	-0.06	-0.03	-0.02
Keebler	-0.06	0.33	-0.03	-0.02
Nabisco	-0.19	-0.17	0.11	-0.10
Particular	-0.07	-0.06	-0.05	0.14

* tabela com valores percentuais

Examinando as elasticidades próprias de preço na Figura 3.7 verifica-se que a marca Sunshine tem a maior elasticidade própria (-3.07), seguida pela Keebler (-2.55), marca genérica (-1.33) e por fim , Nabisco (-1.02). Pode-se verificar também , que as elasticidades próprias de propaganda foram todas maiores que as correspondentes elasticidades de "display". Isso pode ser interpretado da seguinte maneira: a presença de uma propaganda para Sunshine aumenta sua probabilidade de escolha em 0.64 enquanto que o "display" especial aumenta 0.37. Então, uma propaganda parece ser mais efetiva na influência da decisão de escolha de marca que um "display" especial, embora um "display" especial seja usado com mais frequência que propaganda nessa categoria de produto avaliada.

Observando as elasticidades cruzadas , pode-se notar que quando uma unidade de preço da marca Sunshine diminui de 1% a probabilidade de escolher Nabisco diminui 0.29%, e quando o preço da Nabisco diminui de 1% a probabilidade de escolher Sunshine diminui de 2.08%.

Da mesma maneira, quando Sunshine tem alguma propaganda, a probabilidade de escolher Nabisco diminui de 0.06%, enquanto se a Nabisco tem propaganda a probabilidade de escolher Sunshine diminui 0.31%. Esse comportamento assimétrico de competição entre marcas geralmente foi consistente com as descobertas de Blattberg e Wisniewski (1989) [5] de que a competição entre marcas de diferentes preços é assimétrica

3.3.6 - Heterogenidade na Preferência de Marca

Para especificação de efeitos aleatórios semi-paramétrica verificou-se que três pontos suporte foram suficientes para aproximar a distribuição de probabilidade das preferências intrínsecas para todas quatro marcas.

Os pontos suporte estimados e as probabilidades de massa associadas são dadas na Figura 3.8.

Através dessa Figura pode-se verificar como as distribuições de preferência de cada marca se comportam e obter maior clareza nos resultados obtidos.

Figura 3.8 - Distribuição das Preferências Intrínsecas

Marca	Primeiro		Segundo		Terceiro	
	Ponto	Prob.	Ponto	Prob.	Ponto	Prob.
Sunshine	-7.13	0.20	0.81	0.35	2.54	0.45
Keebler	-3.59	0.28	0.78	0.58	3.94	0.14
Nabisco	-2.60	0.27	-0.44	0.27	1.79	0.46
Genérica	-3.68	0.39	0.50	0.36	5.02	0.25

Fica claro da Figura 3.8 que as distribuições foram diferentes para as marcas que estavam sendo consideradas nessa categoria de produtos. Essa descoberta implica que as preferências intrínsecas são específicas de marca e consumidor e então não seria apropriado impor uma distribuição comum para todas as marcas. Essa é a vantagem do uso de uma abordagem flexível tal como a especificação semi-paramétrica.

É possível também observar que , embora a distribuição de preferência para marca Keebler seja de alguma maneira simétrica as outras três distribuições parecem ter padrão mais monotônico. E mais, examinando as variâncias das diferentes distribuições , nota-se que a marca Sunshine teve a maior heterogeneidade nas preferências intrínsecas (13.3) seguida pela marca genérica (11.7), Keebler (6.1) e Nabisco (3.4). Então, a Nabisco teve a menor heterogeneidade em preferências entre os consumidores considerados .

A média da preferência intrínseca para cada marca é dada pela constante específica da marca. Da Figura 3.5 observa-se que a maior média

de preferência foi dada pela Nabisco (4.527), seguida pela Keebler (0.933), Marca Genérica (0) e Sunshine (-0.328) é interessante notar que a ordem das médias das preferências intrínsecas foi inversa a variância das distribuições de heterogeneidade para as quatro marcas.

Um assunto interessante para marketing é a interpretação da distribuição das preferências intrínsecas entre os consumidores. Para melhor compreender esse assunto pode-se primeiro procurar compreender o significado dos pontos suporte e das probabilidades de massa associada.

Considere por exemplo, a distribuição de preferências para marca Keebler . A interpretação seria que qualquer consumidor selecionado aleatoriamente teria um valor de preferência intrínseca de -3.59 com probabilidade de 0.28 , um valor de 0.78 com probabilidade 0.58 e um valor de 3.94 com probabilidade de 0.14. Ou ainda , que 28% dos consumidores teriam uma preferência intrínseca de -3.59 , 58% de 0.78 e finalmente 14% com valor de 3.94 para marca Keebler. Interpretações similares são validas para as outras marcas.

Nota-se que os pontos e as probabilidades de massa associadas foram diferentes para todas as marcas. Então , a especificação semi-paramétrica permite fazer inferências sobre o agrupamento das preferências dos consumidores (a localização e o tamanho relativo dos grupos) para cada marca.

3.3.7 - Conclusões

Fica mostrado nesse artigo que uma das melhores maneiras de se tratar a heterogeneidade entre os consumidores é através da especificação de efeitos aleatórios.

Isso foi mostrado através de um teste e através de resultados de previsão usando uma amostra de validação.

Uma propriedade importante do modelo discreto de escolha é sua habilidade de prever corretamente uma escolha de marca na amostra de validação. Essa análise mostrou que o modelo proposto por Guadagni e Little (1983) [15] perde o seu poder de previsão quando trata-se de uma amostra onde os consumidores não são leais à marca. Em contraste, a habilidade preditiva da especificação de efeitos aleatórios é consistente com a natureza do comportamento de escolha dos consumidores. Ou seja, sua performance foi tão boa no caso de uma amostra de consumidores não leais a marcas quanto foi no caso de consumidores leais.

Esse resultado foi consistente com as descobertas de Dalal e Klein (1988) [10] de que o modelo logito multinomial com uma especificação "correta" para heterogeneidade prevê muito bem as probabilidades de escolha de marca agregadas.

A distribuição das preferências intrínsecas para as quatro marcas consideradas indica que a Nabisco possui a maior média de preferência e a menor heterogeneidade enquanto, a Sunshine tem a menor média e o maior nível de heterogeneidade. Essa descoberta, juntamente com as elasticidades estimadas sugerem que a performance da marca Sunshine no mercado é governada mais pela natureza e extensão das atividades promocionais que pelo próprio mercado consumidor. Esse fato é bastante interessante e caracteriza um produto de baixo envolvimento de compra. Como foi visto na Figura 1.2, esse tipo de produto tem como características o fato da preferência a marca se dar pela rotina da compra e a repetição se dar pela persuasão.

Através desse artigo fica então demonstrada a importância de considerar a heterogeneidade entre os consumidores participantes de um painel e ainda, a importância de realizar análise comparativa entre as abordagens existentes para obter o melhor resultado para os dados disponíveis.

☞ 3.4. - Aplicação 3

3.4.1 - Introdução

Esse exemplo de aplicação foi apresentado no artigo "*Impact of Sales Promotions on When , What and How Much to Buy*" de Sunil Gupta (1988) [16].

Nesse artigo o autor pretende avaliar a eficácia das vendas promocionais através da decomposição do pico de vendas durante o período promocional em , aumento de vendas devido à mudança de marca, aumento de vendas devido ao adiantamento do momento de compra e , por fim , aumento devido à estocagem de mercadoria.

Como mencionado anteriormente , sabe-se que as variáveis do composto de marketing afetam as decisões de compra do consumidor e assim também , afetam as vendas de uma marca. Por isso , o interesse constante sobre os fenômenos ocorridos durante os períodos promocionais.

Fabricantes e intermediários percebem um aumento nas vendas durante o período promocional , mas é preciso investigar exatamente as razões desse aumento de vendas.

Nesse artigo, a idéia do autor foi decompor o pico de vendas de forma conveniente e , assim , avaliar melhor os determinantes dessa ocorrência.

Essa idéia está relacionada com o entendimento do impacto das vendas promocionais nas decisões do consumidor de quando, o que e quanto comprar de determinada mercadoria.

A escolha da marca e a decisão da quantidade a ser comprada são condicionadas ao momento da compra e foram modeladas e estimadas separadamente.

No caso dessa dissertação , o maior interesse é o modelo utilizado para escolha da marca : o modelo logito multinomial.

O modelo foi utilizado por ser baseado na teoria comportamental da utilidade permitindo variáveis explicativas e também, levando em consideração as competições entre marcas concorrentes.

3.4.2 - Dados

Os dados foram fornecidos pelo painel do IRI (Information Resources , Inc) e foram usados para calibração e validação do modelo.

O produto escolhido , como no artigo de Guadagni e Little (1983) [15], foi o café moído cafeinado , que é um produto de baixo envolvimento de compra.

O painel completo é composto de aproximadamente 2000 consumidores com acompanhamento mínimo de compras de 2 anos.

Esse painel contém informações do histórico de compra completo para cada participante , com identificação do consumidor , marca comprada, quantidade comprada , semana e momento de compra , loja , etc ... E , além disso , informações sobre as atividades promocionais da loja e preço das mercadorias para todas as marcas de café disponíveis em todas as lojas do mercado a ser considerado.

A análise foi restrita aos dados sobre venda de café moído cafeinado no mercado de Pittsfield , MA (EUA). Foram eliminados os consumidores de uso raro , e os que não consumiam o produto. Foi considerado consumidor de uso raro aquele que realizou menos de 10 (dez) compras de café em um período correspondente a 2 anos. Adotando esse critério foram tomados 395 consumidores.

Uma amostra aleatória de 100 consumidores foi selecionada dos 395 por limitações computacionais.

O histórico de compra desses 100 consumidores foi combinado com as informações das lojas de maneira que o banco de dados final tivesse informações sobre as marcas concorrentes e as estratégias de marketing adotadas pelas lojas avaliadas.

Os dados das semanas 30 a 137 foram divididos da seguinte maneira: semanas de 30 a 59 formando o período de pré-calibração usado para inicializar algumas variáveis. O período de pré-calibração assegura que algumas variáveis estabilizem antes que sejam usadas com propósito de calibração.

As semanas de 60-100 foram usadas para calibração do modelo e as semanas 101-125 usadas para validação do modelo.

As semanas 126-137 foram eliminadas porque nenhum dos consumidores selecionados realizou compra de café moído cafeinado nesse período. Dessa maneira, foi totalizado um período de 2 anos (96 semanas) que foi usado na análise.

Os dados de calibração contém 1.526 ocasiões de compra (quando o café foi efetivamente comprado) e 4.211 oportunidades de compra (quando o café poderia ter sido comprado).

Os dados de validação contém 859 ocasiões de compra e 2535 oportunidades de compra. Cada semana foi considerada como uma oportunidade de compra para o consumidor adquirir o café. (Se um consumidor fez $n \geq 2$ compras de café em uma semana, supõe-se que aquela semana possuía n oportunidades de compras para aquele consumidor).

Alguns consumidores compraram mais de uma marca em determinadas ocasiões de compra, mas como foram poucas as ocorrências (menos de 1,5% de todas as ocasiões de compra) designou-se aleatoriamente uma das marcas compradas.

Das 1526 ocasiões de compra nos dados de calibração algumas possuíam dados faltantes ("missing values") e foram retirados da análise , sobrando 1462 ocasiões de compra.

Como mencionado no artigo de Guadagni e Little (1983) [15] e mostrado no item 3.2.1 desse capítulo , nota-se que para essa categoria de produto , ao se tratar com a mesma marca , mas com tamanhos diferentes , o produto é considerado diferente por consumidores e intermediários. Por isso , trabalhou-se também nesse caso com marca-tamanho. Ou seja , produtos com a mesma marca mas com tamanhos diferentes serão considerados produtos diferentes.

Das muitas marca-tamanho disponíveis no mercado de Pittsfield foram selecionadas 10 cujas participações no mercado total ascendem 87,5%.

Essas marcas são: Maxwell House 16oz (19.92%) , Folgers 16oz (18.43%) , Chock-Full-O'Nuts 16oz (16.58%) , Hills Brothers 16oz (8.57%) , Masta Blend 13oz (6.12%) , Marca Genérica 16oz (5.63%) , Maxwell House 32oz (3.36%) , Folgers 13oz (2.84%) , Chase & Sanborn 16oz (2.59%) e , Martinson 16oz (2.41%).

A marca genérica , nesse caso , é uma composição do café vendido pelos estabelecimentos com marca própria.

Foi considerada também , uma outra categoria denominada "outras marcas" , caracterizada pela composição de 4 outras marcas restantes de maior participação de mercado.

3.4.3 - Variáveis do Modelo

As variáveis presentes na função utilidade que compõe o modelo de escolha de marca , são listadas a seguir.

X_{0j} - constante específica de marca

$X_{1,ijn}$ - preço regular da mercadoria (cents/ounces)

$X_{2,ijn}$ - redução no preço (cents/ounces)

$X_{3,ijn}$ - propaganda ou display (entre 0 e 1)

$X_{4,ijn}$ - propaganda e display (entre 0 e 1)

$X_{5,ijn}$ - lealdade à marca (entre 0 e 1)

$X_{6,ijn}$ - lealdade ao tamanho (entre 0 e 1)

Onde , i representa o consumidor , j a marca e n a ocasião de compra.

Como no modelo de Guadagni e Little , aqui também foram introduzidas as constantes específicas de marca que tem como objetivo captar particularidades de uma marca que não possam ser explicadas pelas outras variáveis introduzidas no modelo.

Os dados do painel do IRI contém informações semanais do preço praticado para as marcas. O preço praticado é uma composição do preço regular e da redução no preço.

$$\text{Preço Praticado} = \text{Preço Regular} - \text{Redução do Preço}$$

Como nenhum dado possuía o preço regular ele foi inferido do preço praticado. Se o preço praticado de uma marca em uma loja permaneceu o mesmo por quatro semanas consecutivas , assumiu-se que esse era o preço regular da mercadoria. Como diferentes lojas tinham preços diferentes para uma marca , o preço regular foi ponderado pela participação de cada loja nas compras do consumidor.

As variáveis de X_2 a X_4 (redução no preço , propaganda ou "display", propaganda e "display") também foram ponderadas segundo a participação de cada loja nas compras do consumidor.

A correlação entre a Propaganda e o "display" obtida é de $R=0.704$, mostrando então, que são bastante correlacionadas. Então, para evitar a colinearidade foram construídas duas variáveis para cada loja s :

X_{3s} - propaganda ou display = 1, se a marca estava em promoção ou display na loja s e, 0, caso contrário.

X_{4s} - propaganda e display = 1 se a marca estava em promoção e display na loja s e, 0, caso contrário.

X_3 e X_4 foram então construídas como médias ponderadas de X_{3s} e X_{4s} respectivamente, onde os pesos foram as participações das lojas nas compras do consumidor.

A correlação das variáveis X_3 e X_4 (propaganda ou "display", propaganda e "display") obtida é de -0.326.

As variáveis lealdade à marca e lealdade ao tamanho foram definidas como média ponderada exponencialmente das compras passadas, maiores detalhes podem ser obtidos no item 3.3.2 com a aplicação de Guadagni e Little.

No artigo de Guadagni e Little (1983) [15] as constantes estimadas para lealdade à marca e para lealdade ao tamanho foram de 0.875 e 0.812, respectivamente.

No artigo de Ortmeyer (1985) [26] foi mostrado que o ajuste do modelo e os parâmetros estimados são muito robustos a pequenas mudanças (± 0.1). Dessa maneira, ao invés de estimar as constantes, nesse caso, as variáveis de lealdade foram pré-especificadas como 0.8 nos dois casos.

Lealdade à marca e lealdade ao tamanho foram inicializadas como no artigo de Guadagni e Little (1983) [15] descrito no item 3.2 desse mesmo capítulo. Por exemplo, para a primeira ocasião de compra do consumidor i, $X_{5,ij}$ foi considerado igual ao valor da constante 0.8, se a marca comprada

na marca-tamanho era j , caso contrário o valor é $0.2/(n^0 \text{ de marcas}-1)$. Isso assegura que a soma das lealdades entre as marcas seja sempre igual a 1 para um consumidor.

A variável lealdade ao tamanho é inicializada de maneira análoga.

Para todas variáveis, exceto as constantes específicas de marca, os coeficientes foram os mesmos para todas marca-tamanho. Então, o modelo é relativamente parcimonioso.

Na Figura 3.10 apresentada a seguir, são apresentados os sinais esperados para cada variável do modelo de escolha de marca:

Figura 3.10 - Sinais Esperados para Parâmetros.

Variável	Sinal	Comentário
Preço Regular	-	Qto maior preço menor chance de compra
Redução do Preço	+	Promoções aumentam o valor da marca e assim aumentam sua probabilidade de compra
Propaganda e Display	+	
Propaganda ou Display	+	
Lealdade à Marca	+	Qto maior a lealdade maior a probabilidade de escolha
Lealdade ao Tamanho	+	

3.4.4 - Análise e Resultados

O procedimento de máxima verossimilhança foi utilizado para estimar o modelo como apresentado no capítulo 2.

Dezoito parâmetros foram estimados para o modelo de escolha de marca (10 parâmetros para as contantes específicas de marca e mais oito para os parâmetros resposta).

Da mesma forma que Guadagni e Little o modelo foi construído incluindo uma variável de cada vez e o valor da estatística U^2 foi avaliado para ver se a cada etapa ocorre alguma melhora no modelo.(M0 à M5)

Os resultados são apresentados na Figura 3.11 , a seguir , onde , os valores entre parentêses representam as estatísticas t correspondentes.

Nessa Figura não foram apresentadas as constantes específicas de marca , que são as únicas no modelo M0. Esse é o modelo de participação das marcas no mercado e que contém somente as variáveis específicas de marca. Os valores de U^2 foram calculados considerando M0 o modelo nulo , os valores que estiverem entre colchetes foram calculados considerando M1 (modelo com as variáveis lealdade à marca e lealdade ao tamanho) o modelo nulo. O χ^2 para cada modelo foi sempre calculado em relação ao modelo anterior. Por exemplo, o χ^2 do modelo M4 igual a 48.64 foi calculado em relação ao modelo M3.

Figura 3.11 - Modelo de Escolha de Marca

Variável	M0	M1	M2	M3	M4	M5
Lealdade à Marca		3.848 (30.14)	4.373 (28.77)	4.340 (26.91)	4.785 (29.91)	4.555 (28.26)
Lealdade Tamanho		2.207 (10.60)	2.970 (12.79)	2.098 (9.59)	2.096 (9.38)	2.209 (9.80)
Propag. e Display			4.068 (36.89)	2.764 (24.16)	3.868 (26.27)	3.543 (22.81)
Propag. ou Display			2.143 (16.96)	1.528 (11.06)	1.927 (14.72)	1.658 (10.63)
Corte de Preço				0.716 (16.14)	1.157 (15.61)	1.147 (14.85)
PeD*Corte Preço					-0.802 (-8.41)	-0.757 (-7.64)
PouD*Corte Preço					-0.398 (-3.48)	-0.360 (-0.317)
Preço Real						-0.117 (-1.72)
-Log	3020.33	2325.40	1532.57	1378.61	1354.29	1346.03
U ²		0.229	0.491 [0.340]	0.542 [0.406]	0.549 [0.415]	0.552 [0.419]
χ^2		1389.86	1585.66	307.92	48.64	16.52
g.l.		2	2	1	2	1

Observando a Figura 3.11 verifica-se que todos as estimativas dos parâmetros do modelo final M5 foram significantes e com o sinal esperado.

As variáveis do mercado tais como preço e promoção e, as características do consumidor como lealdade à marca e lealdade ao tamanho ajudaram substancialmente a explicar as decisões de escolha de marca dos consumidores. Isso pode ser observado pelo ganho nos valores da estatística U^2 calculada cada vez que uma variável é acrescida ao modelo.

Os resultados obtidos por Gupta foram consistentes com os obtidos por Guadagni e Little.

Uma das mais importantes variáveis no modelo foram lealdade à marca e lealdade ao tamanho. Altos valores da estatística t e um ganho substancial no modelo ajustado detectado através de U^2 sugerem que as variáveis promocionais também desempenham um importante papel nas decisões de escolha do consumidor.

A adição da variável redução do preço no modelo M3 mudou substancialmente os coeficientes de Propaganda ou "Display" (X_2) e , Propaganda e "Display" (X_3). Isso sugere a possível existência de interação entre essas variáveis.

Os termos de interação foram acrescentados ao modelo em M4. Os sinais negativos sugerem uma possível sobreposição ou substituição entre os diferentes instrumentos promocionais. Nota-se ainda, que com a inclusão de termos de interação houve estabilização dos parâmetros de Propaganda e Display e , Propaganda ou Display (os valores dos parâmetros ficaram próximos aos do modelo M2 e os valores da estatística t foram maiores que em M3).

Finalmente, ao ser acrescentado o preço regular não foi notada uma grande mudança nos valores de U^2 , isso pode demonstrar que o preço

regular não teve um forte papel na decisão de escolha da marca pelo consumidor.

3.4.5 - Validação do Modelo

A obtenção de parâmetros significantes e com os sinais esperados já é um indicativo de que o ajuste foi adequado , indicando a validade do modelo proposto. Uma outra avaliação possível para validar o modelo é examinar a habilidade do mesmo em prever corretamente as participações das marcas no mercado que serão vendidas num período futuro.

Um período de 25 semanas , de 101 a 125 , dos mesmos 100 consumidores foi usado para validação.

Os preços reais e as promoções de todas marca-tamanho foram usadas em conjunção com os parâmetros estimados no modelo de calibração para prever se a compra de um produto seria realizada por um consumidor em uma dada semana (o objetivo não era prever quais decisões seriam tomadas).

Se a compra de um produto foi prevista , as condições de marketing daquela semana foram usadas para prever qual a marca e quanto seria comprado. As vendas entre os consumidores foram então agregadas para obtenção do volume de vendas semanais de um produto.

Sabe-se que algumas variáveis , como as de lealdade , precisam ser atualizadas baseadas no histórico de compras do consumidor , dependendo fortemente das compras anteriores por ele realizadas. Nesse caso, foi usado o procedimento de Monte Carlo e a compra das marcas foi simulada com base nas probabilidades de escolha das marcas como prevista pelo modelo. Esse histórico de compras produzido para cada consumidor foi usado para atualizar as variáveis de lealdade. Para evitar qualquer distorção provocada por uma única escolha gerada pelo procedimento de Monte Carlo , ele foi repetido seis (6) vezes e a média dessas iterações foi utilizada.

Os resultados obtidos , segundo Gupta , utilizando os dados no período de calibração foram bastante satisfatórios indicando que o modelo obtido se ajustou bem as condições de escolha de marca do consumidor.

3.4.6 - Resultados Finais

Como mencionado anteriormente , o artigo de Sunil Gupta tem por objetivo não só trabalhar com o modelo de escolha de marca , mas também procurar mostrar como as atividades promocionais afetam outras atitudes de compra do consumidor.

Com esse objetivo , no artigo mencionado , além do modelo de escolha de marca que foi explicitado nos itens anteriores , Sunil Gupta propôs um modelo para o período entre as compras e um modelo para estimar a quantidade a ser comprada.

Por não serem o assunto principal dessa dissertação , esses modelos não são apresentados. Porém uma análise da decomposição dos efeitos promocionais através das elasticidades será apresentada com base nos modelos obtidos pelo autor.

Os três modelos de escolha de marca , do período entre compras e da quantidade de mercadoria a ser comprada , propostos pelo autor sugerem que as variáveis de marketing mais especificamente , as variáveis de promoção , são as que tem grande influência na escolha do consumidor. Nesses três modelos, foi notado que essas variáveis tem uma contribuição relativa maior no modelo de escolha de marca que nos outros dois de tempo entre compras e quantidade comprada.

Isso já seria um indicativo do efeito das promoções , mas uma análise da elasticidade foi realizada para verificar o impacto relativo das vendas promocionais nas decisões da escolha da marca do consumidor, no tempo entre as compras e na quantidade a ser comprada.

O método a ser utilizado vai ser descrito mais detalhadamente a seguir.

Foram feitas duas suposições para a modelagem proposta: uma é que a escolha da marca e a quantidade comprada são condicionadas ao momento da compra, a outra, é que as decisões da marca a ser comprada e a quantidade adquirida são independentes.

Essas suposições permitem escrever as vendas de uma marca j em qualquer intervalo de tempo 0-t como uma função separada:

$$(A) \quad Q_j = p_j Q p_t$$

onde, Q_j = vendas da marca j no intervalo de 0-t,

p_j = prob. de escolha da marca j (condicionada a compra do produto)

Q = vendas da categoria de produto

p_t = prob. do produto ser comprado no período de 0-t (i.e., prob. de incidência de compra).

Se D_j é o nível da quantidade ou das vendas promocionais de uma marca j durante o período de 0-t, então, usando a equação (A), a elasticidade da marca j é dada por:

$$(B) \quad \eta_j = \frac{D_j}{Q_j} \frac{\delta Q_j}{\delta D_j} = \frac{D_j}{p_j} \frac{\delta p_j}{\delta D_j} + \frac{D_j}{p_t} \frac{\delta p_t}{\delta D_j} + \frac{D_j}{Q} \frac{\delta Q}{\delta D_j}$$

ou então,

$$\eta_j = \eta_m + \eta_t + \eta_q$$

onde :

η_j - elasticidade total da marca j

η_m - elasticidade da escolha de marca para marca j

η_t - elasticidade do tempo de compra com relação a quantidade da marca j

η_q - elasticidade da quantidade comprada com relação a quantidade da marca j

Usando os modelos propostos , obtem-se as elasticidades como função dos parâmetros do modelo , quantidade da marca j e , outras variáveis explicativas. Usando essas expressões e os valores médios para todas variáveis durante o período de calibração obtem-se as elasticidades para escolha da marca , tempo de compra e quantidade comprada.

A análise da elasticidade foi feita para as três (3) variáveis promocionais: Propaganda ou "Display" (X3) , Propaganda e "Display" (X4) e Redução no Preço (X2). Para cada uma das marcas que foram analisadas foi realizada a análise de elasticidade para as três variáveis promocionais.

O resultado da análise de elasticidade para todas as marcas e todos instrumentos promocionais mostrou que do total do aumento de vendas mais de 84% foi devido a troca de marca , 14% ou menos foi devido ao adiantamento do período de compra e por fim, menos de 2% foi devido a estocagem de mercadoria.

Essas descobertas indicam que no mercado de café moído cafeinado as promoções são muito efetivas para troca de marca pelos consumidores. Como a marca foi definida como marca-tamanho parte dessa troca pode ocorrer entre tamanhos diferentes da mesma marca. No entanto, o que foi observado é que os dados analisados possuíam poucos produtos com a mesma marca e tamanho diferentes e , uma outra análise da matriz de troca mostrou que esse tipo de troca corresponde somente a 4% do total ocorrido.

Portanto, no caso analisado a maioria das trocas ocorre para marcas concorrentes.

Quase todo o aumento de vendas devido a redução no preço foi devido a troca de marca (98%). Uma razão é que de maneira diferente a Propaganda ou "Display" e Propaganda e "Display" , a redução no preço afeta a decisão do momento de compra do consumidor.

O aumento das vendas promocionais devido antecipação do período de compra e/ou devido estocagem de mercadoria não podem ser considerados como um incremento real as vendas porque pelo menos alguns desses consumidores acabariam comprando a marca em promoção mais tarde. Para todas as marcas o aumento devido a antecipação do tempo de compra foi menor que 14% e devido a estocagem de mercadoria foi menor que 2%.

O pequeno efeito de estocagem , no caso do produto em questão , poderia ser justificado por três possíveis motivos: a percepção do consumidor que a estocagem de café pode fazer com que a mercadoria perca sua característica de frescor , o grande volume que representaria o armazenamento da mercadoria e, a grande intensidade de promoções que ocorre desse produto garantindo que se a compra não for realizada agora mais tarde será possível obter a mesma mercadoria em outra promoção.

Isso sugere que a estocagem pode ser um fator de maior importância em outras categorias de produto talvez, produtos não perecíveis . Essa possibilidade deve ser investigada em estudos de outras mercadorias.

3.4.7. - Comentários

Os resultados obtidos nesse artigo representam mais um avanço para compreensão do comportamento de compra do consumidor pois, oferecem mais ferramentas para uma análise detalhada desse assunto.

E, o modelo logito multinomial mais uma vez, foi apresentado como uma excelente técnica para modelar a escolha de marca, tendo sempre em

vista a necessidade de um banco de dados que disponibilize informações completas sobre a categoria do produto que está sendo avaliada bem como sobre as atividades promocionais garantindo assim, um acompanhamento completo das concorrências existentes no mercado avaliado.

3.5 - Conclusões Finais

Nesse capítulo foram apresentadas 3 (três) aplicações do modelo logito multinomial para analisar o comportamento de compra individual.

Os produtos analisados (café e bolacha) são tipicamente produtos de baixo envolvimento de compra.

É importante verificar nessas 3 (três) aplicações que o modelo obtido , com seus coeficientes e constantes específicas de marca auxiliam bastante na melhor compreensão do comportamento de escolha do consumidor. Com isso , crescem as possibilidades de realizar promoções de mercadorias com objetivos direcionados , e que garantam sucesso na sua realização.

A possibilidade da realização de simulações com modelo obtido garante uma pesquisa sobre as reações que determinadas ações de marketing podem causar no mercado , possibilitando uma avaliação das participações de mercado de marcas concorrentes de um mesmo produto.

Além das vantagens apresentadas com o estudo do modelo logito multinomial , ainda foi possível analisar distribuições das marcas (aplicação 2) , ou mesmo , as elasticidades de preço , promoção (aplicação 1) , e as elasticidades de modelos para detectar a causa do aumento de vendas em um período promocional (aplicação 3).

Cada uma das aplicações forneceu um novo ponto que possibilita esclarecimento sobre a utilidade do modelo e as atividades de marketing.

Como mencionado anteriormente , todo avanço desse estudo se deve a uma coleta de dados tão completa quanto a oferecida pelos "scanner data" ,

ou dados de varredura. Sem isso , com certeza não haveria qualquer possibilidade de uma análise tão complexa sobre o comportamento de escolha do consumidor.

Nos Estados Unidos, muitas empresas realizam coletas de dados completas. Infelizmente , no Brasil , ainda é muito difícil conseguir um volume de dados tão completo e confiável , sem contar , com a grande quantia que seria despendida para obter algo similar que permitisse análises tão complexas.

Seria bastante interessante avaliar o comportamento de escolha do consumidor em nosso país , com certeza , esse comportamento é decorrência direta da cultura de cada povo o que coduziria possivelmente a conclusões diferentes das obtidas nas aplicações apresentadas nesse capítulo.

Capítulo 4 - Considerações Finais

Essa dissertação procurou mostrar , através do modelo logito multinomial , uma maneira de compreender melhor o comportamento de escolha individual de produtos com baixo envolvimento de compra e verificar a participação no mercado avaliado de marcas concorrentes de uma mesma categoria de um produto.

Foi apresentado o modelo logito multinomial cujo desenvolvimento foi feito por meio de uma derivação axiomática. Esse modelo vai considerar um indivíduo confrontando a escolha de um produto ou serviço e associando , a cada uma das alternativas , uma utilidade.

Procurou-se quantificar a função utilidade "criada" pelo indivíduo , o que permitiu o desenvolvimento de um modelo capaz de levar em consideração as características inerentes ao indivíduo bem como , as influências por ele sofridas no decorrer do processo decisório.

A formulação de um modelo com essa característica só é possível com a utilização de um banco de dados completo que fornece um histórico de compras para vários indivíduos.

Algumas sugestões podem ser apresentadas para utilização do modelo desenvolvido nessa dissertação. Uma delas está relacionada ao banco de dados. Embora tenham sido utilizados bancos de dados muito informativos nas aplicações apresentadas, outros atributos poderiam auxiliar bastante se fossem acrescentados entre eles , a informação sobre a atividade promocional caracterizada pelos cupons de desconto (cuponagem) que são distribuídos com muita frequência , e de diversas maneiras , nos Estados Unidos.

Das aplicações apresentadas , nenhuma avaliou esse tipo de promoção porque os dados disponíveis não continham essa informação. No entanto , a cuponagem é praticada com sucesso nos Estados Unidos gerando elevados retornos.

Sabe-se que o sucesso desse tipo de promoção nos Estados Unidos é bastante grande. Uma prova desse sucesso pode ser verificado no esforço que um jornal de grande circulação na cidade de São Paulo vem fazendo para aumentar suas vendas , oferecendo cupons de desconto para os mais variados produtos e serviços. Com essa atividade , procurou-se aumentar as vendas do jornal bem como dos produtos e serviços em promoção.

Resta saber , se no Brasil haverá o retorno desejado.

Isso nos alerta para um outro importante fator que deve ser considerado durante a modelagem do comportamento de escolha individual através do modelo logito multinomial.

Como destacado anteriormente , o comprador sofre diversos estímulos externos e entre eles , estão os estímulos caracterizados pelo ambiente em que o indivíduo realiza suas atividades. Portanto , no momento da concepção do modelo deve-se compreender o ambiente econômico , tecnológico , político e cultural a que o indivíduo está sujeito para incorporá-lo ao modelo de maneira adequada.

Mais especificamente, podemos exemplificar a importância de se considerar os estímulos externos com as diferenças claras por que passam o consumidor americano e consumidor brasileiro.

O consumidor brasileiro , está sujeito a características ambientais (econômica , tecnológica , política e cultural) extremamente diferentes das do consumidor americano , o que por exemplo , pode justificar o fato do consumidor brasileiro não ser influenciado por promoções caracterizadas por uma pequena redução no preço.

Uma pequena redução no preço , que nos Estados Unidos poderia causar um grande aumento na compra de uma marca de determinado produto , aqui no Brasil ,no presente momento, poderia passar despercebido , pois aqui , a inflação é tão alta e as mudanças de preço tão constantes que o indivíduo com frequência perde a noção do valor real da mercadoria.

Uma sugestão seria trabalhar com um percentual a partir do qual se pudesse dizer que o consumidor vai perceber o benefício oferecido. Mas , também para se determinar o valor a partir do qual isso ocorre seria preciso uma avaliação específica para cada categoria de produto , e para cada mercado consumidor.

Também ao ambiente cultural deve ser dada grande atenção , pois é sabido , que determinados produtos são consumidos em maior ou menor intensidade dependendo da cultura da região que se deseja avaliar.

O café , que foi um produto avaliado em duas das aplicações apresentadas nessa dissertação , possui certamente , características de compras bastante diferentes entre os consumidores brasileiros e os americanos. Essas diferenças são baseadas principalmente nas diferenças culturais existentes entre esses dois tipos de consumidores com relação a esse produto e que os leva a consumir de forma diferente.

Com tudo isso , fica claro que quando da modelagem de qualquer produto ou serviço , é preciso um estudo de diversos fatores entre eles , os motivos de consumo , o envolvimento que o consumidor deve dispor para a compra do mesmo , e ainda , o conhecimento das características ambientais que o consumidor , do qual desejamos avaliar o comportamento de escolha, está sujeito.

Com esse entendimento , fica mais fácil definir as variáveis ou atributos de maior importância e que devem fazer parte da função utilidade que vai caracterizar o modelo logito multinomial para escolha individual.

Os atributos considerados de grande importância no modelo de escolha, entre as aplicações que foram apresentadas nessa dissertação , foram os de lealdade à marca e lealdade ao tamanho.

Não é difícil avaliar a importância desses atributos em um modelo de comportamento de escolha pois , basta lembrar que se tratarmos com indivíduos (consumidores) leais , a escolha de uma marca de determinado produto dificilmente vai mudar sendo que , para que isso ocorra fatores

•

muito determinantes devem influenciar o comportamento do consumidor. Por outro lado , quando se trata de indivíduos não leais sabe-se que eles mudarão de marca com grande facilidade de acordo com o que mais for capaz de influenciar sua decisão no momento da compra.

Ou ainda , pode-se encontrar indivíduos muito leais à marca no momento de comprar produtos de alto envolvimento de compra como por exemplo , carros , ou indivíduos pouco leais à marca na compra de caixa de fósforos , um produto de baixo envolvimento de compra.

Nesses casos , o que determina a lealdade é o tipo de envolvimento exigido para compra do produto em questão.

Em alguns casos apresentados, os atributos de lealdade foram os únicos a serem usados para captar a heterogeneidade entre os indivíduos pois , seriam capazes de considerar as diferenças dos indivíduos e ainda a dinâmica do processo.

Deve-se ressaltar porém , que também esses atributos são influenciados pelo ambiente a que estão sujeitos os indivíduos em questão e ao produto que se considera. Por exemplo , não é possível pensar em lealdade quando existe grande falta de mercadoria , estimulada basicamente pela política econômica vigente no mercado que está sendo estudado.

A importância de se considerar as diferenças entre os indivíduos é muito grande por isso , a proposta de Chintagunta , Jain e Vilcassim (1991) [9] sobre o tratamento da heterogeneidade no modelo logito multinomial é extremamente interessante mas também , não se deve esquecer das dificuldades computacionais apresentadas por esse tipo de abordagem.

Foi também , mostrada uma maneira de se contornar os cálculos provocados pela estimação dos parâmetros através da função de verossimilhança. O uso da quadratura Gaussiana diminui a complexidade dos cálculos e fornece novas maneiras de visualizar os pontos de massa e as probabilidades.

Apesar disso , é importante lembrar que muitas vezes os pesquisadores não dispõem de ferramentas que permitam cálculos tão complexos , o que iria desestimular maiores estudos nesse sentido.

Ao se realizar esses estudos , é preciso avaliar as ferramentas de que se dispõe antes de definir principalmente , qual vai ser a abordagem dada a heterogeneidade.

Sem dúvida , a proposta de uma abordagem semi-paramétrica nos levará a resultados "melhores" como destacado ao avaliar-se os vários modelos propostos. Mas , na prática é preciso considerar uma solução que equilibre o que existe de melhor na teoria e o que se tem disponível para realização da pesquisa.

Existem pacotes estatísticos disponíveis no mercado que , apresentam módulos onde , é possível realizar a modelagem do modelo logito multinomial com a ressalva , de que não é viável considerar a heterogeneidade a não ser incorporando a mesma , nas variáveis de preferência , e usando o comportamento de escolha passado do indivíduo.

Um pacote estatístico que permite essa análise é o BMDP módulo PAR (Derivative - Free Nonlinear Regression).

Esse módulo estima os parâmetros de qualquer função que possa ser especificada através da linguagem Fortran.

Os parâmetros são estimados por um algoritmo iterativo através de mínimos quadrados , e podem ser usados para calcular estimativas de máxima verossimilhança. O programa é apropriado para uma grande variedade de funções que não são lineares nos parâmetros e para as quais as derivadas são difíceis de especificar e custosas de se calcular pois são também não lineares.

Na falta de pacotes já preparados para esse objetivo é preciso dispendir algum tempo na realização de programas que possibilitem a análise desejada.

Esse estudo vem sendo desenvolvido desde 1983 , o que significa , que é possível implementá-lo com novos conceitos.

Não só a parte teórica e computacional pode ser implementada , mas também , o banco de dados que possibilita esse estudo.

Infelizmente , no Brasil , ainda não existe uma empresa realizando uma coleta de dados que permita um estudo como esse apresentado nessa dissertação.

Talvez , isso decorra das influências do mercado em que vivemos . Os custos seriam elevados e sem retorno certo fazendo com que as empresas não se disponham a gastar muito em um mercado pouco estável e de dimensões reduzidas.

É preciso mostrar como o aperfeiçoamento do conhecimento do comportamento de escolha individual pode proporcionar benefícios as empresas para que as mesmas aumentem seu interesse em pesquisas como essa.

Ao se avaliar as atividades promocionais e as influências que o consumidor sofre no período anterior a tomada de decisão pode-se certamente aumentar o volume de vendas do seu produto através da realização das promoções corretas.

Realizar promoções corretas ,significa dispendir dinheiro com atividades que irão dar resultados positivos no aumento do volume de vendas obtidos pós promoção.

Pode-se exemplificar com a atividade promocional "display" . Suponhamos que seja detectada empiricamente ,através do modelo , a

•

influência que o "display" possui na venda de determinada mercadoria , isso por exemplo ,leva o representante a ter maior interesse por determinados locais na gôndola de um supermercado , por exemplo.

Os supermercados tem políticas para dispor as mercadorias em suas gôndolas . Se é do conhecimento do representante que o "display" em locais mais visados realmente aumenta o volume de vendas da mercadoria ele não se importa em fazer concessões por algo que lhe dará retorno certo.

Se por outro lado , se detecta que o "display" não é capaz de influenciar o comportamento de escolha individual , o representante que anteriormente se preocupava com a colocação da sua mercadoria nos melhores locais da loja , poderá realizar outras atividades promocionais que lhe garantam retorno em aumento do volume de vendas.

Somente a compreensão da extensão dos resultados do modelo logito multinomial para comportamento de escolha individual pode fazer com que mais empresas se interessem por esses estudos e dispendam dinheiro para uma boa coleta de dados e uma maior exploração do modelo aqui apresentado.

Uma qualidade considerável do modelo proposto é sua parcimônia. Os atributos são os mesmos para todas marcas (ou marca-tamanho) e para todos consumidores a exceção dos atributos específicos , e ainda , o modelo permite prever as participações no mercado de cada marca , seguindo as tendências e mudanças de desempenho realizados ao longo do tempo pelas diferentes lojas.

O modelo poderia ter outras extensões. No caso apresentado , há uma preocupação em descrever o que ocorre dentro das lojas.

Isso pode não ser adequado para o varejista que está talvez , pouco interessado nas competições entre as marcas , exceto talvez com relação a marca da casa , e mais interessado na competição entre outros varejistas.

Da mesma maneira , embora os fabricantes estejam bastante interessados na competição entre as marcas , os atributos de controle do modelo estão bastante relacionados com o varejista e somente indiretamente podem ser influenciados pelo fabricante.

Outras idéias podem surgir para extensão dos resultados aproximando o pesquisador das necessidades dos tomadores de decisão.

A compreensão da categoria de produtos analisada é importante pois algumas suposições do modelo logito multinomial deveriam ser avaliadas para analisar a validade das mesmas em situações mais complicadas onde , por exemplo , a diferenciação do produto é grande ou a procura de variedade é comum , ou ainda , o consumidor que fica exposto a um portfólio de produtos.

Novas maneiras de modelar podem surgir , exigindo talvez idéias originais para resolução desses problemas.

Uma outra aplicação interessante é apresentada por Krishnamurthi e Raj (1988) [20] onde , um modelo para estudar a escolha de marca repetitiva e a decisão da quantidade a ser comprada são modeladas conjuntamente de uma maneira parcimoniosa.

Muito também , ainda pode ser feito pela implementação computacional desse modelo.

Explorar o comportamento de escolha através de um modelo estatístico como o apresentado é um material riquíssimo que pode ampliar o uso da estatística no campo de pesquisa de mercado e trazer certamente benefícios para pesquisadores e estudiosos de ambas as partes.

Referências Bibliográficas

- [1]. **Abramowitz , M. ; Stegun , I. A. ;** (1972) - Handbook of Mathematical Functions - New York : Dover Publications , Inc.
- [2]. **Allenby , Greg ; Rossi , Peter ;** (1990) - "*Quality Perceptions and Asymetric Switching Between Brands* "- working paper , Ohio State University
- [3]. **Ben-Akiva , Moshe E. ;** (1973) - "*Program for Maximum Likelihood Estimation of the Multinomial Logit Model* " - Transportation Systems Division , Department of Civil Engineering , MIT (January)
- [4]. **Ben-Akiva , Moshe E. ; Lerman , Steven R. ;** (1985) - Discrete Choice Analysis : Theory and Applications to Travel Demand , Cambridge , MA ; MIT Press
- [5]. **Blattberg , Robert C. ; Wisniewski , Kenneth J. ;** (1989) - "*Price-Induced Patterns of Competition* " - Marketing Science , 4 (Fall) , 291-309
- [6]. **Butler , J.S. ; Moffit , R. ;** (1982) - "*A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model* " - Econometrica , 50 , 761-764
- [7]. **Canton , Adolpho W. P. ; Grisi , Celso C. de H. ; Santos , Rubens da C.** (1989) - "*Avaliando os Níveis de Envolvimento dos Consumidores* " - Trabalho apresentado no XIII Encontro Anual da ANPAD
- [8]. **Canton , Adolpho W. P. ; Grisi , Celso C. de H. ; Santos , Rubens da C.** (1989) - "*Identificando os Níveis de Envolvimento do Consumidor* " - Mercado Global 79 , jun/jul , 36-39
- [9]. **Chintagunta , Pradeep K. ; Jain , Dipak C. ; Vilcassim , Naufel J.** (1991) - "*Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data*" - Journal of Marketing Research , 28(November) 417-428

- [10]. **Dalal , Siddhartha R. ; Klein , Roger W. ;** (1988) - "*A Flexible Class of Discrete Choice Models*" - *Marketing Science* , 7 (Summer) , 232-251
- [11]. **Davis , Scott ; Inman , Jeffrey J. ; McAlister , Leigh** (1992) - "*Promotion has a Negative Effect on Brand Evaluation - Or Does It ? Additional Disconfirming Evidence*"- *Journal of Marketing Research* , 29 (February) 143-148
- [12]. **Dodds , William B. ; Monroe , Kent B. ; Grewal , Dhruv** (1991) - "*Effects of Price, Brand and Store Information on Buyers Product Evaluation*" - *Journal of Marketing Research* , 28 (August) 307-319
- [13]. **Domencich , T. A. ; McFadden , D.** (1975) - *Urban Travel Demand* - Amsterdam : North Holland
- [14]. **Gensch , Dennis H. ; Recker W.W.** (1979) - "*The Multinomial , Multiattribute Logit Choice Model* " - *Journal of Marketing Research* , 16(February) 124-132
- [15]. **Guadagni , Peter M. ; Little , John D.C.** (1983) - "*A Logit Model of Brand Choice Calibrated on Scanner Data*" - *Marketing Science* , 2 (Summer) 203-238
- [16]. **Gupta , Sunil** (1988) - "*Impact of Sales Promotions on When , What and How much to Buy*" - *Journal of Marketing Research* , 25 (November) 324-355
- [17]. **Heckman , James J. ; Singer , Burton** (1984) - "*A Method of Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data*" - *Econometrica* , 52(2) 271-320
- [18]. **Hsiao , Cheng** (1986) - *Analysis of Panel Data* - Cambridge , U.K. : Cambridge University Press
- [19]. **Kotler , Philip** (1992) - *Administração de Marketing , Análise , Planejamento , Implementação e Controle* - 2ª edição - Editora Atlas S.A.

- [20]. **Krishnamurthi , Lakshman ; Raj , S. P.** (1988) - "*A Model of Brand Choice and Purchase Quantity Price Sensitivities*" - Marketing Science , 7(Winter) 1-20
- [21]. **Luce , R. D. ;** (1959) - Individual Choice Behavior , New York ; John Wiley and Sons
- [22]. **Luce , R. ; Suppes P. ;** (1965) - "*Preference , Utility ,and Subjective Probability*"- In Handbook of Mathematical Psychology (R.Luce , R.Bush and e.Galander , eds.) , Vol. 3 Wiley , New York
- [23]. **Marschak , L. ;** (1960) - "*Binary Choice Constrains on Random Utility Indicators*"- In Stanford Symp. Math. Methods Soc. Sci. (K. Arrow , ed) Stanford Univ. Press , Stanford , California
- [24]. **McFadden , Daniel** (1974) - "*Conditional Logit Analysis of Qualitative Choice Behavior* " in Frontiers of Econometrics , P. Zarembka - Editora New York : Academic Press , Inc. 105-142
- [25]. **McFadden , Daniel ; Richter , M. ;** (1971) - "*On the Extension of a Set Function to a Probability on the Boolean Algebra Generated by a Family of Events , with Applications* "- working paper 14 , MSSB Workshop on the Theory of markets under Uncertainty , Dept. of Econ. , Univ. of California, Berkeley, California
- [26]. **Ortmeyer , Gwen K. ;** (1985) - "*A Model of Consumer Response to Promotions* " - unpublished dissertation , Stanford University
- [27]. **Press , W. ; Fleming , B. ; Tenkolsky , S. ; Vetterling W. ;** (1986) - Numerical Receipes - Cambridge , UK : Cambridge University Press
- [28]. **Punj , Girish N. ; Staelin , Richard** (1978) - "*The Choice Process for Graduate Business Schools*" - Journal of Marketing Research , 15(November) 588-598
- [29]. **Robertson , T.S. ; Zielinski , Joan ; Ward , Scott -** (1984) Consumer Behavior - Glenview , III ; Scott Foresman and Company , United States

- [30]. **Simonson , Itamar ; Winer, Russel S.** (1992) - " *The influence of Purchase Quantity and Display Format on Consumer Preference for Variety*" *Journal of Consumer Research* , 19 (June) 133-138
- [31]. **Urban , G. L. ; Johnson , P.L. ; Brudnick , R. H. ; Hauser , J. R.** ; (1982) - "*Modeling of Competitive Market Structure: A Methodology Based on Forced Switching Probabilities*" - Sloan School Working Paper 1236 - 1281 , MIT (November)
- [32]. **Vilcassim , Naufel J. ; Jain , Dipak C.** ; (1986) - "*Modeling Purchase-Timing and Brand Switching Behavior Incorporating Explanatory Variables and Unobserved Heterogeneity*" - *Journal of Marketing Research* , 28(February) , 29-41
- [33]. **Zufriden , Fred S.** (1986) - "*Multibrand Transition Probabilities as a Function of Explanatory Variables : Estimation by a Least Squares Based Approach* " - *Journal of Marketing Research* , 23(May) 177-183
- [34]. **Zufryden , Fred S. ; Pedrick , James H. ; Sankaralingam , Avu ;** (1993) - "*Zapping and Its Impact on Brand Purchase Behavior*" - *Journal of Advertasing Research* - (January/February) , 58-66