

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**UM MODELO DESAGREGADO DE ESCOLHA
DISCRETA PARA A ANÁLISE DAS PREFERÊNCIAS
DO MERCADO:
UMA ABORDAGEM BAYESIANA**

Tese de Doutorado

Luis Alberto Aguilar Córdova

**Florianópolis
2002**

LUIS ALBERTO AGUILAR CÓRDOVA

**UM MODELO DESAGREGADO DE ESCOLHA
DISCRETA PARA A ANÁLISE DAS PREFERÊNCIAS
DO MERCADO:
UMA ABORDAGEM BAYESIANA**

Tese submetida ao Programa de Pós-
Graduação em Engenharia de
Produção da Universidade Federal de
Santa Catarina - UFSC, como requisito
parcial para a obtenção do grau de
Doutor em Engenharia de Produção.

ORIENTADOR: Prof. Bruno H. Kopittke, Dr.

Florianópolis

2002

Luis Alberto Aguilar Córdoba

**UM MODELO DESAGREGADO DE ESCOLHA DISCRETA PARA
A ANÁLISE DAS PREFERÊNCIAS DO MERCADO:
UMA ABORDAGEM BAYESIANA**


Esta tese foi julgada e aprovada para a obtenção do título de Doutor em Engenharia de Produção no Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina

Florianópolis, 18 de março de 2002



Prof. Ricardo Miranda Barcia, Ph.D.
Coordenador do Curso

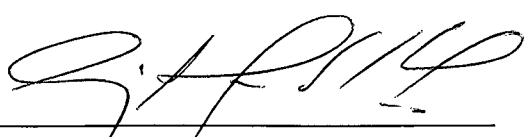
BANCA EXAMINADORA



Prof. Bruno Hartmut Kopittke, Dr.
Orientador – UFSC



Prof. Rubens Araújo de Oliveira, Dr.
Examinador – ESAG / UDESC



Prof. Luiz A. dos Santos Senna, Ph.D.
Examinador – UFRGS



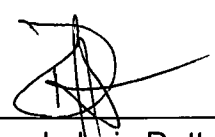
Prof. Carlos Wolowsky Mussi, Dr.
Examinador – UFSC



Prof. Antônio Galvão Novaes, Dr.
Examinador – UFSC



Prof.ª Lia Caetano Bastos, Dra.
Examinadora – UFSC



Prof. Armando Luiz Dettmer, Dr.
Moderador – UFSC

AGRADECIMENTOS

A realização de um trabalho desta natureza representa um esforço de anos, não isento de apoios, companheirismo, entreatura, amizade e carinho. Por esta razão gostaria de agradecer, reconhecidamente:

À Wadaed, minha companheira fiel de todas as horas no desafiante caminho do conhecimento, agradeço profundamente e dedico o resultado do trabalho.

Aos meus pais Germán e Angélica e à minha irmã Maria Dolores por todo apoio, amor e valores humanos transmitidos.

Ao Prof. Dr. Bruno H. Kopittke pela orientação, amizade e apoio irrestrito desde o início desta jornada. Agradeço especialmente a confiança e a oportunidade de realizar este trabalho.

À Universidade Federal de Santa Catarina, aos professores, funcionários e alunos do Curso de Pós-Graduação em Engenharia de Produção, pela acolhida e pelo clima de amizade e cooperação profissional.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo apoio financeiro.

Aos eminentes integrantes da Banca Examinadora pela revisão criteriosa, valiosas sugestões e generoso estímulo.

Aos professores Dr. Renato Marchetti e Dr. Jair Mendes Marques da Universidade Federal de Paraná - UFPR, pelas críticas argutas e valiosas considerações efetuadas.

Ao Prof. Dr. Armando Luiz Dettmer pela leitura crítica.

Ao Juliano Gerber pelo apoio administrativo.

Ao Prof. Dr. Reynaldo Chile por viabilizar as entrevistas na UNIVALI.

Às 156 pessoas que gratuitamente se disponibilizaram para serem entrevistadas.

Ao Brasil, terra hospitaleira e generosa.

E a todos aqueles amigos, que direta ou indiretamente, contribuíram para a realização deste trabalho.

RESUMO

AGUILAR, Luis Alberto¹; 2002. Um Modelo Desagregado de Escolha Discreta para a Análise das Preferências do Mercado: Uma Abordagem Bayesiana. Florianópolis. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, UFSC. 117p.

Este trabalho visa oferecer uma oportunidade de agregar valor ao processo de marketing estratégico das organizações orientadas para o mercado. Apresenta-se uma metodologia, baseada em análise de escolha discreta, consistente com o princípio fundamental do marketing atual: os consumidores são diferentes. Desenvolve-se um modelo hierárquico de inferência bayesiana que permite a estimação de parâmetros individuais de preferências utilizando reduzida informação por entrevistado. O comportamento individual de escolha é descrito através de um modelo multinomial logit ordenado. Os parâmetros estimados alimentam um simulador de preferências, para desenvolver análise competitiva, visando identificar oportunidades de mercado além de formular e avaliar estratégias alternativas de marketing para aproveitar essas oportunidades. A aplicabilidade da metodologia é ilustrada mediante uma aplicação empírica num contexto de marketing educacional.

Palavras-chave: pesquisa de mercado, preferência declarada, modelos bayesianos hierárquicos, simuladores de mercado.

¹ aguilarbr@yahoo.com.br

ABSTRACT

AGUILAR, Luis Alberto¹; 2002. **A Disaggregate Discrete Choice Model for Market Preferences Analysis: A Bayesian Approach.** Florianópolis. Doctor's degree thesis in Production Engineering – Centro Tecnológico, Universidade Federal de Santa Catarina. 117p.

This work aims to offer an opportunity of adding value to the process of strategic marketing of market-driven organizations. A methodology based on discrete choice analysis is presented, which is consistent with the fundamental principle of marketing: consumers are different. A hierarchical model of bayesian inference was developed. This model allows the estimation of individual parameters from the reduced information available in respondents' choices. The individual behavior is described through an ordered multinomial logit model. In order to perform competitive analysis, the estimated parameters are embodied within a preference simulator. The objective is to identify market opportunities, as well as the formulation and assessment of alternative marketing strategies to take advantage of these opportunities. The applicability of the proposed methodology is illustrated through an empirical application, in a context of marketing education.

Key-words: market research, discrete choice modeling, hierarchical bayesian models, market simulators.

¹ aguilarbr@yahoo.com.br

SUMÁRIO

Lista de Figuras	xi
Lista de Tabelas	xii
Capítulo I	
INTRODUÇÃO	1
1.1 Objetivos	3
1.2 A abordagem	4
1.3 Contribuição científica	5
1.4 Limitações do estudo	5
1.5 Estrutura do trabalho	6
Capítulo II	
A MENSURAÇÃO DA ESTRUTURA DE PREFERÊNCIAS DO CONSUMIDOR	8
2.1 Origem	8
2.2 Principais aplicações e importância	9
2.3 Uma classificação	10
2.4 Os modelos conjuntos	11
Modelos conjuntos baseados em preferências: <i>Conjoint Analysis</i>	12
Modelos conjuntos baseados em escolhas: <i>Experimental Choice Analysis</i>	12
2.5 Teoria da escolha probabilística	13
2.6 Formulação básica dos modelos de escolha discreta	14
2.6.1 Modelos de escolha binária	15
2.6.2 Modelos de escolha múltipla	16

Modelos não-ordenados	16
Modelos ordenados	17
2.7 Vantagens e desvantagens dos modelos agregados	19
2.8 Relaxando a condição IID	20
2.9 Na procura de <i>part-worths</i> individuais	23

Capítulo III

A ABORDAGEM BAYESIANA PARA O PROBLEMA DA ESCOLHA

DISCRETA	26
3.1 Inferência bayesiana	26
3.2 Regra de Bayes	27
3.3 O Modelo de regressão bayesiana: Um exemplo	29
3.4 O Amostrador de Gibbs	32
3.5 Modelos bayesianos para escolha discreta	36
3.6 Modelos bayesianos hierárquicos	38
3.7 O Algoritmo Metropolis-Hastings	40

Capítulo IV

O MODELO PROPOSTO	42
4.1 O modelo	43
4.1.1 Inferência dos parâmetros individuais	44
I. Formulação	45
II. Definição de pontos iniciais	46
III. Estimação iterativa dos parâmetros: Amostragem de Gibbs	47
IV. A amostragem dos <i>part-worths</i> individuais: Metropolis-Hastings	49
V. Procedimento de diagnóstico da convergência	50
VI. Validação preditiva do modelo	51
Cenários <i>holdout</i>	52
Critérios de desempenho	52

Hit Rate	52
Mean Absolute Error (MAE)	52
4.1.2 Simuladores de escolha	53
O simulador desenvolvido	54
Capítulo V	
VALIDAÇÃO DO MODELO	56
5.1 Descrição do mercado virtual	56
5.2 Definição do estudo	57
5.3 Delineamento do experimento	57
5.4 Situações simuladas	59
5.4.1 1 ^ª situação: Mercado homogêneo	59
5.4.1 2 ^ª situação: Mercado com dois segmentos	64
Capítulo VI	
APLICAÇÃO EMPÍRICA	76
6.1 Metodologia da pesquisa	76
6.1.1 Objetivo do estudo	77
6.1.2 Definição dos atributos e níveis	77
6.1.3 Formulação do delineamento experimental	79
6.1.4 Definição da população-alvo	81
6.1.5 Determinação do tamanho da amostra	81
Tipo de amostragem	81
Composição da amostra	82
6.1.6 Coleta dos dados	82
Descrição do sistema desenvolvido para a coleta dos dados	84
6.1.7 Estimação dos parâmetros	87
Diagnóstico de convergência	87
Parâmetros agregados	87
6.1.8 Validação dos parâmetros individuais	90
Confirmação das hipóteses	90

6.1.9 Análise competitiva	93
6.1.9.1 Caso Base A	93
6.1.9.2 Caso Base B	99
6.7 Considerações adicionais	102
Capítulo VII	
CONCLUSÕES E RECOMENDAÇÕES	104
REFERÊNCIAS BIBLIOGRÁFICAS	109

LISTA DE FIGURAS

Figura 2.1	Abordagens para a mensuração da estrutura de preferências declaradas	11
Figura 4.1	Simulador de escolhas	55
Figura 5.1	Evolução da norma das diferenças: $D_1 - I$	62
Figura 5.2	Evolução da média do <i>Hit Rate</i>	63
Figura 5.3	Evolução da média do MAE	64
Figura 5.4	<i>Part-worths</i> agregados	66
Figura 5.5	<i>Part-worths</i> individuais	67
Figura 5.6	Parâmetros agregados em cada segmento	68
Figura 5.7	Média dos <i>part-worths</i> , para 2, 3 e 4 <i>clusters</i>	70
Figura 5.8	Mapa da Análise de Correspondência	75
Figura 6.1	Metodologia de pesquisa	77
Figura 6.2	Menu principal	86
Figura 6.3	Um cenário de escolha	86
Figura 6.4	Média dos <i>part-worths</i> , para 2, 3 e 4 <i>clusters</i>	89
Figura 6.5	Probabilidade normal	92
Figura 6.6	Distribuição das probabilidades de escolha	94

LISTA DE TABELAS

Tabela 5.1	Um estudo hipotético	57
Tabela 5.2	Matriz Arranjo Ortogonal L8	58
Tabela 5.3	Blocos incompletos balanceados	58
Tabela 5.4	Alternativas para formação de <i>holdouts</i>	59
Tabela 5.5	Cenários <i>holdout</i>	59
Tabela 5.6	Análise de convergência: Mercado homogêneo	60
Tabela 5.7	<i>Part-worths</i> agregados, por cenário	61
Tabela 5.8	Matriz D para 8 e 2 cenários de escolha	61
Tabela 5.9	Porcentagens de acertos em cada cenário	62
Tabela 5.10	<i>Hit Rate</i> (parâmetros β_i vs parâmetros agregados α)	63
Tabela 5.11	MAE para cada cenário <i>holdout</i>	64
Tabela 5.12	<i>Part-worths</i> agregados α	65
Tabela 5.13	Matriz D para 6 cenários por indivíduo	66
Tabela 5.14	<i>Part-worths</i> agregados α inferidos para os dois segmentos	68
Tabela 5.15	Índices de desempenho preditivo	68
Tabela 5.16	Análise de variância	71
Tabela 5.17	Distribuição de probabilidade multinomial, para simulação de Z_1	72
Tabela 5.18	Frequências observadas	73
Tabela 5.19	Resíduos: diferença entre frequências observada e esperada	74
Tabela 6.1	Atributos e níveis	78
Tabela 6.2	Matriz Arranjo Ortogonal L9	79
Tabela 6.3	Codificação <i>dummy</i>	79
Tabela 6.4	Blocos incompletos balanceados	80
Tabela 6.5	Cenários <i>Holdout</i>	80
Tabela 6.6	Composição da amostra (%)	82
Tabela 6.7	Tempos	85
Tabela 6.8	Análise de convergência	87

Tabela 6.9	Parâmetros agregados	88
Tabela 6.10	<i>Hit Rate</i> e MAE para cada cenário <i>holdout</i>	90
Tabela 6.11	Caso Base A	93
Tabela 6.12	Composição do mercado - Caso Base A	94
Tabela 6.13	Indivíduos alvo - Caso Base A	95
Tabela 6.14	<i>Part-worths</i> dos indivíduos alvo do segmento B	96
Tabela 6.15	Primeira estratégia de marketing	97
Tabela 6.16	Composição do mercado (Primeira estratégia)	97
Tabela 6.17	Segunda estratégia de marketing	98
Tabela 6.18	Composição do mercado (Segunda estratégia)	98
Tabela 6.19	Caso Base B	99
Tabela 6.20	Composição do mercado - Caso Base B	99
Tabela 6.21	<i>Part-worths</i> agregados dos indivíduos “satisfeitos”	99
Tabela 6.22	<i>Part-worths</i> agregados dos indivíduos alvo	100
Tabela 6.23	Indivíduos alvo	100
Tabela 6.24	C ingressa no mercado	100
Tabela 6.25	Composição do mercado: novo cenário (A B e C)	100
Tabela 6.26	<i>Part-worths</i> agregados dos indivíduos “satisfeitos”	101
Tabela 6.27	<i>Part-worths</i> agregados dos indivíduos alvo	101
Tabela 6.28	Primeira reação competitiva	101
Tabela 6.29	Composição do mercado (1 ^{ra} reação)	102
Tabela 6.30	Segunda reação competitiva	102
Tabela 6.31	Composição do mercado (2 ^{da} reação)	102

INTRODUÇÃO

Em uma era em que o conhecimento é fator crítico para o sucesso, a pesquisa de mercado pode ser uma das principais geradoras desse conhecimento. Hoje, existe consenso entre executivos e acadêmicos sobre a importância de desenvolver capacidades superiores de aprendizado organizacional como fonte de vantagem competitiva [Kohli e Jaworski 1990; Nonaka 1991; Slater e Narver 1995; Stata 1989]. Garvin (1993) define as organizações que aprendem como “...organizações com capacidade de criar, adquirir e transferir conhecimento, e modificar sua conduta para revelar novos conhecimentos e percepções”. Diferentes autores relacionam inteligência com a capacidade da organização em transformar informação em mudanças organizacionais [Glazer 1991; Moorman 1995].

Acredita-se amplamente, independentemente da estratégia do negócio, na relevância que tem a geração de inteligência focalizada no mercado. Tem-se tornado uma visão convencional que habilidades para continuamente gerar inteligência sobre as necessidades expressadas e latentes dos consumidores, e sobre como satisfazer essas necessidades, é essencial para criar, em forma continuada, valor superior para o consumidor [Day 1994a]. Segundo Day (1994b), organizações orientadas para o mercado desenvolvem novas inteligências sobre a demanda do mercado e sobre como atendê-la melhor através de habilidades superiores que as capacite para continuamente entender e agir em resposta a eventos e tendências dos seus mercados.

Diferentes pesquisadores fornecem evidências empíricas sugerindo que o aprendizado focalizado no mercado é estreitamente relacionado ao sucesso de novos produtos [Lynn, 1998] e ao crescimento das vendas [Atuahene-Gima 1995; Slater e Narver 1995].

Nas empresas baseadas no conhecimento e orientadas para o mercado, a finalidade da pesquisa de mercado é desenvolver conhecimento sobre o mercado e dessa forma gerar valor para a organização. Mahajan e Wind (1999) declaram que o maior potencial da pesquisa de

marketing não é apenas o uso efetivo de técnicas, nem a abordagem integradora, nem o uso de novas tecnologias de informação. Estes defendem que o maior potencial é assistir à gerência a melhor formular as questões certas para ajudar a empresa entender seus clientes, identificar novas oportunidades, e dar forma à direção estratégica da organização. E é por isto que, a pesquisa de marketing precisa ser reconhecida como parte central do processo de criação de conhecimento da organização.

Pesquisadores de mercado, para fornecer inteligência de mercado para alta gerência, precisam atender os objetivos da empresa com resultados úteis e válidos, e comunicar estes resultados de forma efetiva. Uma das etapas críticas para superar este desafio de marketing é a adequada mensuração e consideração da estrutura de preferências do mercado, que possibilita a compreensão dos fatores mais importantes que influenciam a escolha dos produtos e serviços ofertados pelas organizações. Este é o assunto onde esta tese se insere e efetua suas contribuições.

Tipicamente, o problema da mensuração da estrutura de preferências do mercado tem sido abordado mediante a aplicação de modelos de escolha discreta [Louviere *et al*, 2000] que empregam técnicas estatísticas agregadas, isto é, não estimam funções utilidade para cada indivíduo. Consequentemente, assumem que todos os consumidores possuem a mesma estrutura de preferência pois seus resultados apenas fornecem a tendência central dos dados, sem capturar elementos importantes de heterogeneidade entre os consumidores. Os modelos agregados não atendem em forma satisfatória as necessidades do marketing atual, necessidades que estão intimamente relacionadas com prognósticos de escolhas individuais. É evidente [Pine, 1993; Roberts, 2000] a acelerada mudança no comportamento dos mercados atuais e a conseqüente necessidade do marketing em acompanhar esta evolução. O marketing atual demanda por conhecimento desagregado do mercado, isto é, a identificação das preferências de consumidores individuais assim como também da distribuição destas preferências na população [Allenby e Rossi, 1999; Huber e Train, 2000].

Por outro lado, as vantagens da modelagem de escolha discreta são únicas. Técnicas de pesquisa de mercado cujo mecanismo de obter informação representa diretamente o comportamento do mercado, produzem resultados com maior validade e confiabilidade que aquelas que representam o comportamento através de formas indiretas. Em outras palavras, independentemente do processo de compra de produtos e serviços utilizado pelo consumidor,

em última instância a saída deste processo é a escolha; e é em base à observação desta que a modelagem de escolha discreta se fundamenta.

Esta tese se propõe à indagação de técnicas estatísticas desagregadas, que forneçam funções individuais de utilidade baseadas em escolha para obter conhecimento das preferências, e as formas de empregar este conhecimento sob perspectivas gerencial e estratégica. A hipótese fundamental é que a aplicação do conhecimento das preferências individuais dos consumidores para resolução de problemas estratégicos de marketing, junto a uma forma efetiva de comunicar os resultados, fornece elementos de apoio à decisão gerencial agregando assim valor superior à organização.

1.1 Objetivos

Esta tese tem como objetivo geral apresentar uma metodologia que gera conhecimento do comportamento individual dos consumidores, e utiliza este conhecimento para auxiliar às organizações na identificação de oportunidades de mercado e na formulação e avaliação de estratégias alternativas de marketing para melhor aproveitar essas oportunidades.

Especificamente, esta tese tem os seguintes objetivos:

- Recuperação da heterogeneidade das preferências do mercado mediante um modelo bayesiano hierárquico, baseado em escolhas ordenadas, para inferência de parâmetros individuais.
- Introdução de um teste formal de convergência para o processo estocástico de amostragem por Cadeias de Markov Monte Carlo (MCMC).
- Construção de um mercado virtual para a obtenção de dados de escolhas, com a finalidade de testar a validade preditiva do modelo mediante simulação, empregando cenários *holdout* e critérios de desempenho preditivo.
- Formulação de um simulador de escolha para desenvolver análise competitiva, através de questões “*what if?*”.
- Desenvolvimento de um sistema interativo, assistido por computador, para a coleta dos dados.

1.2 A abordagem

Para atingir os objetivos propostos apresenta-se uma metodologia que se compõe de um modelo de inferência bayesiana hierárquica e um simulador de escolha. Esta abordagem reconhece que os consumidores não são uniformes mas sim heterogêneos, o que justifica a concentração dos esforços estratégicos de marketing, dentro de um contexto competitivo, em componentes específicos e claramente identificados de mercado.

Inicialmente, o modelo objetiva a estimação de parâmetros de preferências em nível individual. Este modelo assume que estes variam na população segundo uma distribuição contínua de probabilidade e o objetivo é a inferência dos parâmetros individuais e a distribuição que descreve a heterogeneidade das preferências na população. Esta inferência é viabilizada por intermédio de uma modelagem bayesiana hierárquica do problema, que agrega a informação proveniente da amostra para completar a escassa informação disponível em nível individual (escolhas), possibilitando desta forma o estudo das preferências de indivíduos específicos. A integração das densidades posteriores marginais dos parâmetros é efetuada utilizando métodos MCMC: Amostrador de Gibbs e Algoritmo Metropolis-Hastings. A essência desta primeira parte da abordagem é semelhante à abordagem de Allenby e Rossi (1999), ainda que as hipóteses assumidas sejam diferentes.

Os parâmetros estimados são utilizados para construir um simulador de escolhas ou de participação de preferências. Este simulador fornece uma estrutura útil que permite obter prognósticos de “demanda” a partir das preferências envolvidas nos parâmetros individuais. Este simulador é imune à propriedade de independência de alternativas irrelevantes (IIA) e outorga probabilidades finitas de escolha a todos os produtos considerados no cenário competitivo. O simulador desenvolvido aproveita a extensa informação obtida no processo de estimação: hiperparâmetros, que descrevem o comportamento em nível agregado, e vetores de parâmetros individuais definidos por suas respectivas distribuições de probabilidade, para prognosticar participação de escolha e medidas de “graus de satisfação” para produtos competitivos. Esta informação é a que em última instância permite identificar oportunidades específicas de mercado além de formular e avaliar estratégias alternativas de marketing.

1.3 Contribuição científica

Tanto o modelo estatístico para inferência de parâmetros em nível individual, quanto o simulador de participação de escolhas, desenvolvidos nesta tese, possuem características próprias não encontradas na literatura. O modelo de inferência, sob uma abordagem bayesiana hierárquica, descreve o comportamento individual de escolha assumindo um modelo logit multinomial ordenado. O simulador, para prognosticar participação de escolha, utiliza a variabilidade intrínseca do processo individual de escolha por intermédio das distribuições marginais posteriores de cada parâmetro, desta forma, é imune perante as dificuldades da hipótese de independência de alternativas irrelevantes (IIA) sem se contrapor com a teoria da utilidade aleatória.

1.4 Limitações do estudo

Antes que tudo, é preciso considerar que o modelo mensura a utilidade total por intermédio de considerações probabilísticas. Por conseguinte, é impossível prognosticar escolha com certeza se as próprias utilidades não foram mensuradas exatamente.

O processo de escolha em um mercado real envolve um alto grau de incerteza e ruído, que não é considerado neste modelo. Fatores geradores de ruído, que interferem ou modificam a simples preferência do consumidor entre “produtos” definidos explicitamente, podem proceder de diferentes fontes, como por exemplo, campanhas de marketing, disponibilidade do produto, conhecimento do consumidor sobre o produto e seus atributos, comprometimento do consumidor com o processo de compra. Por exemplo, em uma categoria de produtos de compra freqüente muitos dos entrevistados podem não ter uma função de utilidade definida e preferir mudar de produto pelo simples fato de mudar.

Uma suposição básica do modelo é que o mercado é estático e em equilíbrio. Os prognósticos do simulador não explicam a disseminação da informação ao longo do tempo, mudança de características ou percepções do produto, ou o fato que o mercado precisa tempo para tender a operar no seu curso. Esta limitação estatística pode se tornar crítica em mercados extremamente dinâmicos, ou na avaliação de inovações descontínuas: produtos realmente novos que empregam tecnologias novas.

No entanto, é preciso destacar que estas limitações não atingem unicamente o modelo aqui proposto, mas também a todas as metodologias de pesquisa de marketing atualmente consolidadas que utilizam dados de preferência declarada. Por outro lado, em plena concordância com Louviere *et al* (2000), as organizações precisam a melhor informação possível em relação a resposta do mercado às estratégias de marketing e, desde a aparição das metodologias conjuntas do final dos anos 60 até hoje, muitas organizações de classe mundial tem-se baseado em algum tipo de dados de preferência declarada para preencher essa necessidade. Deveria ser óbvio, portanto, que estas metodologias não teriam persistido, nem continuariam sendo aperfeiçoadas, se as organizações não tivessem encontrado valor real na sua utilização.

Depreende-se que a questão de fundo, não é o questionamento da filosofia da metodologia conjunta, mas a indagação de se o modelo proposto amplia as capacidades dos modelos atualmente existentes, visando agregar maior valor ao processo de marketing estratégico das organizações.

1.5 Estrutura do trabalho

O Capítulo II aborda o problema da mensuração da estrutura de preferências do mercado, principalmente com modelos de escolha discreta. Descrevem-se as vantagens e desvantagens apresentadas pelos modelos agregados de escolha e as linhas atuais de pesquisa que visam superar as limitações.

O Capítulo III apresenta conceitos teóricos básicos de inferência bayesiana. São ressaltadas, mediante um exemplo, as dificuldades encontradas pela abordagem bayesiana para computar numericamente as integrais requeridas para obter as distribuições posteriores. Na sequência, são apresentados dois métodos de simulação de cadeias de Markov Monte Carlo (MCMC): o Amostrador de Gibbs e o Algoritmo Metropolis-Hastings, que possibilitam superar ditas dificuldades. Finalmente, modelos bayesianos hierárquicos são discutidos.

O Capítulo IV apresenta a abordagem proposta. Na primeira parte, ênfase é dada à inferência dos parâmetros individuais, parte central desta metodologia. Um procedimento de diagnóstico de convergência é apresentado. Dois índices de desempenho preditivo: *Hit Rate* e *Mean Absolute Error* (MAE), para validação cruzada em cenários *holdout*, são definidos para

avaliar a capacidade preditiva do modelo. Na segunda parte do capítulo, discutem-se conceitos em relação a simuladores de escolha e finalmente o simulador desenvolvido para prognosticar participação de preferências é apresentado.

O Capítulo V aborda a validação do modelo, mediante simulação. É descrito inicialmente um algoritmo para a construção de mercados virtuais. Objetiva-se neste capítulo, além de testar a capacidade do modelo na recuperação de parâmetros conhecidos, a avaliação da validade preditiva dos parâmetros estimados. Foram abordadas duas situações simuladas, uma assumindo um mercado homogêneo e outra um mercado com dois segmentos claramente diferenciados.

O Capítulo VI apresenta uma aplicação empírica da abordagem proposta, inserida num contexto de marketing educacional. Cursos de pós-graduação em marketing são visualizados como produtos, e preferências de estudantes são estudadas visando incorporá-las no projeto das ofertas educacionais. É apresentado um sistema baseado em computador desenvolvido para a coleta dos dados. Finalmente, o processo de identificação de oportunidades e a formulação e avaliação de estratégias de marketing é exemplificado através da exploração de possíveis cenários competitivos.

Para finalizar, o capítulo VII apresenta um resumo onde são ressaltadas as principais contribuições e limitações da tese, juntamente com recomendações para futuros trabalhos. Uma seção de referências bibliográficas complementa o texto.

A MENSURAÇÃO DA ESTRUTURA DE PREFERÊNCIAS DO CONSUMIDOR

Diversos fatores influenciam a decisão de compra ou escolha de produtos/serviços. Muitos destes fatores compõem a estrutura de preferências dos consumidores, e a adequada mensuração desta estrutura pode dar resposta a questões estratégicas do marketing e da própria organização.

Este capítulo apresenta os modelos utilizados para a mensuração das preferências do mercado. Maior ênfase é dada aos modelos agregados¹ baseados em escolha, discutindo principalmente suas vantagens e desvantagens. Finalmente, linhas recentes de pesquisa que objetivam superar as limitações da análise agregada são apresentadas.

2.1 Origem

As origens das técnicas hoje utilizadas para o estudo e análise das preferências e escolha dos consumidores remontam-se a Luce e Tukey, um psicólogo matemático e um estatístico, que desenvolveram o primeiro trabalho que motivou o interesse na área de marketing. Estes pesquisadores estiveram inicialmente interessados nas condições sob as quais existem escalas de medição para as variáveis dependentes e independentes, quando apenas é dado: 1^o) informação da ordem dos efeitos conjuntos das variáveis independentes e 2^{do}) uma regra de composição preestabelecida. Eles denominaram sua abordagem *Conjoint Measurement* [Luce e Tukey, 1964]. *Conjoint Measurement* forneceu uma teoria para criar

¹ Nesta tese distingue-se entre os modelos que fornecem apenas um único vetor de parâmetros comuns a toda a população (modelos agregados) daqueles que possibilitam a identificação de parâmetros em nível individual (modelos desagregados).

escalas de medida a partir de preferências sobre objetos conjuntos, gerando grande interesse na época que foi proposta².

No entanto, pesquisadores em marketing, interessados em modelos e técnicas que enfatizassem a transformação de respostas subjetivas -a estímulos objetivos- em parâmetros estimados, utilizaram as formas funcionais propostas pela teoria de *Conjoint Measurement* para relacionar a utilidade do produto aos valores dos atributos do produto e, adotaram o nome de *Conjoint Analysis* para ressaltar esta distinção. Green e Srinivasan (1978) definiram-na como "... qualquer método decomposicional que estima a estrutura das preferências dos consumidores baseado nas avaliações globais de um conjunto de alternativas pre-especificadas em termos de níveis de diferentes atributos".

2.2 Principais aplicações e importância

Desde a aparição do primeiro artigo sobre a aplicação da metodologia conjunta ao estudo da conduta do consumidor [Green e Rao, 1971], os métodos baseados em preferências declaradas tem se tornado a metodologia mais amplamente aplicada para medir e analisar preferências do consumidor. Carroll e Green (1995) e Louviere *et al.* (2000) resumem os atuais desenvolvimentos das técnicas; Wittink e Cattin (1989) pesquisam várias aplicações comerciais nos EEUU e Wittink *et al.* (1994) na Europa.

O fato destas técnicas possuírem a capacidade de responder a uma ampla gama de questões de marketing, tem originado uma quantidade crescente de aplicações práticas. Gerentes de marketing se defrontam com inúmeras dificuldades quando avaliam lucros futuros, vendas, e participação de mercado para novos produtos ou para modificações de produtos existentes ou para diferentes estratégias de marketing. Estas técnicas podem dar resposta a, por exemplo:

1. Prognosticar a lucratividade e/ou a participação de mercado para um novo produto dada a oferta atual dos concorrentes. Deve o novo produto ser introduzido? Em caso positivo, qual a configuração ótima para este?

² Huber (1987) declara que foi visualizada como a "*psychometric conjurer's stone*", uma forma de "transformar a matéria impura das preferências ordinais no ouro das escalas intervalares". Pretendia-se, mediante o estudo da relação funcional entre estímulos de múltiplos atributos e o valor subjetivo percebido pelo indivíduo, colocar a medição de preferências à par com a medição nas ciências exatas.

2. Prognosticar o impacto de novos produtos dos concorrentes sobre o lucro ou a participação de mercado se a empresa não efetuar mudanças na sua posição competitiva. Quais as possíveis reações competitivas? Deve-se modificar o preço ou outro atributo dos produtos da empresa em resposta à competição?
3. Prognosticar deslocamentos de consumidores ou dos produtos atuais da empresa para novos produtos oferecidos por esta (canibalismo), ou de produtos dos concorrentes para o novo produto da empresa.
4. Prognosticar em forma diferenciada, resposta aos itens 1-3, em termos de segmentos de mercado.
5. Prognosticar reações competitivas às estratégias da empresa de introdução de novos produtos.
6. Prognosticar respostas dos segmentos perante planos alternativos de marketing.

Gustafsson *et al.* (1999), indicam que recentemente estas técnicas tem sido introduzidas como uma ferramenta para dar suporte à utilização da Função de Desdobramento da Qualidade (QFD) [Griffin e Hauser, 1993]. Adicionalmente, *Conjoint Analysis* tem sido incluída entre as sete ferramentas de planejamento de produto (*7 PP tools*)³.

2.3 Uma classificação

Nas três últimas décadas, principalmente em relação à quantificação das preferências do mercado por produtos/serviços/conceitos, a pesquisa de marketing tem-se beneficiado de uma série de valiosas contribuições de outras ciências como: econometria, estatística, psicometria e transporte. Isto tem propiciado o surgimento de uma ampla gama de modelos, ferramentas e técnicas para a mensuração de estruturas de preferências com base em dados de preferência declarada PD -conduta esperada em relação a opções hipotéticas- visando superar, assim, as limitações de se trabalhar apenas com dados de preferência revelada PR -conduta observada em relação a opções reais. Louviere *et al.* (2000) declaram que enquanto dados PR descrevem contextos de decisão reais, sem permitir controle experimental dos atributos dos produtos (restrições tecnológicas são fixadas), dados de PD descrevem contextos de decisão

³ *7 PP tools* é um *toolbox* desenvolvido pela *Union of Japanese Scientists and Engineers (JUSE)* para destacar as ferramentas e técnicas úteis no processo de desenvolvimento de novos produtos.

hipotéticos ou virtuais, permitindo maior flexibilidade para mapear funções de utilidade com tecnologias diferentes das existentes. Por outro lado, dados PR possuem boa confiabilidade enquanto que a confiabilidade dos dados PD está condicionada à compreensão e comprometimento do entrevistado com a tarefa de escolha.

Green e Srinivasan (1990) propõem (Figura 2.1) uma taxonomia das diferentes abordagens.

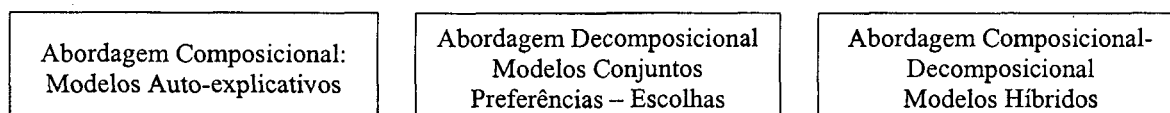


Figura 2.1 - Abordagens para a mensuração da estrutura de preferências declaradas.
Fonte: Green e Srinivasan (1990). Adaptado.

Na abordagem composicional, o consumidor avalia os níveis de cada atributo em forma seqüencial e isolada dos outros atributos. Partindo-se das avaliações em separado dos atributos e seus níveis, deseja-se conhecer a preferência global por um estímulo (produto/serviço/conceito).

A abordagem decomposicional objetiva estimar a importância de cada atributo (utilidades dos níveis dos atributos) a partir da avaliação do efeito conjunto do estímulo. Desta forma, mediante a avaliação de combinações de níveis, quer-se conhecer a contribuição de cada atributo na formação da preferência do entrevistado.

A abordagem híbrida, formada pela união dos modelos de preferência composicionais e decomposicionais, envolve duas tarefas: uma que submete ao entrevistado à avaliação de atributos independentemente dos outros e outra, posterior, onde este desenvolve um exercício conjunto.

2.4 Os modelos conjuntos

Entre os métodos de medição da estrutura de preferências do consumidor apresentados na Figura 2.1, os modelos conjuntos podem ser visualizados como aquelas abordagens que estão baseadas em preferências declaradas em vez de reveladas, e que submetem os consumidores a considerar simultaneamente diferentes atributos do produto.

Na abordagem decomposicional dois métodos principais tem sido sugeridos: uma abordagem baseada em preferências, que precisa que o entrevistado classifique cada alternativa numa escala métrica (*rating*) e outra, baseada em escolha, em que o entrevistado escolhe uma entre diferentes alternativas. Ambos os métodos operacionalizam a mensuração das preferências a partir do princípio comportamental denominado “maximização da utilidade” [Ben-Akiva e Lerman, 1985].

Modelos Conjuntos Baseados em Preferências: *Conjoint Analysis*

A hipótese fundamental de *Conjoint Analysis* quando aplicado a pesquisa de mercado, é que a avaliação de diferentes alternativas, produtos/serviços/conceitos, estão baseadas em valores subjetivos métricos de utilidade alocados a cada uma daquelas alternativas.

Um outro postulado básico de *Conjoint Analysis* é que o valor de utilidade atribuído para cada produto é uma composição dos valores associados a cada nível dos atributos ou características incluídas no produto.

A regra de composição, que permite agregar a utilidade associada a cada atributo para fornecer o valor do produto, é a parte central na aplicação de *Conjoint Analysis*. Esta regra é usualmente conhecida como o modelo de preferência, expressada através de uma função de utilidade. Assumindo um determinado modelo de preferência, *Conjoint Analysis* objetiva estimar a importância das combinações atributos-níveis a partir de medidas explícitas de preferência, em uma escala métrica, para cada um dos diferentes produtos alternativos.

Modelos Conjuntos Baseados em Escolhas: *Experimental Choice Analysis*

Esta outra abordagem, intimamente relacionada com análise de *trade offs* -trocas compensatórias-, originou-se na área de pesquisa de transporte sob a denominação *Discrete Choice Analysis*, que utiliza dados de preferência revelada [McFadden 1974, Ben-Akiva e Lerman 1985]. McFadden (1986) provê uma revisão das técnicas baseadas em escolha para a medição da utilidade, incluindo sua relação com *Conjoint Analysis*.

Segundo Ben-Akiva e Lerman (1985), o problema básico abordado pela análise de escolha discreta é a modelagem da escolha de um conjunto de alternativas mutuamente

exclusivas e coletivamente exaustivas. Esta abordagem utiliza o princípio de maximização da utilidade. Brevemente, o decisor é modelado escolhendo a alternativa com a maior utilidade entre as alternativas disponíveis no momento em que é feita a escolha. O modelo operacional consiste de funções parametrizadas de utilidade em termos de variáveis independentes observáveis e parâmetros desconhecidos, e seus valores são estimados de uma amostra observada de escolhas feitas pelo decisor quando confrontado com uma decisão de escolha.

Modelos multinomiais logit e probit são tipicamente empregados em análises de escolha discreta de dados de preferência revelada. Na maioria dos casos estes modelos são aplicados em nível agregado com dados não experimentais; portanto, uma grande quantidade de observações é necessária para estimação estável dos parâmetros.

Posteriormente, uma versão de *Discrete Choice Analysis*, denominada *Experimental Choice Analysis*, foi proposta [Louviere e Woodworth, 1983; Mahajan *et al.*, 1982]. Esta abordagem utiliza a metodologia de delineamento experimental de *Conjoint Analysis* para obter opções hipotéticas de escolha. No entanto, a metodologia difere de *Conjoint Analysis* no tipo de variável dependente -discreta- e nas técnicas de estimação -máxima verossimilhança-. Em *Experimental Choice Analysis* o entrevistado escolhe a opção preferida do conjunto.

Assim como a análise de preferências reveladas, *Experimental Choice Analysis* utiliza modelagem de escolha probabilística (ex. o modelo logit multinomial nas suas diferentes formas) para estimar parâmetros em nível agregado [Batsell e Louviere, 1991].

Esta última abordagem tem recebido especial atenção na área de pesquisa de transporte, onde é geralmente conhecida como *Stated Preferences Techniques* [Louviere, 1988; Kroes e Sheldon, 1988; Bates, 1988; Morikawa, 1989; Jones, 1991; Davison, 1991; Novaes e Carvalho, 1996].

2.5 Teoria da escolha probabilística

Modelos baseados em maximização da utilidade que fornecem prognósticos determinísticos são denominados modelos determinísticos de utilidade. Segundo esta teoria, não existe incerteza nas escolhas prognosticadas, assume-se que o indivíduo escolhe a alternativa com a maior utilidade e, de acordo com o modelo, não existe a possibilidade de quaisquer outra alternativa ser escolhida.

Na prática, no entanto, é freqüente que indivíduos aparentemente similares efetuem diferentes escolhas quando se deparam com alternativas semelhantes ou inclusive idênticas. De fato, um indivíduo poderia fazer diferentes escolhas quando defrontado com as mesmas alternativas em ocasiões diferentes. Duas limitações de ordem prática tornam os modelos determinísticos inadequados: por um lado, os decisores podem não ter conhecimento exato dos atributos das alternativas disponíveis; e por outro, o analista possui informação limitada devido a omissão de variáveis relevantes, erros de medição e diferenças entre indivíduos.

Por conseguinte, o analista sempre estima a utilidade com erro; um termo de erro de tamanho desconhecido está sempre presente na especificação da função utilidade. Considera-se este erro como uma variável aleatória cujos valores são descritos por uma distribuição de probabilidade. A função utilidade é também uma variável aleatória, composta pela soma de um componente determinístico e um componente aleatório.

A argumentação dos defensores dos modelos probabilísticos é que devido ao fato que as utilidades das alternativas são variáveis aleatórias, não é possível declarar com certeza qual alternativa tem a maior utilidade ou qual delas será escolhida. O máximo que o analista pode fazer é prognosticar a probabilidade de uma alternativa ter a máxima utilidade e, por conseguinte, a probabilidade com a qual será escolhida. Desta forma, o analista representa a conduta de escolha mediante modelos probabilísticos [Luce 1977; Amemiya 1981; Louviere 1984; McFadden 1974, 1980, 1984].

2.6 Formulação básica dos modelos agregados para escolha discreta

Diferentes modelos emergem das hipóteses assumidas em relação à distribuição probabilística do erro. Dentro de uma estrutura geral de modelos de probabilidade, define-se:

$$\text{Prob}(\text{escolher a alternativa } j) = \text{Prob}(y = j) = F [\text{parâmetros, atributos}] \quad (2.1)$$

Apenas por clareza classificam-se os modelos em duas categorias: de escolha binária (duas alternativas) e múltipla (mais de duas alternativas). Adicionalmente, os modelos de escolha múltipla abrangem modelos não-ordenados (o entrevistado efetua uma única escolha) e ordenados (o entrevistado efetua uma ordenação total ou parcial do conjunto de escolha).

2.6.1 Modelos de escolha binária

No caso binário, sejam N variáveis y_j distribuídas Bernoulli com probabilidade de sucesso p_j relacionada com um conjunto de variáveis independentes (atributos) que podem ser contínuas e/ou discretas. Define-se o modelo de regressão como $p_j = F(\beta^T x_j)$ com $j = 1, \dots, N$; onde β é um vetor de parâmetros desconhecidos, x_j é a variável independente (vetor de atributos); e F é uma função de distribuição de probabilidade (cdf) conhecida que relaciona as probabilidades p_j com a estrutura linear $\beta^T x_j$.

Se a função *link* F , é a cdf normal padrão (Φ) deriva-se o modelo probit. Esta hipótese de visualizar o erro como a soma de uma grande quantidade de componentes não observados, porém independentes, é lógica, já que pelo teorema central do limite a distribuição destes tende a ser normal [Ben-Akiva e Lerman, 1985]. Assim,

$$\text{Prob}(y = 1) = \int_{-\infty}^{\beta^T x} \phi(\varepsilon) d\varepsilon = \Phi(\beta^T x) \quad (2.2)$$

Se F é a cdf logística emerge o modelo logit,

$$\text{Prob}(y = 1) = \frac{e^{\beta^T x}}{1 + e^{\beta^T x}} = \Lambda(\beta^T x) \quad (2.3)$$

O modelo logit tem sido o mais utilizado pela sua conveniência analítica. O modelo probit, ainda que intuitivamente razoável e de forte base teórica, não possui uma forma funcional fechada dando lugar à necessidade de se expressar a probabilidade de escolha como uma integral.

É interessante comparar os modelos logit e probit com o modelo de regressão linear do tipo $y = \beta^T x + \varepsilon$. No modelo de regressão linear os coeficientes β são interpretados como a mudança na avaliação de uma alternativa em relação a uma alteração unitária no nível do atributo correspondente. Isto não é verdade para os modelos logit e probit [Greene, 1997]. Nestes, os efeitos marginais, ou seja, a mudança na probabilidade de escolha em relação a uma alteração unitária no nível do atributo, são expressados como:

$$\frac{\partial F(\beta^T x)}{\partial x} = \left\{ \frac{dF(\beta^T x)}{d(\beta^T x)} \right\} \cdot \beta = f(\beta^T x) \cdot \beta \quad (2.4)$$

onde $f(\cdot)$ é a função densidade de probabilidade (fdp) correspondente à distribuição acumulada $F(\cdot)$. Considerando a distribuição normal é:

$$\frac{\partial F(\beta^T x)}{\partial x} = \phi(\beta^T x) \cdot \beta \quad (2.5)$$

onde $\phi(\cdot)$ é a densidade normal padrão.

Para a distribuição logística,

$$\frac{d\Lambda(\beta^T x)}{d(\beta^T x)} = \frac{e^{\beta^T x}}{(1 + e^{\beta^T x})^2} = \Lambda(\beta^T x) \cdot [1 - \Lambda(\beta^T x)] \quad (2.6)$$

então, no modelo logit

$$\frac{\partial F(\beta^T x)}{\partial x} = \Lambda(\beta^T x) \cdot [1 - \Lambda(\beta^T x)] \cdot \beta \quad (2.7)$$

2.6.2 Modelos de escolha múltipla

Os modelos de escolha múltipla são uma generalização dos modelos binários, os indivíduos efetuam escolhas de um conjunto de três ou mais alternativas mutuamente exclusivas. Existem diferentes formas pelas quais analisar este problema [Greene, 1997]. Em geral, os conjuntos de escolha podem se classificar em dois tipos: não-ordenados e ordenados.

Modelos não-ordenados. Nos modelos não-ordenados de escolha o indivíduo i efetua uma escolha única j de um conjunto J . Isto implica que a utilidade U_{ij} é a máxima entre as J utilidades,

$$U_{ij} = \beta^T x_{ij} + \varepsilon_{ij} \quad (2.8)$$

onde U_{ij} é a utilidade da alternativa j para o indivíduo i ; β é um vetor de parâmetros; x_{ij} é o vetor de atributos da alternativa j para o indivíduo i ; e ε_{ij} é o erro.

Portanto, o modelo estatístico especifica a probabilidade que a escolha de j seja feita, isto é:

$$\text{Prob}(U_{ij} > U_{ik}) \quad \forall k \neq j \quad (2.9)$$

A operacionalização do modelo, da mesma forma que no caso binário, é efetivada pela escolha de uma distribuição para o erro. Como antes, dois modelos multinomiais tem sido considerados, logit e probit. Devido à necessidade de se avaliar integrais múltiplas da distribuição normal, o modelo probit tem tido uso limitado neste contexto. O modelo logit, pelo contrário, tem sido amplamente utilizado em muitos campos, economia, marketing, engenharia de transporte, etc.

McFadden (1974) tem demonstrado que se as J perturbações são idêntica e independentemente distribuídas (iid) Gumbel, $F(\varepsilon_{ij}) = e^{-e^{-\varepsilon_{ij}}}$, então:

$$\text{Prob}(y_i = j) = \frac{e^{\beta^T x_{ij}}}{\sum_{k=1}^J e^{\beta^T x_{ik}}} \quad (2.10)$$

Este é denominado modelo logit multinomial (MNL).

Modelos ordenados. Estes modelos de escolha tem sido utilizados para analisar respostas inerentemente ordenadas, ou *ranking*. Este é o caso quando o entrevistado, além da identificação da sua escolha preferida, fornece uma ordenação parcial ou completa das alternativas do conjunto de escolha [Bastos, 1994; Hensher, 1994].

Para explicar a informação obtida dessa ordenação é preciso relacionar o comportamento da ordenação com o comportamento da escolha. Chapman e Staelin (1982) declaram que a justificativa teórica para esta relação é fornecida pelo Teorema de Ranking de Escolha de Luce e Suppe. Este teorema declara que a ordenação de J alternativas é

equivalente a uma sequência de $J - 1$ escolhas independentes,

$$p(1,2,\dots,J) = p(1|\{1,2,\dots,J\}) \cdot p(2|\{2,\dots,J\}) \cdots p(J-1|\{J-1,J\}) = \prod_{j=1}^{J-1} p(j|\{j,\dots,J\}) \quad (2.11)$$

onde $p(1,2,\dots,J)$ é a probabilidade de observar uma ordenação na qual a alternativa 1 é preferida à alternativa 2, a alternativa 2 à 3, e assim sucessivamente, e $p(1|\{1,2,\dots,J\})$ é a probabilidade da alternativa 1 ser escolhida no conjunto $\{1,2,\dots,J\}$.

Sob a hipótese que as $J - 1$ escolhas independentes são representadas através do modelo MNL da expressão (2.10), a probabilidade da ordenação completa [Morikawa, 1989] é:

$$p_i(1,2,\dots,J) = \prod_{j=1}^J \frac{e^{x_j \beta_i}}{\sum_{a=j}^J e^{x_a \beta_i}} \quad (2.12)$$

Um aspecto fundamental do MNL é a hipótese de independência de alternativas irrelevantes (IIA). Esta declara que a probabilidade de escolher uma alternativa em relação a outra não é influenciada pelas outras alternativas do conjunto de escolha. Esta é uma propriedade restritiva que pode ser violada facilmente⁴, segundo Currim (1982), devido a sua falta de consideração de diferenças e similitudes entre os produtos alternativos. O modelo implica que, quando um novo produto ou serviço é introduzido, obtém participação de mercado em forma proporcional de todos os outros produtos e serviços, sem consideração de graus de substituição diferenciada entre estes.

⁴ Um exemplo, embora extremo, conhecido como “*red bus, blue bus problem*”, ilustra este problema. Supõe-se um mercado de transporte com dois produtos, automóveis e ônibus vermelhos, ambos com 50% de participação de mercado. Introduzindo agora um segundo ônibus, cor azul, um simulador com propriedade IIA prognosticará que o ônibus azul toma participações proporcionais do automóvel e do ônibus vermelho, de forma que a participação total de ônibus se tornará 67%. Porém, claramente é mais razoável esperar que o ônibus azul tome participação majoritariamente do ônibus vermelho e que a participação total dos ônibus permaneça perto de 50%.

2.7 Vantagens e desvantagens dos modelos agregados de escolha discreta

A principal vantagem exibida pelos modelos de escolha discreta em relação aos modelos conjuntos métricos *full profile* [Herman, 1988] e os modelos conjuntos híbridos [Johnson, 1987], e provavelmente a causa da sua crescente aceitação e uso na comunidade de marketing, é que esta abordagem é mais realista, expressando melhor o que os consumidores fazem quando avaliam e compram produtos ou serviços no mundo real. Os entrevistados expressam suas preferências simplesmente escolhendo opções de um conjunto, antes de ter que avaliá-las em alguma escala métrica [Huber 1992; Carroll e Green, 1995; Wyner, 1995].

Louviere *et al.* (2000), declaram que o modelo logit multinomial (MNL) continua a ser a estrutura preferida de modelagem de escolha para a grande maioria de usuários. Dentre as causas desta preferência, estes autores citam: a simplicidade na estimação (o modelo fornece uma solução única de parâmetros globalmente ótimos), forma fechada (não é necessária a avaliação de integrais complexas), velocidade de convergência, software acessível e fácil de utilizar, e robustez quando se dispõe de uma grande quantidade de dados.

Por outro lado, escolhas podem ser uma forma ineficiente de se obter informação das preferências, estas apenas indicam qual alternativa é preferida, sem fornecer informação adicional sobre a intensidade da preferência. Devido ao fato que escolhas fornecem pouca informação sobre cada entrevistado, dados de escolha tem sido analisados em nível agregado. Assume-se que todos os entrevistados são essencialmente similares e a heterogeneidade inerente às preferências individuais é confundida com ruído aleatório. Portanto, a diferença dos modelos conjuntos desagregados baseados em preferências, modelos agregados não permitem efetuar segmentação baseada em dados de escolha. Se existe um alto grau de heterogeneidade na amostra pesquisada, estes modelos produzem estimações viesadas⁵. Para superar esta deficiência, incluem-se características dos entrevistados (ex. variáveis sócioeconômicas) na função utilidade. Isto origina duas complicações: primeiro, incrementa-se a complexidade do modelo [Elrod *et al.*, 1992] e segundo, mais importante ainda, precisa-se conhecer a priori quais são as variáveis individuais que realmente discriminam entre segmentos de consumidores com preferências e escolhas diferentes. Esse conhecimento é

⁵ Modelos agregados podem fornecer resultados enganosos. Se metade dos entrevistados possuem uma forte preferência em relação a um atributo e a outra metade também, só que de forma oposta, a análise agregada reportará que este atributo é relativamente de pouca importância, quando de fato ele é extremamente importante só que de forma diferente para os distintos entrevistados.

geralmente não disponível, estas abordagens freqüentemente efetuam hipóteses questionáveis sobre a relação entre as variáveis que identificam o consumidor e as diferenças nas reações deste a diferentes estímulos de marketing [Rossi e Allenby, 1993]. Desta forma podem ser identificados segmentos que não variam em termos de reações comportamentais a estímulos de marketing.

Muitos dos problemas apresentados pelos modelos agregados baseados em logit (ex. MNL) obedecem a violações da propriedade IID das alternativas, que indica que os componentes de erro da utilidade aleatória de cada alternativa são independentemente e identicamente distribuídos. Esta propriedade, segundo Ben-Akiva e Lerman (1985), representa importantes restrições: primeiro, fixa o parâmetro de escala da distribuição do erro, como consequência todas as variâncias dos componentes aleatórios da utilidade precisam ser iguais, e segundo precisa que os erros sejam mutuamente independentes, isto é, não existe correlação entre quaisquer par de alternativas. Assim, a matriz de variância-covariância é diagonal e homocedástica, p. ex.:

$$\begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}$$

2.8 Relaxando a condição IID

Considerável esforço está sendo dedicado para incrementar o realismo comportamental dos modelos de escolha discreta. Este esforço, segundo Louviere *et al.* (2000), concentra-se em relaxar a forte condição IID de formas práticas, comportamentalmente plausíveis e computacionalmente viáveis. Destaca-se ainda que o grau de complexidade da estimação incrementa rapidamente conforme os modelos se afastam do MNL e crescentemente relaxam as hipóteses efetuadas sobre os elementos diagonais e fora da diagonal da matriz de variância-covariância.

McFadden (1981) propôs o modelo *Generalised Extreme Value* (GEV), que generaliza o modelo Multinomial Logit (MNL). O modelo não-IID mais popular e simples é o denominado Nested Logit (NL), que é um caso particular do GEV. O modelo NL relaxa a

condição IID entre subconjuntos de alternativas, porém preserva a condição IID dentro de cada subconjunto. Pode ser visualizado como um conjunto de modelos MNL conectados hierarquicamente por intermédio de um conjunto de relações condicionais. Por exemplo, para uma decisão de escolha em dois níveis, o modelo assume que existe uma matriz de variância-covariância diagonal e homocedástica associada à decisão de escolha genérica, e matrizes de variância-covariância diagonais e homocedásticas associadas a cada subconjunto de alternativas:

$$\begin{bmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_a^2 & 0 \\ 0 & 0 & \sigma_a^2 \end{bmatrix} \begin{bmatrix} \sigma_b^2 & 0 & 0 \\ 0 & \sigma_b^2 & 0 \\ 0 & 0 & \sigma_b^2 \end{bmatrix} \begin{bmatrix} \sigma_c^2 \\ \sigma_c^2 \end{bmatrix}$$

O modelo logit heterocedástico (HL) [Bhat, 1995] relaxa a hipótese de variância igual ou constante do modelo MNL. A diferença do modelo NL, não precisa a partição prévia do conjunto de escolha em subconjuntos mutuamente exclusivos. A matriz de variância-covariância tem elementos zero fora da diagonal e elementos diagonais únicos, ex:

$$\begin{bmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_b^2 & 0 \\ 0 & 0 & \sigma_c^2 \end{bmatrix}$$

O modelo Mixed Logit (ML) [Brownstone e Train, 1999] assume que a função utilidade é composta por uma componente determinística linear nos parâmetros, $\beta^T x$, por uma componente aleatória ε independente e identicamente distribuída Gumbel e por termos aleatórios adicionais agrupados aditivamente em η . Este último termo possibilita a identificação de heterocedasticidade e correlação nas alternativas. Admite-se qualquer distribuição f para η , definida por uma densidade $f(\eta | \Omega)$, onde Ω são os parâmetros que a caracterizam. A função utilidade para a alternativa j , denota-se:

$$U_j = \beta^T x_j + \varepsilon_j + \eta_j \quad (2.13)$$

Dado um valor fixado para η , a probabilidade de escolha condicional de que o indivíduo n escolha a alternativa j corresponde ao modelo Multinomial Logit (MNL), devido

aos erros remanescentes serem iid Gumbel:

$$P_n(j/\eta) = \frac{e^{(\beta^T x_j + \eta_j)}}{\sum_k e^{(\beta^T x_k + \eta_k)}} \quad (2.14)$$

Portanto, para calcular a probabilidade (marginal) de escolha da alternativa j é preciso integrar (2.14) em todos os possíveis valores de η :

$$P_n(j) = \int P_n(j/\eta) \cdot f(\eta|\Omega) \cdot d\eta \quad (2.15)$$

A probabilidade de escolha (2.15) do ML, a diferença do MNL, não possui uma forma fechada; como esta integral não pode ser resolvida analiticamente é necessário utilizar simulação [McFadden e Train, 2000]. Este modelo relaxa a condição IID própria do MNL, permitindo identificar correlação entre as alternativas.

O modelo Multinomial Probit (MNP), a diferença do ML, assume um único termo de erro aleatório com matriz de variância-covariância genérica. Deriva-se de assumir uma função utilidade:

$$U_j = \beta^T x_j + \varepsilon_j \quad \text{com } \varepsilon \sim N(0, \Sigma) \quad (2.16)$$

Louviere *et al.* (2000) declaram que, embora este seja um modelo genérico que pode relaxar totalmente a condição IID, demanda uma substancial complexidade computacional conforme o número de alternativas incrementa. Para calcular a verossimilhança para p escolhas é necessário integrar uma densidade normal de dimensão $p-1$. Técnicas de estimação para este modelo empregam integração Monte Carlo [McFadden, 1989] ou, mais recentemente, métodos bayesianos avançados [Geweke, 1991; Albert e Chib, 1993; McCulloch e Rossi, 1994].

O objetivo principal dos modelos anteriores tem sido relaxar a condição IID, mediante uma estimação mais precisa da matriz de variância-covariância. A saída desses modelos é um vetor de parâmetros agregados e uma matriz de variância-covariância que caracteriza a distribuição dos parâmetros na população.

Em marketing, principalmente para efetuar estudos de segmentação de mercado, modelos de Classe Latente, também denominados *finite mixture models*, tem sido aplicados. Estes modelos assumem que uma quantidade discreta de pontos é suficiente para representar a distribuição conjunta dos parâmetros [Kamakura e Russell, 1989; Chintagunta *et al.*, 1991; Chintagunta 1994]. Estes métodos, a diferença dos anteriormente apresentados, que estimam um único conjunto de parâmetros que melhor caracteriza a média de todos os entrevistados, buscam identificar grupos ou classes de entrevistados que compartilham preferências semelhantes, estimando utilidades (baseadas em logit) para cada classe. Os modelos não assumem que o entrevistado pertence a um determinado segmento ou classe, senão que consideram o consumidor como tendo alguma probabilidade finita de pertencer a cada segmento.

Nessa abordagem, Kamakura e Russell (1989) propõem um modelo probabilístico de escolha, que utiliza dados de preferência revelada (compras observadas) para classificar utilidades de marcas e probabilidades de compra em um número pequeno de segmentos de preferências. DeSarbo *et al.* (1995) generalizam a metodologia anterior para modelos de escolha discreta baseados em classe latente, considerando replicações dos indivíduos sobre conjuntos de escolha.

Allenby e Rossi (1999) declaram que, embora os modelos de Classe Latente possam aproximar as tendências centrais da distribuição mista dos parâmetros, existe crescente evidência sugerindo que *finite mixture models* com pequena quantidade de pontos podem não capturar adequadamente o grau completo de heterogeneidade nos dados.

2.9 Na procura de *part-worths* individuais

Para caracterizar as preferências do mercado de formas mais precisas e aproveitar este conhecimento na identificação de ações lucrativas focalizadas no consumidor, pesquisadores tem-se voltado a procurar formas confiáveis de estimar parâmetros de preferências em nível individual [Allenby e Rossi, 1999; Huber e Train, 2000].

Rossi e Allenby (1993), para estimar parâmetros em nível individual, propõem um método bayesiano que adiciona informação às escolhas individuais através de uma distribuição prévia. Os parâmetros desta distribuição prévia são determinados em nível

agregado, empregando os dados de escolha da amostra. Seja $L_i(\theta | y_i)$ a verosimilhança para o entrevistado i e $p(\theta)$ a distribuição dos *part-worths* em nível agregado, aplicando o teorema de Bayes [seção 3.2] obtém-se a distribuição posterior dos *part-worths* do indivíduo i :

$$p_i(\theta | y_i) \propto l_i(\theta | y_i) \cdot p(\theta) \quad (2.17)$$

Observa-se que, através de um processo de atualização bayesiana, os dados de cada entrevistado “movem” os parâmetros agregados da distribuição prévia para formar o parâmetro individual.

Mais recentemente, em marketing tem sido propostos modelos bayesianos hierárquicos que combinam informação da distribuição agregada das preferências com as escolhas em nível individual para obter os parâmetros individuais. Lenk *et al.* (1996) demonstraram analítica e empiricamente que modelos bayesianos hierárquicos não precisam que as matrizes de projeto em nível individual sejam de posto completo; isto possibilita utilizar poucas escolhas por entrevistado. Eles utilizaram dados conjuntos *full profile* (métricos) e determinaram que a análise bayesiana é capaz de estimar utilidades individuais razoáveis, inclusive quando a quantidade de informação disponível por entrevistado é inferior à quantidade de parâmetros estimados para cada indivíduo. Estes autores declaram: “A generalização desta estrutura bayesiana hierárquica para ACA, modelos híbridos e *conjoint analysis* baseado em escolha é desejável dada a crescente popularidade destes procedimentos conjuntos alternativos.” Allenby e Ginter (1995) e Allenby *et al.* (1998) estendem a abordagem bayesiana hierárquica para modelagem de escolha discreta.

Por outro lado, em economia tem sido desenvolvido um modelo que permite a estimação de parâmetros em nível individual empregando uma abordagem clássica. Revelt e Train (1999) propõem uma abordagem que, segundo estes autores, é semelhante em espírito à hierárquica bayesiana. Estima-se parâmetros agregados da população num contexto de *mixed logit* e posteriormente deriva-se a distribuição de cada entrevistado condicional nas escolhas observadas e na distribuição da população, empregando o teorema de Bayes. Este último procedimento é denominado pelos autores como *conditional of individual tastes* (COIT).

Huber e Train (2000) afirmam que os *part-worths* individuais estimados pelo modelo

COIT com *mixed logit* e pelos modelos hierárquicos bayesianos são equivalentes, sendo a escolha entre ambas metodologias regida pela conveniência de implementação e orientação filosófica em lugar de vantagens pragmáticas. No entanto, Brownstone (2000) afirma que as inferências da abordagem clássica são válidas apenas assintoticamente enquanto que a inferência bayesiana evita ditas aproximações. Da mesma forma, Allenby e Rossi (1999) denominam *approximate bayesian approach* às abordagens utilizadas em Rossi e Allenby (1993) e às que empregam um procedimento COIT e indicam que estas permitem obter apenas parâmetros individuais aproximados.

Depreende-se que a equivalência entre as abordagens bayesiana hierárquica e clássica COIT com *mixed logit* é uma questão ainda em discussão.

Esta tese insere-se na abordagem bayesiana para inferir parâmetros em nível individual. É proposto um modelo bayesiano hierárquico que, a diferença dos modelos apresentados por Allenby *et al.* (1998) e Allenby e Rossi (1999), utiliza dados de escolha ordenada para descrever o comportamento de escolha individual.

No próximo capítulo são apresentados conceitos de inferência bayesiana e algoritmos de simulação estocástica de cadeias de Markov, que fornecem embasamento teórico para a metodologia proposta.

A ABORDAGEM BAYESIANA PARA O PROBLEMA DA ESCOLHA DISCRETA

No capítulo anterior foram apresentados diferentes modelos de escolha discreta que baseiam-se na teoria probabilística da utilidade e assumem alguma distribuição de probabilidade para o erro.

Neste capítulo o problema inferencial das preferências do consumidor é tratado através de uma abordagem filosófica e metodologicamente diferente dos procedimentos clássicos da estatística: a abordagem bayesiana. As hipóteses assumidas pela modelagem bayesiana hierárquica, os métodos de simulação estocástica (Amostrador de Gibbs e Algoritmo Metropolis-Hastings) e os avanços de *hardware* possibilitam a inferência de parâmetros em nível individual utilizando reduzida quantidade de informação.

3.1 Inferência bayesiana

Inferência estatística aborda o problema de inferir propriedades de uma distribuição desconhecida a partir de dados gerados por aquela distribuição. O tipo mais comum de inferência implica a aproximação da distribuição desconhecida através da escolha de uma família restrita de distribuições, as quais são em geral, especificadas parametricamente.

A posição bayesiana é que a incerteza seja descrita probabilisticamente e seja calibrada através de resultados experimentais. Pereira e Viana (1988), destacam a simplicidade conceitual da abordagem: para uma quantidade desconhecida de interesse, θ , conhecendo-se H , calcula-se as probabilidades condicionais de θ dado H utilizando as regras do cálculo de probabilidades. Esta receita está, na verdade, encobrindo um processo sequencial. Se H

representa o conhecimento do analista até um determinado momento, no instante seguinte deverá conter também as informações adicionais que ele poderia estar adquirindo naquele momento. O objetivo da inferência bayesiana é justamente estudar as formas com as quais os resultados experimentais, adicionados a H , alteram $p(\theta | H)$. Da mesma forma, Greene (1997) indica que o processo da estimação da metodologia bayesiana não é a dedução de valores de parâmetros fixados mas a de continuamente atualizar e dar forma às próprias crenças subjetivas do estado do mundo.

Gelman *et al* (1995) definem a análise bayesiana de dados como um conjunto de métodos práticos para efetuar inferências utilizando modelos probabilísticos para quantidades observadas e para quantidades sobre as quais deseja-se aprender. A característica essencial destes métodos é seu uso explícito da probabilidade para quantificar a incerteza nas inferências baseadas em análise estatística de dados. Estes autores declaram que o processo de análise bayesiana de dados pode ser decomposto nas três etapas seguintes:

1. Definição de um modelo completo de probabilidade, isto é, uma distribuição de probabilidade conjunta para todas as quantidades observáveis e não observáveis do problema¹. O modelo deve ser consistente com o conhecimento sobre o problema científico básico e o processo de coleta de dados.
2. Condicionamento sobre os dados observados: computação e interpretação da distribuição posterior apropriada, ou seja, a distribuição de probabilidade condicional das quantidades não observadas dadas as quantidades observáveis.
3. Avaliação do ajuste do modelo e das implicações da distribuição posterior resultante.

3.2 Regra de Bayes

A diferença fundamental entre a estatística clássica e a bayesiana é que esta última considera que os parâmetros são variáveis aleatórias. A inferência bayesiana está baseada na distribuição posterior dos parâmetros (θ) condicional nos dados observados (y), $p(\theta | y)$.

¹ Quantidades observáveis são os dados observados. Quantidades não observáveis, são aqueles para os quais inferência estatística é efetuada. Estes podem ser de dois tipos: 1^o quantidades potencialmente observáveis tais como observações futuras de um processo e 2^{do} quantidades que não são diretamente observáveis, isto é, os parâmetros que governam o processo hipotético que produz os dados observados, por ex. os coeficientes de regressão [Gelman *et al.*, 1995].

Para efetuar afirmações probabilísticas sobre $p(\theta | y)$, a abordagem bayesiana precisa de um modelo que forneça a distribuição de probabilidade conjunta para θ e y , $p(\theta, y)$ [Gelman *et al*, 1995]. Esta distribuição é especificada como o produto de duas densidades: a distribuição prévia $p(\theta)$, mediante a qual bayesianos expressam seu conhecimento prévio sobre θ antes de observar os dados, e a distribuição de amostragem ou função de verossimilhança $p(y|\theta)$. Assim,

$$p(\theta, y) = p(\theta) \cdot p(y|\theta) \quad (3.1)$$

De acordo com operações usuais de funções de distribuição de probabilidades, obtém-se a regra de Bayes:

$$p(\theta|y) = \frac{p(\theta) \cdot p(y|\theta)}{p(y)} \quad (3.2)$$

onde $p(y) = \sum_{\theta} p(\theta) \cdot p(y|\theta)$ é a soma sobre todos os valores possíveis de θ , ou

$p(y) = \int_{\theta} p(\theta) \cdot p(y|\theta) \cdot d\theta$ no caso de θ contínuo. Devido a que o fator $p(y)$ não depende de θ e,

para y fixado, pode ser considerado uma constante, pode-se rescrever (3.2) como:

$$p(\theta|y) \propto p(\theta) \cdot p(y|\theta) \quad (3.3)$$

O lado direito da expressão (3.3) é a densidade posterior não-normalizada ou *kernel*² da função densidade de probabilidade (fdp) posterior condicional de θ em y . Estas simples expressões resumem a essência da inferência bayesiana. A tarefa fundamental de qualquer

² A avaliação direta da expressão (3.2) é usualmente difícil já que a constante $p(y)$ não pode ser determinada analiticamente. Exceto para uma pequena quantidade de modelos simples, é necessário utilizar métodos numéricos para obter a densidade posterior. No entanto, esta dificuldade pode ser superada sem precisar conhecer $p(y)$. Qualquer função proporcional a $p(\theta|y)$ é uma densidade *kernel* posterior para θ . (Em geral, qualquer função proporcional a uma densidade de probabilidade é uma densidade *kernel*.) A partir de uma densidade *kernel*, é frequentemente fácil identificar a distribuição da variável aleatória em questão. Para evitar ambiguidade (3.2) é denominada densidade posterior normalizada: $\int_{\theta} p(\theta|y) \cdot d\theta = 1$, enquanto que, em geral em

(3.3) verifica-se: $\int_{\theta} p(\theta) \cdot p(y|\theta) \cdot d\theta \neq 1$.

aplicação específica é desenvolver o modelo $p(\theta, y)$ e realizar os cálculos necessários para representar $p(\theta | y)$ de forma apropriada.

Toda a inferência sobre θ vem da inspeção da densidade posterior. Regiões de alta densidade posterior (regiões HPD)³ são interpretadas como regiões fixadas que contém o parâmetro aleatório com determinada probabilidade; seu equivalente frequentista são os intervalos de confiança definidos em relação à função verossimilhança.

Uma das mais importantes diferenças práticas entre a abordagem bayesiana e a abordagem clássica é a natureza das integrais necessárias para a implementação de cada uma delas. Em ambas, apesar das suas diferenças fundamentais, a distribuição de amostragem para os dados (a função verossimilhança) desempenha um rol chave; porém a computação clássica integra a distribuição de amostragem sobre o espaço amostral, enquanto que a computação bayesiana integra sobre o espaço paramétrico. Em outras palavras, o frequentista condiciona nos parâmetros e integra sobre os dados; o bayesiano condiciona nos dados e integra sobre os parâmetros [Bernardo e Smith 1994; Tanner 1993].

3.3 O modelo de regressão bayesiana: um exemplo

Como uma introdução aos modelos de regressão bayesiana com amostragem MCMC, considera-se um modelo de regressão linear com distribuições prévias informativas. Seja o modelo de regressão linear normal múltipla, definido como:

$$y = X\beta + \varepsilon \quad (3.4)$$

$$\varepsilon \sim N(0, \sigma^2 I_n) \quad (3.5)$$

onde y é um vetor ($n \times 1$) de variáveis dependentes; X representa uma matriz ($n \times k$) de variáveis exploratórias e ε é um vetor ($n \times 1$) de variáveis aleatórias independentes, normalmente distribuídas e de variância constante σ^2 .

³ Obtida a fdp posterior $p(\theta | y)$ é possível, geralmente, calcular a probabilidade de que o vetor de parâmetros θ esteja em uma sub-região particular \bar{R} , do espaço paramétrico: $p(\theta \in \bar{R} | y) = \int_{\bar{R}} p(\theta | y) d\theta$. Esta probabilidade mede o grau de confiança de que $\theta \in \bar{R}$ dadas a amostra e a informação prévia.

Os parâmetros a serem estimados neste modelo são β e σ^2 . Assumindo estes dois parâmetros como independentes, a distribuição prévia conjunta verifica:

$$p(\beta, \sigma^2) = p(\beta) \cdot p(\sigma^2) \quad (3.6)$$

Completa-se o modelo com as distribuições prévias independentes para β e σ^2 ,

$$\beta \sim N(\mathbf{r}, \mathbf{T}) \quad (3.7)$$

$$\sigma^2 \sim \text{Inv} - \chi^2(v, s^2) \quad (3.8)$$

onde \mathbf{r} é um vetor ($k \times 1$) contendo as medias prévias; \mathbf{T} é uma matriz ($k \times k$) contendo as variâncias e covariâncias prévias e v e s^2 são os graus de liberdade e o fator de escala respectivamente. Geweke (1997) declara que o modelo poderia ser completado com qualquer distribuição prévia que expresse adequadamente o conhecimento prévio. Porém, as expressões (3.7) e (3.8) possuem propriedades analíticas atrativas por serem distribuições prévias conjugadas⁴.

As expressões analíticas para as distribuições prévias seguem-se,

$$p(\beta) = \frac{|\mathbf{T}^{-1}|^{1/2}}{(2\pi)^{k/2}} e^{-1/2 \cdot [(\beta - \mathbf{r})^T \mathbf{T}^{-1} (\beta - \mathbf{r})]} \quad (3.9)$$

$$p(\sigma^2) = \frac{(vs^2)^{v/2} (\sigma^2)^{-(v+2)/2}}{2^{v/2} \Gamma(v/2)} e^{-vs^2/2\sigma^2} \quad (3.10)$$

onde $\Gamma(v/2)$ é a função gamma⁵.

⁴ A propriedade que a distribuição posterior mantenha a mesma forma paramétrica que a distribuição prévia é denominada *conjugacy*. Por exemplo, a distribuição prévia Beta é uma família conjugada para a verossimilhança binomial. A família conjugada é matematicamente conveniente porque a distribuição posterior segue uma forma paramétrica conhecida. Além disso, distribuições prévias conjugadas tem a vantagem prática de serem interpretáveis como dados adicionais [Gelman *et al*, 1995].

⁵ A função gamma é definida como: $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$ [Bronshtein e Semendiaev, 1974].

A função verossimilhança ou a densidade condicional dos dados é:

$$L(\beta, \sigma^2) = p(\mathbf{y}|\mathbf{X}, \beta, \sigma^2) = \frac{e^{-1/2\sigma^2 \cdot [(\mathbf{y}-\mathbf{X}\beta)^T (\mathbf{y}-\mathbf{X}\beta)]}}{(2\pi\sigma^2)^{n/2}} \quad (3.11)$$

Seguindo a metodologia bayesiana habitual, combina-se a função de verossimilhança com as prévias $p(\beta)$ e $p(\sigma)$ para produzir a densidade posterior:

$$p(\beta, \sigma^2 | \mathbf{y}, \mathbf{X}) = \frac{|\mathbf{T}^{-1}|^{1/2}}{(2\pi)^{(k+n)/2}} \cdot \frac{(v_s^2)^{v/2} (\sigma^2)^{-(v+n+2)/2}}{2^{v/2} \Gamma(v/2)} \quad (3.12)$$

$$\cdot e^{-(v_s^2)/2\sigma^2} \cdot e^{-1/2 \cdot [(\beta-\mathbf{r})^T \mathbf{T}^{-1} (\beta-\mathbf{r})]} \cdot e^{-1/2\sigma^2 \cdot [(\mathbf{y}-\mathbf{X}\beta)^T (\mathbf{y}-\mathbf{X}\beta)]}$$

efetuando operações algébricas, obtém-se:

$$p(\beta, \sigma^2 | \mathbf{y}, \mathbf{X}) = \frac{|\mathbf{T}^{-1}|^{1/2}}{(2\pi)^{(k+n)/2}} \cdot \frac{(v_s^2)^{v/2} (\sigma^2)^{-(v+n+2)/2}}{2^{v/2} \Gamma(v/2)} \quad (3.13)$$

$$\cdot e^{-(v_s^2)/2\sigma^2} \cdot e^{-R/2} \cdot e^{-1/2 \cdot [(\beta-\bar{\beta})^T \mathbf{V}^{-1} (\beta-\bar{\beta})]}$$

onde:

$$\mathbf{V} = \left(\mathbf{T}^{-1} + \frac{\mathbf{X}^T \mathbf{X}}{\sigma^2} \right)^{-1} \quad (3.14)$$

$$\bar{\beta} = \left(\mathbf{T}^{-1} + \frac{\mathbf{X}^T \mathbf{X}}{\sigma^2} \right)^{-1} \cdot \left(\mathbf{T}^{-1} \mathbf{r} + \frac{\mathbf{X}^T \mathbf{y}}{\sigma^2} \right) \quad (3.15)$$

$$R = \frac{\mathbf{y}^T \mathbf{y}}{\sigma^2} + \mathbf{r}^T \mathbf{T}^{-1} \mathbf{r} - \bar{\beta}^T \mathbf{V}^{-1} \bar{\beta} \quad (3.16)$$

Observa-se que para obter as densidades marginais de β e σ^2 deve-se integrar a expressão (3.13) em σ^2 e β respectivamente: $\int_{\sigma^2} p(\beta, \sigma^2) d\sigma^2$ e $\int_{\beta} p(\beta, \sigma^2) d\beta$. Porém, como $\bar{\beta}$ e V dependem de σ^2 , estas integrais devem resolver-se numericamente.

Devido aos progressos de hardware, técnicas de amostragem que fazem uso intensivo do poder computacional tem sido desenvolvidas, e algumas destas aplicadas à inferência bayesiana. Dentre estas destacam-se os métodos de Cadeias de Markov Monte Carlo (MCMC), principalmente o Amostrador de Gibbs e o Algoritmo Metrópolis-Hastings [Neal, 1993; Tanner, 1993; Geweke, 1999], que provêm uma forma de computar numericamente as integrais requeridas para se obter distribuições marginais posteriores, mediante simulação.

3.4 O Amostrador de Gibbs

A idéia básica do amostrador de Gibbs é tornar um problema multivariado numa seqüência de problemas univariados, entre os quais itera-se para produzir uma cadeia de Markov. A distribuição de equilíbrio é a distribuição posterior desejada [Cassella e George, 1992].

O amostrador de Gibbs, também denominado *alternating conditional sampling*, data do trabalho de Geman e Geman (1984) em tratamento de imagens. Posteriormente, Tanner e Wong (1987), Gelfand *et al* (1990), Gelfand and Smith (1990), Zeger e Karim (1991), têm contribuído no desenvolvimento da teoria desta metodologia de amostragem e demonstrando seu uso em uma ampla gama de aplicações estatísticas, onde distribuições posteriores intratáveis previamente dificultavam a análise bayesiana.

Esta técnica possibilita gerar indiretamente variáveis aleatórias de uma distribuição marginal, sem precisar calcular sua densidade de probabilidade. Este mecanismo baseia-se em propriedades elementares das cadeias de Markov [Cassella e George, 1992]. O amostrador pode ser visualizado como uma implementação prática do fato de que o conhecimento das distribuições condicionais é suficiente para a determinação da distribuição conjunta, se ela existir.

Intuitivamente: seja um vetor de parâmetros θ e sua densidade conjunta $p(\theta)$; em muitas aplicações, esta distribuição sobre todos os parâmetros é muito complexa não sendo possível amostrar diretamente dela. Porém, frequentemente, particionando o vetor de parâmetros em m conjuntos, $\theta = [\theta_1, \theta_2 \dots \theta_m]$, as densidades condicionais de cada subvetor θ_i em todos os outros possuem formas conhecidas, sendo possível amostrar delas.

O algoritmo pode ser resumido como:

```

Inicializar  $\theta^0$ 
Repete {
    Amostra  $\theta_1^{t+1} \sim p(\theta_1 | \theta_2^t, \dots, \theta_m^t)$ 
    Amostra  $\theta_2^{t+1} \sim p(\theta_2 | \theta_1^{t+1}, \theta_3^t, \dots, \theta_m^t)$ 
    :
    Amostra  $\theta_m^{t+1} \sim p(\theta_m | \theta_1^{t+1}, \theta_2^{t+1}, \dots, \theta_{m-1}^{t+1})$ 
     $t = t + 1$ 
} até  $t = N$ 

```

Uma grande quantidade de amostras (M) são descartadas para permitir que a cadeia atinja o equilíbrio, subseqüentes amostras são consideradas representativas da distribuição posterior.

Geman e Geman (1984) demonstraram que este processo estocástico, que amostra a seqüência completa das distribuições condicionais, representa uma cadeia de Markov com a distribuição correta de equilíbrio. Portanto, para t grande o suficiente, as amostras finais $N-M$ são efetivamente amostras das densidades dos m parâmetros de interesse. Logo, qualquer característica da população, inclusive a própria densidade, pode ser obtida. Por ex., para calcular a média de θ_1 , pode-se utilizar o fato que:

$$\lim_{N \rightarrow \infty} \frac{1}{N-M} \sum_{i=M+1}^N \theta_1^i = \int_{-\infty}^{\infty} \theta_1 p(\theta_1) d\theta_1 = E[\theta_1] \quad (3.17)$$

Voltando ao exemplo da seção 3.3, o modelo de regressão linear normal múltipla, o vetor de parâmetros é $\theta = [\beta, \sigma^2]$. Na amostragem de Gibbs é necessário identificar as distribuições de probabilidade condicionais: $p(\beta | \sigma^2)$ e $p(\sigma^2 | \beta)$.

Se a expressão (3.13) é interpretada unicamente como função de β , essa função é uma densidade *kernel* posterior de β condicional em σ^2 :

$$p(\beta | \sigma^2, \mathbf{y}, \mathbf{X}) \propto e^{-1/2 \cdot [(\beta - \bar{\beta})^T \mathbf{V}^{-1} (\beta - \bar{\beta})]} \quad (3.18)$$

onde \mathbf{V} e $\bar{\beta}$ são dados por (3.14) e (3.15) respectivamente:

$$\mathbf{V} = \left(\mathbf{T}^{-1} + \frac{\mathbf{X}^T \mathbf{X}}{\sigma^2} \right)^{-1}, \quad \bar{\beta} = \mathbf{V} \cdot \left(\mathbf{T}^{-1} \mathbf{r} + \frac{\mathbf{X}^T \mathbf{y}}{\sigma^2} \right)$$

Portanto, a distribuição posterior condicional de β é uma normal multivariada,

$$\beta \mid \sigma^2, \mathbf{y}, \mathbf{X} \sim N(\bar{\beta}, \mathbf{V}) \quad (3.19)$$

Observa-se a simetria entre a distribuição prévia e os dados quando são combinados em (3.18). A matriz inversa de variância-covariância posterior \mathbf{V}^{-1} é a soma da matriz inversa de variância-covariância prévia \mathbf{T}^{-1} e do termo $\mathbf{X}^T \mathbf{X} / \sigma^2$. Este último termo é a matriz inversa de variância-covariância posterior quando não há informação prévia, $\mathbf{T}^{-1} = 0$. A média da densidade posterior condicional é a soma ponderada da média da densidade prévia, \mathbf{r} , e de um termo que depende dos dados. Observa-se outra vez que, quando não há informação prévia os dados determinam a média.

Interpretando a equação (3.12) unicamente como função de σ^2 , obtém-se a distribuição *kernel* de σ^2 condicional em β :

$$p(\sigma^2 | \beta, \mathbf{y}, \mathbf{X}) \propto (\sigma^2)^{-(v+n+2)/2} \cdot e^{-1/2 \sigma^2 \cdot [v s^2 + (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta)]} \quad (3.20)$$

observa-se que a densidade posterior condicional de σ^2 é uma *scaled-inverted* χ^2 ,

$$\sigma^2 | \beta, \mathbf{y}, \mathbf{X} \sim \text{Inv} - \chi^2 \left(n + \nu, \frac{\nu s^2 + (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta)}{n + \nu} \right) \quad (3.21)$$

logo⁶:

$$\frac{\nu s^2 + (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta)}{\sigma^2} | \beta, \mathbf{y}, \mathbf{X} \sim \chi^2(n + \nu) \quad (3.22)$$

Outra vez, há uma evidente simetria entre a distribuição prévia e os dados. Deve-se observar também que ambas as distribuições obtidas não são condicionais apenas nos dados e na informação prévia, (3.18) depende de σ^2 e (3.20) de β .

A amostragem de Gibbs utiliza estas duas distribuições posteriores condicionais, iniciando com valores arbitrários para os parâmetros β^0 e $(\sigma^2)^0$ e envolve as computações seguintes:

1. Computar a média e a variância usando (3.15) e (3.14) condicionais no valor inicial de $(\sigma^2)^0$, e utilizar estes para obter de (3.19) um vetor normal multivariado β^1 .
2. Amostrar um novo valor para σ^2 , $(\sigma^2)^1$ da expressão (3.21), utilizando o valor β^1 determinado no passo 1 e seguindo o procedimento descrito na expressão (3.22).

Retorna-se ao passo 1 utilizando os novos valores β^1 e $(\sigma^2)^1$ em lugar dos valores iniciais β^0 e $(\sigma^2)^0$, e realiza-se outra passagem através do amostrador. Isto produz um novo conjunto de valores β^2 e $(\sigma^2)^2$.

Gelfand e Smith (1990) estabelecem que, sob condições suficientemente amplas, continuas passagens através do amostrador produzem sucessões de valores β^k e $(\sigma^2)^k$, cuja distribuição converge à densidade posterior conjunta de interesse, $p(\beta, \sigma^2)$. Dadas ambas sucessões, aproxima-se o valor esperado dos parâmetros β e σ^2 em base aos momentos da distribuição dos valores amostrados:

⁶ Para obter uma amostra θ da distribuição $\text{Inv} - \chi^2(\nu, s^2)$, primeiro amostra-se x da distribuição $\chi^2(\nu)$ e logo deixa-se $\theta = \nu s^2 / x$ [Gelman *et al.*, 1995].

$$\frac{1}{t} \sum_{i=1}^t \beta^i \xrightarrow{t \rightarrow \infty} E[\beta] \quad (3.23)$$

$$\frac{1}{t} \sum_{i=1}^t (\sigma^2)^i \xrightarrow{t \rightarrow \infty} E[\sigma^2] \quad (3.24)$$

Observa-se como o amostrador de Gibbs possibilita a amostragem de uma densidade de probabilidade multivariada, utilizando apenas densidades de subconjuntos de vetores condicionais em todos os outros. A atratividade desta abordagem é que provê uma solução para o problema bayesiano de integração múltipla quando as densidades condicionais são identificáveis, de forma conhecida e passíveis de amostragem. Possibilitando, assim, a integração de distribuições posteriores complicadas através da amostragem seqüencial de distribuições condicionais.

3.5 Modelos bayesianos para escolha discreta

O modelo de regressão múltipla descrito anteriormente é apropriado para *conjoint analysis* quando a variável dependente é métrica. Porém, estudos conjuntos baseados em escolha tem variável dependente multinomial. Na literatura bayesiana a inferência de parâmetros agregados, da mesma forma que na inferência clássica, tem sido abordada empregando modelos logit e probit.

Nos modelos de escolha discreta, a variável dependente é associada a uma variável latente: a utilidade que o indivíduo outorga à alternativa em questão. Esta variável não é diretamente observável e deve portanto, na aplicação do amostrador de Gibbs, ser amostrada condicionando-a nas escolhas observadas.

Allenby e Lenk (1994) propõem um modelo que relaciona as escolhas observadas y aos atributos X, através de uma regressão normal logística. Aplicando a regra de Bayes,

$$p(\beta, \sigma^2 | \mathbf{X}, \mathbf{y}) \propto p(\beta, \sigma^2) p(\mathbf{y} | \mathbf{X}, \beta, \sigma^2) \quad (3.25)$$

e assumindo que as prévias para β e σ^2 são independentes e introduzindo a utilidade latente \mathbf{z} , obtém-se:

$$p(\beta, \sigma^2 | \mathbf{X}, \mathbf{y}) \propto p(\beta) p(\sigma^2) p(\mathbf{y} | \mathbf{z}) p(\mathbf{z} | \mathbf{X}, \beta, \sigma^2) \quad (3.26)$$

Este modelo é semelhante ao da seção 3.3, porém aqui aparece o fator $p(\mathbf{y} | \mathbf{z})$ que relaciona a variável latente \mathbf{z} com a escolha discreta \mathbf{y} . Portanto, para aplicar o amostrador de Gibbs é necessária uma terceira distribuição condicional que possibilite amostrar a variável latente \mathbf{z} . A sequência do amostrador segue-se:

1. Amostrar β : $\beta | \sigma^2, \mathbf{y}, \mathbf{X}, \mathbf{z}$
2. Amostrar σ^2 : $\sigma^2 | \beta, \mathbf{y}, \mathbf{X}, \mathbf{z}$
3. Amostrar \mathbf{z} : $\mathbf{z} | \beta, \sigma^2, \mathbf{y}, \mathbf{X}$

As duas primeiras expressões são idênticas a (3.19) e (3.21) respectivamente, considerando a redundância de \mathbf{y} . Para a terceira densidade, os autores propõem amostragem por rejeição. De:

$$\mathbf{z} | \beta, \sigma^2, \mathbf{X}, \mathbf{y} \sim N(\mathbf{X}\beta, \sigma^2) \quad (3.27)$$

gera-se um valor \mathbf{z} candidato que é aceito com uma probabilidade igual à probabilidade de escolha,

$$prob_a = \frac{e^{z(a)}}{\sum_j e^{z(j)}} \quad se \quad y(a) = 1 \quad (3.28)$$

Allenby *et al* (1995) estendem a abordagem anterior propondo um método bayesiano para incorporar conhecimento prévio, de caráter ordinal e intervalar, na análise de estudos conjuntos baseados em *rating* e escolhas. Restrições nos parâmetros são implementadas aceitando unicamente os valores amostrados que as atendem,

$$\beta|\sigma^2, \mathbf{y}, \mathbf{X}, \mathbf{z} \sim N(\bar{\beta}, \mathbf{V}) \cdot I_{\text{restrições de ordem}} \quad (3.29)$$

onde $I_{\text{restrições de ordem}}$ é uma função indicadora que é igual a um se a restrição é satisfeita e zero em caso contrário.

Albert e Chib (1993) propõem um modelo probit binário onde a variável observada y é associada com a utilidade latente z , de forma que, para o indivíduo i , $z_i < 0$ corresponde a $y_i = 0$ e $z_i \geq 0$ corresponde a $y_i = 1$. Os autores mostram que a distribuição posterior de z condicional em todos os outros parâmetros é uma normal truncada:

$$z_i|\beta, \sigma^2 \sim \begin{cases} N(y_i, \sigma_i^2), \text{ truncada à esquerda por } 0 \text{ se } y_i = 1 \\ N(y_i, \sigma_i^2), \text{ truncada à direita por } 0 \text{ se } y_i = 0 \end{cases} \quad (3.30)$$

Estas expressões indicam que é possível substituir valores de $y_i = 1$ por amostras de uma normal truncada à esquerda por zero, e valores de $y_i = 0$ por amostras de uma normal truncada à direita por zero. Desta forma, o modelo probit de regressão binária em y_i é conectado com um modelo de regressão normal linear na variável latente z_i .

Os autores generalizam para modelos de resposta multinomial: categorias ordenadas e não ordenadas.

Todas as abordagens bayesianas até aqui apresentadas estimam modelos agregados, nos quais a heterogeneidade é confundida com ruído aleatório. Os parâmetros inferidos fornecem apenas a tendência central dos dados e não permitem a inferência de *parth-worths* individuais. Portanto, em relação à recuperação de heterogeneidade, apresentam desvantagens similares as dos modelos agregados de máxima verossimilhança.

3.6 Modelos bayesianos hierárquicos

A abordagem bayesiana para inferência depende de distribuições prévias para os parâmetros do modelo. Estas distribuições prévias podem depender de parâmetros desconhecidos, que podem seguir por sua vez uma distribuição prévia num segundo nível.

Esta sequência de parâmetros e prévias constitui um modelo hierárquico, onde a *kernel* da posterior é:

$$L(y|\theta_1) \cdot p(\theta_1|\theta_2) \cdot p(\theta_2|\theta_3) \cdots p(\theta_{n-1}|\theta_n) p(\theta_n) \quad (3.31)$$

A hierarquia precisa parar em algum ponto, com todos os parâmetros prévios remanescentes assumidos conhecidos. Antes de fazer esta hipótese, a abordagem bayesiana empírica utiliza os dados observados para estimar os parâmetros do nível final e prossegue como na análise bayesiana usual [Carlin e Louis, 1996].

A modelagem hierárquica bayesiana aplicada à mensuração da estrutura de preferências do consumidor possibilita superar um problema típico em marketing: a existência de informação limitada em relação às preferências individuais dos consumidores impossibilita a estimação de parâmetros em nível individual. Esta limitante tem forçado o uso de modelos agregados e de métodos de segmentação a priori.

Allenby e Ginter (1995), propõem um modelo hierárquico no qual os entrevistados devem dar uma avaliação numérica às alternativas. Em nível individual a resposta de um indivíduo é dada por uma equação de regressão normal,

$$y_{h,i} = \mathbf{X}_{h,i} \cdot (\alpha + \beta_i) + \varepsilon_{h,i} \quad (3.32)$$

onde $y_{h,i}$ é a avaliação da alternativa h feita pelo entrevistado i ; $\mathbf{X}_{h,i}$ é o vetor de variáveis independentes; α é um coeficiente de efeitos fixos constante para todos os entrevistados; β_i é um coeficiente específico para cada entrevistado e $\varepsilon_{h,i}$ é o erro iid normal, $N(0, \sigma^2)$. Para estimar os parâmetros β_i empregando o procedimento descrito na seção 3.3 seria necessária uma grande quantidade de dados individuais.

Na abordagem hierárquica proposta pelos autores, é especificada uma segunda equação que descreve a variação de β_i entre os indivíduos,

$$\beta_i = \Gamma \cdot \mathbf{w}_i + \xi_i \quad (3.33)$$

onde \mathbf{w}_i é o vetor de covariáveis que levam em conta a heterogeneidade; Γ é a matriz que relaciona β_i ao valor de \mathbf{w}_i e ξ_i é um componente de heterogeneidade com distribuição $N(0, \mathbf{D})$.

Esta última equação representa a variação dos coeficientes de regressão dos indivíduos na população. Os parâmetros são estimados empregando o amostrador de Gibbs.

Os autores estendem este modelo para escolha discreta adicionando uma terceira equação, que representa a probabilidade que o entrevistado i escolha a alternativa h :

$$prob(h)_i = \frac{e^{y_{h,i}}}{\sum_j e^{y_{h,j}}} \quad (3.34)$$

Nesta abordagem precisa-se definir a priori as variáveis socioeconômicas w_i , que caracterizam a heterogeneidade em (3.33). Na prática, este conhecimento pode ser não disponível ou ambíguo, devido ao fato que a verdadeira fonte de heterogeneidade não é realmente conhecida a priori.

Em Allenby *et al* (1998) é proposto outro modelo hierárquico para obtenção de parâmetros individuais, que apresenta a vantagem de não requerer a especificação a priori das covariáveis que caracterizam o consumidor. O modelo, dentro de um contexto multinomial logit, assume que a probabilidade de escolha da alternativa j para o entrevistado i é dada por (3.34). A heterogeneidade através dos entrevistados é considerada assumindo um modelo de mistura de componentes normais:

$$\beta_i \sim \sum_k \phi_k \cdot N(\bar{\beta}_k, \mathbf{D}_k) \quad (3.35)$$

onde k indica a quantidade de componentes normais, e ϕ_k é a massa de cada componente. Cada componente é modelado com média e matriz de variância-covariância diferentes.

Antes de prosseguir apresenta-se outro método MCMC, que será de utilidade no modelo proposto.

3.7 O Algoritmo Metropolis-Hastings

O algoritmo Metropolis-Hastings é um método de simulação estocástica que permite amostrar de qualquer densidade de probabilidade $p(x)$, sem que seja necessário conhecer as

distribuições condicionais. Isto é especialmente relevante quando p é a distribuição posterior num contexto bayesiano [Smith e Roberts, 1993; Tierney, 1994; Chib e Greenberg, 1995].

Considera-se a geração de uma amostra aleatória x da função densidade alvo $p(x)$. O algoritmo Metropolis-Hastings utiliza outra função de distribuição, denominada densidade geradora do candidato, $p^*(x)$. Esta densidade deve possuir duas condições: 1^o permitir a geração de amostras aleatórias (x_c) de forma fácil e rápida e, 2^o deve ser definida no mesmo domínio que a densidade alvo $p(x)$.

O algoritmo é implementado como segue:

```

Inicializa  $x^0$ 
Repete {
    1. Dado  $x^{i-1}$ , gerar uma amostra aleatória  $x_c$  de  $p^*$ 
       ( $x | x^{i-1}$ ) e uma amostra aleatória uniforme  $u$  do
       intervalo entre zero e um.
    2. Calcular a probabilidade de aceitação
       
$$r = \min \left( \frac{p(x_c) \cdot p^*(x^{i-1} | x_c)}{p(x^{i-1}) \cdot p^*(x_c | x^{i-1})}, 1 \right)$$

       Se  $u \leq r$ , definir  $x^{i-1} = x_c$ .
       De outra forma, definir  $x^i = x^{i-1}$ .
} até  $t = N$ 

```

Da mesma forma que no amostrador de Gibbs, as primeiras M amostras são descartadas e subseqüentes amostras são consideradas representativas da distribuição alvo.

Os conceitos teóricos apresentados neste capítulo: modelos bayesianos hierárquicos e algoritmos de simulação estocástica, são a base teórica para o modelo proposto de inferência dos parâmetros individuais apresentado no próximo capítulo.

O MODELO PROPOSTO

Após a definição dos objetivos de pesquisa, o delineamento experimental e a coleta dos dados de entrevistas, o estudo da estrutura de preferências do consumidor prossegue com a análise e interpretação dos resultados. Estas duas últimas etapas são o foco deste trabalho, onde é proposta uma abordagem bayesiana de escolha discreta que possibilita a estimação de parâmetros de preferência em nível individual, e utiliza este conhecimento para identificar oportunidades de mercado e formular e avaliar estratégias de marketing.

A modelagem de escolha discreta aborda a mensuração da estrutura de preferências do consumidor, principalmente devido a problemas de custo, utilizando limitada informação (escolhas) em nível individual. Esta limitante tem impossibilitado inferir estruturas de preferências para cada integrante da amostra pesquisada. Assim, modelos agregados assumem que o mercado é homogêneo. Se existe evidência que isto não é verdade, estes modelos tentam recuperar heterogeneidade incluindo variáveis que caracterizam o consumidor na função utilidade. No entanto, os indivíduos podem não ser diferentes em função deste tipo de características, os consumidores são basicamente diferentes nas suas preferências e escolhas. A heterogeneidade do mercado é explicada principalmente pelas escolhas que os consumidores efetuam, escolhas que são função de preferências: marcas preferidas, atributos com maior valor, políticas de preços, planos alternativos de marketing, etc., e do contexto no qual as escolhas são efetuadas. Eventualmente, pode existir uma associação entre grupos de consumidores com preferências similares e algumas variáveis que melhor caracterizem os grupos identificados.

Neste contexto, torna-se de extrema importância a identificação da heterogeneidade do mercado. O conhecimento das preferências individuais dos consumidores, assim como da distribuição populacional destas preferências possibilita à organização projetar produtos atrativos para consumidores específicos, reconhecer oportunidades para marketing direto e identificar grupos de consumidores com preferências similares.

Os *part-worths* (utilidades-parte) das preferências do consumidor refletem o valor que este aloca a cada atributo do produto/serviço/conceito. A estimação destes parâmetros em nível individual, possibilita a construção de simuladores de escolha que melhor emulam o comportamento real do mercado, permitindo efetuar estudos de análise competitiva sob diferentes cenários hipotéticos. Simuladores de escolha procuram avaliar, dentro de um contexto competitivo, a viabilidade de produtos/serviços/conceitos. Estimativas de escolha em situações novas (novos produtos, reposicionamento de produtos existentes, planos alternativos de marketing, etc.) são aprimoradas com a informação das preferências em nível individual. Simuladores de escolha baseados em preferências médias podem resultar em decisões gerenciais incorretas. As distorções tornam-se críticas quando o ambiente competitivo considera produtos similares. Por exemplo, um modelo agregado logit, devido à hipótese IIA, não recupera efeitos de substituição diferenciada entre estes produtos.

A abordagem proposta neste trabalho integra os conceitos apresentados no capítulo anterior, modelagem bayesiana hierárquica e simulação estocástica, com simuladores de escolha, visando atender melhor às necessidades do marketing atual. Esta abordagem reconhece que os consumidores não são uniformes mas sim heterogêneos, o que justifica a concentração dos esforços estratégicos de marketing, dentro de um contexto competitivo, em componentes específicos e claramente identificados de mercado.

4.1 O modelo

Inicialmente, o modelo objetiva a estimação de parâmetros em nível individual. Esta estimação é viabilizada através de uma modelagem bayesiana hierárquica do problema, que efetua um *pooling* da informação proveniente da amostra e possui a capacidade de estudar as preferências de indivíduos específicos. Este modelo assume que os parâmetros de preferências individuais variam através da população segundo uma distribuição contínua de probabilidade. O objetivo é a inferência dos parâmetros em nível individual e a distribuição que descreve a heterogeneidade das preferências na população. Duas importantes hipóteses são assumidas: uma distribuição normal multivariada, descrita por hiperparâmetros agregados, para caracterizar a distribuição da heterogeneidade das preferências e, um modelo logit multinomial ordenado para descrever o comportamento individual de escolha.

Posteriormente, os parâmetros estimados são utilizados para construir um simulador de escolha ou de participação de preferências. Este simulador fornece uma estrutura útil que permite obter prognósticos de “demanda” a partir das preferências envolvidas nos parâmetros individuais. Este simulador é imune à propriedade IIA e outorga probabilidades finitas de escolha a todos os produtos considerados no cenário competitivo. O simulador desenvolvido aproveita a extensa informação obtida no processo de estimação: hiperparâmetros, que descrevem o comportamento em nível agregado, e vetores de parâmetros individuais definidos por suas respectivas distribuições de probabilidade, para prognosticar participação de escolha e medidas de “graus de satisfação” para produtos competitivos. Esta informação é a que em última instância permite a identificação de oportunidades de mercado, e a formulação e avaliação de estratégias alternativas de marketing.

A seguir descreve-se, detalhadamente o modelo desenvolvido.

4.1.1 Inferência dos parâmetros individuais

Em uma abordagem bayesiana de inferência [seções 3.1, 3.2], estimar um parâmetro envolve a determinação da distribuição posterior de probabilidade condicional em dados observados. Desta forma, para cada indivíduo i , a distribuição posterior do vetor de parâmetros β_i verifica:

$$p(\beta_i | \mathbf{X}, \mathbf{y}) \propto L(\beta_i | \mathbf{y}) \cdot p(\beta_i) \quad (4.1)$$

isto é, verossimilhança vezes prévia. Dado que a quantidade de informação em nível individual é escassa, devido a se dispor de poucas escolhas individuais (verossimilhança) e de pouca informação prévia individual ($p(\beta_i)$), a aplicação da análise bayesiana tradicional para um indivíduo é inviável.

Aborda-se o problema da estimação de parâmetros individuais mediante um modelo bayesiano hierárquico, que permite incorporar informação referente a todos os entrevistados à equação 4.1. Consideram-se dois níveis de hierarquia. O nível inferior caracteriza o comportamento individual mediante um modelo probabilístico de escolha. Assume-se que as escolhas são governadas por um modelo logit multinomial ordenado. Esta modelagem possibilita a incorporação de uma quantidade maior de informação, através da consideração

da ordem de escolha no processo de análise, que a formulação de Allenby *et al.* (1998) onde considera-se um modelo logit multinomial.

O nível superior caracteriza a informação prévia, $p(\beta_i)$, através da consideração de hiperparâmetros agregados, $\theta = [\alpha, \mathbf{D}]$, que descrevem a heterogeneidade entre os indivíduos:

$$p(\beta_i) = p(\beta_i|\theta) \cdot p(\theta) \quad (4.2)$$

I. Formulação

Assume-se que os parâmetros individuais, no nível superior da hierarquia, seguem uma distribuição normal multivariada,

$$\beta_i \sim N(\alpha, \mathbf{D}) \quad (4.3)$$

onde β_i é o vetor ($k \times 1$) de parâmetros para o indivíduo i ; e α e \mathbf{D} são, respectivamente o vetor ($k \times 1$) de médias e a matriz ($k \times k$) de variância-covariância da distribuição dos parâmetros individuais. De acordo com a abordagem bayesiana empírica os hiperparâmetros, parâmetros das distribuições de α e \mathbf{D} , são estimados a partir dos próprios dados (a seção 4.1.1-III detalha este procedimento).

Desta forma, para cada indivíduo i , o modelo hierárquico é representado por:

$$p(\beta_i, \alpha, \mathbf{D}|\mathbf{X}, \mathbf{y}) \propto L(\beta_i|\mathbf{y}) \cdot p(\beta_i|\alpha, \mathbf{D}) \cdot p(\alpha) \cdot p(\mathbf{D}) \quad (4.4)$$

As escolhas são descritas por um modelo multinomial logit ordenado, a probabilidade de uma ordenação feita pelo indivíduo i é dada por:

$$p_i(1, 2, \dots, (J - J_P)) = \prod_{j=1}^{J-J_P} \frac{e^{x_j \beta_i}}{\sum_{a=j}^J e^{x_a \beta_i}} \quad (4.5)$$

onde a alternativa 1 é preferida à alternativa 2, a alternativa 2 à 3, etc. em um total de J alternativas efetuando $J - J_P$ escolhas, e \mathbf{x}_a representa a linha (alternativa) a da matriz de variáveis independentes \mathbf{X} .

Este modelo envolve a estimação de muitos parâmetros: um vetor β_i de *part-worths* para cada indivíduo, um vetor α de médias da distribuição dos *part-worths*, e uma matriz \mathbf{D} de variância-covariância dessa distribuição. A estimação desta grande quantidade de parâmetros, $nk + \frac{k(k+3)}{2}$, (n : entrevistados, k : atributos) é viabilizada através da decomposição do problema num conjunto de problemas mais simples, como é apresentado a seguir.

Os parâmetros β_i , α e \mathbf{D} são estimados mediante um processo iterativo, empregando o amostrador de Gibbs:

1. Amostrar α : $\alpha \mid \{\beta_i\}, \mathbf{D}, \mathbf{y}, \mathbf{X}$
2. Amostrar \mathbf{D} : $\mathbf{D} \mid \{\beta_i\}, \alpha, \mathbf{y}, \mathbf{X}$
3. Amostrar $\{\beta_i, i=1\dots n\}$: $\beta_i \mid \alpha, \mathbf{D}, \mathbf{y}, \mathbf{X}$

O processo é robusto e seus resultados não dependem dos valores de início. No entanto, para incrementar a velocidade de convergência, inicia-se com estimativas de parâmetros que aproximem de forma razoável os valores finais.

II. Definição de pontos iniciais

Definem-se pontos de início que, embora estejam sobredispersados tenham uma boa chance de pertencer à região de alta densidade do espaço paramétrico; para não retardar em muito a taxa de convergência. Uma regra simples dada por Kass *et al*, (1997) é: inicializar cinco cadeias paralelas para os parâmetros θ , de forma que a cadeia j inicia em $\mu + (j - 3)\sigma$, $j = 1, \dots, 5$, onde μ e σ são a média e o desvio padrão prévios de θ_i . Aqui segue-se esta regra, só que antes de empregar estimativas prévias para μ e σ , utiliza-se estimativas de métodos mais simples como o da máxima verossimilhança para iniciar as 3 primeiras cadeias. Ainda,

para criar maior dispersão dos pontos de início, as duas cadeias restantes são inicializadas arbitrariamente em uns e zeros. A sucessão de valores destas 5 cadeias permitem efetuar o diagnóstico de convergência.

III. Estimação iterativa dos parâmetros: Amostragem de Gibbs

Especificados os valores iniciais, cada iteração consiste de três etapas que envolvem a geração recursiva de amostras aleatórias das distribuições condicionais posteriores dos parâmetros do modelo:

1. α é normalmente distribuído com média igual à media dos β_i e matriz de variância-covariância igual a \mathbf{D} dividida pelo número de entrevistados. Assim, utilizando as estimativas atuais dos β_i e \mathbf{D} , gera-se uma nova estimativa de α :

$$\alpha|\{\beta_i\}, \mathbf{D}, \mathbf{X}, \mathbf{y} \sim N\left(\bar{\beta}, \frac{\mathbf{D}}{n}\right) \quad (4.6)$$

onde $\bar{\beta}$ é a média dos *part-worths* individuais β_i e n é a quantidade de entrevistados.

2. Utilizando as estimativas atuais de β_i e α , amostra-se uma nova estimativa para \mathbf{D} de uma distribuição inversa Wishart¹. A distribuição Wishart é a distribuição prévia conjugada da matriz inversa de variância-covariância para uma distribuição normal multivariada [Gelman *et al*, 1995], e uma generalização multivariada da distribuição Gamma. Desta forma, isto pode ser visto como uma generalização da seção 3.3 onde se estima a variância σ^2 no modelo de regressão linear normal, já que a distribuição inversa- χ^2 é um caso especial da distribuição inversa-Gamma.

A matriz de variância-covariância de amostras iid de uma normal multivariada possui distribuição Wishart. Portanto, simulações normais multivariadas podem ser utilizadas para gerar amostras de uma distribuição Wishart [Gelman *et al*, 1995]. Em geral, procede-se:

¹ A variável \mathbf{W} é distribuída *inverted-Wishart*, com v graus de liberdade e matriz de escala \mathbf{S} de dimensão ($k \times k$), simétrica e definida positiva, $p(\mathbf{W}) = \text{Inv-Wishart}_v(\mathbf{W} | \mathbf{S}^{-1})$:

$$p(\mathbf{W}) = \left[2^{vk/2} \cdot \pi^{k(k-1)/4} \cdot \prod_{i=1}^k \Gamma\left(\frac{v+1-i}{2}\right) \right]^{-1} \cdot |\mathbf{S}|^{v/2} \cdot |\mathbf{W}|^{-(v+k+1)/2} \cdot e^{-0.5 \cdot \text{tr}(\mathbf{S}\mathbf{W}^{-1})}$$

com: $E(\mathbf{W}) = (v-1-k)^{-1} \cdot \mathbf{S}$

- Simular n amostras independentes u de uma distribuição normal multivariada, $N(0, \mathbf{S})$, onde n é o número de entrevistados.
- Calcular a matriz de variância-covariância dessas amostras, $\mathbf{D}^* = \sum_{i=1}^n u \cdot u^T$, onde $\mathbf{D}^* \sim \text{Wishart}(n, \mathbf{S})$.

A matriz de escala \mathbf{S} é obtida, a cada iteração, dos valores atuais de β_i e α :

$$\mathbf{S} = \sum_{i=1}^n (\alpha - \beta_i) \cdot (\alpha - \beta_i)^T \quad (4.7)$$

Finalmente, $\mathbf{D} = (\mathbf{D}^*)^{-1}$ é distribuído *inverted-Wishart*.

- Utilizando as estimativas atuais de α e \mathbf{D} , gera-se novas estimativas para os β_i . Para cada indivíduo, a *kernel* da distribuição condicional de β_i é dada por:

$$p(\beta_i | \alpha, \mathbf{D}, \mathbf{X}, \mathbf{y}) \propto e^{-1/2 \cdot [(\beta_i - \alpha)^T \mathbf{D}^{-1} (\beta_i - \alpha)]} \cdot \prod_C \prod_{j=1}^{J-J_p} \frac{e^{\mathbf{x}_j \beta_i}}{\sum_{a=j}^J e^{\mathbf{x}_a \beta_i}} \quad (4.8)$$

onde C é o número de conjuntos ou cenários de escolha. Observa-se que esta *kernel* não possui uma forma conhecida, impossibilitando a amostragem direta. Emprega-se o algoritmo Metropolis-Hastings, que permite a geração de *part-worths* individuais combinando as informações disponíveis em nível agregado (normal multivariada) e individual (escolha logística multinomial ordenada). Este procedimento é detalhado na seção seguinte.

Em cada uma destas etapas reestima-se um parâmetro ou conjunto de parâmetros (α , \mathbf{D} ou β_i) condicionalmente em todos os outros, dados seus valores atuais. Continua-se o processo até atingir convergência para os três conjuntos de parâmetros. Após a convergência, continua-se iterando para obter amostras dos parâmetros para cada um dos entrevistados β_i , assim como de α e \mathbf{D} .

Desta forma, a distribuição marginal posterior de cada parâmetro é caracterizada pela sucessão de valores amostrados. Os valores finais dos *part-worths* individuais β_i e dos parâmetros agregados α e \mathbf{D} são determinados promediando os valores guardados.

IV. A amostragem dos *part-worths* individuais: Metropolis-Hastings

Dado que a distribuição alvo, distribuição dos parâmetros para um indivíduo (4.8), não possui forma conhecida, implementa-se uma “caminhada aleatória Metropolis” (*random walk Metropolis*) [Carlin e Louis, 1996]. Desta forma são gerados valores candidatos que tem chance de serem aceitos e que percorrem o espaço paramétrico da densidade dos parâmetros para cada indivíduo.

O procedimento é o seguinte:

- Amostrar um vetor d de uma distribuição normal multivariada com média zero e matriz de variância-covariância \mathbf{D} , $N(0, \mathbf{D})$.
- Calcular o valor candidato do parâmetro β_C a partir do valor atual β_A : $\beta_C = \beta_A + d$.
- Aceitar o candidato com probabilidade r :

$$r = \min \left(\frac{e^{-1/2 \cdot [(\beta_C - \alpha)^T \mathbf{D}^{-1} (\beta_C - \alpha)]} \cdot \prod_C \prod_{j=1}^{J-J_P} \frac{e^{\mathbf{x}_j \beta_C}}{\sum_{a=j}^J e^{\mathbf{x}_a \beta_C}}}{e^{-1/2 \cdot [(\beta_A - \alpha)^T \mathbf{D}^{-1} (\beta_A - \alpha)]} \cdot \prod_C \prod_{j=1}^{J-J_P} \frac{e^{\mathbf{x}_j \beta_A}}{\sum_{a=j}^J e^{\mathbf{x}_a \beta_A}}}, 1 \right) \quad (4.9)$$

O primeiro termo em (4.9) é a razão da densidade posterior (4.8) avaliada nos valores candidato e atual do parâmetro, β_C e β_A . Observa-se que desta forma r combina informação agregada, dada pela razão das densidades dos *part-worths* avaliada em β_C e β_A , com informação individual, dada pela razão das verossimilhanças da escolha logit multinomial ordenada para β_C e β_A . Esta faculdade de partilhar informação é o que permite a este modelo bayesiano hierárquico produzir estimativas razoáveis dos parâmetros individuais ainda quando os dados disponíveis são insuficientes para uma análise individual.

V. Procedimento de diagnóstico da convergência

Entre os diferentes fatores de caráter implementacional associados aos métodos MCMC (escolha do amostrador, comprimento da corrida, escolha dos valores iniciais e problemas de estimação e eficiência) o problema da convergência é um fator crítico; desde que as taxas de convergência dos algoritmos em diferentes distribuições alvo variam consideravelmente, dependendo do problema específico [Mengersen e Tweddie, 1996; Brooks e Roberts, 1998].

Brooks e Roberts (1998), definem os procedimentos de diagnóstico de convergência como métodos para avaliar o comprimento de corrida da cadeia de Markov a fim de se obter observações da distribuição estacionária da cadeia. Com este objetivo diferentes métodos tem sido propostos [Cowles e Carlin, 1996; Brooks e Roberts, 1998]. Porém este é ainda um tema em desenvolvimento, principalmente para incrementar a velocidade de convergência [Chauveau e Vandekerkhove, 1999]. Kass *et al.* (1997) declaram que em aplicações práticas, usuários utilizam métodos informais como curvas de nível e gráficos de dispersão.

A metodologia MCMC baseia-se no fato que a medida que $n \rightarrow \infty$, o n -ésimo passo da *kernel* de transição do amostrador converge à densidade alvo. Portanto, pode-se efetuar uma longa corrida de uma única cadeia, ignorar uma fase inicial de transição e formar uma amostra das observações restantes. Quando o tempo de convergência é conhecido, a implementação de apenas uma cadeia é suficiente. Gelman *et al.* (1995) propõem uma solução para quando os tempos de convergência são desconhecidos: correr diferentes cadeias independentes e formar amostras daquelas observações, ignorando a fase inicial de cada uma. Formalmente, para cada parâmetro estimado θ , denomina-se as amostras de J sequências paralelas de longitude n como $\theta_{i,j}$ ($i = 1, \dots, n; j = 1, \dots, J$). Calcula-se as variâncias entre e dentro das sequências, B e W :

$$B = \frac{n}{J-1} \sum_{j=1}^J \left(\bar{\theta}_{\cdot,j} - \bar{\theta} \right)^2 \quad (4.10)$$

$$W = \frac{1}{J} \sum_{j=1}^J s_j^2 \quad (4.11)$$

onde:

$$\bar{\theta}_{.,j} = \frac{1}{n} \sum_{i=1}^n \theta_{i,j}; \quad \bar{\theta} = \frac{1}{J} \sum_{j=1}^J \bar{\theta}_{.,j} \quad (4.12)$$

$$s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (\theta_{i,j} - \bar{\theta}_{.,j})^2 \quad (4.13)$$

Pode-se estimar a variância marginal posterior do parâmetro θ , $var(\theta | y)$:

$$\hat{var}(\theta|y) = \frac{n-1}{n} W + \frac{1}{n} B \quad (4.14)$$

Gelman *et al.* (1995) monitoram a convergência da simulação iterativa estimando o fator pelo qual a escala da distribuição atual para θ poderia ser reduzida se as simulações continuassem até o limite $n \rightarrow \infty$. Esta redução potencial de escala é dada por (4.15) e declina para 1 conforme $n \rightarrow \infty$.

$$\sqrt{\hat{R}} = \sqrt{\frac{\hat{var}(\theta|y)}{W}} \quad (4.15)$$

Se a redução potencial de escala é alta, existe razão para acreditar que simulações adicionais podem aprimorar a inferência sobre a distribuição alvo. Gelman *et al.* (1995) recomendam continuar as simulações até obter valores para $\sqrt{\hat{R}}$ abaixo de 1.2, para todos os parâmetros θ de interesse. Posteriormente, coletar as $J \times n$ amostras da segunda metade das sequências e considerá-las como amostras da distribuição alvo.

VI. Validação preditiva do modelo

Green e Srinivasan (1990) declaram que, acima de tudo *Conjoint Analysis* é uma técnica para prognosticar escolha, ou pelo menos a intenção desta. Proporcionam ainda evidência empírica que sugere a validação desta como uma técnica preditiva. A melhor forma de validação é a comparação dos prognósticos contra resultados reais. No entanto, num contexto de marketing, este teste é difícil de se aplicar, devido aos efeitos perturbadores das

variáveis do plano de marketing, por ex. publicidade e promoções. Portanto, a validação preditiva do modelo precisa de testes de validação cruzada, isto é, a identificação da capacidade do modelo para prognosticar a escolha em um conjunto *holdout* de alternativas².

Cenários *holdout*

As alternativas em um cenário *holdout* devem ser parecidas, de forma a constituir um “teste ácido” para o modelo de inferência. A seguir é apresentada uma regra de formação das alternativas *holdouts* [Elrod *et al*, 1992],

$$\begin{bmatrix} a_1 + d_1 & a_2 & a_3 & \dots \\ a_1 & a_2 + d_2 & a_3 & \dots \\ a_1 & a_2 & a_3 + d_3 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} \quad (4.16)$$

onde $d_k = \gamma_k f_k$; $[f_1, \dots, f_k]$ são constantes arbitrárias todas positivas ou todas negativas; γ_k é uma variável indicadora $\gamma_k = 1$ se a preferência pelo atributo k é monótona crescente, caso contrário $\gamma_k = -1$; e $[a_1, \dots, a_k]$ são constantes arbitrárias.

Critérios de desempenho

Visando avaliar de forma quantitativa o desempenho preditivo do modelo, utiliza-se de dois critérios:

Hit Rate. Percentagem de vezes que o modelo prognostica de forma correta cada escolha do entrevistado no cenário *holdout*.

Mean Absolute Error (MAE). Média do valor absoluto da diferença entre as participações de mercado prognosticadas e as participações identificadas para cada alternativa nos cenários *holdout*.

² Com a finalidade de testar a validade preditiva do modelo de preferências dos indivíduos, são considerados cenários adicionais de escolha ou *holdouts*. Estes são conjuntos de alternativas que não são utilizadas na estimação dos *part-worths* [Hair, 1998]. Os parâmetros estimados são posteriormente utilizados para prognosticar as preferências entre as alternativas *holdout*, objetivando avaliar a validade e confiabilidade das estimativas originais. Em geral, quanto melhor prognosticadas as escolhas *holdout*, melhor a possibilidade de predição na vida real.

4.1.2 Simuladores de escolha

Geralmente, as metodologias conjuntas que fornecem parâmetros individuais de preferências utilizam algum tipo de simulador de escolha para “agregar” as preferências individuais e obter estimativas de participação de escolha para os produtos considerados. Estes simuladores assumem que os consumidores escolhem ou tendem a escolher aquelas alternativas que oferecem a maior utilidade total para eles [Finkbeiner, 1988; Huber *et al*, 1999].

Existem dois tipos básicos de simuladores de escolha conjunta [Bretton-Clark, 1992]. O mais empregado é o denominado Modelo *First Choice* ou Modelo de Máxima Utilidade, que assume que o consumidor sempre escolhe o produto com a maior utilidade para ele. Embora este modelo seja atrativo por ser simples e intuitivo, apresenta alguns inconvenientes: na prática tende a sobrestimar participação para produtos populares e subestimar participação para produtos menos populares. Esta propensão é devida às fortes suposições básicas do modelo, que implicitamente assume que não existe erro nos dados de utilidade e que todos os aspectos importantes do produto tem sido incluídos no estudo. No entanto, esta suposição é claramente incorreta e inconsistente com os princípios utilizados na estimação das funções utilidade.

Por outro lado, o Modelo Probabilístico assume que a probabilidade com que um consumidor escolhe um produto é proporcional a sua utilidade relacionada com as utilidades de todos os outros produtos do cenário de escolha. Ao contrário do modelo *First Choice*, o Modelo Probabilístico reconhece explicitamente a presença de erro na medição da função utilidade. Ele assume que a probabilidade de compra de um produto incrementa-se conforme sua utilidade aumenta em relação às utilidades dos outros produtos do cenário de simulação. Ainda que sob este modelo o consumidor pode escolher um produto com relativamente baixa utilidade para ele, em concordância com a teoria probabilística da utilidade, os resultados obtidos podem ser viesados. Este problema é consequência da hipótese de independência de alternativas irrelevantes (IIA) [seção 2.6.2]. Isto é, o Modelo Probabilístico considera apenas a utilidade relativa dos produtos e não a semelhança entre estes. Por exemplo, se um consumidor tem marcada preferência por veículos importados em relação aos nacionais, adicionar carros nacionais ao cenário da simulação, deveria ter pouca influência na probabilidade de escolha do carro importado. No entanto, o modelo probabilístico continuará

a diminuir a probabilidade de escolha do carro importado conforme mais carros nacionais sejam adicionados ao conjunto de escolha. Em essência o Modelo Probabilístico tende a subestimar participação para produtos de alta preferência e superestimar participação para produtos de menor preferência.

Huber *et al.* (1999) propõem o simulador *Randomized First Choice* (RFC), que combina as características desejáveis dos modelos *First Choice* e Probabilístico. O RFC adiciona variância a cada *part-worth* (e_A) e/ou à utilidade de cada produto (e_P),

$$U_i = X_i \cdot (\beta + e_A) + e_P \quad (4.17)$$

As participações de preferências são calculadas de forma semelhante ao *First Choice*. Cada entrevistado é amostrado muitas vezes para estabilizar as participações estimadas.

O simulador desenvolvido

O simulador proposto neste trabalho, ao contrário do simulador RFC, não precisa agregar variabilidade às funções utilidade. Este simulador considera a variabilidade intrínseca do processo individual de escolha por intermédio das distribuições marginais posteriores de cada parâmetro individual, caracterizadas pela sucessão dos valores amostrados [seção 4.1.1]. A Figura 4.1 apresenta esquematicamente o simulador de escolha proposto.

Deve-se observar que os indivíduos apresentam probabilidades de escolha para cada produto considerado em um cenário competitivo específico. Neste trabalho, interpreta-se estas probabilidades como uma medida da satisfação do indivíduo com o produto escolhido. Para operacionalizar esta informação, aloca-se cada indivíduo ao segmento correspondente ao produto que este escolhe o maior número de vezes. Posteriormente, visando identificar oportunidades de mercado, estuda-se a composição de cada segmento: grupos de indivíduos que apresentam “baixa” probabilidade de escolha -níveis menores de satisfação com a oferta atual- constituem-se em possíveis nichos atrativos de mercado. Admite-se que os consumidores mais propensos a mudar para um novo produto são aqueles que valoram melhor algum atributo ou combinação de atributos da nova oferta ou aqueles que menos valoram os produtos atuais.

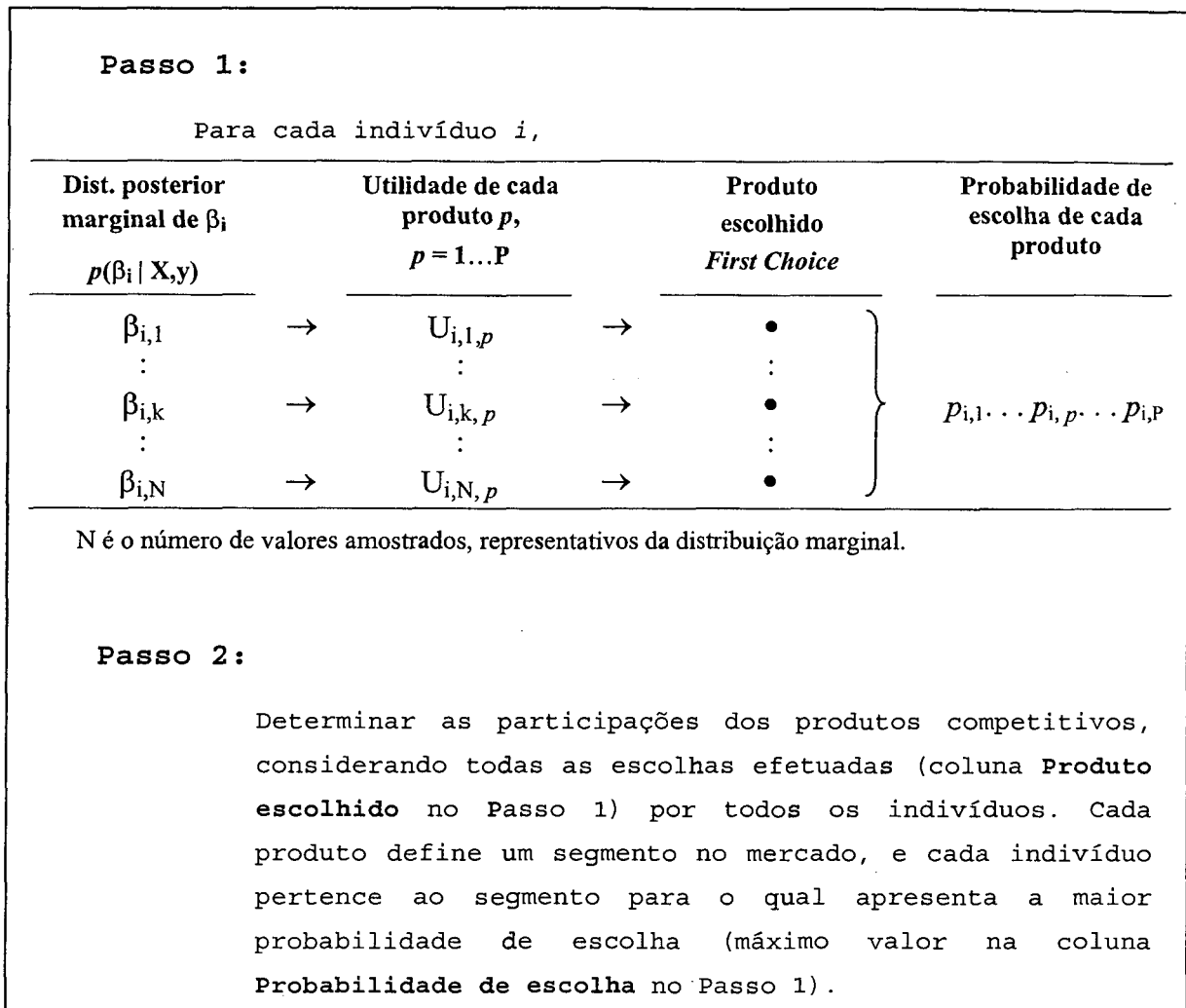


Figura 4.1 – Simulador de escolha.

A simulação de uma sequência de cenários, efetuando diferentes hipóteses competitivas em relação às ofertas da organização e dos concorrentes, possibilita avaliar estratégias de marketing em função das participações de escolha obtidas.

O próximo capítulo aborda a validação preditiva, mediante simulação, da metodologia proposta; e o capítulo subsequente ilustra sua aplicabilidade por intermédio de uma execução prática.

VALIDAÇÃO DO MODELO

Neste capítulo apresenta-se uma validação primária do modelo proposto efetuada mediante simulação. As aplicações consideradas utilizam dados artificiais gerados com simulações Monte Carlo [Vriens *et al* 1996]. Validação através de simulação apresenta duas vantagens de ordem prática: 1^o a flexibilidade de se trabalhar com dados sintéticos possibilita o teste do modelo em uma ampla gama de situações e 2^o conhece-se a priori qual a configuração exata do mercado virtual pesquisado: heterogeneidade na amostra, número de segmentos, tamanho dos segmentos, e quais os parâmetros que governam a distribuição de cada um dos segmentos. Por conseguinte, pode-se avaliar com alto grau de precisão o desempenho da capacidade do modelo na recuperação dos parâmetros conhecidos. A avaliação da capacidade preditiva é efetuada empregando os índices *Hit Rate* e MAE dos cenários de escolha *holdout*.

Todos os algoritmos tem sido implementados em Matlab 5.3.

5.1 Descrição do mercado virtual

Tem sido implementados algoritmos que permitem simular escolhas individuais logit ordenadas, segundo estrutura de preferências e configuração do mercado pré-especificados. O algoritmo que produz os dados sintéticos segue-se:

Definir

- QS = quantidade de segmentos;
- T_{seg} = tamanho de cada segmento;

Para $seg = 1 \dots QS$

Definir

- Vetor de parâmetros agregados α_{seg} ;
- Matriz de variância-covariância D_{seg} ;

Para cada entrevistado $i = 1 \dots T_{seg}$

- Obter β_i amostrando de $N(\alpha_{seg}, \mathbf{D}_{seg})_i$
- Para cada cenário de escolha:
 - Determinar a probabilidade de escolha logística das alternativas.
 - Escolher uma alternativa segundo as probabilidades definidas no passo anterior.
 - Repetir os dois passos anteriores até esgotar o conjunto de escolha.

Fim

Fim

5.2 Definição do estudo

Seja um produto/serviço/conceito qualquer, apresentado na Tabela 5.1.

Tabela 5.1 - Um estudo hipotético

Atributo	Níveis	Código <i>dummy</i>
At. 1	N. 1	1
	N. 2	0
At. 2	N. 1	1
	N. 2	0
At. 3	N. 1	1
	N. 2	0
At. 4	N. 1	1
	N. 2	0
At. 5	N. 1	1
	N. 2	0
At. 6	N. 1	1
	N. 2	0

5.3 Delineamento do experimento

O estudo hipotético se constitui num fatorial com 6 atributos e 2 níveis, $2^6 = 64$ alternativas possíveis. Pode-se utilizar o arranjo ortogonal L8 [Taguchi, 1991]¹ apresentado na Tabela 5.2.

¹ Experimentos ortogonais são considerados eficientes devido a sua capacidade de avaliar diversos fatores e adquirir grande quantidade de informação com número mínimo de testes ou ensaios [Ross, 1991].

Tabela 5.2 - Matriz Arranjo Ortogonal L8.

Alternativas	Atributos					
	1	2	3	4	5	6
1	0	0	0	0	0	0
2	0	0	0	1	1	1
3	0	1	1	0	0	1
4	0	1	1	1	1	0
5	1	0	1	0	1	0
6	1	0	1	1	0	1
7	1	1	0	0	1	1
8	1	1	0	1	0	0

Para diminuir o cenário de escolha, foram ainda montados 14 blocos incompletos balanceados [Montgomery, 1984; Souza, 1999] com quatro alternativas, de forma que par de alternativas compareça juntas o mesmo número de vezes em todos os grupos.

Tabela 5.3 - Blocos incompletos balanceados.

Grupos	Alternativas
01	1 2 5 6
02	1 3 5 7
03	1 4 5 8
04	3 4 7 8
05	2 4 6 8
06	2 3 6 7
07	3 4 5 6
08	1 2 7 8
09	1 3 6 8
10	2 4 5 7
11	1 4 6 7
12	2 3 5 8
13	1 2 3 4
14	5 6 7 8

Adicionalmente, com a finalidade de testar a validade preditiva do modelo proposto, definem-se cenários de escolha *holdout*. O conjunto de alternativas utilizado, formado segundo a regra (4.16), é apresentado na Tabela 5.4.

Tabela 5.4 - Alternativas para formação de *holdouts*.

Alternativas	Atributos					
	1	2	3	4	5	6
9	1	1	1	1	1	0
10	1	1	1	1	0	1
11	1	1	1	0	1	1
12	1	1	0	1	1	1
13	1	0	1	1	1	1
14	0	1	1	1	1	1

Opta-se pelos cenários *holdout* apresentados na sequência:

Tabela 5.5 – Cenários *holdout*

<i>Holdouts</i>	Alternativas
15	9 10 11 12
16	11 12 13 14
17	10 11 12 13
18	9 10 13 14

5.4 Situações simuladas

Com a finalidade de testar a validade do modelo proposto tem sido simuladas duas situações.

5.4.1 1^{ra} situação: *Mercado homogêneo*

Utilizando o algoritmo descrito na seção 5.1, simulam-se 200 entrevistas para um mercado virtual homogêneo, isto é, $QS = 1$, $T_{seg} = 200$; com os seguintes hiperparâmetros:

$$\alpha_{seg} = [0.89 \ 0.40 \ 0.67 \ 1.20 \ 0.92 \ 0.78]$$

$$\mathbf{D}_{seg} = \mathbf{I}_6, \text{ matriz identidade de dimensão 6.}$$

Estudaram-se seis casos, que correspondem a diferente número de cenários de escolha por indivíduo (8, 6, 5, 4, 3 e 2 cenários por indivíduo), com a finalidade de identificar a capacidade preditiva do modelo com reduzido número de cenários.

Para cada caso foi verificada a convergência, considerando cinco sequências de amostragem MCMC com diferentes valores iniciais para os *part-worths*:

- sequência 1: MV (estimador de máxima verossimilhança),
- sequência 2: $MV + 2\sigma$,
- sequência 3: $MV - 2\sigma$,
- sequência 4: uns,
- sequência 5: zeros.

Para cada sequência consideram-se 2000 amostras, as primeiras 1000 amostras são desconsideradas nos cálculos. A Tabela 5.6 apresenta os valores do índice $\sqrt{\hat{R}}$ obtido em cada caso. Deve-se verificar a convergência de cada componente do vetor de *part-worths*. Observa-se que, ainda com reduzido número de cenários de ordenação por indivíduo, a convergência é atingida, $\sqrt{\hat{R}} < 1.2$.

Tabela 5.6 – Análise de convergência: Mercado homogêneo.

Caso	Cenários de escolha por indivíduo	$\sqrt{\hat{R}}$					
		β_1	β_2	β_3	β_4	β_5	β_6
1	8	1.0077	1.0057	1.0080	1.0197	1.0096	1.0093
2	6	1.0580	1.0489	1.0628	1.0505	1.0831	1.0509
3	5	1.0258	1.0156	1.0251	1.0420	1.0642	1.0249
4	4	1.0656	1.0357	1.0179	1.0341	1.0610	1.0622
5	3	1.0292	1.0391	1.0080	1.0218	1.0201	1.0664
6	2	1.1038	1.0506	1.0509	1.1581	1.1215	1.1151

A Tabela 5.7 apresenta os *part-worths* agregados, α , obtidos com o modelo de inferência proposto. Os parâmetros estimados aproximam-se muito bem dos parâmetros reais, como é indicado pelos valores da norma euclidiana da diferença entre os vetores de parâmetros estimado e real, sem existir maior variabilidade aparente em relação ao número de cenários por indivíduo. Deve-se ressaltar que, embora os parâmetros agregados obtidos com o modelo proposto se aproximam dos parâmetros reais, o interesse principal é testar a validade preditiva do modelo, o que será abordado posteriormente.

Tabela 5.7 - *Part-worths* agregados, por cenário.

Cenários de escolha por indivíduo	α						Norma euclidiana
	α_1	α_2	α_3	α_4	α_5	α_6	
	0.89	0.40	0.67	1.20	0.92	0.78	
8	1.028	0.575	0.930	1.286	0.883	1.117	0.490
6	1.074	0.426	0.911	1.282	0.874	1.163	0.498
5	1.071	0.527	0.900	1.262	0.795	1.151	0.509
4	0.991	0.478	0.757	1.286	0.749	0.993	0.325
3	0.934	0.496	0.880	1.318	0.821	1.247	0.545
2	1.054	0.428	0.807	1.374	0.891	1.151	0.464

A Tabela 5.8 apresenta a matriz de variância-covariância \mathbf{D} apenas para os casos 1 e 6. Para 8 cenários por indivíduo os elementos diagonais da matriz \mathbf{D}_8 são mais próximos de um, e os elementos fora da diagonal mais próximos de zero, quando comparados com os respectivos elementos da matriz \mathbf{D}_2 .

Tabela 5.8 - Matriz \mathbf{D} para 8 e 2 cenários de escolha.

$\mathbf{D}_8 =$	1.043	0.160	-0.067	0.034	0.131	-0.060
	0.160	1.183	0.104	0.143	-0.036	0.160
	-0.067	0.104	0.880	0.150	0.207	0.121
	0.034	0.143	0.150	0.968	0.159	0.081
	0.131	-0.036	0.207	0.159	1.343	0.135
	-0.060	0.161	0.121	0.081	0.135	1.086
$\mathbf{D}_2 =$	1.317	0.129	0.274	-0.151	-0.118	-0.116
	0.129	1.329	-0.180	-0.011	-0.168	-0.368
	0.274	-0.180	1.376	0.286	0.390	-0.040
	-0.151	-0.011	0.286	1.542	0.388	0.426
	-0.118	-0.168	0.390	0.388	1.502	0.024
	-0.116	-0.368	-0.040	0.426	0.024	1.446

A Figura 5.1 apresenta a evolução da norma da diferença entre a matriz de variância-covariância \mathbf{D} e a matriz identidade \mathbf{I} , em função do número de cenários apresentados a cada entrevistado. Observa-se que a estimativa de \mathbf{D} é aprimorada quando dispõe-se de maior informação por entrevistado.

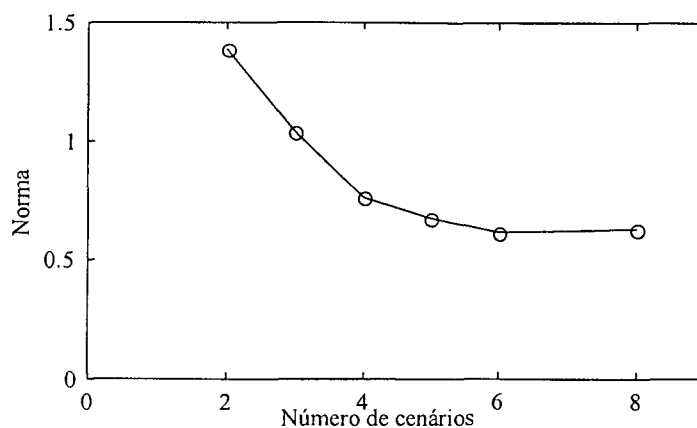


Figura 5.1 - Evolução da norma das diferenças: $D_i - I$.

Com a finalidade de observar a qualidade de ajuste, a Tabela 5.9 apresenta as porcentagens de acertos, obtidas com o modelo proposto (MP), na reprodução das escolhas utilizadas na inferência dos parâmetros.

Tabela 5.9 – Porcentagens de acertos em cada cenário.

<i>Caso</i>	<i>Nº de cenários</i>	<i>MP</i>
1	8	0.684
2	6	0.721
3	5	0.731
4	4	0.755
5	3	0.783
6	2	0.830

O desempenho preditivo do modelo é estudado mediante o cálculo dos índices *Hit Rate* e MAE para os cenários de escolha *holdout* da Tabela 5.5. A Tabela 5.10 apresenta o *Hit Rate* obtido com os parâmetros individuais β_i e com os parâmetros agregados α , ambos estimados com o modelo proposto, em cada cenário *holdout*. Adicionalmente, a Figura 5.2 mostra o gráfico da média do *Hit Rate* dos conjuntos *holdout* para cada caso.

Tabela 5.10 – *Hit Rate* (parâmetros individuais β_i vs parâmetros agregados α).

Caso	Cenários de escolha	Holdout 15		Holdout 16		Holdout 17		Holdout 18	
		β_i	α	β_i	α	β_i	α	β_i	α
1	8	0.465	0.350	0.485	0.335	0.565	0.360	0.485	0.370
2	6	0.390	0.350	0.485	0.335	0.470	0.360	0.430	0.370
3	5	0.460	0.350	0.440	0.335	0.455	0.360	0.455	0.370
4	4	0.415	0.350	0.385	0.335	0.430	0.360	0.365	0.370
5	3	0.390	0.350	0.380	0.335	0.450	0.360	0.405	0.370
6	2	0.310	0.230	0.385	0.335	0.430	0.360	0.390	0.370

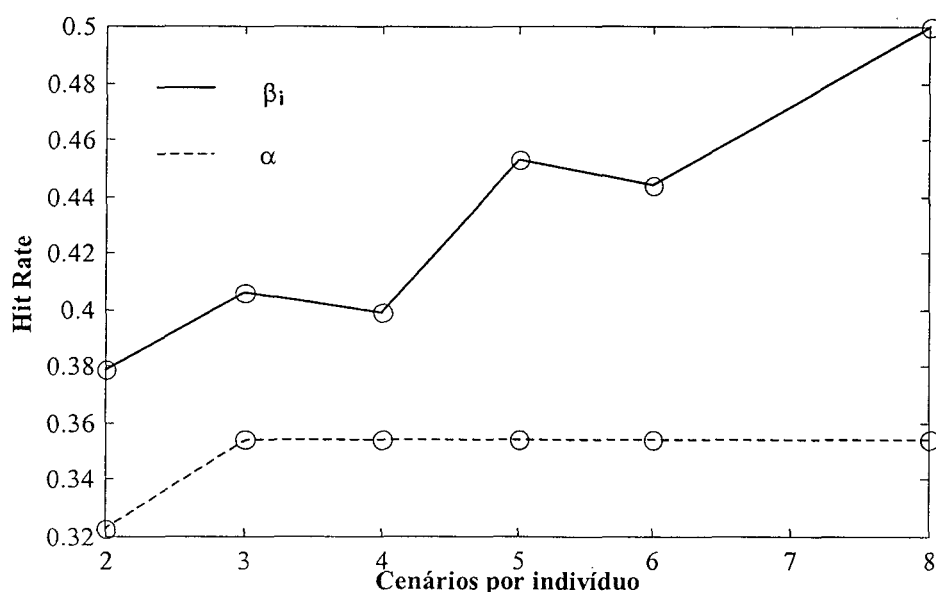


Figura 5.2 - Evolução da média do *Hit Rate*.

A Figura 5.2 evidencia claramente a superioridade preditiva dos parâmetros individuais. Ainda, a estimação desagregada faz uma melhor utilização da informação disponível -quantidade de escolhas- já que o *Hit Rate* é crescente com o número de cenários.

A Tabela 5.11 apresenta o MAE em cada cenário *holdout*. A Figura 5.3 mostra o gráfico da média do MAE dos cenários *holdout* para cada caso. Observa-se que existe uma tendência à estabilização do índice à medida que se incrementa a quantidade de ordenações por indivíduo.

Tabela 5.11 – MAE para cada cenário *holdout*.

<i>Caso</i>	Cenários de escolha	<i>Holdout 15</i>	<i>Holdout 16</i>	<i>Holdout 17</i>	<i>Holdout 18</i>
1	8	0.0285	0.0236	0.0222	0.0246
2	6	0.0286	0.0296	0.0240	0.0184
3	5	0.0245	0.0222	0.0199	0.0260
4	4	0.0367	0.0179	0.0301	0.0226
5	3	0.0321	0.0299	0.0223	0.0334
6	2	0.0415	0.0268	0.0342	0.0197

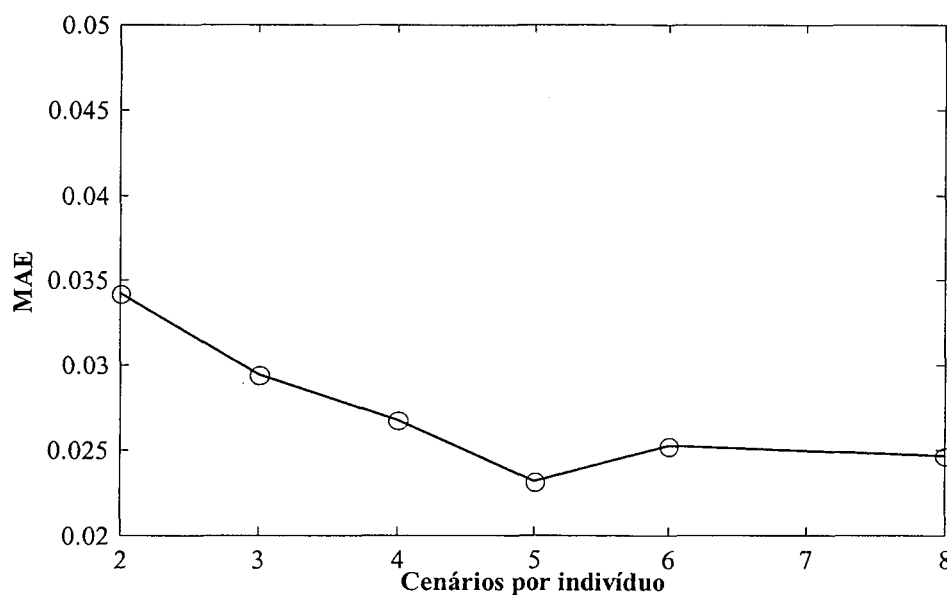


Figura 5.3 - Evolução da média do MAE.

Ressalta-se que a recuperação da heterogeneidade das preferências por intermédio de parâmetros individuais, inclusive se tratando de mercados “homogêneos”, fornece melhores resultados preditivos que a utilização de parâmetros agregados.

5.4.2 2^{da} situação: *Mercado com dois segmentos*

O mercado virtual nesta situação assume a existência de dois grupos de consumidores claramente definidos. Procura-se testar a capacidade do modelo em identificar grupos com preferências diferentes e os parâmetros que governam a distribuição destes. Consideram-se seis cenários de escolha por indivíduo.

Para $seg = 1$, $T_{seg} = 100$ indivíduos; com os hiperparâmetros:

$$\alpha_1 = [-0.90 \quad 0.40 \quad 0.70 \quad 1.20 \quad 0.90 \quad 0.80]$$

$$\mathbf{D}_1 = \mathbf{I}_6$$

Para $seg = 2$, $T_{seg} = 100$ indivíduos; com os hiperparâmetros:

$$\alpha_2 = [+0.90 \quad 0.40 \quad 0.70 \quad 1.20 \quad 0.90 \quad 0.80]$$

$$\mathbf{D}_2 = \mathbf{I}_6$$

Os parâmetros agregados α estimados com o modelo proposto (MP) são apresentados na Tabela 5.12 e na Figura 5.4. Observa-se que o *part-worth* α_1 é próximo a zero, reportando que o atributo 1 é de pouca importância relativa. Este resultado é esperado, já que ao efetuar a agregação de dois segmentos com preferências opostas em relação a esse atributo, os efeitos são compensados. Isto ilustra o fato, indicado na seção 2.7, que modelos agregados produzem estimadores viesados quando existe heterogeneidade na amostra. Destaca-se que o modelo proposto estima uma função utilidade para cada indivíduo, onde α é apenas a média dos *part-worths* individuais.

Tabela 5.12 - *Part-worths* agregados α .

	α_1	α_2	α_3	α_4	α_5	α_6
Real	+/- 0.90	0.40	0.70	1.20	0.90	0.80
MP	-0.016	0.373	0.679	0.988	0.869	0.842

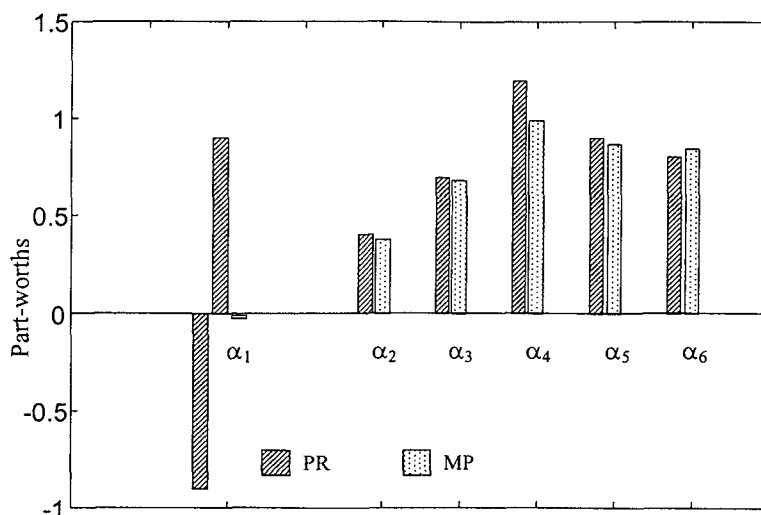


Figura 5.4 – *Part-worths* agregados.
PR = parâmetros reais, MP = modelo proposto.

A Tabela 5.13 apresenta a matriz de variância-covariância dos *part-worths* individuais obtida com o modelo proposto. Observa-se que os elementos fora da diagonal são próximos de zero. Os elementos diagonais, exceto o $D(1,1)$, são próximos de 1 e o elemento $D(1,1)$ é próximo de 2. Deve-se ressaltar que esta matriz representa a heterogeneidade entre os indivíduos, assim os elementos diagonais medem a variabilidade das preferências do mercado em relação a um atributo. A alta variância identificada no elemento $D(1,1)$ evidencia a existência de heterogeneidade no mercado em relação a esse atributo.

Tabela 5.13 - Matriz D para 6 cenários por indivíduo.

$D =$	2.084	0.069	0.167	0.099	-0.058	-0.043
	0.069	1.100	0.086	-0.032	0.121	-0.050
	0.167	0.086	0.857	-0.072	0.012	0.019
	0.099	-0.032	-0.072	0.908	-0.153	-0.101
	-0.058	0.121	0.012	-0.153	0.863	0.129
	-0.043	-0.050	0.019	-0.101	0.129	1.237

A Figura 5.5 apresenta graficamente os parâmetros estimados β_i para os 200 indivíduos entrevistados. Evidencia-se claramente a existência de heterogeneidade nas preferências do mercado em relação ao atributo 1, já que para $\beta(1)$ os indivíduos $i = 1...100$ estão agrupados em torno de um valor diferente dos indivíduos $i = 101...200$.

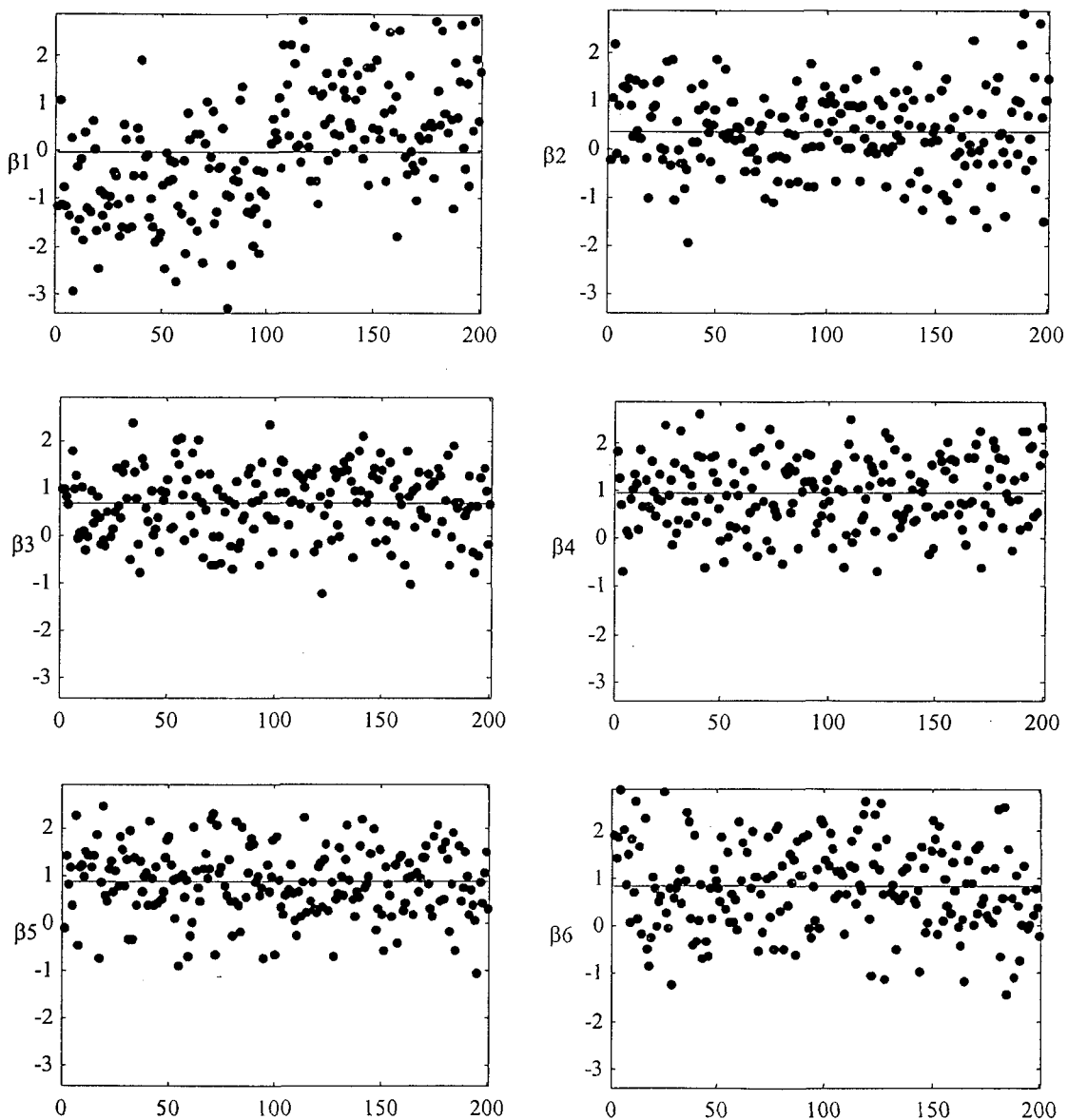


Figura 5.5 – *Part-worths* individuais.

A Tabela 5.14 e a Figura 5.6 apresentam os parâmetros agregados para cada segmento. Observa-se claramente que os valores estimados reproduzem as características reais dos segmentos.

Tabela 5.14 - *Part-worths* agregados α inferidos para os dois segmentos.

		α_1	α_2	α_3	α_4	α_5	α_6
seg = 1	Real	- 0.90	0.40	0.70	1.20	0.90	0.80
	MP	- 0.788	0.381	0.616	0.943	0.930	0.872
seg = 2	Real	0.90	0.40	0.70	1.20	0.90	0.80
	MP	0.755	0.364	0.742	1.033	0.807	0.812

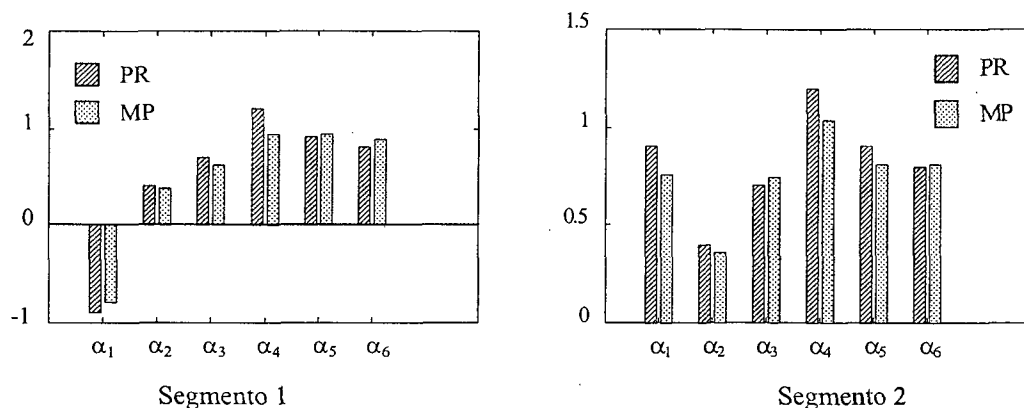


Figura 5.6 – Parâmetros agregados em cada segmento, PR = parâmetros reais, MP = modelo proposto.

A Tabela 5.15 apresenta os índices de desempenho *Hit Rate* e MAE para cada cenário *holdout*.

Tabela 5.15 – Índices de desempenho preditivo.

<i>Holdout</i>	<i>Hit Rate</i>	MAE
15	0.445	0.0216
16	0.525	0.0338
17	0.480	0.0442
18	0.510	0.0286
Média	0.490	0.0321

Os parâmetros agregados apresentados na Tabela 5.14 foram inferidos utilizando:

$$\bar{\alpha}_{seg} = \frac{1}{T_{seg}} \cdot \sum_{i=T_{seg-1}+1}^{T_{seg}} \beta_i, \quad seg = 1, \dots, QS \quad (5.1)$$

isto é,

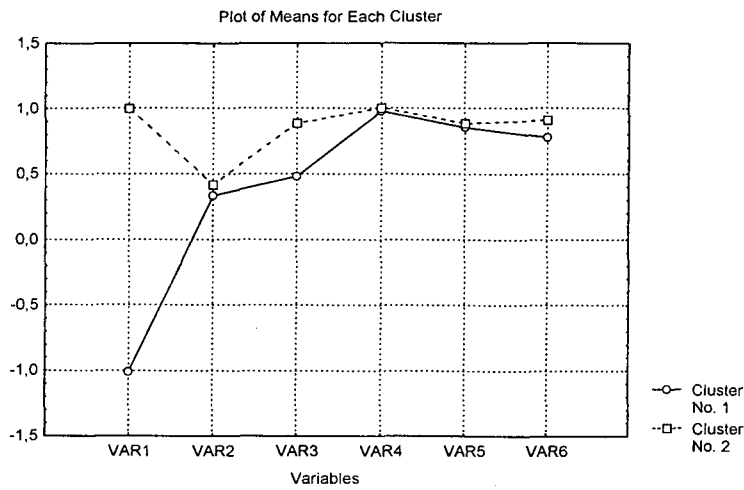
$$\bar{\alpha}_{seg=1} = \frac{1}{100} \cdot \sum_1^{100} \beta_i, \quad \bar{\alpha}_{seg=2} = \frac{1}{100} \cdot \sum_{101}^{200} \beta_i \quad (5.2)$$

Porém, em um estudo real, a quantidade de segmentos (QS), o tamanho de cada um deles (T_{seg}) e os indivíduos que pertencem a cada segmento são, logicamente, desconhecidos a priori. O objetivo deste exemplo é ilustrar a aplicabilidade do modelo proposto em um estudo de segmentação de mercado.

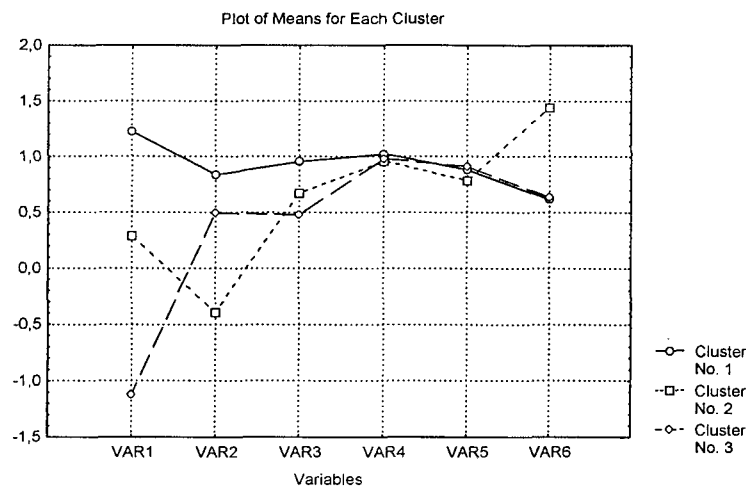
O interesse é agrupar os indivíduos em função do comportamento de escolha. Busca-se identificar grupos de consumidores que apresentem alta homogeneidade no comportamento de escolha dentro de cada segmento e alta heterogeneidade entre os segmentos identificados. Para tanto, pode aplicar-se um procedimento de *cluster analysis*¹, neste exemplo emprega-se o método *k-means* [Bussab *et al*, 1990; Malhotra, 1995; Johnson e Wichern, 1998]. O algoritmo *k-means* precisa da estipulação da quantidade de *clusters* a serem formados. Dado que isto não é conhecido a priori é necessário considerar diferentes possibilidades, nesta aplicação considera-se a segmentação em 2, 3 e 4 grupos.

Observa-se claramente na Figura 5.7 que os *clusters*, em quaisquer das situações consideradas, são discriminados pela primeira variável (atributo 1).

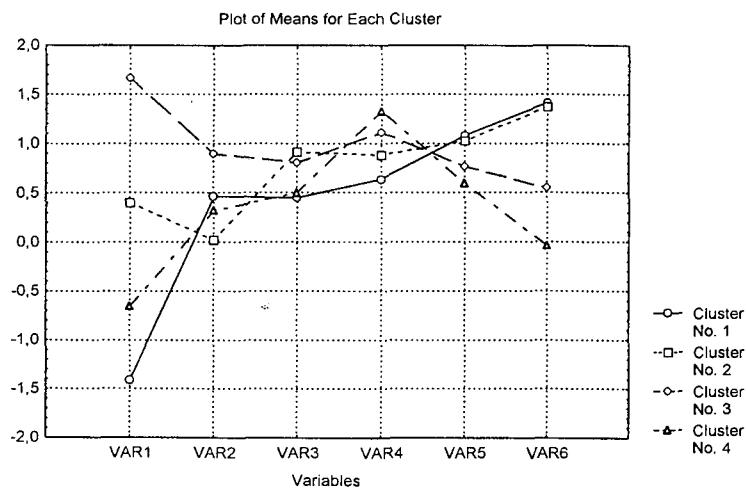
¹ *Cluster Analysis*, também denominada análise de segmentação ou de agrupamento, é uma técnica multivariada cuja finalidade é agrupar objetos com base nas características que estes possuem [Hair *et al*, 1998; Gore, 2000]. *Cluster analysis* classifica objetos de forma que os perfis destes no mesmo grupo são muito similares e perfis de objetos em distintos grupos são bastante diferentes, em relação a algum critério predeterminado.



Cluster	Tamanho
1	101
2	99



Cluster	Tamanho
1	63
2	52
3	85



Cluster	Tamanho
1	44
2	63
3	41
4	52

Figura 5.7 – Média dos *part-worths*, para 2, 3 e 4 clusters.

A Tabela 5.16 apresenta os resultados da análise de variância para as 3 situações consideradas. A magnitude do valor F é outro indicador do grau de discriminação das variáveis entre os *clusters*. Observa-se que em todos os casos o valor F para o atributo 1 é o maior.

Tabela 5.16 – Análise de variância.

	B	W	F	p
	Inércia inter-classes	Inércia intra-classes	$\frac{B}{W} \cdot \frac{g.l.W}{g.l.B}$	Nível de significância
2 clusters				
	g.l.=1	g.l.=198		
1	201,8435	113,9459	350,7366	0,000000
2	0,3527	140,3477	0,4976	0,481399
3	8,1288	92,0130	17,4920	0,000043
4	0,0362	108,4736	0,0661	0,797306
5	0,0510	100,5966	0,1005	0,751605
6	0,8418	162,7354	1,0242	0,312763
3 clusters				
	g.l.=2	g.l.=197		
1	206,6302	109,1592	186,4532	0,000000
2	45,6412	95,0591	47,2933	0,000000
3	8,1853	91,9564	8,7678	0,000225
4	0,1192	108,3906	0,1083	0,897368
5	0,5274	100,1203	0,5188	0,596027
6	25,4120	138,1652	18,1166	0,000000
4 clusters				
	g.l.=3	g.l.=196		
1	233,8583	81,9310	186,4830	0,000000
2	19,6234	121,0770	10,5888	0,000002
3	8,0514	92,0903	5,7121	0,000908
4	12,6077	95,9022	8,5890	0,000022
5	7,9097	92,7379	5,5724	0,001090
6	75,4273	88,1499	55,9039	0,000000

A decisão de quantos segmentos considerar é de fato uma questão estratégica e esta ainda poderá ser posteriormente assistida pelo simulador de fatias de preferências. No entanto, seja qual for a decisão tomada, o algoritmo identificará segmentos com a máxima distinção possível. Nesta aplicação opta-se por dois segmentos.

Os segmentos identificados na análise de agrupamento podem, eventualmente, ser descritos em termos de um conjunto de variáveis sócioeconômicas, que possibilitam rotular grupos de consumidores com preferências diferentes. Estes rótulos caracterizam o perfil dos clientes atuais e/ou potenciais da organização. O adequado conhecimento do perfil dos consumidores e dos motivos que os impulsionam à escolha, ajudam às organizações não apenas na configuração de melhores ofertas, mas também na formulação de melhores planos de marketing para o posicionamento psicológico de seus produtos e serviços.

Para completar o presente exemplo geram-se características sócioeconômicas para cada entrevistado. Consideram-se apenas 3 características para descrever os consumidores: Z_1^k ($k=1, 2, 3, 4$), Z_2^k ($k=1, 2, 3$) e Z_3^k ($k=1, 2$). Supõe-se que unicamente a primeira variável discrimina os segmentos, com a distribuição de probabilidades mostrada na Tabela 5.17. Os níveis para Z_2 e Z_3 foram simulados com distribuição uniforme.

Tabela 5.17 - Distribuição de probabilidade multinomial, para simulação de Z_1 .

Nível	1	2	3	4
1 ^o segmento [1,100]	1/12	2/12	3/12	6/12
2 ^o segmento [101,200]	6/12	3/12	2/12	1/12

Explora-se a possível interdependência entre os grupos identificados e as variáveis sócioeconômicas aplicando uma técnica de análise multivariada de dados: Análise de Correspondência [Kaciak e Louviere, 1990; Hoffman *et al.*, 1994; Malhotra, 1995; Johnson e Wichern, 1998]. O objetivo desta é a redução da quantidade de dados a serem analisados, utilizando procedimentos de estatística multivariada, de forma que seja possível analisar uma quantidade maior de grupos de consumidores e de variáveis sócioeconômicas simultaneamente, a partir de um espaço com dimensões reduzidas.

A Tabela 5.18 apresenta as frequências observadas, isto é, a quantidade de indivíduos discriminada por *clusters* e características sócioeconômicas. A partir desta tabela pode-se testar a hipótese de independência entre linhas e colunas. A estatística de contraste admite que, para amostras grandes, a correspondência segue uma distribuição assintótica χ^2 com $(linhas-1)*(colunas-1)$ graus de liberdade:

$$\sum_{colunas} \sum_{linhas} \frac{(fo - fe)^2}{fe} \sim \chi^2 \quad (5.3)$$

Tabela 5.18 – Frequências observadas.

	C ₁	C ₂	Z ₁ ¹	Z ₁ ²	Z ₁ ³	Z ₁ ⁴	Z ₂ ¹	Z ₂ ²	Z ₂ ³	Z ₃ ¹	Z ₃ ²	Total
C ₁	101	0	21	12	25	43	34	42	25	52	49	404
C ₂	0	99	31	31	21	16	39	32	28	43	56	396
Z ₁ ¹	21	31	52	0	0	0	18	19	15	25	27	208
Z ₁ ²	12	31	0	43	0	0	14	15	14	24	19	172
Z ₁ ³	25	21	0	0	46	0	18	15	13	20	26	184
Z ₁ ⁴	43	16	0	0	0	59	23	25	11	26	33	236
Z ₂ ¹	34	39	18	14	18	23	73	0	0	36	37	292
Z ₂ ²	42	32	19	15	15	25	0	74	0	32	42	296
Z ₂ ³	25	28	15	14	13	11	0	0	53	27	26	212
Z ₃ ¹	52	43	25	24	20	26	36	32	27	95	0	380
Z ₃ ²	49	56	27	19	26	33	37	42	26	0	105	420
Total	404	396	208	172	184	236	292	296	212	380	420	3200

Total χ^2 : 1464,23 Graus de liberdade: 100 Significância p = 0,000

O teste χ^2 , apresentado na última linha da tabela anterior, indica que existe relação estatisticamente significativa entre as linhas e colunas ($p < 0,001$). Portanto, é apropriado efetuar uma análise de correspondência.

Pode-se efetuar uma exploração dos resíduos (diferenças entre as frequências observada e esperada) para que sejam identificadas as principais associações (dependências) entre as categorias linha/coluna das variáveis. A Tabela 5.19 apresenta estes resíduos e as associações existentes podem ser facilmente identificadas. As células com resíduos positivos

mais elevados encontram-se mais fortemente associadas e, por analogia, aquelas células com resíduos negativos maiores indicam fortes dissociações existentes.

Tabela 5.19 – Resíduos: diferença entre frequências observada e esperada.

	C_1	C_2	Z_1^1	Z_1^2	Z_1^3	Z_1^4	Z_2^1	Z_2^2	Z_2^3	Z_3^1	Z_3^2
C_1	50	-50	-5	-10	2	13	-3	5	-2	4	-4
C_2	-50	50	5	10	-2	-13	3	-5	2	-4	4
Z_1^1	-5	5	38	-11	-12	-15	-1	0	1	0	0
Z_1^2	-10	10	-11	34	-10	-13	-2	-1	3	4	-4
Z_1^3	2	-2	-12	-10	35	-14	1	-2	1	-2	2
Z_1^4	13	-13	-15	-13	-14	42	1	3	-5	-2	2
Z_2^1	-3	3	-1	-2	1	1	46	-27	-19	1	-1
Z_2^2	5	-5	0	-1	-2	3	-27	47	-20	-3	3
Z_2^3	-2	2	1	3	1	-5	-19	-20	39	2	-2
Z_3^1	4	-4	0	4	-2	-2	1	-3	2	50	-50
Z_3^2	-4	4	0	-4	2	2	-1	3	-2	-50	50

As relações de interdependência são melhor visualizadas no mapa da análise de correspondência, Figura 5.8, onde todas as variáveis são representadas simultaneamente logo após ter sido escalonadas em unidades correspondentes. A representação gráfica, principal contribuição da Análise de Correspondência, possibilita visualizar como em um mapa as proximidades/distâncias entre os grupos e as variáveis sócioeconômicas, permitindo deprender rapidamente as relações existentes entre estes.

Tem-se considerado os dois maiores autovalores obtidos, que representam o 61,5% da informação explicada pelos eixos (autovetores). Este mapa é interpretado como um mapa geográfico, onde distâncias menores representam associações mais fortes e distâncias maiores representam dissociações ou repulsões entre as variáveis. Analisando o mapa depreende-se que Z_1^4 e C_1 (*cluster 1*) por um lado e Z_1^1 e C_2 (*cluster 2*) por outro estão fortemente associadas, o que é coerente com a situação simulada.

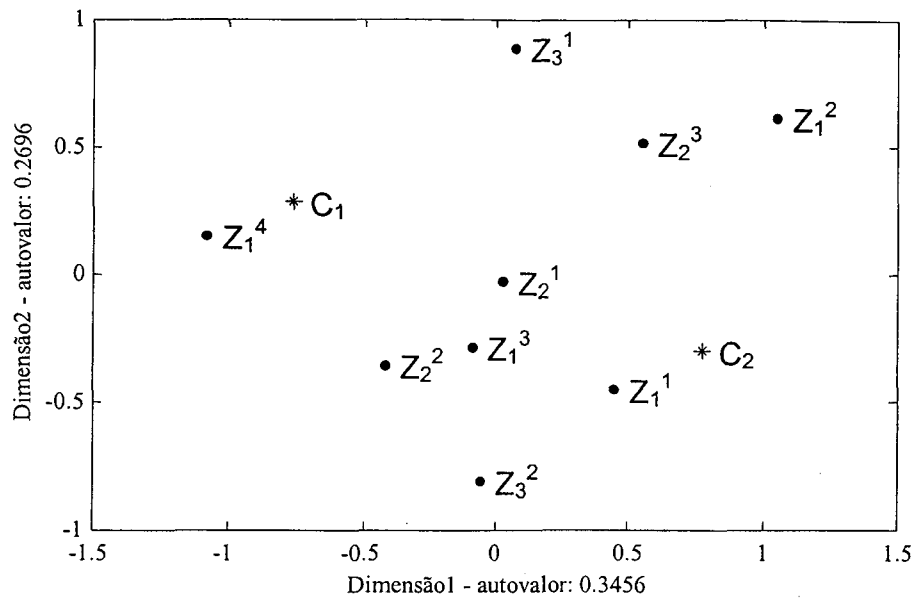


Figura 5.8 – Mapa da Análise de Correspondência

Neste capítulo foi verificada a capacidade do modelo de recuperar parâmetros conhecidos. Adicionalmente foi ilustrada uma das formas pelas quais o modelo pode ser integrado com técnicas de análise multivariada para efetuar estudos de segmentação de mercado. No próximo capítulo apresenta-se uma aplicação empírica do modelo.

APLICAÇÃO EMPÍRICA

O estudo de como as pessoas aprendem tem recebido muita atenção, porém a identificação de como as pessoas *desejam* aprender poderia potencialmente agregar valor ao processo de aprendizado. Cohen *et al* (2001) colocam a seguinte questão: porque não tratar os estudantes como consumidores com preferências distintas em relação a que, como e quando eles aprendem? E porque não avaliar essas preferências antes de criar soluções de aprendizado?

Nesse sentido esta aplicação prática, inserida em um contexto de marketing educacional, visualiza cursos de ensino como produtos que tem como clientes os estudantes que poderiam tomar estes cursos. Na literatura encontram-se alguns trabalhos que abordam o problema de projeto de cursos educacionais utilizando *conjoint analysis* métrico [Zufryden 1983; Dubas e Strong 1993; Katzenstein *et al* 1994; Tarasewich e Nair 2000]. Shukla e Bruno (1992) declaram que *conjoint analysis* é um método sofisticado que pode ser utilizado por educadores para obter *insigth* prático da real estrutura de preferências dos usuários de serviços educacionais.

6.1 Metodologia da pesquisa

Para realizar-se um estudo de pesquisa utilizando a metodologia desenvolvida devem ser consideradas as etapas apresentadas na Figura 6.1. É preciso ressaltar que a única finalidade desta aplicação empírica é ilustrar a aplicabilidade da metodologia apresentada nesta tese. Não houve portanto, maior preocupação com alguns fatores importantes que precisam ser considerados em qualquer aplicação real de um estudo conjunto: definição dos atributos e níveis a serem pesquisados e seleção, tamanho e composição da amostra pesquisada.

-
1. Definição dos objetivos da pesquisa.
 2. Definição dos atributos e níveis.
 3. Formulação do delineamento experimental.
 4. Definição da população alvo.
 5. Determinação do tamanho da amostra.
 6. Coleta dos dados conjuntos.
 7. Estimação dos parâmetros.
 8. Validação dos parâmetros individuais.
 9. Análise competitiva através do simulador de escolhas.
-

Figura 6.1 - Metodologia de pesquisa

6.1.1 Objetivo do estudo

O propósito desta aplicação é estudar as preferências dos estudantes em relação a cursos de pós-graduação em marketing, visando a incorporação destas no projeto de ditos cursos.

6.1.2 Definição dos atributos e níveis

Os atributos podem ser identificados por meio de discussões com a administração, análise de dados secundários, pesquisa qualitativa e análise fatorial. Após a identificação dos atributos importantes, é preciso selecionar seus níveis adequados. A quantidade de níveis de atributos determina a quantidade de parâmetros a serem estimados e conseqüentemente influencia a quantidade de alternativas a serem avaliadas pelos entrevistados. Dado que esta aplicação tem um caráter puramente ilustrativo, os atributos e níveis do estudo, descritos na Tabela 6.1, tem sido definidos pelo autor da tese.

Tabela 6.1 – Atributos e níveis.

Atributo	Níveis	Descrição
Nome do curso	Marketing Research (Res)	Capacitar profissionais na procura e análise sistemática e objetiva de informação relevante para a identificação e solução de problemas de marketing. Considerável ênfase é dada ao fornecimento de prática na implementação de métodos quantitativos e qualitativos de pesquisa de marketing.
	Marketing Internacional (Int)	Capacitar profissionais na identificação de oportunidades de internacionalização das empresas num âmbito de uma economia globalizada. Ênfase é dada a formulação e implementação de planos de ação mercadológica em negócios internacionais.
	Marketing Eletrônico (Ele)	Capacitar profissionais no planeamento, programação e operações de gestão de marketing com o uso de recursos de informática e da Internet para executar as atividades de obter e de servir à demanda na nova economia da era da informação.
Certificação	Especialização (Esp)	Curso lato sensu, oferece uma abordagem bastante especializada e voltada para o mercado de trabalho Duração: 12 meses. Confere-se o certificado de especialista.
	Mestrado Profissional (Pro)	Prioridade à prática. Maior ênfase é dada a disciplinas de carácter prático e metodológico. A dissertação é uma pesquisa aplicada, produto ou protótipo. Duração: 24 meses, 12 de aulas e 12 para a dissertação. Período parcial. Calendário letivo flexível. Confere-se o título de mestre.
	Mestrado Acadêmico (Aca)	Voltado para a pesquisa científica, sem deixar de ter uma visão aplicada. Período integral. Duração: 24 meses, 12 de aulas e 12 para a dissertação. Período integral Confere-se o título de mestre.
Método de ensino	Convencional (Con)	O curso é conduzido mediante aulas expositivas e discussões de assuntos ligados aos objetivos específicos, incentivando-se constantemente a participação e as contribuições dos alunos.
	Experiencial (Exp)	Em adição às aulas expositivas, utiliza-se: Estudos de caso: O estudante avalia o problema e faz sugestões para a tomada de decisão. Jogos de empresas ¹ : O estudante é imerso num ambiente empresarial simulado e toma decisões considerando as principais variáveis que agem nesse ambiente.
	Real (Rea)	Em adição as aulas expositivas, os estudantes examinam uma situação real e formulam um plano para resolver um problema ou atingir um objetivo. Elaboração de um artigo relatando os resultados obtidos.
Habilidades requeridas	Leves (Lev)	Capacidade de redação técnica e de pesquisa bibliográfica. Condições de utilizar conceitos matemático-estatísticos básicos aplicados à Administração. Conhecimento de planilhas eletrônicas.
	Moderadas (Mod)	Em adição às anteriores: Conhecimento de algum pacote estatístico. Conceitos da teoria das probabilidades e inferência estatística. Análise de regressão linear.
	Altas (Alt)	Em adição às anteriores: Disposição a adquirir conhecimentos estatísticos avançados: Análise multivariada aplicada ao marketing. Programação de computadores. Alta capacidade de trabalho em grupo.

¹ Kopittke (1989); Kopittke *et al* (1999).

6.1.3 Formulação do delineamento experimental

O propósito é definir, sem perda de realismo, a combinação dos níveis de todos os atributos incluídos no experimento, que permita a estimação eficiente dos efeitos de interesse em uma base não correlacionada.

Este estudo se constitui em um fatorial com 4 atributos e 3 níveis, 3^4 alternativas possíveis. Utilizou-se o arranjo ortogonal descrito na Tabela 6.2. A Tabela 6.3 apresenta a codificação empregada.

Tabela 6.2 - Matriz Arranjo Ortogonal L9.

Alternativas	Atributos							
	Res β_1	Ele β_2	Aca β_3	Pro β_4	Rea β_5	Exp β_6	Alt β_7	Mod β_8
1	0	0	0	0	0	0	0	0
2	0	0	0	1	0	1	0	1
3	0	0	1	0	1	0	1	0
4	0	1	0	0	0	1	1	0
5	0	1	0	1	1	0	0	0
6	0	1	1	0	0	0	0	1
7	1	0	0	0	1	0	0	1
8	1	0	0	1	0	0	1	0
9	1	0	1	0	0	1	0	0

Tabela 6.3 – Codificação *dummy*.

Curso	Certificação	Método	Habilidades
M. Research (Res) 1 0	Mest. Acadêmico (Aca) 1 0	Real (Rea) 1 0	Altas (Alt) 1 0
M. Eletrônico (Ele) 0 1	Mest. Profissional (Pro) 0 1	Experiencial (Exp) 0 1	Moderadas (Mod) 0 1
M. Internacional (Int) 0 0	Especialização (Esp) 0 0	Convencional (Con) 0 0	Leves (Lev) 0 0

Segundo procedimento proposto por Souza (1999), com a finalidade de diminuir a complexidade da tarefa (quantidade de alternativas a serem avaliadas pelo entrevistado em cada cenário de escolha), foram montados blocos incompletos de seis alternativas, apresentados na Tabela 6.4. Nesta aplicação, 72% dos entrevistados ordenaram 6 grupos, os restantes apenas 3, em média cada entrevistado realizou 5.15 ordenações.

Tabela 6.4 - Blocos incompletos balanceados.

Grupos	Alternativas
01	1 2 4 5 7 8
02	2 3 5 6 8 9
03	1 3 4 6 7 9
04	1 2 5 6 7 9
05	1 3 4 5 8 9
06	2 3 4 6 7 8
07	1 3 5 6 7 8
08	1 2 4 6 8 9
09	2 3 4 5 7 9
10	4 5 6 7 8 9
11	1 2 3 4 5 6
12	1 2 3 7 8 9

Adicionalmente, com a finalidade de identificar a validade preditiva dos resultados, foram definidos seis cenários *holdout*, apresentados na Tabela 6.5, aplicando a regra 4.16.

Tabela 6.5 – Cenários *holdout*.

<i>Holdout</i>	Curso	Certificação	Método	Habilidades
1	M. Internacional	Especialização	Convencional	Moderadas
	M. Research	M. Acadêmico	Convencional	Moderadas
	M. Research	Especialização	Experiencial	Moderadas
	M. Research	Especialização	Convencional	Leves
2	M. Eletrônico	M. Profissional	Real	Leves
	M. Internacional	M. Acadêmico	Real	Leves
	M. Internacional	M. Profissional	Convencional	Leves
	M. Internacional	M. Profissional	Real	Altas
3	M. Research	M. Profissional	Convencional	Altas
	M. Eletrônico	Especialização	Convencional	Altas
	M. Eletrônico	M. Profissional	Real	Altas
	M. Eletrônico	M. Profissional	Convencional	Moderadas
4	M. Eletrônico	M. Acadêmico	Experiencial	Moderadas
	M. Research	M. Profissional	Experiencial	Moderadas
	M. Research	M. Acadêmico	Real	Moderadas
	M. Research	M. Acadêmico	Experiencial	Altas
5	M. Research	M. Acadêmico	Experiencial	Altas
	M. Internacional	Especialização	Experiencial	Altas
	M. Internacional	M. Acadêmico	Convencional	Altas
	M. Internacional	M. Acadêmico	Experiencial	Leves
6	M. Internacional	Especialização	Real	Leves
	M. Eletrônico	M. Profissional	Real	Leves
	M. Eletrônico	Especialização	Convencional	Leves
	M. Eletrônico	Especialização	Real	Moderadas

6.1.4 Definição da população-alvo

A população-alvo para esta aplicação empírica são estudantes que possuem a informação procurada e sobre os quais devem ser feitas as inferências, especificamente:

- a) **Elemento:** Estudantes de graduação e pós-graduação que manifestaram possível interesse em cursos de pós-graduação em marketing.
- b) **Unidade amostral:** Universidades:
- Universidade Federal de Santa Catarina (UFSC),
- Escola Superior de Administração de Negócios (ESAG)
- Universidade do Vale de Itajai (UNIVALI).
- c) **Âmbito:** Estado de Santa Catarina, Brasil.
- d) **Época:** Agosto - Setembro, 2001.

6.1.5 Determinação do tamanho da amostra

Adota-se a regra 6.1 para a determinação do tamanho da amostra em modelagem de escolha discreta em nível agregado [Orme, 1998]. Ao todo foram realizadas 156 entrevistas.

$$\frac{(\text{N}^\circ \text{ Entrevistados}) (\text{Tarefas} / \text{Entrevistado}) (\text{Alternativas} / \text{Tarefa})}{(> \text{N}^\circ \text{ de níveis do atributo})} > 500 \quad (6.1)$$

onde:

$$\text{N}^\circ \text{ de Entrevistados} = 156$$

$$\text{Tarefas} / \text{Entrevistado} = 5.15$$

$$\text{Alternativas} / \text{Tarefa} = 6$$

$$> \text{N}^\circ \text{ de níveis do atributo} = 3$$

Tipo de amostragem: Não probabilística

Os estudantes foram selecionados por estarem disponíveis no local e no momento em que a pesquisa estava sendo realizada.

Composição da amostra: Tabela 6.6

Tabela 6.6 – Composição da amostra (%).

Sexo	Masculino: 63	Feminino: 37				
Idade	< 19 : 9	19 – 24 : 62	25 – 30 : 10	> 30 : 19		
Estado civil	Solteiro: 83	Casado : 17				
Ocupação	Estudante: 81					
	Nível	Graduação: 88	Mestrado: 6	Doutorado : 6		
	Bolsista	Sim: 25	Não: 75			
	Professor: 8					
	Outra: 11					
	Experiência	< 2 : 20	2 – 5 : 27	6 – 10 : 23	> 10 : 30	
Trabalha na área	Sim: 24	Não: 76				
	Experiência	< 2 : 53	2 – 5 : 31	6 – 10 : 13	> 10 : 3	
Fez pós em Mkt	Sim: 8	Não: 92				
Instituição	UFSC: 21	ESAG: 31	UNIVALI: 48			

6.1.6 Coleta dos dados conjuntos

Uma das etapas fundamentais na aplicação das metodologias de análise conjunta refere-se à coleta dos dados. Porém, nota-se que este componente tem recebido pouca atenção por parte dos pesquisadores. A maioria dos esforços de aprimoramento da técnica de análise conjunta tem-se direcionado ao desenvolvimento dos modelos matemáticos utilizados para a análise dos dados e observa-se nesta área um alto grau de sofisticação. Pouca pesquisa tem sido feita em relação aos diferentes fatores que afetam a validade e qualidade dos dados utilizados. Um destes fatores é o formato de apresentação das tarefas de escolha. Pesquisas, como Jones e Polak (1994), evidenciam que a forma em que são apresentadas as tarefas de escolha, o cenário de escolha e as alternativas a serem escolhidas, influenciam os resultados e portanto a validade dos dados.

Wittink *et al* (1994) declaram que a entrevista pessoal é o meio de coleta de dados mais utilizado. Este método está usualmente baseado em impressos em papel ou cartões. Em cada cartão são representadas as alternativas sobre as quais os entrevistados declaram as suas preferências segundo algum método predeterminado, tais como: colocando as alternativas em ordem de preferência (*ranking*), submetendo-as a uma escala de avaliação (*rating*), ou escolhendo a opção preferida dentro do conjunto de alternativas disponíveis (*choice*).

Green e Srinivassan (1990) declaram que cartões representando as alternativas, com descrições concisas dos níveis dos atributos, são o método mais popular de apresentação das alternativas. Eles acharam um crescente uso de material gráfico e citam que este tipo de apoio torna a tarefa mais interessante para o respondente e provê formas mais fáceis e menos ambíguas de representar a informação. Representações de alternativas em duas e três dimensões são virtualmente indispensáveis em estudos conjuntos associados com aparência e estética, tais como projeto da embalagem e estilo do produto. Estes autores indicam também que *conjoint analysis* está crescentemente sendo utilizada empregando estímulos sensoriais em áreas como alimentos, bebidas, fragrâncias, etc. Ainda ressaltam que o crescente uso de material pictórico e protótipos reais ampliará o escopo de *conjoint analysis* e incrementará seu realismo na representação das condições do mercado.

Por outro lado, Jones e Polak (1994) declaram que apesar dos métodos de entrevista pessoal baseados em papel tenham sido majoritariamente utilizados e provavelmente continuem sendo o método mais popular nos próximos anos, eles impõem algumas limitações sobre a complexidade, velocidade e qualidade dos dados resultantes da pesquisa. Para estes autores estas limitações estão se tornando mais evidentes devido a: a) crescente demanda por processamento de dados mais rápido e eficiente, e b) as questões formuladas ao mercado estão se tornando mais sofisticadas. Limitações que, segundo os autores, podem ser superadas com os métodos de entrevista pessoal assistida por computador.

Pesquisas envolvendo representação gráfica das alternativas tem sido desenvolvidas. Vriens (1995) pesquisou a influência das representações verbal e gráfica na qualidade dos dados conjuntos. Como vantagens da representação gráfica o autor indica a possibilidade de incluir fatores estéticos na descrição dos produtos hipotéticos, imagens tendem a ser processadas simultaneamente enquanto que representações verbais são processadas sequencialmente, e representações gráficas incrementam o realismo da tarefa de avaliação.

Klabbers *et al* (1996) propõem um sistema multimídia para experimentos de preferência declarada. O sistema permite a utilização de diferentes formatos de apresentação (textos, imagens, som e combinações), possibilitando a medição da influência do formato empregado.

Dijkstra *et al* (1996) apresentam uma estrutura que combina *conjoint analysis* e realidade virtual para estudar como seres humanos navegam em ambientes espaciais. Este tipo

de abordagem oferece a grande possibilidade de imersão e interação com o ambiente virtual de escolha, o que de fato, introduz uma elevada dose de realismo na tarefa de escolha.

Urban *et al* (1996; 1997) desenvolveram uma metodologia, *Information Acceleration*, para facilitar o processo de aprendizado na avaliação de novos produtos. Estes autores integram *conjoint analysis* com computadores multimídia para representar produtos, pessoas e cenários de escolha.

Observa-se uma necessidade crescente por métodos de coleta de dados que obtenham informação do consumidor via canais multisensoriais. Zaltman e Higie (1995) declaram que importantes oportunidades para aprender dos consumidores estão sendo desperdiçadas por ignorar canais não verbais de comunicação como parte integrante do processo de pesquisa.

Percebe-se portanto que os métodos tradicionais de coleta de dados não atendem plenamente as exigências do nível atual de pesquisa de mercado. Por outro lado, os avanços acelerados no aperfeiçoamentos de hardware e software estão possibilitando a aplicação de ferramentas computacionais na coleta de dados.

Para efetuar a aplicação da metodologia proposta nesta tese foi desenvolvido um sistema interativo assistido por computador para a coleta dos dados conjuntos. Esta ferramenta de entrevista pessoal possui a possibilidade de utilizar a potencialidade dos recursos multimídia incorporando outros canais de comunicação além da palavra escrita como som, gráficos, vídeo, animação em duas e três dimensões e interatividade. É possível criar um ambiente virtual para coletar dados conjuntos do entrevistado permitindo que este interaja livremente com o contexto de escolha.

O sistema de coleta de dados foi desenvolvido em *Authorware 5*.

Descrição do sistema desenvolvido para a coleta dos dados

Inicialmente (Figura 6.2), o sistema apresenta apelos para participar da pesquisa, explica os objetivos e os benefícios que os resultados teriam para um melhor atendimento das necessidades dos alunos. Se o entrevistado aceitar o convite a participar, o sistema explica, através de animação, a mecânica da tarefa a ser desenvolvida. Um módulo de ajuda, que explica o significado de cada um dos atributos que estão sendo avaliados assim como seus

respectivos níveis, pode ser acessado interativamente em qualquer momento da entrevista. Visa-se superar qualquer dificuldade que o entrevistado possa ter na identificação daquilo que está avaliando.

Feita a introdução inicial, solicita-se o ingresso de variáveis sócio-econômicas do entrevistado. Estas variáveis são codificadas e automaticamente armazenadas.

Posteriormente, o entrevistado é defrontado com uma sequência de cenários de escolha gerados automaticamente. Cada cenário (Figura 6.3) representa um bloco de experimento da Tabela 6.4, onde é apresentado um conjunto de alternativas de cursos de pós-graduação em marketing descrito pelos níveis dos seus atributos. O sistema solicita que o entrevistado escolha a alternativa preferida até esgotar o conjunto de escolha. A sequência das alternativas escolhidas são armazenadas em um arquivo externo no formato requerido pelos algoritmos desenvolvidos para a análise.

Adicionalmente, registra-se a quantidade de vezes que o entrevistado solicita ajuda, vezes que o entrevistado cancela a escolha por causa de erros por ele cometidos e os tempos total e parciais de execução da tarefa (Tabela 6.7). Estes dados podem ser de utilidade para aprimorar o sistema de coleta em relação a aspectos ergonômicos, complexidade da tarefa, sobrecarga de informação e utilização de recursos multimídia em aplicações futuras.

Tabela 6.7 – Tempos (seg)

1º esc	2º esc	3º esc	4º esc	5º esc	1º esc holdout	2º esc holdout	Total tarefa (min)
19.72	9.79	8.80	6.05	5.02	11.91	6.96	13.90

Todas as entrevistas foram efetuadas pelo autor da tese. Nas duas primeiras instituições foi utilizado um computador portátil. Na terceira instituição o aplicativo de coleta de dados foi instalado no centro de cômputo da universidade.



Figura 6.2 - Menu principal.



Figura 6.3 - Um cenário de escolha.

6.1.7 Estimação dos parâmetros

Nesta seção são apresentados os resultados obtidos aplicando o modelo de inferência bayesiana desenvolvido para estimação de *part-worths* individuais [Seção 4.1.1]. Este modelo fornece os parâmetros agregados, isto é, o vetor α (8x1) de médias dos *part-worths* e a matriz de covariâncias \mathbf{D} (8x8); os parâmetros individuais que correspondem a 156 vetores β_i (8x1), e a distribuição posterior marginal de cada parâmetro individual β_i , representadas por 156 matrizes (2000x8).

Diagnóstico de convergência

A convergência foi verificada após uma fase inicial de transição de 3000 amostras (Tabela 6.8), que demandou um tempo de 3.5 hrs. aprox. empregando um processador Pentium II 300Mhz, 64M de RAM. Contudo, nesta aplicação foram corridas 8,000 iterações das quais as últimas 2,000 foram consideradas amostras das distribuições alvo.

Tabela 6.8 – Análise de convergência.

	Res	Ele	Aca	Pro	Rea	Exp	Alt	Mod
	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8
$\sqrt{\hat{R}}$	1.0008	1.0021	1.0019	1.0098	1.0572	1.0574	1.0087	1.1706

Parâmetros agregados

Os parâmetros agregados, vetor de médias α e a matriz de variância-covariância \mathbf{D} , estimados com o modelo proposto são apresentados na Tabela 6.9.

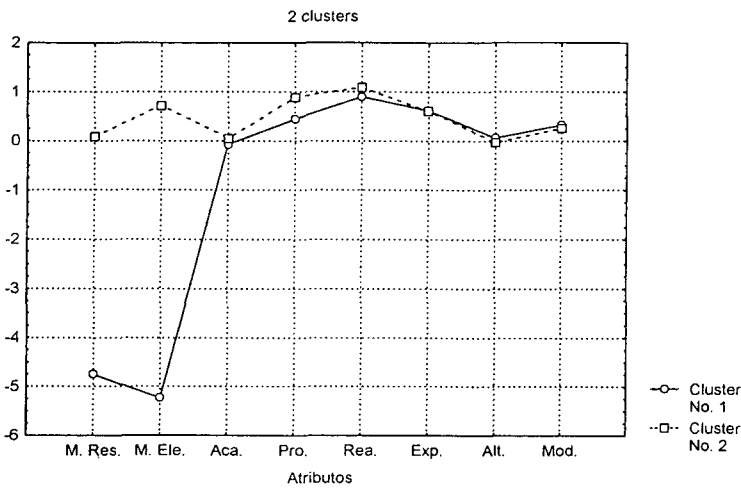
Tabela 6.9 – Parâmetros agregados
(Desvio padrão posterior) Correlações em itálico

	Médias (α)		Matriz de variância-covariância (D)						
Marketing Research	-1.56 (0.34)	12.15 (1.78)	<i>0.49</i>	<i>-0.06</i>	<i>-0.01</i>	<i>-0.01</i>	<i>0.00</i>	<i>0.03</i>	<i>0.04</i>
Marketing Eletrônico	-1.31 (0.30)	6.92 (1.54)	16.19 (2.47)	<i>-0.11</i>	<i>0.07</i>	<i>0.03</i>	<i>-0.11</i>	<i>-0.11</i>	<i>-0.17</i>
Mestrado Acadêmico	0.01 (0.23)	-0.52 (0.79)	-1.08 (0.89)	5.92 (0.83)	<i>0.73</i>	<i>0.00</i>	<i>-0.03</i>	<i>0.02</i>	<i>0.05</i>
Mestrado Profissional	0.74 (0.13)	-0.06 (0.62)	0.50 (0.69)	3.28 (0.53)	3.43 (0.48)	<i>0.12</i>	<i>0.17</i>	<i>0.09</i>	<i>0.10</i>
Método Real	1.03 (0.13)	-0.04 (0.48)	0.15 (0.53)	-0.01 (0.30)	0.26 (0.22)	1.35 (0.24)	<i>0.58</i>	<i>0.25</i>	<i>0.20</i>
Método Experiencial	0.61 (0.08)	-0.01 (0.45)	-0.44 (0.45)	-0.07 (0.26)	0.33 (0.21)	0.70 (0.22)	1.07 (0.25)	<i>0.07</i>	<i>0.23</i>
Habilidades Altas	0.00 (0.14)	0.14 (0.51)	-0.66 (.55)	0.06 (0.38)	0.24 (0.28)	0.42 (0.19)	0.11 (0.16)	2.14 (0.30)	<i>0.70</i>
Habilidades Moderadas	0.28 (0.07)	0.12 (0.33)	-0.64 (0.40)	0.10 (0.25)	0.18 (0.19)	0.22 (0.12)	0.22 (0.11)	0.94 (0.16)	0.85 (0.12)

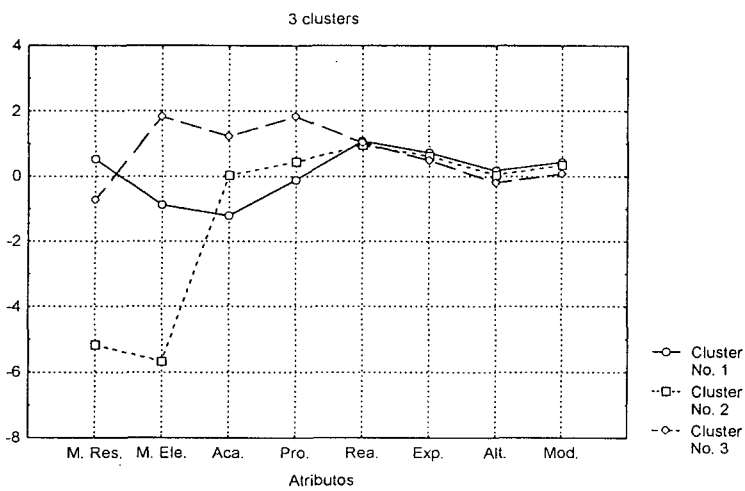
Na Tabela 6.9 observa-se que não existem interações importantes entre os níveis dos atributos, já que os elementos fora da diagonal (covariâncias entre níveis de atributos diferentes) são em geral próximos de zero. As altas variâncias (elementos diagonais de D) identificadas nos atributos **Curso** e **Certificação** sugerem a possível existência de grupos de indivíduos com preferências distintas em relação a estes atributos. Por exemplo, o intervalo de confiança bayesiano HPD 95% para o curso **Marketing Eletrônico** é [-9.20 6.58]. Apenas com o intuito de confirmar esta evidência de alta heterogeneidade do mercado, aplica-se uma análise de clusters considerando o agrupamento em 2, 3 e 4 grupos.

Observa-se claramente na Figura 6.4 que nos três agrupamentos os clusters são discriminados principalmente pelo atributo **Curso**. Para os agrupamentos em 3 e 4 grupos, o atributo **Certificação** também se constitui em variável discriminadora dos grupos resultantes.

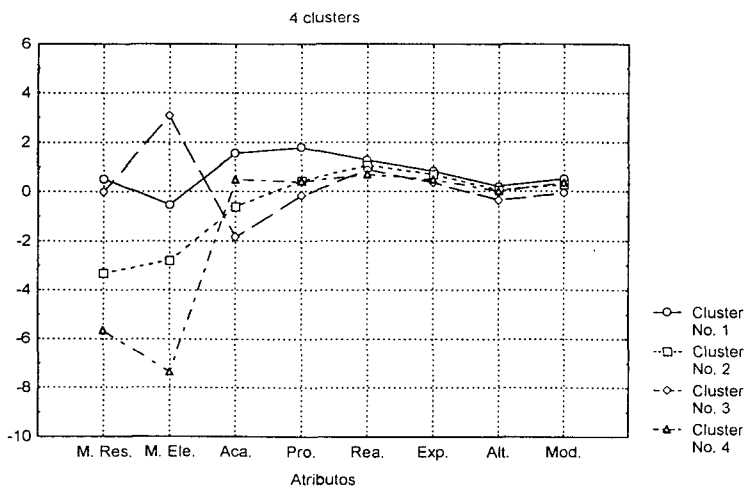
É evidente que neste mercado existe um alto grau de heterogeneidade nas preferências, que não está sendo adequadamente representado nos parâmetros agregados estimados α e D . Portanto, para estudar melhor este mercado é necessário analisar os *part-worths* individuais.



Cluster	Tamanho
1	53
2	103



Cluster	Tamanho
1	56
2	45
3	55



Cluster	Tamanho
1	54
2	37
3	39
4	26

Figura 6.4 – Média dos *part-worths* para 2, 3 e 4 clusters.

6.1.8 Validação dos parâmetros individuais

Verifica-se que 38% dos parâmetros individuais (β_i) estimados são estatisticamente significativos. Allenby e Ginter (1995) reportam uma percentagem de 25% para um estudo com 946 entrevistados.

A Tabela 6.10 apresenta os resultados da validação preditiva dos parâmetros individuais, são reportados os índices *Hit Rate* e MAE para os cenários *holdout* da Tabela 6.5. Observa-se que os índices de desempenho aqui obtidos são superiores aos índices obtidos com dados de simulação no capítulo anterior.

Para o modelo proposto dois índices MAE são reportados: no primeiro considera-se o valor esperado do parâmetro individual β_i , e no segundo a distribuição posterior marginal completa de β_i , calculado empregando o simulador desenvolvido (Figura 4.1). Destaca-se que quando a variabilidade intrínseca das preferências individuais é considerada, erros de prognóstico de participação são diminuídos de forma considerável. Nesta aplicação esta diminuição do MAE é de 18% aprox.

Tabela 6.10 - *Hit Rate* e MAE para cada cenário *holdout*.

<i>Holdout</i>	Modelo proposto (MP)			Modelo concorrente	
	<i>Hit Rate</i> MP	MAE MP	MAE (simulador MP)	<i>Hit Rate</i>	MAE
1	0.7564	0.0256	0.0288	0.6795	0.0865
2	0.5064	0.0385	0.0271	0.4744	0.1282
3	0.6090	0.0481	0.0464	0.4295	0.1122
4	0.5833	0.0353	0.0235	0.4936	0.0929
5	0.6486	0.0270	0.0198	0.5135	0.0856
6	0.6964	0.0402	0.0310	0.5893	0.0958
média	0.6334	0.0358	0.0294	0.5300	0.0999

Confirmação das hipóteses

A confirmação das hipóteses assumidas pelo modelo de inferência bayesiano hierárquico desenvolvido nesta tese é efetuada em dois estágios.

O nível inferior da hierarquia, que descreve o comportamento individual de escolha, assume um modelo logit multinomial ordenado. Para confirmar esta hipótese, a capacidade

preditiva dos parâmetros individuais estimados foi confrontada com a de um modelo concorrente. Este último é um modelo bayesiano hierárquico [Allenby *et al*, 1998; Allenby e Rossi, 1999] que descreve o comportamento individual de escolha através de um modelo logit multinomial não-ordenado (uma única escolha). Observa-se na Tabela 6.10 que o modelo proposto efetua um melhor prognóstico das escolhas nos cenários *holdout* (*Hit Rate*), e origina um erro menor no prognóstico da participação de preferências implicadas nos parâmetros individuais (MAE).

O nível superior da hierarquia do modelo, que descreve a variabilidade das preferências na população, assume que os parâmetros individuais são distribuídos normalmente. Verifica-se esta hipótese utilizando o comando `normplot` do Matlab [MathWorks, 1999]. Este comando provê um gráfico útil que possibilita avaliar se dados são gerados por uma distribuição normal. Se os dados fossem originados por uma distribuição normal o gráfico será linear, outras funções de densidade de probabilidade introduziriam curvatura no gráfico.

A Figura 6.5 apresenta os gráficos obtidos com o comando `normplot` para cada atributo nos 156 indivíduos. Observa-se que os gráficos são lineares confirmando que os parâmetros individuais de preferência provêm de uma distribuição normal.

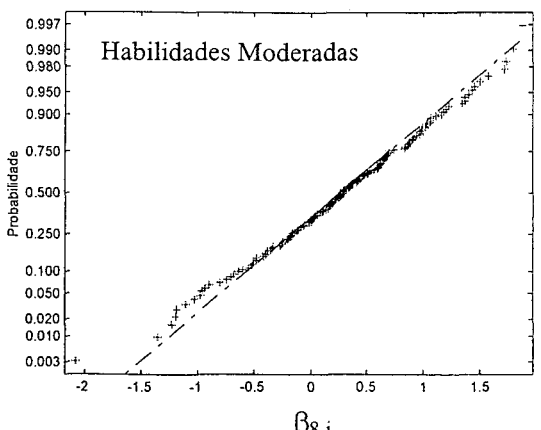
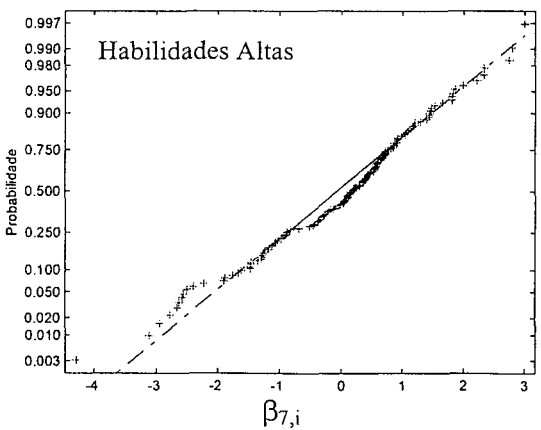
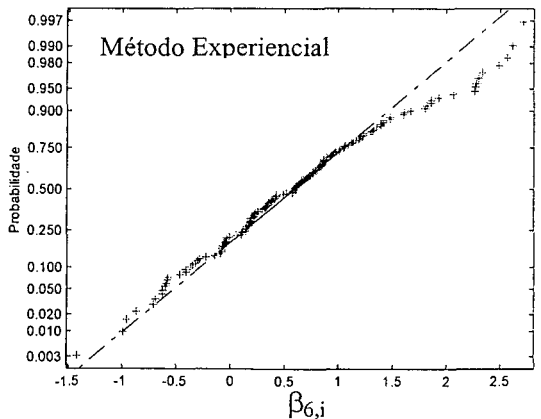
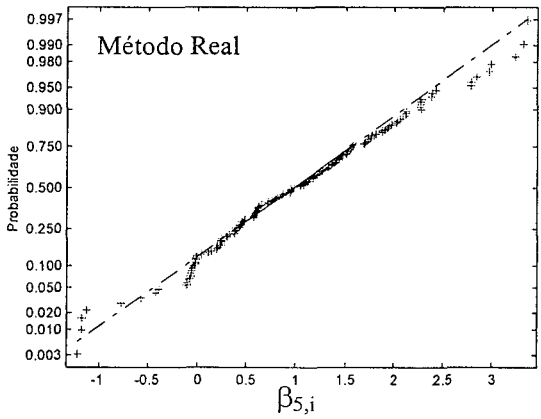
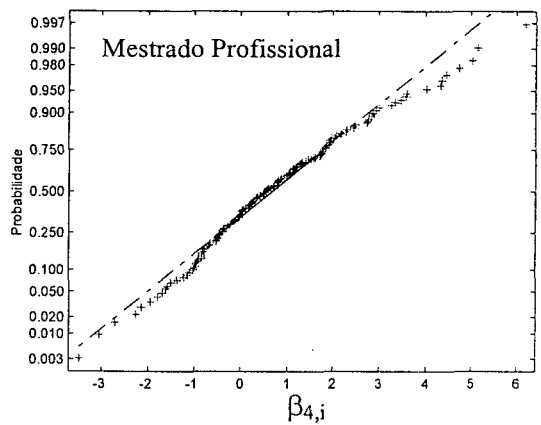
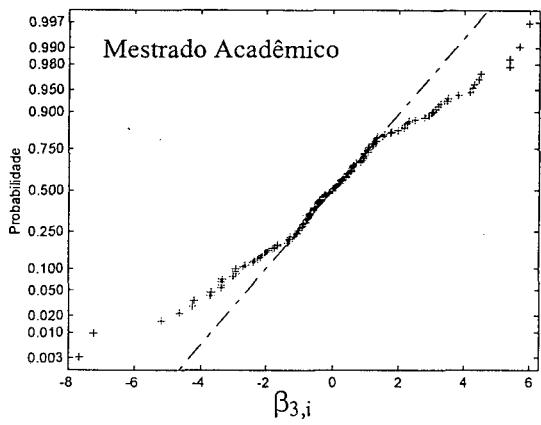
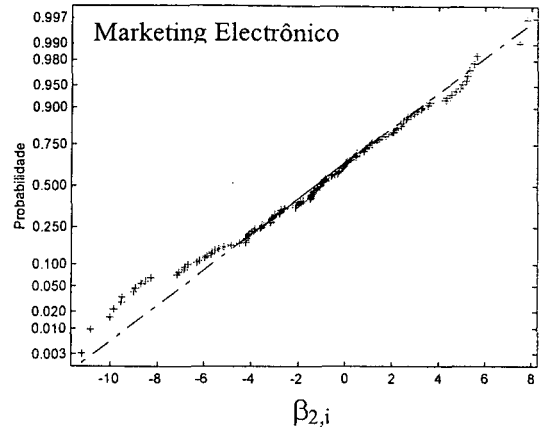
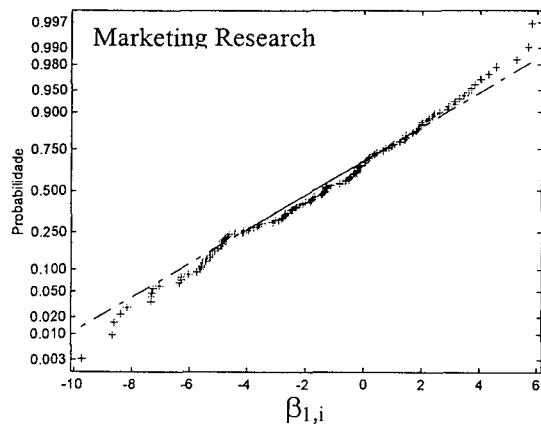


Figura 6.5 - Probabilidade normal

6.1.9 Análise competitiva

A última etapa desta metodologia busca identificar, a partir da grande quantidade de informação fornecida pelo modelo de inferência, ações lucrativas focalizadas no consumidor. Procura-se converter o conhecimento desagregado das preferências do mercado em informação acionável sobre grupos de consumidores que potencialmente respondam positivamente às ofertas da organização. Para tanto, é necessária a consideração do ambiente competitivo no qual serão introduzidas as ofertas da organização. A adequada análise deste ambiente competitivo, através do simulador de participação de preferências desenvolvido, deve levar a atingir o objetivo último desta tese: identificar oportunidades específicas de mercado além de formular e avaliar estratégias de marketing para aproveitar essas oportunidades.

Para ilustrar, definem-se 2 casos base hipotéticos e exploram-se possíveis cenários competitivos.

6.1.9.1 Caso Base A

Seja um mercado com 3 cursos de pós-graduação em marketing ministrados por 3 instituições de ensino superior (Tabela 6.11).

Tabela 6.11 – Caso Base A.

Instituição	Curso	Certificação	Método	Habilidades
A	M. Internacional	Mest. Acadêmico	Convencional	Moderadas
B	M. Eletrônico	Mest. Profissional	Experiencial	Moderadas
C	M. Research	Mest. Profissional	Convencional	Moderadas

O objetivo da instituição **A** é incrementar sua participação no mercado procurando atrair indivíduos com preferência por **Marketing Internacional**. A Tabela 6.12 apresenta as participações de mercado obtidas com o simulador de escolhas. Observa-se que **B** detém a maior fatia.

Tabela 6.12 – Composição do mercado - Caso Base A.

Instituição	Participação (%)
A	38.38
B	40.02
C	21.60

Os indivíduos apenas “medianamente satisfeitos” com o curso escolhido são os mais propensos a mudar de segmento, de forma que eles constituem o principal alvo das ações de marketing da instituição A. Assume-se que existe uma relação direta entre o grau de satisfação do indivíduo i com a probabilidade com a qual escolhe um determinado curso. A Figura 6.6 apresenta a quantidade de indivíduos em função de sua probabilidade de escolha. Ressalta-se que quanto maior a concentração dos indivíduos na vizinhança de probabilidade 1 maior é a satisfação com o produto. Por exemplo, empresas que apresentam alta percentagem de clientes insatisfeitos, isto é, consumidores com probabilidades de escolhas distribuídas de forma aproximadamente uniforme, são altamente sensíveis às ações competitivas dos concorrentes.

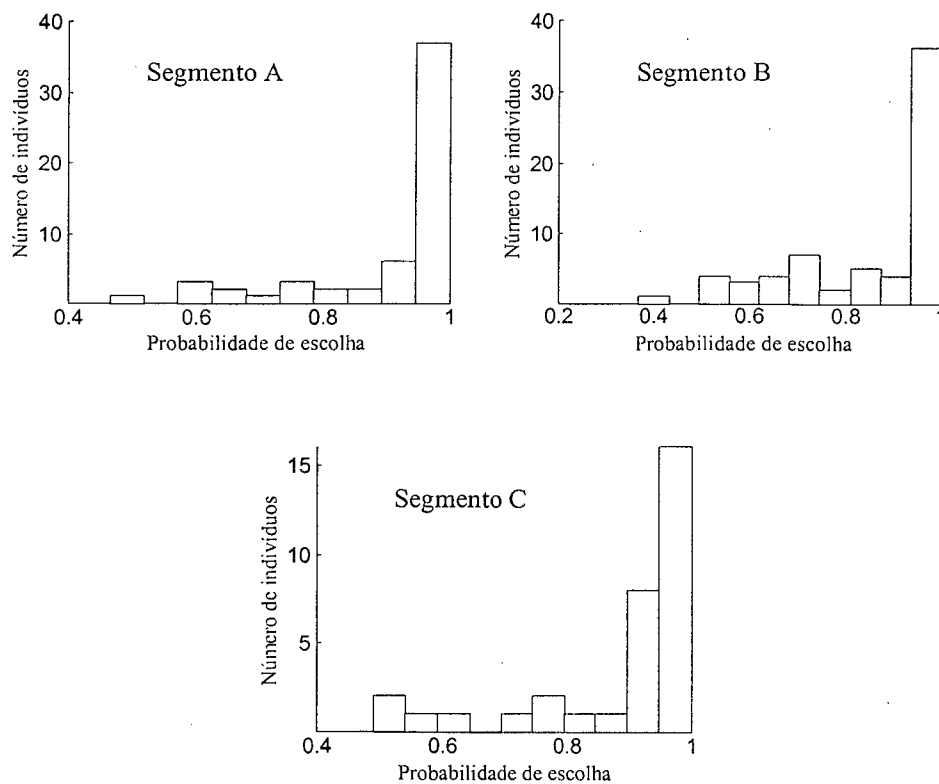


Figura 6.6 – Distribuição das probabilidades de escolha.

Nesta aplicação define-se os indivíduos “medianamente satisfeitos” para cada segmento, ou indivíduos alvo, àqueles que escolhem um curso com uma probabilidade de escolha menor que 0.7. Visualiza-se estes grupos de indivíduos (Tabela 6.13) como possíveis nichos de mercado cuja atratividade pode ser interessante analisar. Observa-se que **B** apresenta a maior quantidade de indivíduos alvo.

Tabela 6.13 – Indivíduos alvo - Caso Base A.

Instituição	Nº Ind. alvo (%)	Indivíduos
A	6 (10.53)	45 63 95 112 124 138
B	13 (19.70)	30 42 60 79 88 101 102 107 113 117 141 151 153
C	4 (12.12)	31 71 86 137

A instituição **A** estuda especificamente o padrão de preferências dos estudantes que escolheram o curso ofertado por **B** com baixa probabilidade, buscando identificar uma oportunidade de mercado nesse possível nicho. A Tabela 6.14 apresenta os parâmetros de preferência deste pequeno grupo de consumidores.

Tabela 6.14 – Part-worths dos indivíduos alvo do segmento B.
(Desvio padrão posterior)

Indivíduos	Res β_1	Ele β_2	Aca β_3	Pro β_4	Rea β_5	Exp β_6	Alt β_7	Mod β_8
30	-1.36 (0.56)	-0.85 (0.61)	0.39 (0.53)	0.51 (0.48)	1.04 (0.40)	1.05 (0.40)	-0.23 (0.52)	0.42 (0.38)
42	-1.35 (0.57)	-1.20 (0.55)	0.42 (0.56)	1.12 (0.59)	-1.17 (0.51)	0.76 (0.46)	-0.93 (0.55)	0.06 (0.49)
60	-0.20 (0.50)	-0.29 (0.55)	-0.95 (0.57)	-0.80 (0.55)	-0.07 (0.43)	0.19 (0.51)	-1.62 (0.57)	0.35 (0.46)
79	-3.98 (1.17)	-2.13 (0.87)	-1.31 (0.97)	-0.02 (0.69)	2.12 (0.77)	1.01 (0.67)	0.39 (0.72)	0.22 (0.64)
88	-0.19 (0.75)	-0.95 (0.81)	4.29 (1.11)	4.36 (1.13)	1.33 (0.70)	1.17 (0.52)	1.84 (0.62)	1.37 (0.59)
101	-1.31 (0.60)	-1.06 (0.69)	4.14 (0.95)	4.46 (0.87)	0.76 (0.51)	0.93 (0.64)	-2.40 (0.79)	-0.55 (0.52)
102	-0.21 (0.62)	0.21 (0.58)	-3.70 (0.89)	-1.69 (0.62)	0.03 (0.55)	-0.07 (0.52)	0.3 (0.52)	0.64 (0.45)
107	-5.23 (1.34)	-3.16 (1.04)	-2.02 (1.00)	0.55 (0.71)	1.72 (0.78)	0.86 (0.61)	0.67 (0.79)	0.30 (0.41)
113	-0.65 (0.87)	-0.28 (0.93)	1.32 (0.86)	1.02 (.67)	2.79 (0.68)	0.73 (0.60)	1.21 (0.74)	1.07 (0.47)
117	-5.01 (1.38)	-1.91 (0.88)	0.43 (0.84)	1.94 (1.00)	1.52 (0.91)	0.93 (0.73)	-0.32 (0.62)	0.43 (0.50)
141	-1.32 (0.71)	0.81 (0.62)	0.46 (0.63)	0.60 (0.55)	0.30 (0.48)	-0.59 (0.61)	1.18 (0.57)	0.98 (0.45)
151	-3.12 (0.92)	-0.13 (0.73)	3.19 (0.85)	3.02 (0.82)	2.39 (0.75)	0.85 (0.57)	0.20 (0.64)	-0.06 (0.54)
153	-2.79 (1.12)	-3.13 (1.09)	-3.72 (1.15)	-0.94 (0.71)	0.75 (0.72)	1.07 (0.68)	0.57 (0.90)	0.11 (0.61)
Médias (α)	-2.06	-1.08	0.23	1.09	1.04	0.69	0.07	0.41

Estes são indivíduos com preferência por **Marketing Internacional**, porém são exigentes em suas preferências pelos atributos **Certificação** e **Método**. Observa-se que a grande maioria deste grupo rejeita os níveis **Acadêmico** e **Convencional** o que impede a escolha do curso ofertado por **A**. E esta é uma oportunidade de mercado que o simulador possibilita identificar.

A seguir consideram-se duas possíveis estratégias de marketing que a instituição de ensino **A** poderia adotar para aumentar sua participação de mercado.

1. Reposicionar o curso atual.
2. Introduzir um novo curso no mercado, mantendo o anterior.

Antes de comprometer recursos da organização, **A** precisa ter um prognóstico da resposta do mercado em relação às duas estratégias alternativas consideradas. A seguir são apresentados os resultados obtidos aplicando o simulador de escolha.

1. **Primeira estratégia:** Reposicionar o curso atual.

A instituição **A** considera aprimorar sua oferta educacional atual adotando um método de ensino **Experiencial** (Tabela 6.15).

Tabela 6.15 – Primeira estratégia de marketing.

Instituição	Curso	Certificação	Método	Habilidades
A	M. Internacional	Mest. Acadêmico	Experiencial	Moderadas
B	M. Eletrônico	Mest. Profissional	Experiencial	Moderadas
C	M. Research	Mest. Profissional	Convencional	Moderadas

A Tabela 6.16 apresenta as participações de mercado resultantes. Observa-se que **A** toma uma grande participação de **B** enquanto que a participação de **C** diminui em menor proporção. Lembra-se que **A** formulou sua estratégia para atingir a um nicho específico do mercado: àqueles estudantes insatisfeitos com o curso de **B**, e com esta estratégia consegue atrair a 85% desses estudantes. Verifica-se que dos 13 indivíduos insatisfeitos do segmento **B** no Caso Base A, 11 deslocaram-se para o segmento **A**.

Tabela 6.16 – Composição do mercado (Primeira estratégia).

Instituição	Participação (%)	Nº Ind. Alvo	(%)
A	47.26	16 / 79	20.25
B	33.29	6 / 47	12.77
C	19.45	4 / 30	13.33

O simulador permite reconhecer o impacto diferenciado das estratégias de marketing. Esta característica é importante devido ao fato que frequentemente estratégias de marketing são formuladas para atingir a grupos de consumidores com necessidades claramente identificadas, e os simuladores de mercado precisam reconhecer e avaliar o impacto diferenciado das estratégias de marketing sobre as participações dos produtos competitivos.

2. **Segunda estratégia:** Introduzir um novo curso (**A'**) no mercado, mantendo o anterior (Tabela 6.17).

Tabela 6.17 – Segunda estratégia de marketing.

Instituição	Curso	Certificação	Método	Habilidades
A	M. Internacional	Mest. Acadêmico	Convencional	Moderadas
B	M. Eletrônico	Mest. Profissional	Experiencial	Moderadas
C	M. Research	Mest. Profissional	Convencional	Moderadas
A'	M. Internacional	Mest. Profissional	Convencional	Moderadas

A Tabela 6.18 apresenta as participações resultantes. A participação total da instituição **A**, dada pelos dois cursos que oferece, é 47.22 %.

Tabela 6.18 – Composição do mercado (Segunda estratégia).

Instituição	Participação (%)	Nº Ind. Alvo (%)
A	19.89	9 em 28 32.14
B	33.82	12 em 56 21.43
C	18.97	4 em 30 13.33
A'	27.33	10 em 42 23.81

Observa-se que parte importante da própria demanda de **A** se desloca para o novo curso **A'** (canibalismo). O simulador reconhece portanto efeitos de substituição diferenciada, isto é, uma nova oferta toma maior participação de marcas similares e menor participação daquelas marcas que são visualizadas pelo consumidor como sendo diferentes. Esta característica do simulador é importante devido ao fato que efeitos de substituição diferenciada são muito comuns em por exemplo extensões de linhas de produto, onde frequentemente novos produtos tomam maior participação dos produtos da própria linha de produção da organização. Gerentes de produto reconhecem a existência deste problema e precisam administrar sua carteira de produtos de forma a maximizar participação de mercado obtida de produtos concorrentes e minimizar perdas internas de participação.

Sob as duas estratégias de marketing, as participações de mercado resultantes para a instituição **A** são semelhantes. Porém, a primeira estratégia origina um nível de satisfação maior, isto é, uma quantidade maior de indivíduos com alta probabilidade (>0.7) de escolha. Desta forma, pode-se optar pela primeira estratégia de marketing (reposicionar o produto

existente) que tem 79.75% indivíduos satisfeitos contra 72.86% da segunda estratégia. É claro que o processo de tomada de decisão envolve fatores adicionais, como por exemplo, o custo de implementação e operacionalização de cada estratégia. Porém, o conhecimento gerado pela metodologia deve servir como apoio à decisão gerencial.

6.1.9.2 Caso Base B

Considera-se o mercado descrito na Tabela 6.19.

Tabela 6.19 – Caso Base B.

Instituição	Curso	Certificação	Método	Habilidades
A	M. Internacional	Mest. Acadêmico	Convencional	Moderadas
B	M. Eletrônico	Mest. Acadêmico	Convencional	Moderadas

Uma terceira instituição de ensino superior **C**, especializada em **Marketing Research**, pretende ingressar neste mercado se conquistasse uma participação maior que 20%. As participações de cada produto no mercado atual são apresentadas na Tabela 6.20. O produto A detém a maior fatia.

Tabela 6.20 – Composição do mercado - Caso Base B.

Instituição	Participação (%)
A	62.31
B	37.69

A Tabela 6.21 apresenta os parâmetros agregados para os indivíduos com alta preferência (“satisfeitos”, probabilidade de escolha maior que 0.7) pelo produto escolhido. A Tabela 6.22 corresponde aos indivíduos com probabilidade de escolha menor que 0.7.

Tabela 6.21 – Part-worths agregados dos indivíduos “satisfeitos”.

Médias	Res α_1	Ele α_2	Aca α_3	Pro α_4	Rea α_5	Exp α_6	Alt α_7	Mod α_8
A	-2.8328	-3.8894	0.1669	0.6832	1.0231	0.7243	0.0760	0.3690
B	0.0289	2.6793	-0.4533	0.6685	0.9399	0.3782	-0.2688	0.0237

Tabela 6.22 – *Part-worths* agregados dos indivíduos alvo.

Médias	Res α_1	Ele α_2	Aca α_3	Pro α_4	Rea α_5	Exp α_6	Alt α_7	Mod α_8
A	-0.7896	-0.2012	0.4862	1.3814	1.7188	0.9640	0.7087	0.7112
B	1.6475	0.1038	0.7855	1.1039	0.7640	0.4346	0.0978	0.4223

Os indivíduos satisfeitos (Tabela 6.21) apresentam alta preferência pelo tipo de curso escolhido, e são menos exigentes em relação aos atributos **Certificação** e **Método** que os indivíduos da Tabela 6.22. Observa-se que os indivíduos medianamente satisfeitos com o produto **B** possuem forte preferência por **Marketing Research**. Adicionalmente, os indivíduos medianamente satisfeitos com os produtos **A** e **B** preferem certificação **Mestrado Profissional** e método de ensino **Real**. A questão é: este grupo de indivíduos, caracterizado na Tabela 6.23, se constitui em um nicho de mercado atrativo (demanda suficiente) para a instituição **C**?

Tabela 6.23 – Indivíduos alvo.

Instituição	Nº Ind. Alvo (%)	Indivíduos
A	9 (9.28)	24 29 60 85 89 112 113 128 151
B	7 (11.86)	31 50 73 94 102 106 118

A oferta educacional de **C** é apresentada na Tabela 6.24, e a nova distribuição do mercado, resultante da entrada de **C**, na Tabela 6.25.

Tabela 6.24 – **C** ingressa no mercado.

Instituição	Curso	Certificação	Método	Habilidades
A	M. Internacional	Mest. Acadêmico	Convencional	Moderadas
B	M. Eletrônico	Mest. Acadêmico	Convencional	Moderadas
C	M. Research	Mest. Profissional	Real	Moderadas

Tabela 6.25 – Composição do mercado: novo cenário (A B e C).

Instituição	Participação (%)
A	36.88
B	18.22
C	44.90

A instituição **C** antes de tomar sua decisão considera ainda possíveis reações competitivas dos concorrentes. Nas Tabelas 6.26 e 6.27 verifica-se novamente que os indivíduos medianamente satisfeitos são mais exigentes em relação aos atributos **Certificação** e **Método**, inclusive alguns escolhem o produto **C** ainda que sua preferência por **Marketing Internacional** é maior.

Tabela 6.26 – Part-worths agregados dos indivíduos “satisfeitos”.

Médias	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8
A	-4.8977	-5.0772	0.0695	0.3000	0.6448	0.4352	-0.0614	0.2478
B	-1.9839	3.3742	0.4230	0.8503	0.6983	0.0004	-0.6453	-0.1585
C	0.9891	-0.4731	-0.6766	0.7176	1.3757	0.9812	0.3689	0.5374

Tabela 6.27 – Part-worths agregados dos indivíduos alvo

Médias	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8
A	-1.9828	-1.6214	1.8249	1.9553	1.4362	0.8379	-0.1828	0.1864
B	1.7115	2.8371	1.1612	1.0491	0.6264	0.2749	-0.4004	0.0169
C	-1.1894	-1.2705	-0.0199	1.1768	1.2290	0.2815	-0.707	-0.2139

Uma análise prévia dos concorrentes tem permitido a **C** identificar que eles não tem capacidade de implementar métodos reais de ensino. Portanto consideram-se as seguintes reações:

1. **Primeira reação competitiva:** **A** e **B** adotam um método de ensino **Experiencial** (Tabela 6.28). A Tabela 6.29 apresenta as participações resultantes.

Tabela 6.28 – Primeira reação competitiva.

Instituição	Curso	Certificação	Método	Habilidades
A	M. Internacional	Mest. Acadêmico	Experiencial	Moderadas
B	M. Eletrônico	Mest. Acadêmico	Experiencial	Moderadas
C	M. Research	Mest. Profissional	Real	Moderadas

Tabela 6.29 – Composição do mercado (1^{ra} reação).

Instituição	Participação (%)	Nº	Ind.Alvo (%)
A	41.87	12	(17.65)
B	20.56	4	(12.90)
C	37.58	10	(17.54)

2. **Segunda reação competitiva:** A e B em adição à utilização do método de ensino **Experiencial**, outorgam certificação de **Mestrado Profissional** (Tabela 6.30). A Tabela 6.31 apresenta as participações resultantes.

Tabela 6.30 – Segunda reação competitiva.

Instituição	Curso	Certificação	Método	Habilidades
A	M. Internacional	Mest. Profissional	Experiencial	Moderadas
B	M. Eletrônico	Mest. Profissional	Experiencial	Moderadas
C	M. Research	Mest. Profissional	Real	Moderadas

Tabela 6.31 – Composição do mercado (2^{da} reação).

Instituição	Participação	Nº	Ind.Alvo (%)
A	45.93	9	12.33
B	25.06	5	12.82
C	29.01	7	15.91

A instituição C, para tomar a decisão com uma maior margem de segurança, considera sua participação de mercado composta unicamente dos clientes satisfeitos com sua oferta, isto é, aqueles com probabilidade de escolha de C maior a 0.7. A participação de mercado prognosticada para C, considerando possíveis as respostas competitivas dos concorrentes é 24.52%, que é maior ao 20% de participação pretendida. Conseqüentemente, C pode optar por entrar no mercado.

6.2 Considerações adicionais

A aplicabilidade da ferramenta desenvolvida tem sido ilustrada na análise precedente. É pertinente efetuar algumas observações, embora não conclusivas, em relação à tendência das preferências educacionais. Observa-se que as pessoas procuram serviços educacionais que

as capacitem para a atividade profissional de uma forma predominantemente aplicada. Adicionalmente, há uma clara preferência por métodos de ensino que os coloquem em situações mais próximas à realidade profissional. Depreende-se a necessidade de procurar por métodos de ensino mais eficazes e das instituições acadêmicas se aproximar ao setor empresarial.

Em relação à coleta dos dados, o sistema de entrevista assistido por computador desenvolvido apresenta vantagens em relação aos impressos em papel. Por um lado, o formato de apresentação é mais interessante e flexível, o que tem um efeito motivador sobre o entrevistado. Os cenários de escolha são apresentados automaticamente, de forma que o sistema poderia reagir *on-line* às respostas do entrevistado com a finalidade de, por exemplo, incorporar mecanismos de controle para evitar inconsistências e respostas erradas.

O sistema otimiza tempo e custo: sendo os dados ingressados e armazenados eletronicamente, os arquivos de saída seguem um formato compatível com o algoritmo de análise, isto elimina a necessidade de codificação dos dados e evita erros resultantes da codificação manual. Reduz o trabalho do entrevistador, já que o sistema está programado para realizar todo o trabalho administrativo e o entrevistador apenas é responsável pelo contato inicial com o entrevistado.

Podem ser criados cenários de escolha mais realistas, que facilitem compreender melhor a tarefa a ser desenvolvida pelo entrevistado.

Por outro lado, existem também algumas desvantagens que precisam ser consideradas: o manuseio do computador portátil é mais difícil e menos flexível que os cartões. Existe maior dificuldade para corrigir as ações de escolha, cartões praticamente não tem este problema. Adicionalmente, algumas poucas pessoas podem sentir-se inibidas de participar da pesquisa por desconhecimento dos recursos da informática.

CONCLUSÕES E RECOMENDAÇÕES

Esta tese visou oferecer uma oportunidade de agregar valor ao processo de marketing estratégico das organizações orientadas para o mercado. Com essa finalidade foi desenvolvida uma metodologia consistente com o princípio fundamental do marketing atual: os consumidores são diferentes. Eles diferem nos produtos e serviços que escolhem, no valor que alocam a cada atributo que compõe esses produtos e serviços, na sensibilidade a diferentes variáveis, por exemplo o preço, e na forma em que reagem a diferentes estímulos de marketing, isto é, consumidores possuem preferências diferentes. Geralmente, empresas que incorporam de forma efetiva as preferências do mercado nas suas ofertas adquirem vantagens competitivas em relação aos seus concorrentes.

Devido aos custos elevados da personalização de produtos e serviços as organizações atuais, em geral, não têm condições de atender individualmente cada consumidor. No entanto observa-se uma tendência crescente do marketing em focalizar seus esforços em segmentos cada vez menores de consumidores que respondem, ou poderiam responder, positivamente às ofertas das organizações. Porém, a identificação das oportunidades representadas por estes nichos de mercado pode ser uma tarefa árdua para o marketing, dado que envolve a identificação da heterogeneidade das preferências e das diferentes sensibilidades existentes no mercado.

Nesse panorama atual, torna-se premente a necessidade do marketing por modelos que permitam, antes do comprometimento de recursos das organizações, caracterizar as preferências do mercado de forma mais precisa e aproveitar este conhecimento na identificação de ações lucrativas focalizadas no consumidor. No intuito de preencher essa necessidade, a comunidade de marketing tem recentemente direcionado seus esforços na procura de modelos confiáveis para estimar parâmetros de preferências em nível individual,

que ajudem às organizações a projetar produtos e serviços que satisfaçam as necessidades de grupos reduzidos de consumidores com preferências semelhantes.

Esta tese teve como objetivo geral apresentar uma metodologia que gera conhecimento do comportamento individual dos consumidores, e utiliza este conhecimento para auxiliar as organizações na identificação de oportunidades de mercado e na formulação e avaliação de estratégias alternativas de marketing para melhor aproveitar essas oportunidades.

Especificamente, desenvolveu-se um modelo bayesiano de inferência, que permite a mensuração da inerente heterogeneidade das preferências do mercado, e um simulador de escolhas para prognosticar respostas de grupos reduzidos de consumidores perante produtos/serviços/conceitos alternativos. A principal característica da metodologia apresentada é a capacidade de estimar parâmetros de preferências em nível individual, assim como a distribuição destas preferências na população. A adequada utilização deste conhecimento pode auxiliar a formular planos estratégicos de marketing e ofertar produtos e serviços a grupos específicos de consumidores com necessidades e preferências semelhantes. Para tanto, o simulador de escolhas possibilita desenvolver análise competitiva para diferentes cenários alternativos avaliando o impacto de diferentes estratégias hipotéticas de marketing, da organização ou de seus competidores, sobre a participação de escolha dos produtos e serviços considerados num determinado ambiente competitivo.

A contribuição desta tese no que tange ao modelo bayesiano de inferência de parâmetros de preferência, é a descrição do comportamento individual de escolha através de um modelo logit multinomial ordenado. Tanto na validação com dados sintéticos como na aplicação empírica efetuada, verificou-se a capacidade do modelo na obtenção de parâmetros individuais razoáveis com escassa informação individual (na aplicação empírica utilizaram-se aproximadamente cinco ordenações por indivíduo).

É interessante observar que inclusive se tratando de mercados “homogêneos”, o desempenho preditivo obtido considerando os parâmetros individuais foi superior ao desempenho obtido com os parâmetros agregados, estimados ambos com o mesmo modelo.

A integração numérica das distribuições posteriores tem sido abordada através de métodos baseados em simulação estocástica e integração de Monte Carlo (Amostrador de Gibbs e Algoritmo Metropolis-Hastings). Este procedimento, além de elegante, mostrou-se

bastante robusto tanto nas simulações efetuadas com dados sintéticos como com dados reais, permitindo obter resultados confiáveis.

Em relação ao simulador para prognosticar fatias de escolha, a contribuição efetuada é a consideração explícita da variabilidade intrínseca do processo individual de escolha por intermédio das distribuições marginais posteriores de cada parâmetro. Desta forma, não é necessária a adição de variância de forma artificial no cálculo das utilidades dos produtos competidores. O simulador é imune perante as dificuldades da hipótese de independência de alternativas irrelevantes (IIA) sem se contrapor com a teoria da utilidade aleatória. Verificou-se empiricamente que o simulador possui as capacidades de avaliar o impacto diferenciado das estratégias de marketing sobre as participações de produtos competitivos e de detectar padrões de substituição diferenciada entre as diferentes alternativas consideradas.

Adicionalmente, foi desenvolvido um sistema interativo assistido por computador para a coleta dos dados conjuntos, que possui a possibilidade de incorporar outros canais de comunicação além da palavra escrita. Na aplicação empírica ficou evidenciado o efeito motivador deste tipo de ferramenta.

A principal limitação do modelo de inferência desenvolvido emerge do fato de fazer uso intensivo do computador. O procedimento de diagnóstico de convergência precisa de cinco processos estocásticos independentes, em cada um destes processos uma grande quantidade de amostras iniciais são descartadas e subseqüentes amostras são armazenadas para efetuar a análise. Estudos conjuntos relativamente grandes podem requerer várias horas de computação para cada processo estocástico, assim como também grande capacidade de armazenamento e de memória RAM. No entanto, é preciso considerar que o acelerado avanço de *hardware* deve reduzir drasticamente esta limitação. Adicionalmente, é preciso considerar que todos os algoritmos foram implementados em Matlab 5.3, que é um interpretador de comandos. A implementação em linguagens compiladas com certeza deve diminuir os tempos de processamento requeridos.

Algumas sugestões para a continuidade deste trabalho podem ser efetuadas.

Em relação à modelagem, sugere-se a extensão deste modelo para permitir por respostas de “não-escolha” ou “nenhuma das anteriores” visando incrementar o realismo comportamental da tarefa. A estimativa do parâmetro “não-escolha” pode ser posteriormente

utilizado no simulador para prognosticar a proporção de consumidores que não escolheriam quaisquer dos produtos simulados no cenário competitivo. Estrategicamente, grupos de consumidores cujas necessidades não são preenchidas pelas ofertas competitivas existentes constituem-se em oportunidades para a introdução de novas ofertas.

Ainda em relação à modelagem, a inferência bayesiana possui a característica de possibilitar a incorporação de conhecimento prévio do analista em relação ao problema abordado. Nesse sentido sugere-se a incorporação de conhecimento prévio, de caráter ordinal e/ou intervalar, sobre determinados níveis dos atributos. Esta incorporação é quase direta aplicando amostragem por rejeição, é interessante comparar o impacto da estimação com e sem restrições nos índices de desempenho preditivo.

A pesquisa de marketing não deve ser visualizada como um evento discreto, mas como um processo contínuo que fornece informação importante para o marketing e para sistemas de apoio à decisão. Com essa visão sugere-se a adaptação do modelo para fornecer em forma contínua informação de mudanças e tendências das preferências do mercado em relação a determinadas categorias de produtos, assim como estudos de satisfação. Especificamente, antes de efetuar um único estudo conjunto pode ser mais útil implementar uma sequência de estudos menores que incorporem sucessivamente o conhecimento adquirido de mercado. A característica estocástica deste modelo é teoricamente atrativa e provavelmente permita identificar estruturas de autoregressão que introduzem a dimensão tempo no estudo conjunto.

No que tange ao procedimento automatizado de coleta dos dados conjuntos, sugere-se a incorporação de mecanismos de controle para evitar inconsistências e respostas erradas, e um melhor aproveitamento dos recursos multimídia visando incrementar o realismo dos cenários de escolha apresentados.

Em relação à integração com outras técnicas, sugere-se a incorporação no modelo de outras técnicas multivariadas de marketing, como *Multidimensional Scaling* e Análise de Correspondências, para visualizar graficamente através de mapas perceptuais, as saídas dos cenários competitivos considerados no simulador. Por exemplo, para um possível nicho de mercado visualizar probabilidade de escolha vs. produtos escolhidos e relacionar com variáveis que caracterizam o consumidor.

Finalmente, não obstante as varias simulações efetuadas com dados sintéticos e a aplicação empírica realizada, é preciso destacar a necessidade por maiores aplicações do modelo, que forneçam maiores evidências em relação a sua aplicabilidade, limitações e oportunidades de aprimoramento.

Em marketing, como em muitas outras áreas do conhecimento, não existe uma única abordagem que resolva todos os problemas nem um determinado problema pode ser abordado através de apenas um único modelo. Portanto, é preciso abordar um problema específico sob diferentes metodologias e contrastar os resultados obtidos à luz das hipóteses efetuadas. Doses prévias de ceticismo, geralmente altas, são sempre frutíferas. Menores graus de incerteza posterior são o objetivo último.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBERT, H. James & Siddhartha CHIB (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, Vol.88, No.422, (June), p. 669-679.
- ALLENBY, Greg M. & James L. GINTER (1995), "Using Extremes to Design Products and Segment Markets," *Journal of Marketing Research*, 32, (November), p. 392-403.
- ALLENBY, Greg M. & Peter E. ROSSI (1999), "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics*, 89, p. 57-78.
- ALLENBY, Greg M. & Peter J. LENK (1994), "Modeling Household Purchase Behavior with Logistic Normal Regression," *Journal of the American Statistical Association*, Vol.89, No.428, (December), p.1218-1231.
- ALLENBY, Greg M., Neeraj ARORA & James L. GINTER (1995), "Incorporating Prior Knowledge into the Analysis of Conjoint Studies," *Journal of Marketing Research*, 32, (May), p. 152-162.
- ALLENBY, Greg M., Neeraj ARORA & James L. GINTER (1998), "On the Heterogeneity of Demand," *Journal of Marketing Research*, Vol. XXXV, (August), p. 384-389.
- AMEMIYA, T. (1981), "Qualitative Response Models: A Survey," *Journal of Economic Literature*, 19, p. 1483-1536.
- ATUAHENE-GIMA, K. (1995), "An Exploratory Analysis of the Impact of Market Orientation on New Product Performance: A Contingency Approach," *Journal of Product Innovation Management*, 12, (September), p. 275-293.
- BASTOS, Lia C. (1994), *Planejamento da rede escolar : uma abordagem utilizando preferência declarada*. Tese de doutorado em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, Brasil.
- BATES, John (1988), "Econometric Issues in Stated Preference Analysis," *Journal of Transport Economics and Policy*, (January), p. 59-69.
- BATSELL, Richard R. & Jordan J. LOUVIERE (1991), "Experimental Choice Analysis," *Marketing Letters*, 2 (August), p. 199-214.
- BEN-AKIVA, M. & S. R. LERMAN (1985), *Discrete Choice Analysis*. MIT Press, Cambridge, MA.
- BERNARDO, José M. & Adrian F. SMITH (1994), *Bayesian Theory*. New York: Wiley.
- BHAT, Chandra (1995), "A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice," *Transportation Research*, 29B(6), p. 471-483.

- BRETTON-CLARK (1992), *SIMGRAF Version 2. Manual*.
- BRONSHTEIN, I. & K. SEMENDIAEV (1973), *Manual de Matemáticas*. 2nd ed., Moscú: Editorial Mir.
- BROOKS, Stephen P. & Gareth O. ROBERTS (1998), "Assessing Convergence of Monte Carlo Algorithms," School of Mathematics, University of Bristol; Statistical Laboratory, University of Cambridge. England. Technical Report. p. 35.
- BROWNSTONE (2000), "Discrete Choice Modeling for Transportation," 9^o IATBR Travel Behavior Conference, (July), Australia.
- BROWNSTONE, D. & K. TRAIN (1999), "Forecasting New Product Penetration with Flexible Substitution Patterns," *Journal of Econometrics*, 89, p. 109-129.
- BUSSAB, Wilton, Édina MIAZAKI & Dalton de ANDRADE (1990), *Introdução à Análise de Agrupamentos*, Associação Brasileira de Estatística. ABE. 9^o Simposio Nacional de Probabilidade e Estatística. São Paulo.
- CARLIN, Bradley P. & Thomas A. LOUIS (1996), *Bayesian and Empirical Bayes Method for Data Analysis*. London: Chapman & Hall.
- CARROLL, Douglas J. & Paul E. GREEN (1995), "Psychometric Methods in Marketing Research: Part I, Conjoint Analysis," *Journal of Marketing Research*, Vol. 32 (November), p. 385-391.
- CASELLA, George & Edward I. GEORGE (1992), "Explaining the Gibbs Sampler," *The American Statistician*, Vol 46, No 3, (August), p. 167-174.
- CHAPMAN, Randall G. & Richard STAELIN (1982), "Exploiting Rank Ordered Choice Set Data Within the Stochastic Utility Model," *Journal of Marketing Research*, Vol. 19, (August), p. 288-301.
- CHAUVEAU, Didier & Pierre VANDEKERKHOVE (1999), "Improving Convergence of the Hastings-Metropolis Algorithm with a Learning Proposal," *Analyse et Mathématiques Appliquées*. Université de Marne-la-Vallée. Technical Report.
- CHIB, S. & GREENBERG, E. (1995), "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, Vol.49, No.4, p. 327-335.
- CHINTAGUNTA, Pradeep K. (1994) "Heterogeneous Logit Model Implications for Brand Positioning," *Journal of Marketing Research*, 31, (May), p. 304-311.
- CHINTAGUNTA, Pradeep K., Dipak C. JAIN & Naufel J. VILCASSIN (1991) "Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data," *Journal of Marketing Research*, 28, (November), p. 417-428.
- COHEN, Stephen L., David W. DOVE & Edward L. BACHELDER (2001), "Time to Treat Learners as Consumers," *Training & Development*, (January), p. 54-57.

- COWLES, M. K. & Bradley P. CARLIN (1996), "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative review," *Journal of the American Statistical Association*, 91, p. 883-904.
- CURRIM, Imran S. (1982), "Predictive Testing of Consumer Choice Models Not Subject to Independence of Irrelevant Alternatives," *Journal of Marketing Research*, 19, (May), p. 208-222.
- DAVISON, Peter (1991), *Stated Preference Techniques*, PTRC: Course Introduction to Stated Preference Techniques, U.K. p. 11.
- DAY, George (1994a), "Continuous Learning About Markets," *California Management Review*, 36, (Summer), p. 9-31.
- DAY, George (1994b), "The Capabilities of Market-Driven Organizations," *Journal of Marketing*, 58 (4), p. 37-52.
- DeSARBO, W. S., V. RAMASWAMY & S. H. COHEN (1995), "Market Segmentation with Choice-Based Conjoint Analysis," *Marketing Letters*, 6, p. 137-148.
- DIJKSTRA, J., W.A.H. ROELEN, e H.J.P. TIMMERMANS (1996), "Conjoint Measurement in Virtual Environments: a Framework," in H.J.P. Timmermans (ed.), 3rd Design and Decision Support Systems in Architecture and Urban Planning Conference, Vol. 1: Architecture Proceedings , pp. 59-71.
- DUBAS, K. M. & J. T. STRONG (1993), "Course Design Using Conjoint Analysis," *Journal of Marketing Education*, 15, p. 31-36.
- ELROD, Terry, Jordan LOUVIERE, & Krishnakumar S. DAVEY (1992), "An Empirical Comparison of Ratings-Based and Choice-Based Conjoint Models," *Journal of Marketing Research*, v29, n3, (August), p. 368-377.
- FINKBEINER, Carl T. (1988), "Comparison of Conjoint Choice Simulators," Sawtooth Software Conference Proceedings, p.75-103.
- GARVIN, David A. (1993), "Building a Learning Organization," *Harvard Business Review*, 71, (4), p. 78-91.
- GELFAND, A.E., S.E. HILLS, H.A. RACINE-POON & A.F.M. SMITH (1990), "Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling," *Journal of the American Statistical Association*, Vol.85, No.412, p. 972-985.
- GELFAND, Alan E. & Adrian F. SMITH, (1990), "Sampling-based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, p. 398-409.
- GELMAN, Andrew, John B. CARLIN, Hal S. STERN & Donald B. RUBIN (1995), *Bayesian Data Analysis*. London: Chapman & Hall,.
- GEMAN, S. & D. GEMAN (1984), "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, p.721-741.

- GEWEKE, J. F. (1991), "Efficient Simulation from the Multivariate Normal and Student-t Distributions Subject to Linear Constraints and the Evaluation of Constraint Probabilities," in *Computer Science and Statistics: Proceedings of the Twenty-Third Symposium on the Interface*, Alexandria, Va: American Statistical Association, p. 571-578.
- GEWEKE, J. F. (1997), *Applied Econometrics II*, Course Lecture Notes, p. 47.
- GEWEKE, J. F. (1999), "Using Simulation Methods for Bayesian Econometric Models: Inference, Developments and Communications," *Econometric Review*, 18 (1).
- GLAZER, Rashi (1991), "Marketing in an Information-Intensive Environment: Strategic Implications of Knowledge as an Asset," *Journal of Marketing*, 55, (4), p. 1-19.
- GORE, Paul A. (2000), "Cluster Analysis," in TINSLEY Howard E. & Steven D. BROWN (eds), *Handbook of Applied Multivariate Statistics and Mathematical Modeling*. Academic Press, p. 297-321.
- GREEN Paul E., Donald S. TULL & Gerald ALBAUM (1988), *Research for Marketing Decisions*. NJ: Prentice Hall.
- GREEN, Paul E. & V. SRINIVASAN (1978), "Conjoint Analysis in Consumer Behavior: Issues and Outlook," *Journal of Consumer Research*, 5, p. 103-23.
- GREEN, Paul E. & V. SRINIVASAN (1990), "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing*, 54 (October), p. 3- 19.
- GREEN, P. E. & V. R. RAO (1971), "Conjoint Measurement for Quantifying Judgmental Data," *Journal of Marketing Research*, 8, p. 355-363.
- GREENE, William H. (1997), *Econometric Analysis*. New Jersey: Prentice Hall, Inc.
- GRIFFIN Abbie & John R. HAUSER (1993), "The Voice of the Customer," *Marketing Science*, Vol. 10, No. 1, (Winter), p. 1-27.
- GUSTAFFSON, Anders, Fredrik EKDHAL & Bo BERGMAN (1999), "Conjoint Analysis: a Useful Tool in the Design Process," *Total Quality Management*, Vol. 10, No. 3, p. 327-343.
- HAIR, J. F., R. E. ANDERSON, R. L. TATHAM & W. C. BLACK (1998), *Multivariate Data Analysis*, NJ: Prentice Hall, Englewood Cliffs.
- HENSHER, David A. (1994), "Stated preference analysis of travel choice : the state of practice," *Transportation*, Netherlands, n. 21, p. 107-133.
- HERMAN, Steve (1988), "Software for Full-Profile Conjoint Analysis," in *Proceeding of the Sawtooth Conference on Perceptual Mapping, Conjoint Analysis, and Computer Interviewing*. M. Metegrano, ed. Ketchum, ID: Sawtooth Software, p. 117-130.
- HOFFMAN, D. L., J. De LEEUW & R. V. ARJUNJI (1994), "Multiple Correspondence Analysis," in BAGOZZI, R. P. (ed.) *Advanced Methods of Marketing Research*. Massachusetts: Blackwell Publishers. p. 260-294.

- HUBER, Joel & Kenneth TRAIN (2000), "On the Similarity of Classical and Bayesian Estimates of Individual Mean Part-Worths," Working Paper N° E00-289, Department of Economics, University of California, Berkeley.
- HUBER, Joel (1987), "Conjoint Analysis: How We Got Here And Where We Are," Published in Sawtooth Software Conference Proceedings.
- HUBER, Joel (1992), "Learning Effects In Preference Tasks: Choice-Based Versus Standard Conjoint," Published in Sawtooth Software Conference Proceedings.
- HUBER, Joel, Bryan ORME & Richard MILLER (1999), "Dealing with Product Similarity in Conjoint Simulations," Sawtooth Software Conference Proceedings, (February).
- JOHNSON, Richard (1987), "Adaptive Conjoint Analysis," in *Proceeding of the Sawtooth Conference on Perceptual Mapping, Conjoint Analysis, and Computer Interviewing*. M. Metegrano, ed. Ketchum, ID: Sawtooth Software, p. 253-265.
- JOHNSON, Richard A. & Dean W. WICHERN (1998), *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall.
- JONES, P. e J. POLAK (1994), "Computer-based Personal Interviewing: State-of-the-Art and Future Prospects," *Journal of the Market Research Society*, v35n3, p.221-233.
- JONES, Peter (1991), An Overview of Stated Preference Techniques, PTRC: Course Lecture Notes, U.K. p. 14.
- KACIAK, E. & J. LOUVIERE (1990), "Multiple Correspondence Analysis of Multiple Choice Experimental Data," *Journal of Marketing Research*, 27, p. 455-466.
- KAMAKURA, W. A. & G. J. RUSSELL (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26, p. 379-390.
- KASS, Robert E., Bradley P. CARLIN, Andrew GELMAN & Radford M. NEAL (1997), "Markov Chain Monte Carlo in Practice: A Roundtable Discussion," Joint Statistical Meetings, Technical Report. p. 26.
- KATZENSTEIN, H., S. KAVIL, V. MUMMALANENI & K. DUBAS (1994), "Design of an Ideal Direct Marketing Course from the Students' Perspective," *Journal of Direct Marketing*, 8, p. 66-72.
- KLABBERS, M.D., H. OPPEWAL & H.J.P. TIMMERMANS (1996), "ESCAPE: (Multimedia) Engine for Stated Choice and Preference Experiments," Working Paper paper 3rd Design and Decision Support Systems in Architecture and Urban Planning Conference.
- KOHLI, Ajay K. & Bernard J. JAWORSKI (1990), "Market Orientation: The Construct, Research Propositions, and Managerial Implications," *Journal of Marketing*, 54 (2), p. 1-18.
- KOPITTKE, Bruno H. (1989), "Simulação Empresarial: Faça o seu Jogo," Belo Horizonte, Anais do XII ENAMPAD, Vol 1, p 313-319.

- KOPITTKE, Bruno H, DETTMER, Armando, HERMENEGILDO, Jorge & MECHELN, Pedro (1999), "Jogo de Empresa e Educação à Distância - Uma Experiência Desenvolvida Via Vídeo Conferência," Rio de Janeiro, Anais do XIX ENEGEP.
- KROES, Eric P. & Robert J. SHELDON (1988), "Stated Preference Methods – An Introduction," *Journal of Transport Economics and Policy*, (January), p. 11-20.
- LENK, P. J., W. S. DeSARBO, P. E. GREEN & M. R. YOUNG (1996), "Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs," *Marketing Science*, Vol. 15, 2, p. 173-191.
- LOUVIERE, J. J. & G. WOODWORTH (1983), "Design and Analysis of Simulated Consumer Choice on Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, 20, p. 350-367.
- LOUVIERE, Jordan J. (1988), "Conjoint Analysis Modeling of Stated Preferences – A Review of Theory, Methods, Recent Developments and External Validity," *Journal of Transport Economics and Policy*, (January), p. 93-119.
- LOUVIERE, Jordan J., David A. HENSHER & Joffre D. SWAIT (2000), *Stated Choice Methods. Analysis and Applications*. Cambridge: University Press.
- LOUVIERE, Jordan J. (1984), "Using Discrete Experiments and Multinomial Logit Choice Models to Forecast Trial in a Competitive Retail Environment: A Fast Food Restaurant Illustration," *Journal of Retailing*, 60, p. 81-107.
- LUCE, R. D. & J. W. TUKEY (1964), "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement," *J. Math. Psychology*, 1, p.1-27.
- LUCE, R. D. (1977), "The Choice Axiom after Twenty Years," *Journal of Mathematical Psychology*, 15, p. 215-233.
- LYNN, Gary (1998), "New Product Team Learning: Developing and Profiting From Your Knowledge Capital," *California Management Review*, 40 (4), p. 74-93.
- MAHAJAN, V., P. E. GREEN & S. M. GOLDBERG (1982), "A Conjoint Model for Measuring Self and Cross-Price/Demand Relationships," *Journal of Marketing Research*, 19, p. 334-342.
- MAHAJAN, Vijay & Jerry WIND (1999), "Rx for Marketing Research," *Marketing Research*, 11, p. 6-13.
- MALHOTRA, N. (1995), *Marketing Research: an Applied Orientation*. 2nd ed. New Jersey: Prentice-Hall.
- MATHWORKS (1999), *Statistical Toolbox. For Use with Matlab*. User's Guide, Version 2. The MathWorks, Inc.
- McCULLOCH, Robert E. & Peter E. ROSSI (1994), "An Exact Likelihood Analysis of the Multinomial Probit Model," *Journal of Econometrics*, 64, p. 207-240.

- McFADDEN, D. & K. TRAIN (2000), "Mixed MNL Models for Discrete Response," *Journal of Applied econometrics*, Vol. 15, No. 5, p. 447-470.
- McFADDEN, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in P. ZAREMBKA (ed.), *Frontiers in Econometrics*, New York: Academic Press, p. 105-142.
- McFADDEN, D. (1980), "Econometric Models for Probabilistic Choice among Products," *Journal of Business*, 53, p. 13-29.
- McFADDEN, D. (1981), "Econometric Models of Probabilistic Choice," in Manski, C. & McFadden, D. (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, Mass: MIT Press, p. 198-272.
- McFADDEN, D. (1984), "Econometric Analysis of Qualitative Response Models," in GRILICHES Z. & M. D. INTRILIGATOR (eds.), *Handbook of Econometrics*, Vol. 2, Amsterdam: North-Holland.
- McFADDEN, D. (1986), "The Choice Theory Approach to Market Research," *Marketing Science*, 5, p. 275-297.
- McFADDEN, D. (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica* 57, p. 995-1026.
- MENGERSEN, K. L. & R. L. TWEEDDIE (1996), "Rates of Convergence of the Hastings and Metropolis Algorithms," *The Annals of Statistics*, 24 (1), p. 101-121.
- MONTGOMERY, Douglas C. (1984), *Design and Analysis of Experiments*, New York: Wiley.
- MOORMAN, Christine (1995), "Organizational Market Information Processes: Cultural Antecedents and New Product Outcomes," *Journal of Marketing Research*, 32 (3), p. 318-335.
- MORIKAWA, T. (1989), *Incorporating Stated Preference Data in Travel Demand Analysis*, Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. p 203.
- NEAL, Radford M. (1993), "Probabilistic Inference Using Markov Chain Monte Carlo Methods," Department of Computer Science. University of Toronto. Technical Report, CRG-TR-93-1. p. 144.
- NONAKA, Ikujiro (1991), "The Knowledge-Creating Company," *Harvard Business Review*, 69 (November-December), p. 96-104.
- NOVAES, A. G. & M. CARVALHO (1996), "Market Share Analysis of Transport Services with Stated Preference Data", *Coletânea Politécnica – Revista Brasileira de Tecnologia e Ciência*, (Maio/Agosto), Vol 1, No. 1, p. 78-91.
- ORME, Bryan (1998), "Sample Size Issues for Conjoint Analysis Studies," Sawtooth Software Conference Proceedings.

- PEREIRA, Carlos B. & Marlos A. VIANA (1988), "Elementos de Inferência Bayesiana," 5º *Simpósio Nacional de Probabilidade*. São Paulo.
- PINE, Joseph (1993), *Mass Customization*. Boston: Harvard University Press.
- REVELT, David & Kenneth TRAIN (1999), "Customer-Specific Taste Parameters and Mixed Logit," Working Paper N° E00-274, Department of Economics, University of California, Berkeley.
- ROBERTS, John H. (2000), "Developing New Rules for New Markets," *Journal of the Academy of Marketing Science*, 28, (Winter) p. 31-44.
- ROSS, Phillip J. (1991), *Aplicações das Técnicas Taguchi na Engenharia da Qualidade*. São Paulo: Makron, McGraw-Hill.
- ROSSI, Peter E. & Greg M. ALLENBY (1993), "A Bayesian Approach to Estimating Household Parameters," *Journal of Marketing Research*, 30, (May), p. 171-182.
- SHUKLA, P. & J. BRUNO (1992), "Use of Conjoint Analysis and Marketing Approaches in Education Surveys," *Education*, (Spring) p. 451-459.
- SLATER, Stanley R & John NARVER (1995), "Does Competitive Environment Moderate the Market Orientation Performance Relationship?," *Journal of Marketing*, 58 (1), p. 46-55.
- SMITH, A. F. & G. O. Roberts (1993), "Bayesian Computation Via the Gibbs Sampler and Related Markov Chain Monte Carlo," (with discussion), *Journal of the Royal Statistical Society, Series B*, 55, p. 3-24.
- SOUZA, Osmar A. (1999), *Delineamento Experimental em Ensaios Fatoriais Utilizados em Preferência Declarada*. Tese de doutorado em Engenharia de Produção. Universidade Federal de Santa Catarina, Florianópolis, Brasil.
- STATA, Ray (1989), "Organizational Learning-The Key to Management Innovation," *Sloan Management Review*, 30 (Spring), p. 63 -74.
- TAGUCHI, Genichi (1988), *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Costs*, New York: Kaus I. P.
- TANNER, M. A. (1993), *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. 2nd ed. New York: Springer Verlag.
- TANNER, M.A. & W.H. WONG (1987), "The Calculation of Posterior Distributions by Data Augmentation," with discussion, *Journal of the American Statistical Association*, Vol.82, N°398, p. 528-550.
- TARASEWICH, Peter & Suresh NAIR (2000), "Course Design using Instructor and Student Preferences," *The Journal of Business and Economic Studies*, 6, p. 40-54.
- TIERNEY, Luke. (1994), "Exploring Posterior Distributions Using Markov Chains," Technical Report 560, School of Statistics. University of Minnesota. p. 27.

- URBAN, Glen L., Bruce D. WEINBERG & John R. HAUSER (1996), "Prelaunch Forecasting of Really New Products," *Journal of Marketing*, 60, p. 47-60.
- URBAN, Glen L., John R. HAUSER, William J. QUALLS, Bruce D. WEINBERG, Jonathan D. BOHLMANN & Roberta A. CHICOS (1997), "Information Acceleration: Validation and Lessons From the Field," *Journal of Marketing Research*, 34, (February), p. 143-153.
- VRIENS, M., M. WEDEL & T. WILMS (1996), "Metric Conjoint Segmentation Methods: A Monte Carlo Comparison," *Journal of Marketing Research*, 33, (February) p. 73-85.
- WITTINK, Dick & Philippe CATTIN, (1989), "Commercial Use of Conjoint Analysis: An Update," *Journal of Marketing*, 53 (July), p. 91- 96.
- WITTINK, Dick, Marco VRIENS & Wim BURHENNE, (1994), "Commercial Use of Conjoint in Europe: Results and Critical Reflections," *International Journal of Research in Marketing*, 11, p. 41-52.
- WYNER, Gordon (1995), "Trade-off techniques and marketing issues," *Marketing Research*, 7, (Fall).
- ZALTMAN, Gerald e Robin HIGIE (1995), "Seeing the Voice of the Customer: Metaphor-Based Advertising Research," *Journal of Advertising Research*, (July/August), p.35-51.
- ZEGER, S.L. & M. R. KARIM (1991), "Generalized Linear Models with Random Effects: A Gibbs Sampling Approach, " *Journal of the American Statistical Association*, Vol. 86, No.413, p. 79-86.
- ZELLNER, Arnold (1971), *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley.
- ZUFRYDEN, F. S. (1983), "Course Evaluation and Design Optimization: A Conjoint Analysis-Based Application," *Interfaces*, 13, (April) p. 87-94.