**Um modelo abrangente
para
análise de preferência conjunta**

Paula Stefanoni Iwamizu

**DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO DE MESTRE
EM
CIÊNCIAS**

**Área de Concentração: Estatística
Orientador: Prof. Dr. Rinaldo Artes**

- São Paulo, Dezembro de 2006 -

**Um modelo abrangente
para
análise de preferência conjunta**

Este exemplar corresponde à redação final da
dissertação de mestrado devidamente corrigida e
defendida por Paula Stefanoni Iwamizu
e aprovada pela comissão Julgadora.

São Paulo, 13 de Dezembro de 2006.

Banca examinadora:

Prof. Dr. Rinaldo Artes (orientador) – IBMEC/SP

Prof. Dr. Luis Gustavo Esteves – IME/USP

Prof. Dr. José Afonso Mazzon – FEA/USP

“Eternidade é, na verdade, o amor vivendo sempre em nós.”

Livro de cantos da Arquidiocese de Jales

Ao Sadao, meu avô

Agradecimentos

Agradeço primeiro ao Guilherme, que esteve presente em cada segundo do desenvolvimento dessa dissertação, que me apoiou em todas as decisões tomadas e que torceu por mim. Saiba que, hoje, é você quem me inspira a estudar e crescer. Tenho orgulho de você. Obrigada.

Minhas professoras do *Cacilda Becker*, obrigada pelo carinho e por terem me envolvido neste mundo das ciências exatas. Acredito que são responsáveis por eu ter escolhido esta área e é por isso que agradeço. Além das disciplinas, agradeço pelas festas juninas, passeios, músicas, olimpíadas etc.

Sem dúvidas agradeço, com carinho, aos meus familiares, que me respeitaram e sempre que possível, me ajudaram em tudo. Dona Luzia (minha mãe), seu Paulo (meu pai), Ju e Rô, obrigada.

Agradeço aos Stefanoni's e à família Iwamizu pela confiança e compreensão. Batian, obrigada pelo incentivo e pelos almoços durante o colegial. Tia Ró, obrigada.

Aos meus amigos do *Básico 2000*, muito obrigada. Foi um prazer estudar e me divertir com vocês; só por isso já teria valido a pena encarar a graduação, contem comigo. É um prazer ainda maior ver cada um de vocês crescendo junto comigo. Macris e Carol, obrigada (*Amizade talvez seja isso*). Agradeço também ao Fábio por me “dar” seu livro bayesiano, saiba que me ajudou muito. A todos os amigos que fiz neste período e aos que já conhecia, obrigada.

Agradeço às pessoas do meu trabalho, *SD&W*, pela confiança depositada toda vez que precisei alterar meus horários e pela torcida.

Devo agradecer aos professores do IME que, rigorosos ou não, me fizeram estudar e aprender a estudar. Finalmente, agradeço ao meu orientador, professor Rinaldo Artes, por ter acreditado e confiado em mim. Obrigada pela paciência e pelo bom-humor com que sempre me recebeu. Não me arrependi em nenhum momento dessa escolha e tenho muito respeito pelo profissional que é. Obrigada.

Por fim, agradeço a Deus.

Resumo

Estudamos, nessa dissertação, um modelo bayesiano desenvolvido para a Análise de Preferência Conjunta, proposto por Marshall e Bradlow (2002). Este modelo incorpora dados obtidos a partir dos métodos de coleta de composição e de decomposição e, apesar de ser estruturado para avaliações feitas por atribuição de notas, pode ser estendido para casos nos quais as avaliações coletadas na fase de decomposição tenham sido feitas através de soma-constante, ordenamento ou escolha discreta.

Através da abordagem bayesiana, o modelo utiliza as informações coletadas no método de composição como informação *a priori* sobre as utilidades não-observadas (coeficientes que correspondem às contribuições parciais dos níveis dos atributos). Na construção da função de verossimilhança, utiliza as informações coletadas no método de decomposição. Com isso, a distribuição *a posteriori* resultante, além de reunir toda a informação disponível, leva a um estimador que pondera essas informações. A obtenção das estimativas é feita pelo método de amostragem de Gibbs.

Iniciamos a dissertação com uma revisão sobre os conceitos básicos de Análise de Preferência Conjunta e, após introduzir o Modelo Bayesiano de Marshall e Bradlow, apresentamos uma aplicação deste modelo a dados reais. Nessa aplicação, verificamos a vantagem de utilizar esse modelo com um número reduzido de estímulos.

Abstract

In this work, we study a bayesian approach to conjoint analysis models, originally presented by Marshall and Bradlow (2002). This model uses data obtained from both compositional and decompositional methods and, despite it was first designed for evaluations under a rating scale, it can be generalized to constant-sum scale, ranking scale or discrete choice evaluations.

Using bayesian approach, this model uses the information obtained from the compositional method to form a prior distribution for the unobserved vector of true parthworths (coefficients which correspond to the partial contributions of the attribute levels) and the information obtained from the decompositional method to constitute the likelihood function. The estimator of the parthworths is given by the resultant posterior mean that weight each data source according to the information contained. The estimation of its parameters is obtained from the Gibbs sampling method.

We start this work reviewing some basic concepts of Conjoint Analysis. After introducing the Bayesian (unified) Model proposed by Marshall and Bradlow, we present an application of this model, in which we can verify its performance when we have a small number of full-profile evaluations. As an illustration, we analysed the conjoint preference of the respondents using the estimatives obtained from it.

Apresentação

O principal objetivo desta dissertação é apresentar o modelo bayesiano desenvolvido para a Análise de Preferência Conjunta proposto por Marshall e Bradlow (2002).

No primeiro capítulo, apresentamos uma introdução sobre a Análise de Preferência Conjunta, a fim de situar o leitor neste contexto antes de introduzir efetivamente o Modelo Bayesiano. Neste capítulo, abordamos sua evolução e alguns conceitos nela utilizados, introduzimos a idéia dos principais métodos de coleta de dados (de composição e de decomposição) e, por fim, citamos alguns dos modelos usuais relacionados a esta técnica.

O modelo de Marshall e Bradlow é apresentado no segundo capítulo. Neste momento, expomos a idéia da abordagem bayesiana dada pelos autores e desenvolvemos todas as etapas do modelo. Além disso, apresentamos tanto seus casos particulares como algumas de suas propriedades; em seguida, detalhamos o processo de estimação dos parâmetros, também proposto pelos autores.

No capítulo seguinte, apresentamos algumas extensões do modelo, que são, na verdade, adaptações do mesmo para situações que não obedecem às premissas feitas no seu desenvolvimento. Avaliações que não são obtidas através de atribuição de notas, são exemplos dessas situações.

No quarto capítulo consta uma aplicação. Apresentamos nele, o planejamento de nossa coleta de dados, o ajuste de alguns modelos a esse conjunto de dados e a comparação dos resultados obtidos através desses modelos. Uma dessas comparações se refere ao número de estímulos avaliados no Modelo Bayesiano. Investigamos o efeito da redução do número de avaliações de estímulos no desempenho do modelo.

Encerramos essa dissertação com uma Análise de Preferência Conjunta feita sobre o conjunto de dados. Com essa análise comparamos, novamente, os resultados obtidos através de duas versões do Modelo Bayesiano de Marshall e Bradlow.

Anexamos a esta dissertação, alguns apêndices que podem ser úteis ao leitor, como, por exemplo, uma introdução aos métodos de estimação via simulações aqui utilizados e um guia que descreve como obtivemos as estimativas dos Modelos Bayesianos através da interface BRugs (implementada no pacote estatístico R).

Sumário

1	Análise de Preferência Conjunta	5
1.1	Introdução	5
1.2	Etapas de uma Análise de Preferência Conjunta	8
1.2.1	Definição e construção dos estímulos	8
1.2.2	Avaliação da preferência do consumidor	10
1.2.3	Métodos de coleta de dados	13
1.2.4	Modelos de preferência	16
1.2.5	Apresentação dos estímulos aos entrevistados	22
1.2.6	Métodos de estimação	22
2	Modelo Bayesiano	25
2.1	Apresentação do Modelo Bayesiano	25
2.1.1	Distribuição <i>a priori</i> de μ_i	26
2.1.2	Função de Verossimilhança	27
2.1.3	Distribuição <i>a posteriori</i> de μ_i	30
2.2	Casos particulares	31
2.3	Extensão	32
2.4	Propriedades	33
2.5	Estimação dos parâmetros	34
3	Extensões do Modelo Bayesiano	37
3.1	Avaliações obtidas sob escalas distintas	37
3.2	Avaliações da fase de decomposição feitas por soma-constante	41

3.3	Avaliações da fase de decomposição feitas por ordenamento	44
3.4	Avaliações da fase de decomposição feitas por escolha discreta	47
3.5	Comentários	49
4	Comparação dos modelos	51
4.1	Definição do produto e planejamento do questionário	51
4.2	Definição dos Modelos ajustados	54
4.2.1	Modelo utilidade do método de decomposição	55
4.2.2	Modelo de composição	56
4.2.3	Modelos bayesianos	58
4.2.4	Estimação dos coeficientes dos Modelos Bayesianos	61
4.3	Resultados da comparação	65
5	Aplicação da Análise de Preferência Conjunta	67
5.1	Cálculo das Importâncias dos atributos	67
5.2	Análise de Agrupamento	70
5.3	Comentários	73
6	Comentários Finais	75
6.1	Futuros estudos	75
A	Demonstrações de resultados	77
A.1	Distribuição <i>a posteriori</i> de $\mu_i, f(\mu_i y_i)$	77
A.2	Distribuições condicionais completas dos parâmetros σ^2 e Ω	80
A.3	Alguns resultados matriciais	84
A.4	Matriz de variâncias-covariâncias do Modelo Bayesiano	85
B	Questionário aplicado e respostas dadas pelos indivíduos	87
C	Métodos de estimação	93
C.1	Método Monte Carlo	94
C.2	Amostrador Gibbs	94
C.3	Método de diagnóstico de convergência para múltiplas cadeias	97

D Gráficos	101
D.1 Modelo Bayesiano I	101
D.2 Modelo Bayesiano II	113
D.3 Modelo Bayesiano III	124
D.4 Modelo Bayesiano IV	135
E Estimativas obtidas no ajuste dos modelos	147
E.1 Modelo de Decomposição	147
E.2 Modelo de Composição	153
E.3 Modelo Bayesiano I	156
E.4 Modelo Bayesiano II	163
F Estatísticas para avaliação do ajuste do modelo	171
G Implementação dos Modelos Bayesianos no BRugs	173

Capítulo 1

Análise de Preferência Conjunta

Neste primeiro capítulo apresentamos uma introdução à Análise de Preferência Conjunta, abordando suas idéias, nomenclaturas, métodos de coleta de dados e alguns de seus modelos. Queremos com isso, resgatar tópicos que julgamos necessários para o entendimento do modelo de Marshall e Bradlow (2002) que é o foco dessa dissertação.

Sugerimos como referência para os que desejam se iniciar em Análise de Preferência Conjunta, a dissertação de mestrado de Friedmann (1998).

1.1 Introdução

A Análise de Preferência Conjunta é uma técnica conhecida na área de marketing, uma vez que lida com a modelagem da preferência de consumidores por produtos. Nessa técnica, a preferência do consumidor é decomposta segundo a importância dada a características do produto; essa decomposição é feita através de ajuste de modelos individuais para os elementos da amostra. A partir dos parâmetros estimados para o modelo escolhido é possível simular a aceitação do mercado a um produto pré-definido; determinar qual seria o produto “ideal” na concepção de cada consumidor; comparar, dentre os atributos que definem o produto, quais apresentam maior impacto na preferência dos consumidores; agrupar os consumidores segundo a importância dada aos componentes do produto (segmentar o mercado consumidor); enfim, permite-nos entender, de uma maneira geral, o mecanismo de escolha do consumidor.

Devido a essas qualidades, a área de marketing é uma das principais áreas de aplicação dessa

técnica (veja Wittink e Cattin, 1989), pois, através dela, é possível capturar as características do produto que direcionam um consumidor a preferir um certo produto e não outro.

A aplicação dessa técnica não está restrita aos produtos do tipo físico, como *automóveis*, *sapatos*, *sabão em pó* etc; ela pode ser aplicada, também, na análise de serviços, como *atendimento pós-venda*, *métodos de aula e professores*, *suportes e benefícios oferecidos por uma empresa* etc. Ou seja, sua aplicação é viável desde que se possa decompor o produto, ou serviço, em componentes (*atributos*) e variações desses componentes (*níveis dos atributos*), possibilitando a construção de produtos hipotéticos, chamados *estímulos*.

As técnicas de Análise de Preferência Conjunta podem variar quanto ao método de coleta dos dados (composição, decomposição); à forma de mensuração da preferência (notas, ordenamento, soma constante etc); às escalas utilizadas na mensuração (0 a 10, 1 a 100 etc); e aos modelos utilizados para obter a relação existente entre a preferência global dada aos produtos hipotéticos e as contribuições parciais dos níveis dos atributos (veja Green e Srinivasan, 1990). Vale ressaltar, que o modelo que irá estabelecer essa relação está diretamente relacionado ao método de obtenção dos dados.

Os métodos de composição, segundo Green e Krieger (1996), são métodos de coleta de dados encontrados desde o final de 1970 (ver Dawes e Corrigan, 1974). Nesse tipo de abordagem, os entrevistados avaliam os componentes e suas variações (atributos e níveis de atributos). A partir dessas avaliações parciais, compõem-se as avaliações globais dos estímulos através de alguma regra (modelo). Por exemplo, soma-se a importância atribuída a cada componente do produto. Veja a Seção 1.2.3.

Estes métodos, bem como os modelos a eles associados, na realidade não pertencem originalmente às técnicas de Análise de Preferência Conjunta, pois não decompõem a preferência dada aos estímulos mas sim as compõem. Porém, dado que estabelecem uma relação entre a preferência global e as contribuições parciais dos atributos, estes métodos compartilham da mesma idéia logo, são citados como alternativa ou então mesclados às técnicas de Análise de Preferência Conjunta (Green e Krieger, 1996).

O que chamou a atenção da área de marketing, para o uso de Análise de Preferência Conjunta, foi o artigo desenvolvido por Luce e Tukey, em 1964. O primeiro artigo detalhado sobre a aplicação da Análise de Preferência Conjunta no estudo do processo de escolha dos consumidores foi produzido por Green e Rao, em 1971. Esse artigo deu margem a várias extensões e aplicações (por exemplo, Green et al., 1972, Srinivasan e Shocker, 1973). Nessas abordagens, o entrevistado avalia produtos hipotéticos (estímulos) criados pelas possíveis variações do produto de interesse, e, a partir dessas avaliações,

infere-se a importância de cada atributo do produto. São os chamados métodos de decomposição, que apresentamos na Seção 1.2.3.

Como a base de dados necessária para desenvolver o método de decomposição é formada por avaliações feitas a estímulos, pode haver problemas relacionados ao número de estímulos a serem avaliados pelo entrevistado. A tarefa de avaliar um número elevado de estímulos pode ser desgastante para ele e, conseqüentemente, gerar resultados não confiáveis (Malhotra, 1986). Suponha 4 atributos com 4, 4, 3 e 3 níveis, respectivamente. Neste caso, o entrevistado deveria avaliar 144 ($4 \times 4 \times 3 \times 3$) estímulos (cada um representando um produto hipotético formado pela combinação dos níveis dos seus atributos); mesmo que ele aceite essa tarefa, as informações obtidas dificilmente seriam confiáveis.

Uma solução para reduzir o número de estímulos a serem avaliados pelo entrevistado é, por exemplo, a aplicação de planejamentos fatoriais fracionários ou planejamentos em blocos (Neter et al., 1996 e Dean e Voss, 1999).

Surgem, também com o propósito de reduzir o trabalho do entrevistado, os modelos mistos, que utilizam informações obtidas via método de composição e de decomposição. A vantagem deste método é que o número de estímulos avaliados no método de decomposição pode ser reduzido. Citamos alguns modelos mistos na Seção 1.2.4.

Com a popularidade da técnica de Análise de Preferência Conjunta, muitos estudos e pesquisas foram desenvolvidos e aprofundados. Um resumo sobre Análise de Preferência Conjunta, que fornece uma visão geral dessa técnica, pode ser encontrado em Green e Srinivasan (1990). Esse artigo é uma atualização do que foi publicado em 1978 (Green e Srinivasan, 1978).

Devido à gama de escalas, modos de avaliação, métodos de coleta de dados e modelos de preferência, é interessante a proposta de um único modelo, que abranja os outros como casos particulares. O modelo unificado, que será aqui estudado, foi proposto por Marshall e Bradlow (2002) e tem essa intenção. Os autores apresentam uma abordagem unificada para Análise de Preferência Conjunta através de um enfoque bayesiano. Este modelo é introduzido na Seção 2.1.

Nas seções seguintes descrevemos algumas etapas necessárias para a condução de uma Análise de Preferência Conjunta.

1.2 Etapas de uma Análise de Preferência Conjunta

Apresentamos nesta seção as definições e tópicos que caracterizaram a técnica de Análise de Preferência Conjunta, para isso, utilizamos como base as etapas descritas em Green e Srinivasan (1990) e seguimos a seguinte sequência:

1. Definição e construção dos estímulos
2. Avaliação da preferência do consumidor
3. Métodos de coleta de dados
4. Modelos de preferência
5. Apresentação dos estímulos aos entrevistados
6. Métodos de estimação

1.2.1 Definição e construção dos estímulos

Em Análise de Preferência Conjunta é possível estudar a preferência do consumidor por produtos físicos como *automóveis, sapatos, celulares* etc; ou conceitos como *Métodos de ensino, Serviços de um banco, Atendimento ao consumidor* etc. Basta que seja possível decompor este produto em componentes.

Os componentes (atributos) que caracterizam o produto são escolhidos pelo pesquisador. Essa escolha deve, simultaneamente, atender seus interesses (atributos específicos e estratégicos), contemplar os atributos importantes para o consumidor e apresentar o menor número possível de atributos. Para que nenhum atributo importante deste produto seja descartado, uma pesquisa prévia (entrevista) com consumidores pode ser interessante (vide Friedmann, 1998).

O cuidado com a definição dos atributos (e níveis) é fundamental, pois é sobre eles que a preferência dada pelo consumidor ao produto será decomposta.

Para ilustrar, suponha que o produto de interesse seja um *automóvel* e que desejamos estudar a importância dada pelos consumidores, na hora da compra, ao *modelo* e à *cor* deste produto. Temos, então, definidos como atributos os componentes *modelo* e *cor*.

Os atributos devem ser passíveis de variações; mesmo que sejam do tipo *presença/ausência* do próprio atributo. Estas variações, denominadas *níveis*, nos permitem inferir a importância de cada

atributo na composição da preferência dada ao produto. No exemplo do *automóvel*, definimos como níveis de atributos as seguintes variações: Gol, Corsa, Palio e Celta para o atributo *modelo*; e verde, vermelho e azul para o atributo *cor*.

Uma maneira de se obter os estímulos é combinando os níveis de todos os atributos em estudo, como pode ser visto na Tabela 1.1; é devido a esse fato que são chamados de estímulos de “perfil-completo”. Note que esta tabela apresenta todas as possíveis combinações obtidas com estes dois atributos ($4 \times 3 = 12$).

Tabela 1.1: Estímulos compostos

Estímulo	Modelo	Cor
1	<i>Gol</i>	<i>verde</i>
2	<i>Gol</i>	<i>vermelho</i>
3	<i>Gol</i>	<i>azul</i>
4	<i>Corsa</i>	<i>verde</i>
5	<i>Corsa</i>	<i>vermelho</i>
6	<i>Corsa</i>	<i>azul</i>
7	<i>Palio</i>	<i>verde</i>
8	<i>Palio</i>	<i>vermelho</i>
9	<i>Palio</i>	<i>azul</i>
10	<i>Celta</i>	<i>verde</i>
11	<i>Celta</i>	<i>vermelho</i>
12	<i>Celta</i>	<i>azul</i>

Uma outra forma de construir estímulos é através de matrizes *Trade off*, que consideram dois atributos por vez (Johnson, 1974). Esta forma pode ser utilizada nos casos em que há mais de dois atributos definidos na composição do produto, o que nos remete a estímulos de “perfis-parciais”. Assuma, por exemplo, que, além de *modelo* e *cor*, tenhamos definido o atributo *Número de portas* para o exemplo do *automóvel*; com os seguintes níveis: *2 portas* e *4 portas*. Poderíamos, então, construir e apresentar os estímulos através de matrizes *Trade off* como ilustra a Figura 1.1.

As matrizes *Trade off* facilitam a construção e avaliação dos estímulos, principalmente quando há um elevado número de atributos e níveis. Por apresentar perfil-parcial, não será possível obter, dos

Figura 1.1: Exemplo de apresentação de estímulos através de matrizes Trade off

		Número de portas	
		2 portas	4 portas
Modelo	Gol		
	Corsa		
	Palio		
	Celta		

		Número de portas	
		2 portas	4 portas
Cor	Vermelho		
	Verde		
	Azul		

		Cor		
		Vermelho	Verde	Azul
Modelo	Gol			
	Corsa			
	Palio			
	Celta			

entrevistados, a avaliação da preferência dada ao produto considerando todos os atributos simultaneamente, porém, poderá reduzir o trabalho dos mesmos no momento da avaliação.

1.2.2 Avaliação da preferência do consumidor

A preferência do consumidor pode ser mensurada através de avaliações feitas a estímulos que lhe são apresentados e/ou através de avaliações feitas diretamente aos atributos e níveis que compõem o produto. Nesta seção, discutimos como mensurar essa preferência dada pelos entrevistados.

Os estímulos que serão apresentados e a maneira que a preferência aos produtos será estimada dependem do *método de coleta de dados* e do *modelo de preferência* adotados (apresentamos estes tópicos na próxima seção).

A avaliação, seja de estímulos ou atributos, pode ser realizada de diferentes modos e com diferentes escalas. Dentre os modos mais utilizados, temos a atribuição de notas, a soma-constante, a atribuição de postos (ordenamento) e a comparação pareada (veja Green e Srinivasan, 1990).

Na avaliação por notas, pede-se ao entrevistado que atribua notas aos estímulos (ou atributos

ou níveis de atributos) de acordo com sua preferência por cada um deles. Neste caso, quanto maior for sua preferência pelo produto, maior deve ser a nota dada. A escala da nota deve ser definida previamente, conforme o interesse do pesquisador. Imagine que um entrevistado prefira o “Gol verde” ao “Corsa azul”; espera-se, então, uma nota maior para o “Gol verde” do que para o “Corsa azul”. A Tabela 1.2 ilustra essa situação.

Tabela 1.2: Exemplo de avaliação feita por atribuição de notas

<i>Atribuir notas de 0 a 100</i>	
Estímulo	Nota
<i>Gol verde</i>	100
<i>Corsa azul</i>	40

Para a avaliação com soma-constante, pede-se ao entrevistado que distribua um total de pontos (100 por exemplo) dentre os estímulos (ou atributos ou níveis de atributos) que recebeu para avaliar. Essa distribuição deve expressar sua preferência por cada um deles, e apresentar uma soma constante. Novamente, quanto maior a preferência por um determinado estímulo, mais pontos este deve receber. Desse modo, o “Gol verde” poderia ter recebido 80 e o “Corsa azul” 20 pontos, se essa pessoa tivesse que avaliar apenas esses dois estímulos e preferisse o “Gol verde” ao “Corsa azul”. Como exemplo, veja a Tabela 1.3, que apresenta uma avaliação a estímulos feita através de soma-constante; neste caso a soma deve totalizar 100 pontos.

Tabela 1.3: Exemplo de avaliação feita por soma-constante

<i>Distribuir 100 pontos</i>	
Estímulo	Pontos
<i>Gol verde</i>	80
<i>Corsa azul</i>	20
Total de pontos	100

A escolha da escala das notas ou do total da soma-constante é arbitrária e pode variar conforme o interesse do pesquisador ou facilidade que o entrevistado terá com a escala.

A avaliação através de postos é obtida quando o entrevistado ordena os estímulos de acordo

com sua preferência (Chapman e Staelin, 1982). Desse modo, cada estímulo recebe um posto, que representará a preferência do entrevistado por ele. Se definirmos que o estímulo de maior preferência do entrevistado recebe o menor posto, teremos, no exemplo, posto 1 para o “Gol verde” e posto 2 para o “Corsa azul”, pois o entrevistado prefere o “Gol verde” ao “Corsa azul”. Suponha que um entrevistado tenha ordenado 7 estímulos, de maneira que o primeiro seja seu preferido, o segundo seja o preferido dentre os restantes e, assim, sucessivamente. A Tabela 1.4 representa essa ordenação.

Tabela 1.4: Exemplo de avaliação feita por ordenamento

<i>Ordenar conforme a preferência</i>	
Estímulo	Posto
<i>Gol verde</i>	1
<i>Gol vermelho</i>	2
<i>Celta azul</i>	3
<i>Palio verde</i>	4
<i>Palio vermelho</i>	5
<i>Corsa verde</i>	6
<i>Corsa azul</i>	7

Uma outra maneira de avaliar os estímulos é através de comparação pareada. Nesse caso, os estímulos são apresentados aos pares e o entrevistado deve indicar qual é o estímulo preferido em cada par. Na realidade, essa escolha pode ser feita sobre um conjunto de n estímulos, sendo que n pode ser igual ou maior que dois. Esse tipo de avaliação é chamado escolha discreta (veja Green e Krieger, 1996 e McFadden, 1974). Na Tabela 1.5 apresentamos um exemplo no qual o estímulo “Gol verde” é preferido ao “Corsa azul” e, no segundo quadro, o “Gol vermelho” é o estímulo preferido.

O método de avaliação dos estímulos, ou dos atributos, ou mesmo dos níveis de atributos, está relacionado com o método de coleta desses dados. Veremos na seção seguinte alguns métodos de coleta de dados.

Tabela 1.5: Exemplo de avaliação feita por escolha discreta

<i>Comparação pareada</i>	
Estímulo	Estímulo preferido
<i>Gol verde</i>	X
<i>Corsa azul</i>	
<i>Escolha discreta</i>	
Estímulo	Estímulo preferido
<i>Gol verde</i>	
<i>Gol vermelho</i>	X
<i>Corsa azul</i>	
<i>Celta verde</i>	

1.2.3 Métodos de coleta de dados

O método de coleta de dados define como será feita a análise de preferência conjunta, bem como o modelo que será ajustado aos dados. Esses métodos são classificados basicamente em três grupos; o de composição, o de decomposição e o misto (que combina os dois anteriores). Nesta seção introduzimos, de maneira simplificada, o conceito desses três métodos de coleta de dados.

Métodos de composição

Nos métodos de composição, o entrevistado avalia os níveis dos atributos conforme sua preferência por eles e, em seguida, avalia os atributos segundo a importância de cada um na sua concepção.

No caso em que a importância dos atributos é obtida através de soma-constante e a dos níveis através de atribuição de notas, a avaliação global do estímulo pode ser obtida com a soma das avaliações feitas para os níveis que compõem esse estímulo, ponderadas pela importância dada ao atributo correspondente (Fishbein, 1967). As parcelas dessa soma são chamadas *utilidades* e representam as contribuições parciais dos níveis dos estímulos à preferência do consumidor.

Na Tabela 1.6 apresentamos um exemplo em que o entrevistado atribuiu notas de 0 a 10 aos modelos de automóvel Gol, Corsa, Palio e Celta, de acordo com a preferência por cada um deles e notas de 0 a 10 às cores Verde, Vermelho e Azul. Em seguida, ele distribuiu um total de 10 pontos entre os atributos *Modelo* e *Cor* conforme a importância de cada um deles na sua preferência por automóveis (veja Tabela 1.7).

Tabela 1.6: Exemplo de avaliações feitas aos níveis dos atributos

Modelo	Nota		Cor	Nota
<i>Gol</i>	10		<i>Verde</i>	9
<i>Corsa</i>	6		<i>Vermelho</i>	2
<i>Palio</i>	3		<i>Azul</i>	7
<i>Celta</i>	9			

Tabela 1.7: Exemplo de avaliações feitas aos atributos

Atributo	Pontos
<i>Modelo</i>	7
<i>Cor</i>	3
Total de pontos	10

O modelo usual, que será definido na seção seguinte, é a soma das avaliações feitas aos níveis dos atributos ponderadas pela avaliação dada ao atributo correspondente.

Note que a tarefa do entrevistado não é complicada; porém, deve-se ter cuidado com a aplicação desse método, pois ele pode gerar resultados distorcidos, uma vez que alguns atributos, quando avaliados de maneira isolada, podem apresentar uma importância maior do que apresentariam se estivessem sendo avaliados em conjunto com outros atributos (Green e Srinivasan, 1990). Como exemplo deste fato, podemos citar o caso de atributos relacionados ao meio ambiente, como reciclagem, ou então responsabilidade social. É provável que, se forem analisados isoladamente, recebam notas altas, pois entendemos que o entrevistado os julgará importantes, mas, no momento da decisão da compra de um produto, talvez não tenham tanta importância.

Este tipo de coleta de dados é interessante quando os atributos e níveis de atributos definidos geram um elevado número de estímulos, que seria inviável de ser avaliado pelo entrevistado. Perceba que apenas os níveis de atributos e os atributos são avaliados, o que é relativamente mais fácil e prático (veja Srinivasan e Park, 1997).

Métodos de decomposição

Nos métodos de decomposição os entrevistados avaliam estímulos que lhe são apresentados, sejam eles de perfil-completo, nos quais estímulos formados pela combinação de níveis de todos os atributos; ou através de matrizes *Trade off* (Johnson, 1974), como apresentamos na Figura 1.1.

As avaliações, que podem ser feitas através de atribuição de nota, ordenamento, soma-constante etc, são “decompostas”, através de um modelo apropriado, em contribuições parciais dos níveis dos atributos definidos para esse produto. Trabalhamos nesta dissertação apenas com as avaliações feitas a estímulos de perfil-completo, a Tabela 1.8 ilustra uma atribuição de notas feitas aos estímulos apresentados na Tabela 1.1.

Tabela 1.8: Exemplo de avaliações da preferência a estímulos

Estímulo	Modelo	Cor	Avaliação
1	<i>Gol</i>	<i>verde</i>	100
2	<i>Gol</i>	<i>vermelho</i>	82
3	<i>Gol</i>	<i>azul</i>	80
4	<i>Corsa</i>	<i>verde</i>	50
5	<i>Corsa</i>	<i>vermelho</i>	35
6	<i>Corsa</i>	<i>azul</i>	40
7	<i>Palio</i>	<i>verde</i>	32
8	<i>Palio</i>	<i>vermelho</i>	10
9	<i>Palio</i>	<i>azul</i>	27
10	<i>Celta</i>	<i>verde</i>	90
11	<i>Celta</i>	<i>vermelho</i>	74
12	<i>Celta</i>	<i>azul</i>	80

Considerando nosso exemplo sobre automóveis, observa-se, por exemplo, a preferência de um consumidor aos estímulos “Gol verde”, “Gol azul”, “Corsa verde” etc; e a partir dessas observações, estimam-se as avaliações (preferências) parciais desse consumidor ao modelo de automóvel “Gol”, modelo de automóvel “Corsa”, à cor de automóvel “verde” etc, ou seja, decompõe-se a preferência global observada em contribuições parciais referentes aos modelos de automóveis e às suas cores.

O número de estímulos a serem avaliados está relacionado com o modelo de Análise de Preferência Conjunta que será ajustado.

Métodos mistos

Os métodos mistos combinam os dois tipos de coleta de dados, ou seja, o entrevistado, além de avaliar os atributos e os níveis dos atributos conforme o método de composição, também avalia um subconjunto de estímulos construídos conforme um método de decomposição (veja Green, 1984 e Green e Krieger, 1996).

A idéia de mesclar os dois métodos permite que se avalie apenas um subconjunto de estímulos, com isso, atinge seu principal objetivo, que é reduzir o trabalho do entrevistado.

1.2.4 Modelos de preferência

Os *modelos de preferência* são funções matemáticas que permitem estabelecer uma relação entre a preferência geral associada ao produto (avaliações feitas pelos consumidores) e as contribuições parciais de cada nível de atributo que o compõem.

Isso nos permitirá, por exemplo, inferir quão valioso é cada nível de atributo na composição da *preferência* do consumidor por aquele produto. Apresentamos aqui alguns destes modelos, que, geralmente, são estimados para cada indivíduo em separado.

Modelo para dados coletados sob o método de composição

No método de composição, avaliações são feitas para os níveis dos atributos e para os próprios atributos. O modelo usual para compor a avaliação de um estímulo é o aditivo e consiste em somar as

avaliações dadas aos níveis que compõem o estímulo, ponderadas pela importância dada aos atributos correspondentes (Fishbein, 1967).

Considere as avaliações das tabelas 1.6 e 1.7. Neste caso, para obter a utilidade de um nível, basta multiplicar a nota atribuída a ele pelo número de pontos atribuído ao atributo correspondente. Por exemplo, a utilidade do nível *Gol* do atributo *Modelo* é dada por

$$[\text{Nota do nível } Gol \times \text{Pontos do atributo } Modelo = 10 \times 7 = 70]$$

e a do nível *Verde* do atributo *Cor* é dada por

$$[\text{Nota do nível "Verde"} \times \text{Pontos do atributo } Cor = 9 \times 3 = 27].$$

A avaliação da preferência a um estímulo é, então, composta pela soma das utilidades dos níveis que o definem. A Tabela 1.9 apresenta as avaliações compostas para este exemplo.

Tabela 1.9: Exemplo de composição de avaliações da preferência a estímulos

Estímulo	Modelo	Cor	Composição	Avaliação composta
1	<i>Gol</i>	<i>verde</i>	70 + 27	97
2	<i>Gol</i>	<i>vermelho</i>	70 + 6	76
3	<i>Gol</i>	<i>azul</i>	70 + 21	91
4	<i>Corsa</i>	<i>verde</i>	42 + 27	69
5	<i>Corsa</i>	<i>vermelho</i>	42 + 6	48
6	<i>Corsa</i>	<i>azul</i>	42 + 21	63
7	<i>Palio</i>	<i>verde</i>	21 + 27	48
8	<i>Palio</i>	<i>vermelho</i>	21 + 6	27
9	<i>Palio</i>	<i>azul</i>	21 + 21	42
10	<i>Celta</i>	<i>verde</i>	63 + 27	90
11	<i>Celta</i>	<i>vermelho</i>	63 + 6	69
12	<i>Celta</i>	<i>azul</i>	63 + 21	84

Em Srinivasan e Park (1997), por exemplo, encontra-se um modelo de composição que incorpora uma informação adicional, que pode ser facilmente coletada, e que se refere aos atributos e níveis que o entrevistado considera “inaceitáveis”.

Modelos para dados coletados sob método de decomposição

Nos casos em que o método adotado é o de decomposição, os modelos dependem da forma de avaliação utilizada (atribuição de nota, soma-constante etc) e da natureza dos níveis dos atributos, pois estes podem ser quantitativos ou qualitativos.

Quando os atributos são quantitativos, temos, por exemplo, o *modelo vetor* (Carrol, 1972), que é um modelo linear e pressupõe que o valor da preferência aumenta ou diminui linearmente conforme o nível do atributo. Suponha que o produto seja *televisor* e que o atributo seja *resolução da imagem*. É intuitivo supor que quanto melhor a resolução, maior será a preferência pelo produto. O primeiro gráfico da Figura 1.2 ilustra essa situação.

Figura 1.2: Exemplo de modelos de Análise de Preferência Conjunta

Um outro exemplo, para atributos com níveis quantitativos, é o *modelo ponto-ideal* (Green e Carmone, 1970), que assume que, para cada indivíduo, existe um estímulo de preferência “máxima” (ideal) e que, quanto maior a distância até esse estímulo ideal, menor será a preferência. Como exemplo dessa situação, suponha que o produto seja *telefone celular*. Não seria estranho supor que o atributo *peso do aparelho* obedecesse esta relação (veja segundo gráfico da Figura 1.2).

Um outro modelo relacionado ao método de decomposição é o *modelo utilidade*, no qual os coeficientes (parâmetros) que o definem equivalem às utilidades dos níveis dos atributos. Como o

modelo bayesiano de Marshall e Bradlow, que será desenvolvido nesta dissertação, é baseado neste modelo, apresentamos a seguir as equações que o caracterizam.

O modelo utilidade é aplicável a atributos qualitativos e nada mais é do que um modelo de regressão linear (Neter et al., 1996), que considera os atributos como os fatores em um modelo de análise de variâncias (ANOVA) e as avaliações dadas aos estímulos (tratamentos) como variável resposta, ou seja, a “preferência” é considerada a variável dependente. Assumimos, assim, que a preferência é composta por contribuições parciais associadas aos níveis dos atributos deste produto. Note que, quando o atributo é de natureza quantitativa, ele pode ser discretizado.

Em um modelo de ANOVA, os coeficientes, que caracterizam os níveis dos atributos, devem ser submetidos a algumas restrições para garantir a sua identificabilidade, como por exemplo: médias de casela, casela de referência ou efeitos de fator (Neter et al., 1996).

No exemplo do *automóvel*, definimos 2 atributos com 4 e 3 níveis cada. Assumindo que não há interação entre os níveis dos fatores, poderíamos considerar um modelo utilidade sob restrições de efeito de fator da seguinte forma

$$y_{ijk} = \alpha_i + \gamma_{ij} + \delta_{ik} + \epsilon_{ijk}, \quad (1.1)$$

$$i = 1, 2, \dots, n; \quad j = 1, 2, 3, 4 \text{ e } k = 1, 2, 3$$

sob as restrições

$$\sum_{j=1}^4 \gamma_{ij} = 0, \text{ e } \sum_{k=1}^3 \delta_{ik} = 0, \text{ para } i = 1, \dots, n$$

no qual,

y_{ijk} é a preferência do indivíduo i associada ao estímulo de nível j do atributo *Modelo* e nível k do atributo *Cor*;

α_i a preferência média referente ao indivíduo i ;

γ_{ij} variação na média da preferência do indivíduo i , atribuída ao nível j do atributo *Modelo*, fixado o nível do atributo *Cor*;

δ_{ik} a variação na média da preferência do indivíduo i , atribuída ao nível k do atributo *Cor*, fixado o nível do atributo *Modelo*;

ϵ_{ijk} são erros aleatórios independentes com distribuição $N(0; \sigma^2)$, onde σ^2 é a sua variância e n número de indivíduos.

Note que γ_{ij} e δ_{ik} expressam as preferências dadas pelo indivíduo i aos níveis j e k dos atributos *Modelo* e *Cor*, respectivamente. Elas são as contribuições parciais que chamamos de utilidades dos respectivos níveis.

Este modelo pode ser expresso matricialmente. Para isso, considere q o número de estímulos possíveis obtidos com estes dois atributos, neste exemplo, $q = 4 \times 3 = 12$ e seja p o número de parâmetros a serem estimados sob a restrição de identificabilidade adotada, neste caso, $p = (4 - 1) + (3 - 1) + 1 = 6$. A equação matricial fica dada por

$$y_i = X\mu_i + \epsilon_i, \quad i = 1, 2, \dots, n \quad (1.2)$$

na qual,

y_i é um vetor ($q \times 1$) com as avaliações feitas para cada um dos estímulos, dadas pelo indivíduo i ;
 X uma matriz de planejamento ($q \times p$) que define a combinação de níveis dos estímulos avaliados;
 μ_i um vetor ($p \times 1$) com os parâmetros (utilidades) a serem estimados;
 ϵ_i um vetor ($q \times 1$) com os erros aleatórios atribuídos para cada y_{ijk} , tal que, $\epsilon_i \sim N_q(0; \sigma^2 I)$ independentes e
 I uma matriz identidade de dimensão q .

Os parâmetros a serem estimados, sob a restrição adotada, neste exemplo, são

$$\mu_i = \begin{bmatrix} \alpha_i \\ \gamma_{i1} \\ \gamma_{i2} \\ \gamma_{i3} \\ \delta_{i1} \\ \delta_{i2} \end{bmatrix}, \quad i = 1, \dots, n \quad (1.3)$$

sendo que, μ_i ($p \times 1$) é chamado vetor de utilidades do indivíduo i .

A matriz X é uma matriz de planejamento fatorial ($q \times p$) que define os estímulos avaliados. Neste exemplo, ela identifica os estímulos dados pela Tabela 1.1 sob a restrição de efeitos de fator e não considera as interações entre atributos, como segue

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & -1 & -1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & -1 & -1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 0 \\ 1 & -1 & -1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix} \quad (1.4)$$

O número de estímulos a serem avaliados é definido pelo planejamento fatorial adotado e depende da relação que existe entre os atributos (se há ou não interações entre eles). Porém, conforme o número de atributos e níveis envolvidos aumenta, maior será o número mínimo de estímulos a serem avaliados pelo entrevistado. Para aliviar essa tarefa, técnicas de redução do número de tratamentos como planejamento em blocos, blocos incompletos ou fatoriais fracionários, são utilizados, da mesma forma que são aplicados em delineamentos de modelos de regressão utilizados em ANOVA (Neter et al., 1996).

Modelos para dados coletados sob o método misto

Quando há informações provenientes de métodos distintos, o modelo é desenvolvido de maneira que agregue estas duas informações e é isso que os diferencia.

Dentre os modelos mistos (híbridos) temos os que trabalham com uma combinação linear dessas duas informações obtidas pelos dois métodos, estimando tanto coeficientes individuais como coeficientes gerais da amostra (Green, Goldberg e Montemayor, 1981 e Green e Krieger, 1996). A informação obtida na fase de composição é incorporada como se fosse uma variável adicional no modelo que é estruturado com as informações coletadas pelo método de decomposição.

Temos também o Adaptive Conjoint Analysis (ACA) de Johnson (1987). Este modelo é baseado em programas computacionais iterativos que, a partir das informações obtidas pelo método de composição, selecionam, para cada indivíduo, os estímulos a serem avaliados pelo método de decomposição. Este modelo, em particular, trabalha com estímulos de perfis-parciais ao invés de estímulos de perfil-completo.

Por fim, citamos os modelos bayesianos, que utilizam as informações obtidas na fase de composição como informação *a priori* sobre as contribuições parciais, que, por sua vez, serão estimadas através das avaliações feitas aos estímulos de perfil-completo (veja, por exemplo, Lenk, DeSarbo, Green and Young, 1996 e Allenby, Arora e Ginter, 1995).

O modelo que será estudado nesta dissertação pertence à essa classe de modelos bayesianos e uma de suas vantagens é que apresenta o modelo de composição e o de decomposição (utilidade) como casos particulares.

1.2.5 Apresentação dos estímulos aos entrevistados

Após definir quais estímulos serão avaliados pelos entrevistados, é preciso escolher de que forma estes serão apresentados.

As formas mais comuns de apresentação de estímulos são: a verbal (explicando ou lendo uma descrição do produto, por telefone por exemplo), através de cartões com a descrição do produto e através de representação gráfica (figuras ou fotografias) do mesmo (Green e Wind, 1975). Para facilitar a tarefa do entrevistado, os produtos hipotéticos também podem ser representados por protótipos, construídos especialmente para essa finalidade. Por exemplo, Wind et al. (1989) usavam combinações de protótipos, fotografias e descrições verbais. Um protótipo pode, também, ser substituído pela computação gráfica, caso a entrevista seja feita através da internet (Rao e Hauser, 2002). O tipo de apresentação do estímulo dependerá principalmente do produto, dos recursos disponíveis e do público que será entrevistado.

1.2.6 Métodos de estimação

Nesta seção apresentamos alguns métodos de estimação. Não é o objetivo desta dissertação se aprofundar sobre os métodos de estimação convencionais de modelos de Análise de Preferência

Conjunta. Iremos apenas citar alguns dos métodos mais comuns e faremos as devidas referências.

O método de estimação varia conforme o método de coleta aplicado, o modelo utilizado e a forma de avaliação adotada. Em Hauser e Rao (2002) os autores citam os processos de estimação comuns em Análise de Preferência Conjunta, ao mesmo tempo que introduzem as variações dessa técnica. Na dissertação de Artes (1991), há uma comparação feita entre diversos métodos de estimação.

Estimação para dados coletados sob o método de composição

A estimação das utilidades neste caso é relativamente simples, pois as utilidades são obtidas diretamente através da regra assumida, como, por exemplo, o modelo de composição de Fishbein (1967) que apresentamos na Seção 1.2.4.

Estimação para dados coletados sob o método de decomposição

Para os modelos de preferência do tipo vetor, ponto-ideal ou utilidade, nos quais os atributos podem ser considerados como fatores, a estimação dos parâmetros pode ser feita através do Método de Mínimos Quadrados. Considere o modelo utilidade dado por (1.2), neste caso, o estimador de mínimos quadrados (EQM) do vetor de utilidades é

$$\hat{\mu}_i = [X'X]^{-1}X'y_i \quad (1.5)$$

no qual,

y_i é um vetor ($q \times 1$) com as avaliações feitas para cada um dos estímulos, dadas pelo indivíduo i ;
 X uma matriz de planejamento ($q \times p$) que define a combinação de níveis dos estímulos avaliados e
 μ_i vetor ($p \times 1$) com os parâmetros (utilidades) a serem estimados;

Uma vez estimadas as contribuições parciais dos níveis dos atributos, é possível, através de modelos convenientes, analisar a importância de cada um desses níveis, bem como a importância de cada atributo. Temos a possibilidade de, com esse método, estimar também as interações entre os níveis dos atributos, caso existam.

Nos casos em que as avaliações são realizadas através de ordenamento (atribuição de postos), pode-se utilizar a Análise de Variância Monotônica (MONANOVA), que irá encontrar a transformação

monotônica dos postos que melhor se ajuste ao modelo (veja Kruskal, 1965 e Wittink e Cattin, 1989). Chapman e Staelin (1982) propõem, como estimação dos parâmetros, o método de simulação de Monte Carlo (Hastings, 1970). Em Albert e Chib (1993), também há uma proposta de estimação via métodos de simulação, porém pelo amostrador de Gibbs.

Quando as avaliações são feitas através de comparação pareada, um método de estimação pode ser o *Linear Mapping* (LINMAP). Nele, os estímulos são alocados em um conjunto de pares ordenados formado com as avaliações observadas e o objetivo da análise é encontrar o vetor de utilidades que menos infrinja as restrições estabelecidas pela alocação dos estímulos (veja Artes, 1991). Para casos de avaliação feita por escolha discreta, tanto Albert e Chib (1993) como Allenby, Arora e Ginter (1995), propõem como método de estimação das utilidades o amostrador de Gibbs, que é um método de simulação.

Estimação para dados coletados sob o método misto

Os parâmetros dos modelos mistos podem ser estimados pelo método de mínimos quadrados, utilizando as avaliações obtidas pelo método de composição, por exemplo, como variáveis do modelo utilidade; neste caso, realiza-se a estimação em duas etapas (veja Artes, 1991). No modelo ACA de Johnson (1987), a estimação é feita primeiro com os dados coletados via método de composição, a partir do resultado a estimação dos atributos mais importantes são refinadas com avaliações feitas a estímulos “escolhidos” através do resultado da estimação anterior. Esta estimação é feita com método de mínimos quadrados combinado com abordagem bayesiana, pois “atualiza” a estimativa das utilidades conforme coleta mais informações do entrevistado (veja também, Green e Srinivasan, 1990).

Por fim, citamos a estimação dos parâmetros dos modelos bayesianos. Em geral, a estimação pode ser feita a partir de métodos de simulação, como o método de Metropolis-Hastings (Hastings, 1970) ou o amostrador de Gibbs (veja Apêndice C). O estimador de *James-Stein* é outro exemplo que pode ser utilizado no processo de estimação para estes modelos.

Finalizamos aqui a introdução sobre a Análise de Preferência Conjunta que acreditamos ser suficiente para que o leitor se interesse do assunto e acompanhe o desenvolvimento do modelo que será estudado adiante.

Capítulo 2

Modelo Bayesiano

Neste capítulo, apresentamos o modelo misto bayesiano, proposto por Marshal e Bradlow (2002), que agrega as informações obtidas através dos métodos de composição (avaliações obtidas na fase auto-explicativa) e de decomposição (avaliações de perfil-completo). Apresentamos, também, algumas propriedades deste modelo e o processo de estimação utilizado, baseado em técnicas de simulação.

2.1 Apresentação do Modelo Bayesiano

O Modelo Bayesiano é estruturado para avaliações de estímulos (perfil-completo) feitas por notas e baseia-se no modelo utilidade. Assume-se, também, que não há interação entre os fatores, o que leva à adoção de um modelo aditivo.

Considere, para ilustrar, o exemplo do *automóvel* apresentado na Seção 1.2. Nele foram definidos os atributos *Modelo* e *Cor* (os possíveis estímulos estão apresentados na Tabela 1.1). Supondo que a preferência dada aos estímulos seja representada por notas e que não haja interação entre os níveis dos atributos, considere o modelo utilidade definido por (1.1) e (1.2).

O modelo utiliza a informação obtida pelo método de composição como informação *a priori* sobre o vetor de utilidades μ_i e a informação obtida no método de decomposição, vetor y_i , é utilizada para gerar a função de verossimilhança, $g(y_i|\mu_i)$ (definida pelo modelo utilidade). Trabalhando com os dados dessa maneira, chega-se em um estimador de μ_i , proveniente da função densidade de probabilidade *a posteriori* de μ_i , $f(\mu_i|y_i)$, que conterà toda informação obtida nas duas coletas. Essa estrutura

é conhecida como modelo bayesiano empírico (DeGroot, 1970).

Essa relação é alcançada através do teorema de Bayes da seguinte maneira

$$f(\mu_i|y_i) \propto h(\mu_i)g(y_i|\mu_i)$$

no qual, $h(\mu_i)$ é a função densidade de probabilidade *a priori* para μ_i . O teorema de Bayes e uma breve introdução sobre inferência bayesiana podem ser encontrados, por exemplo, em DeGroot (1970) e Paulino, Turkman e Murteira (2003).

Assumimos que $y_i|\mu_i$ possui distribuição Normal. Se assumirmos que a distribuição *a priori* de μ_i é Normal, estaremos diante de um modelo Hierárquico Normal-Normal (ver Zellner, 1996) e com isso, a distribuição *a posteriori* de μ_i também será dada por uma distribuição Normal. As seções seguintes tratam de cada uma dessas componentes: a distribuição *a priori* $h(\mu_i)$, a verossimilhança $g(y_i|\mu_i)$ e a distribuição *a posteriori* $f(\mu_i|y_i)$.

2.1.1 Distribuição *a priori* de μ_i

Marshall e Bradlow propõem que a distribuição *a priori* do vetor utilidade não observado μ_i , (1.3), seja dada por

$$[\mu_i|u_i, \Omega] \sim N_p(u_i; \Omega) \quad , i=1,2,\dots,n \quad (2.1)$$

sendo que, u_i é o vetor utilidade, $(p \times 1)$, obtido a partir das avaliações feitas na fase de composição desse indivíduo e Ω a matriz $(p \times p)$ de variâncias-covariâncias das utilidades. Inicialmente, assumimos Ω conhecida e igual para todos os indivíduos. Denote por $h(\mu_i)$ a função densidade de probabilidade desta distribuição.

Como exemplo, retome os dados das tabelas 1.6 e 1.7. Temos as avaliações compostas do indivíduo i conforme a Tabela 1.9. O vetor de utilidades observadas, u_i , que utilizamos como informação *a priori* sobre μ_i é obtido a partir dessas avaliações e da matriz de planejamento (1.4) do modelo utilidade adotado. Neste exemplo, para obter u_i , aplicamos a seguinte relação

$$\begin{bmatrix} 97 \\ 76 \\ 91 \\ 69 \\ 48 \\ 63 \\ 48 \\ 27 \\ 42 \\ 90 \\ 69 \\ 84 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & -1 & -1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & -1 & -1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 0 \\ 1 & -1 & -1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} u_{i1} \\ u_{i2} \\ u_{i3} \\ u_{i4} \\ u_{i5} \\ u_{i6} \end{bmatrix}.$$

Resolvendo o sistema acima, o vetor de utilidades observadas deste indivíduo é dado por

$$u_i = \begin{bmatrix} 67 \\ 21 \\ -7 \\ -28 \\ 9 \\ 12 \end{bmatrix}.$$

Note que, neste exemplo, este vetor foi obtido a partir de avaliações que variam numa escala de 0 a 100.

2.1.2 Função de Verossimilhança

Suponha que os entrevistados tenham avaliado, através de atribuição de notas, um total de m estímulos, sendo que $0 \leq m \leq q$, no qual, q é o número total de possíveis estímulos. Seja y_i o vetor $m \times 1$ com estas avaliações de perfil-completo dadas pelo indivíduo i . Essa informação será expressa pela função de verossimilhança dos dados, denotada por $g(y_i|\mu_i)$.

O Modelo Bayesiano proposto tem como base o modelo utilidade definido em (1.2) e assume que $y_i|\mu_i$ segue a seguinte distribuição

$$[y_i|\mu_i, \sigma^2] \sim N_m(X_i\mu_i; \sigma^2 I_m), \quad i=1,2,\dots,n, \quad (2.2)$$

sendo que, X_i é uma matriz de delineamento ($m \times p$); μ_i é o vetor de utilidades não observadas já definido; σ^2 é a variância residual, igual para todos os indivíduos e, por enquanto, conhecida; e I_m é a matriz identidade ($m \times m$). A função $g(y_i|\mu_i)$ é, então, a função densidade de probabilidade da distribuição (2.2).

A matriz X_i define um subconjunto de estímulos, ou seja, é um subconjunto de linhas da matriz de delineamento utilizada para definir o planejamento fatorial do modelo utilidade adotado. Não precisa ter posto completo (este fato será comentado na Seção 2.4) e pode ser igual para todos os indivíduos, caso todos tenham recebido os mesmos m estímulos para avaliar.

Suponha que, no exemplo do *automóvel*, o indivíduo i tenha avaliado o subconjunto de quatro estímulos que estão destacados na Tabela 2.1.

Tabela 2.1: Estímulos escolhidos para serem avaliados pelo indivíduo i

Estímulo	Modelo	Cor
1	Gol	verde
2	Gol	vermelho
3	Gol	azul
4	Corsa	verde
5	Corsa	vermelho
6	Corsa	azul
7	Palio	verde
8	Palio	vermelho
9	Palio	azul
10	Celta	verde
11	Celta	vermelho
12	Celta	azul

Estes estímulos escolhidos equivalem às seguintes linhas de (1.4) (matriz de planejamento fatorial deste exemplo)

$$X = \begin{bmatrix} \boxed{1} & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & -1 & -1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ \boxed{1} & 0 & 1 & 0 & -1 & -1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ \boxed{1} & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & -1 & -1 \\ \boxed{1} & -1 & -1 & -1 & 1 & 0 \\ 1 & -1 & -1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

ou seja, a matriz X_i , escolhida para este exemplo, é dada por

$$X_i = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & -1 & -1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & -1 & -1 & -1 & 1 & 0 \end{bmatrix}.$$

Suponha que o indivíduo i tenha atribuído notas de 0 a 100 para cada um dos m estímulos que lhe foram apresentados, de acordo com sua preferência. A Tabela 2.2 exemplifica essas notas.

Tabela 2.2: Exemplo de avaliações de estímulos da fase perfil-completo

Estímulo	Modelo	Cor	Nota
1	<i>Gol</i>	<i>verde</i>	100
2	<i>Corsa</i>	<i>azul</i>	40
3	<i>Palio</i>	<i>vermelho</i>	10
4	<i>Celta</i>	<i>verde</i>	90

Neste caso, o vetor y_i , que corresponde às avaliações de perfil-completo dadas ao subconjunto de estímulos apresentado, deste indivíduo é dado por

$$y_i = \begin{bmatrix} 100 \\ 40 \\ 10 \\ 90 \end{bmatrix}.$$

No exemplo do *automóvel*, tanto y_i como as avaliações compostas a partir das informações obtidas na fase auto-explicativa, variam no intervalo de 0 a 100. Além disso, y_i deve conter avaliações feitas por atribuição de notas. Essas restrições de escala e forma de avaliação dos estímulos serão contornadas no Capítulo 3, quando esse modelo for estendido.

Não há restrições sobre o subconjunto de m estímulos, mas este pode ser obtido a partir de algum método de redução do número de “tratamentos”, como um planejamento fatorial fracionário ou um delineamento em blocos (Neter et al., 1996).

2.1.3 Distribuição *a posteriori* de μ_i

Uma vez definidas a função densidade de probabilidade *a priori* de μ_i , e a função de verossimilhança, podemos calcular a função densidade de probabilidade *a posteriori* $f(\mu_i|y_i)$, que irá conter toda informação trazida pelas coletas de dados feitas.

Segundo Zellner (1996), a função densidade de probabilidade *a posteriori* de μ_i , dados σ^2 e Ω (por enquanto são conhecidos), pode ser expressa por

$$[\mu_i|y_i, \sigma^2, \Omega] \sim N_p \left(A_i^{-1} \Omega^{-1} u_i + A_i^{-1} \frac{X_i' y_i}{\sigma^2}; A_i^{-1} \right), \quad i=1,2,\dots,n \quad (2.3)$$

sendo,

$$A_i^{-1} = \left[\Omega^{-1} + \frac{X_i' X_i}{\sigma^2} \right]^{-1}.$$

Veja maiores detalhes sobre a obtenção desta distribuição no Apêndice A.

A distribuição *a posteriori* das utilidades, $f(\mu_i|y_i)$, além de agregar as informações dos dois tipos de coleta (informações obtidas na fase auto-explicativa u_i e avaliações de perfis-completos y_i) num só modelo, também pondera as informações conforme suas matrizes de variâncias-covariâncias.

2.2 Casos particulares

A distribuição *a posteriori* do vetor de utilidades, obtida com o Modelo Bayesiano, é dada por (2.3) e pode ser reescrita como

$$[\mu_i|y_i, \sigma^2, \Omega] \sim N_p \left(\left[\Omega^{-1} + \frac{X_i' X_i}{\sigma^2} \right]^{-1} \left[\Omega^{-1} u_i + \frac{X_i' y_i}{\sigma^2} \right]; \left[\Omega^{-1} + \frac{X_i' X_i}{\sigma^2} \right]^{-1} \right), i=1,2,\dots,n.$$

Admita que a informação obtida pelo método de composição não seja confiável, ou até mesmo, que ela não exista (não tenha sido coletada). Com isso, passamos a ter uma distribuição *a priori* não-informativa para μ_i , ou seja, a matriz de precisão Ω^{-1} tende a zero. Se levarmos esse fato ao extremo, $\Omega^{-1} \rightarrow 0$, chegamos à seguinte situação

$$[\mu_i|y_i, \sigma^2, \Omega] \sim N_p \left(\left[\frac{X_i' X_i}{\sigma^2} \right]^{-1} \left[\frac{X_i' y_i}{\sigma^2} \right]; \left[\frac{X_i' X_i}{\sigma^2} \right]^{-1} \right), i=1,2,\dots,n$$

e, conseqüentemente,

$$[\mu_i|y_i, \sigma^2] \sim N_p \left([X_i' X_i]^{-1} X_i' y_i; [X_i' X_i]^{-1} \sigma^2 \right), i=1,2,\dots,n.$$

Note que, $[X_i' X_i]^{-1} X_i' y_i$ é o estimador de mínimos quadrados de μ_i no caso de uma regressão linear. A variância resultante, $[X_i' X_i]^{-1} \sigma^2$, equivale à variância desse estimador de mínimos quadrados. Ou seja, se a informação é tão vaga a ponto de não ser informativa, ou se realmente não coletamos dados pelo método de composição, o Modelo Bayesiano se reduz ao modelo que seria utilizado no caso de coletas feitas apenas pelo método de decomposição.

Da mesma maneira, se a informação que coletamos via método de decomposição é vaga, então a precisão da função de verossimilhança tenderá a zero. Se chegarmos ao extremo de não coletar as avaliações de perfil-completo, o modelo se reduzirá à abordagem padrão utilizada quando apenas

avaliações da fase auto-explicativa são coletadas (método de composição). Note que, se $\sigma^{-2} \rightarrow 0$, então

$$[\mu_i | y_i, \sigma^2, \Omega] \sim N_p \left(\left[\Omega^{-1} + \frac{X_i' X_i}{\sigma^2} \right]^{-1} \left[\Omega^{-1} u_i + \frac{X_i' y_i}{\sigma^2} \right]; \left[\Omega^{-1} + \frac{X_i' X_i}{\sigma^2} \right]^{-1} \right), i=1,2,\dots,n$$

se reduzirá a

$$[\mu_i | y_i, \sigma^2, \Omega] \sim N_p \left([\Omega^{-1}]^{-1} [\Omega^{-1} u_i]; [\Omega^{-1}]^{-1} \right), i=1,2,\dots,n.$$

Logo,

$$[\mu_i | y_i, \sigma^2, \Omega] \sim N_p (u_i; \Omega), i=1,2,\dots,n.$$

Dessa forma, μ_i será estimado diretamente pelos valores observados u_i . Mostramos aqui que o Modelo Bayesiano pondera as informações ao ponto de se reduzir, em situação extrema, aos modelos particulares dos casos em que apenas um dos métodos de coleta é aplicado.

2.3 Extensão

Na construção da função densidade de probabilidade *a posteriori* de μ_i , consideramos, num primeiro momento, os parâmetros σ^2 e Ω conhecidos.

Para livrar o modelo dessa premissa, assumimos para esses parâmetros distribuições conjugadas naturais, para que a relação obtida (Normal-Normal) não se altere. O modelo que nos confere essa vantagem é o Normal-Wishart (Zelner, 1996). Uma suposição que devemos assumir é a de que σ^2 e Ω sejam independentes.

As distribuições conjugadas naturais (*a priori*) de σ^2 e Ω são, respectivamente, dadas por

$$\sigma^2 \sim I\chi^2_{(g_0)} \quad (2.4)$$

e

$$\Omega \sim IW_p(D_0, d_0) \quad (2.5)$$

na qual, $I\chi^2_{(g_0)}$ representa a distribuição Qui-quadrado Invertida com g_0 graus de liberdade e $IW_p(D_0, d_0)$ a distribuição de Wishart Invertida com d_0 graus de liberdade e parâmetro de escala D_0 ($p \times p$).

Assim como os vetores μ_i , os valores de σ^2 e Ω também serão estimados. Antes de descrever o procedimento de estimação proposto por Marshall e Bradlow para os parâmetros do Modelo Bayesiano, citamos na seção seguinte algumas propriedades deste modelo.

2.4 Propriedades

Quando adotamos o método de perfil-completo, trabalhamos com um modelo análogo ao modelo de regressão linear, que pode ser dado matricialmente por

$$y_i = X_i \mu_i + \epsilon_i, \quad i=1,2,\dots,n$$

sendo que, y_i é um vetor ($m \times 1$) com as avaliações dos estímulos feitas pelo indivíduo i , X_i é uma matriz de delineamento com dimensões ($m \times p$), μ_i é um vetor de coeficientes ($p \times 1$) e ϵ_i são os erros aleatórios que seguem distribuição Normal de média 0 e variância σ^2 , ver Neter et al. (1996).

Adotando o método de mínimos quadrados, temos que a distribuição do estimador de μ_i será dada por

$$[\hat{\mu}_i | \sigma^2] \sim N_p \left(\mu_i; (X_i' X_i)^{-1} \sigma^2 \right), \quad i=1,2,\dots,n.$$

Note que, quando o método de perfil-completo é adotado, a matriz $(X_i' X_i)^{-1}$ precisa existir; logo, X_i deve ter posto completo (Mardia, Kent e Bibby, 1979).

A vantagem do Modelo Bayesiano está no fato de μ_i ser estimado a partir de

$$[\mu_i | y_i, \sigma^2, \Omega] \sim N_p \left(A_i^{-1} \left[\Omega^{-1} u_i + \frac{X_i' y_i}{\sigma^2} \right]; A_i^{-1} \right), \quad i=1,2,\dots,n$$

na, qual

$$A_i^{-1} = \left[\Omega^{-1} + \frac{X_i' X_i}{\sigma^2} \right]^{-1}$$

sendo que, neste caso, a matriz A_i^{-1} pode existir mesmo que X_i não tenha posto completo (Lenk, Desarbo, Green e Young, 1996).

Dessa maneira, um número bem reduzido de estímulos pode ser avaliado pelo indivíduo, pois não precisamos coletar o mínimo necessário de observações para que X_i tenha posto completo (Marshall e Bradlow, 2002). Na aplicação desenvolvida no Capítulo 5 exemplificamos esse fato.

Uma outra propriedade interessante se refere à precisão das matrizes de variâncias-covariâncias dos estimadores de μ_i segundo os diferentes métodos de coleta dos dados. Temos que a matriz de variâncias-covariâncias do estimador de μ_i referente ao Modelo Bayesiano é menor¹ que as matrizes de variâncias-covariâncias dos estimadores de μ_i quando estes são obtidos pelos métodos de composição ou decomposição. Dessa forma, obtemos estimativas mais precisas, no sentido de apresentar “menor” matriz de variâncias-covariâncias, quando optamos pelo Modelo Bayesiano. Provamos esta propriedade na Seção A.4 do Apêndice A.

2.5 Estimação dos parâmetros

O método de estimação sugerido por Marshall e Bradlow é o Método de Monte Carlo via Cadeias de Markov (MCMC), mais especificamente, o amostrador de Gibbs, que é um caso particular de MCMC. Esse procedimento se baseia na geração sucessiva de valores aleatórios a partir de distribuições condicionais completas que compõem a distribuição conjunta de interesse (veja Apêndice C.2).

No Modelo Bayesiano, nos interessa a distribuição *a posteriori* de μ_i , que será o estimador bayesiano do vetor de utilidades de cada indivíduo (DeGroot, 1970). Uma vez observados os vetores y_i e u_i , precisamos amostrar valores das seguintes distribuições condicionais completas:

$$\begin{aligned} &[\mu_i | y, u, \sigma^2, \Omega] \\ &[\sigma^2 | y, u, \mu, \Omega] \\ &[\Omega | y, u, \mu, \sigma^2] \end{aligned} \tag{2.6}$$

nas quais, $y = (y'_1, \dots, y'_n)'$, $u = (u'_1, \dots, u'_n)'$ e $\mu = (\mu'_i, \dots, \mu'_n)'$.

Os vetores $[\mu_i | y_i, u_i, \sigma^2, \Omega]$ seguem a distribuição já definida em (2.3) e as distribuições condicionais completas de σ^2 e Ω são dadas por distribuições Wishart Invertida, como segue

¹Assumimos que uma matriz é “menor” que outra, ($A \leq B$, se a diferença entre elas, $(B - A)$, é uma matriz positiva semi-definida.

$$[\sigma^2|y, \mu, \Omega] \sim IW_1(g, G), \quad (2.7)$$

$$g = g_0 + nm \text{ e } G = G_0 + \sum_{i=1}^n (y_i - X_i \mu_i)' (y_i - X_i \mu_i)$$

e

$$[\Omega|y, \mu, \sigma^2] \sim IW_p(d, D), \quad (2.8)$$

$$d = d_0 + n \text{ e } D = D_0 + \sum_{i=1}^n (\mu_i - u_i) (\mu_i - u_i)'$$

sendo que, g e d são os graus de liberdade. G e D são matrizes de escala.

As distribuições condicionais *a posteriori* de σ^2 e Ω são também Wishart Invertidas, pois foi assumida uma estrutura hierárquica Normal-Wishart (Zellner, 1996). A obtenção das distribuições condicionais completas está presente no Apêndice A, Seção A.2. Para que estas distribuições sejam não-informativas pode-se tomar $g_0 = 1$, $G_0 = 0$, $d_0 = 1$ e $D_0 = 0$.

Neste método de estimação, partimos de valores iniciais para os parâmetros $\mu_i^{(0)}$, $\sigma^{2(0)}$ e $\Omega^{(0)}$, e, a cada iteração t , $t = 1, 2, \dots, T$, geramos realizações dos parâmetros a partir de suas distribuições condicionais, atualizadas a cada passo. Por exemplo, na iteração de número $t + 1$, geramos valores a partir de

$$[\mu_i|y_i, u_i, \sigma^{2(t)}, \Omega^{(t)}]$$

$$[\sigma^2|y, u, \mu^{(t)}, \Omega^{(t)}]$$

$$[\Omega|y, u, \mu^{(t)}, \sigma^{2(t)}]$$

Existem várias maneiras de se conduzir uma estimação através de geração de Cadeias de Markov. Em Marshall e Bradlow (2002), os autores sugerem a utilização de múltiplas e independentes cadeias com valores iniciais distintos, veja Gelman e Rubin (1992). O número total de realizações de cada cadeia, T , deve ser suficiente para que haja convergência da mesma. Além de cada cadeia ter de convergir, as cadeias geradas devem coincidir nessa convergência. O método de diagnóstico de convergência de cadeias de Markov, escolhido pelos autores, é o de Gelman e Rubin (1992), apresentado no Apêndice C, seção C.3. Os valores gerados após a convergência das cadeias são utilizados para inferir os parâmetros.

A inferência é feita pelo método de Monte Carlo, que permite estimar o valor esperado dos parâmetros através de sua aproximação de integrais (veja Apêndice C.1). As médias das distribuições *a posteriori* dos vetores μ_i correspondem aos respectivos estimadores do Modelo Bayesiano.

Suponha que tenham sido construídas c cadeias independentes de Markov e que as t últimas iterações de cada uma sejam consideradas para compor as estimativas. O estimador do vetor de utilidades do indivíduo i , μ_i , neste caso, será dado por

$$\hat{\mu}_i = \frac{1}{ct} \sum_{j=1}^c \sum_{k=1}^t \mu_i^{(j,k)} \quad \text{ou} \quad \hat{\mu}_i = \frac{1}{ct} \sum_{j=1}^c \sum_{k=1}^t E \left[\mu_i | y_i, u_i, \sigma^{2(j,k)}, \Omega^{(j,k)} \right]$$

sendo que, o indexador (j, k) indica que o valor ou vetor gerado pertence à k -ésima iteração da j -ésima cadeia.

O programa *WinBugs*² traz a implementação computacional do amostrador de Gibbs. Uma introdução sobre este amostrador de Gibbs pode ser vista no Apêndice C e os passos utilizados na utilização do programa *WinBugs* são apresentados no Apêndice G.

²<http://www.mrc-bsu.cam.ac.uk/bugs>

Capítulo 3

Extensões do Modelo Bayesiano

O modelo descrito na Seção 2.1 assume que as avaliações tenham sido realizadas por atribuição de notas, e que μ_i e u_i estejam na mesma escala. Como vimos na Seção 1.2.2, as avaliações podem ser feitas de diversas maneiras e assumindo escalas diferentes. Marshall e Bradlow (2002), propõem algumas adaptações deste modelo a essas outras formas de mensuração, tomando sempre o Modelo Bayesiano como base; ainda neste artigo, Marshall e Bradlow aplicam as adaptações do modelo a dados reais. Nas seções seguintes apresentamos essas adaptações e o método de estimação utilizado em cada caso.

3.1 Avaliações obtidas sob escalas distintas

Admita que as avaliações de perfil-completo e as obtidas pelo método de composição tenham sido realizadas através de atribuição de notas, porém em escalas distintas. Podemos, por exemplo, ter um caso no qual as avaliações de perfil-completo tenham sido feitas com uma escala entre 0 e 100; e que, na fase auto-explicativa, notas de 1 a 5 tenham sido dadas aos níveis dos atributos e um total de 2 pontos tenha sido distribuído entre estes atributos, logo, temos avaliações, da fase auto-explicativa (compostas), variando no intervalo de 2 a 10. Devido à diferença entre as escalas, não podemos utilizar o vetor u_i , obtido com as avaliações da fase auto-explicativa como informação *a priori* sobre μ_i .

Para adequar a informação trazida pelo vetor u_i , introduzimos um vetor, α , que irá transformar a escala deste vetor à mesma do vetor μ_i . O Modelo Bayesiano é, então, adaptado da seguinte forma

$$[\mu_i|Z_i, \alpha, \Omega] \sim N_p(Z_i\alpha; \Omega) \quad , i=1,2,\dots,n \quad (3.1)$$

no qual, $Z_i = (l, u_i)$, l é um vetor $(p \times 1)$ com todos os elementos iguais a 1 e α é um vetor (2×1) de coeficientes desconhecidos.

Assuma, por exemplo, que as avaliações dos estímulos de perfil-completo tenham sido obtidas como descrito na Tabela 2.2, ou seja, variam entre 0 e 100, e que, na fase auto-explicativa, os indivíduos tenham dado notas de 1 a 5 aos níveis dos atributos *Modelo* e *Cor*, como na Tabela 3.1. Além disso, tenham distribuído, dentre os atributos, um total de 2 pontos, segundo a importância dada a cada um deles, Tabela 3.2.

Tabela 3.1: Exemplo de avaliação feita aos níveis na fase auto-explicativa - alteração na escala

Modelo	Nota	Cor	Nota
<i>Gol</i>	5	<i>Verde</i>	4,5
<i>Corsa</i>	3	<i>Vermelho</i>	1
<i>Palio</i>	1,5	<i>Azul</i>	3,5
<i>Celta</i>	4,5		

Tabela 3.2: Exemplo de avaliação feita aos atributos na fase auto-explicativa - alteração na escala

Atributo	Pontos
<i>Modelo</i>	1,2
<i>Cor</i>	0,8
Total de pontos	2

Após compor as notas referentes a cada um dos possíveis estímulos hipotéticos, obtemos notas variando entre 2 e 10. As notas compostas deste entrevistado são apresentadas na Tabela 3.3.

Tabela 3.3: Exemplo de composição de avaliações de estímulos da fase auto-explicativa - alteração na escala

Estímulo	Modelo	Cor	Composição	Avaliação
1	<i>Gol</i>	<i>verde</i>	6 + 3,6	9,6
2	<i>Gol</i>	<i>vermelho</i>	6 + 0,8	6,8
3	<i>Gol</i>	<i>azul</i>	6 + 2,8	8,8
4	<i>Corsa</i>	<i>verde</i>	3,6 + 3,6	7,2
5	<i>Corsa</i>	<i>vermelho</i>	3,6 + 0,8	4,4
6	<i>Corsa</i>	<i>azul</i>	3,6 + 2,8	6,4
7	<i>Palio</i>	<i>verde</i>	1,8 + 3,6	5,4
8	<i>Palio</i>	<i>vermelho</i>	1,8 + 0,8	2,6
9	<i>Palio</i>	<i>azul</i>	1,8 + 2,8	4,6
10	<i>Celta</i>	<i>verde</i>	5,4 + 3,6	9
11	<i>Celta</i>	<i>vermelho</i>	5,4 + 0,8	6,8
12	<i>Celta</i>	<i>azul</i>	5,4 + 2,8	8,2

O vetor u_i obtido com essas avaliações é

$$u_i = \begin{bmatrix} 6,6 \\ 1,8 \\ -0,6 \\ -2,4 \\ 1,2 \\ -1,6 \end{bmatrix}.$$

Para igualar as escalas, trabalhamos com os vetores Z_i e α , dados respectivamente por

$$Z_i = \begin{bmatrix} 1 & 6,6 \\ 1 & 1,8 \\ 1 & -0,6 \\ 1 & -2,4 \\ 1 & 1,2 \\ 1 & -1,6 \end{bmatrix} \quad e \quad \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}.$$

Note que α é o mesmo para todos os indivíduos, pois estamos assumindo que os indivíduos tenham recebido as mesmas instruções no momento da coleta dos dados.

As distribuições condicionais dadas em (2.6) ainda são válidas; apenas devemos substituir u_i por $Z_i\alpha$, que nada mais é que a informação u_i na mesma da avaliação feita de perfil-completo. As distribuições que eram condicionadas a u_i agora devem ser condicionadas a Z_i e α .

Assumimos que α é um vetor aleatório com distribuição Normal; com isso, temos novamente uma relação do tipo hierárquica Normal-Normal. Ou seja, assumindo distribuição Normal *a priori* para α , teremos *a posteriori* de α também seguindo distribuição Normal. Dessa maneira, tome

$$\alpha \sim N_2(\alpha_0; \Sigma_0). \quad (3.2)$$

A distribuição *a posteriori* de α (ver Zellner, 1996) será

$$[\alpha|\mu, \Omega] \sim N_2 \left(B^{-1}\Sigma_0^{-1}\alpha_0 + B^{-1} \left(\sum_{i=1}^n Z_i' \Omega^{-1} \mu_i \right); B^{-1} \right)$$

sendo,

$$B^{-1} = \left[\Sigma_0^{-1} + \sum_{i=1}^n Z_i' \Omega^{-1} Z_i \right]^{-1}.$$

Devemos incluir essa distribuição condicional no amostrador de Gibbs. Ou seja, a cada iteração, devemos gerar valores dela e das distribuições condicionais completas de μ_i , σ^2 e Ω , já definidas, substituindo apenas u_i por $Z_i\alpha$. No processo de estimação, para obter a iteração $t + 1$, geramos valores de

$$[\mu_i|y_i, Z_i, \sigma^{2(t)}, \Omega^{(t)}, \alpha^{(t)}]$$

$$[\sigma^2|y, Z, \mu^{(t)}, \Omega^{(t)}, \alpha^{(t)}]$$

$$[\Omega|y, Z, \mu^{(t)}, \sigma^{2(t)}, \alpha^{(t)}]$$

$$[\alpha|y, Z, \mu^{(t)}, \sigma^{2(t)}, \Omega^{(t)}]$$

nas quais, $y = (y'_1, \dots, y'_n)'$, $Z = (Z'_1, \dots, Z'_n)'$ e $\mu = (\mu'_i, \dots, \mu'_n)'$.

Com isso, podemos obter também uma sequência de valores $\alpha^{(0)} \dots \alpha^{(T)}$, bem como para os demais parâmetros, sendo T o total de iterações realizadas. Vale lembrar que, da mesma maneira que fizemos para os demais parâmetros, devemos gerar valores iniciais para o vetor $\alpha^{(0)}$ para cada uma das cadeias de Markov construídas.

3.2 Avaliações da fase de decomposição feitas por soma-constante

Até o momento, tratamos de casos nos quais os entrevistados dão notas aos estímulos que lhes são apresentados. Agora suponha que o entrevistado atribua um certo total de pontos a esses produtos hipotéticos. É razoável pensar que essa tarefa seja simples. Porém, isso só é verdade se o número de estímulos a serem avaliados, m , for pequeno. Se esse número for, por exemplo, 8, a tarefa já fica complicada.

Uma alternativa utilizada para facilitar este trabalho, é apresentar os estímulos em blocos, e pedir que a soma dos pontos seja respeitada dentro de cada um desses blocos. Estes mesmos 8 estímulos podem ser apresentados, por exemplo, em 2 blocos com 4 estímulos cada, como mostra a Tabela 3.4.

Tabela 3.4: Exemplo de apresentação de estímulos em blocos com avaliação feita por soma-constante

Bloco	Estímulo	Pontos	Bloco	Estímulo	Pontos
1	<i>Gol verde</i>	20	2	<i>Palio verde</i>	50
1	<i>Corsa azul</i>	10	2	<i>Corsa vermelho</i>	15
1	<i>Palio vermelho</i>	40	2	<i>Celta azul</i>	30
1	<i>Celta verde</i>	30	2	<i>Gol azul</i>	5
Total de pontos		100	Total de pontos		100

Assuma, então, que os m estímulos hipotéticos a serem avaliados (avaliação de perfil-completo) pelos entrevistados lhes sejam apresentados em b blocos com r estímulos cada ($m = b \times r$), e que, y_{ijk} seja a medida da preferência ao estímulo k do j -ésimo bloco e referente ao indivíduo i , que não observamos diretamente.

Se a avaliação observada é dada por soma-constante, a informação que temos sobre cada estímulo, pertencente a um determinado bloco, é o quanto ele é preferido com relação aos demais estímulos presentes neste mesmo bloco.

Trabalhamos com essa informação na forma percentual, ou seja, calculamos a porcentagem que cada estímulo representa na soma total do seu bloco, e esta informação é denotada por y_{ijk}^* . Por exemplo, através dos pontos distribuídos pelo indivíduo i aos estímulos que lhe foram apresentados no bloco 1, chegamos aos valores y_{ijk}^* dados pela Tabela 3.5.

Tabela 3.5: Distribuição das notas dadas na Tabela 3.4

Bloco	Estímulo	Pontos	Informação
1	<i>Gol verde</i>	20	$y_{i11}^* = 0,2$
1	<i>Corsa azul</i>	10	$y_{i12}^* = 0,1$
1	<i>Palio vermelho</i>	40	$y_{i13}^* = 0,4$
1	<i>Celta verde</i>	30	$y_{i14}^* = 0,3$
Total de pontos		100	$\sum_{k=1}^r y_{i1k}^* = 1$

Bloco	Estímulo	Pontos	Informação
2	<i>Palio verde</i>	50	$y_{i21}^* = 0,5$
2	<i>Corsa vermelho</i>	15	$y_{i22}^* = 0,15$
2	<i>Celta azul</i>	30	$y_{i23}^* = 0,3$
2	<i>Gol azul</i>	5	$y_{i24}^* = 0,05$
Total de pontos		100	$\sum_{k=1}^r y_{i2k}^* = 1$

Dessa maneira, assumimos que as informações observadas, y_{ijk}^* , obedecem à uma relação do tipo

$$y_{ijk}^* = \frac{y_{ijk}}{\sum_{k=1}^r y_{ijk}}, \quad i = 1, 2, \dots, n, j = 1, \dots, b \text{ e } k = 1, \dots, r \quad (3.3)$$

sendo y_{ijk} as medidas de preferência que não observamos diretamente, contudo, quanto maior y_{ijk} , maior a preferência pelo estímulo.

Para a obtenção de y_{ijk} , os autores utilizam um algoritmo baseado em Tanner e Wong (1987). Basicamente, em cada passo, é necessária a geração do valor da soma $\sum_{k=1}^r y_{ijk}$ para cada bloco. Para tanto, valores $\tilde{y}_{ijk}$ são obtidos a partir de (2.2). Em seguida, aplica-se a relação (3.3) em $\sum_{k=1}^r \tilde{y}_{ijk}$, para a obtenção dos valores y_{ijk} . Dessa forma, estaremos utilizando a informação disponível y_{ijk}^* .

As demais distribuições, que são condicionadas inclusive aos valores y_{ijk} , não precisam ser alteradas e, a cada iteração, temos estes valores de y_{ijk} para completá-las. No momento da estimação dos parâmetros pelo amostrador de Gibbs, a $(t+1)$ -ésima iteração será composta pelos valores gerados por

$$\begin{aligned}
& \left[y_i | y_i^*, u_i, \mu_i^{(t)}, \sigma^{2(t)}, \Omega^{(t)} \right] \\
& \left[\mu_i | y_i^{(t)}, y_i^*, u_i, \sigma^{2(t)}, \Omega^{(t)} \right] \\
& \left[\sigma^2 | y^{(t)}, y^*, u, \mu^{(t)}, \Omega^{(t)} \right] \\
& \left[\Omega | y^{(t)}, y^*, u, \mu^{(t)}, \sigma^{2(t)} \right]
\end{aligned} \tag{3.4}$$

sendo que, $y^* = (y_1^*, \dots, y_n^*)'$.

Para obter simulações da distribuição dada pela primeira linha de (3.4), basta seguir os seguintes passos:

- (1) Gerar os valores $\tilde{y}_{ijk}^{(t)}$ a partir de (2.2) atualizada com os valores gerados na iteração t

$$[y_i | \mu_i^{(t)}, \sigma^{2(t)}] \sim N_m \left(X_i \mu_i^{(t)}; \sigma^{2(t)} I_m \right), i = 1, 2, \dots, n. \tag{3.5}$$

- (2) Calcular, para cada indivíduo, as r somas que se referem a cada bloco de estímulos a ele apresentado

$$\sum_{k=1}^r \tilde{y}_{ijk}^{(t)}, \quad i = 1, 2, \dots, n; \quad j = 1, \dots, b \text{ e } k = 1, \dots, r.$$

- (3) Calcular os $y^{(t)}$ utilizando estas somas e os valores observados y_{ijk}^* , através da relação (3.3).

Voltando ao exemplo do *automóvel*, suponha que os entrevistados tenham recebido 2 blocos com 4 estímulos cada e que os tenham avaliado distribuindo 100 pontos por bloco. Assuma as notas dadas pelo indivíduo i dispostas na Tabela 3.5.

Suponha que, na t -ésima iteração do método de estimação tenhamos gerado, a partir de (3.5), as seguintes somas para os blocos 1 e 2 do indivíduo i :

$$\sum_{k=1}^4 \tilde{y}_{i1k}^{(t)} = 240 \quad \text{e} \quad \sum_{k=1}^4 \tilde{y}_{i2k}^{(t)} = 320$$

então, a partir de (3.3), podemos calcular os valores de $y_{ijk}^{(t)}$. Por exemplo, o primeiro valor do bloco 1, $y_{i11}^{(t)}$, é obtido como segue

$$y_{i11}^* = \frac{y_{i11}^{(t)}}{\sum_{k=1}^4 \tilde{y}_{i1k}^{(t)}} \Rightarrow 0,2 = \frac{y_{i11}^{(t)}}{240} \Rightarrow y_{i11}^{(t)} = 48.$$

São esses valores, calculados a cada iteração, que irão completar as distribuições condicionais (3.4).

Vale lembrar que os elementos gerados para compor cada uma das somas $\sum_{k=1}^r \tilde{y}_{ijk}$ devem ser positivos. Essa restrição é satisfeita rejeitando-se os valores gerados que não atendam a essa condição.

Uma última ressalva está relacionada com a variância dos componentes não observados (y_{ijk}), que não está completamente identificada. Então, impomos a restrição de σ^2 ser igual a 1.

3.3 Avaliações da fase de decomposição feitas por ordenamento

Mais uma vez, assuma que os m estímulos (perfil-completo) a serem avaliados pelos entrevistados sejam apresentados em b blocos com r estímulos cada ($m = b \times r$). Suponha, ainda, que o entrevistado tenha avaliado os estímulos ordenando-os conforme sua preferência por cada um deles, e que essa ordenação tenha sido feita em cada um dos blocos. Em outras palavras, suponha que o entrevistado tenha atribuído, por bloco, postos aos estímulos.

Como as avaliações que são observadas são postos dentro de um bloco, a informação trazida por elas é a relação que existe entre os estímulos que pertencem a um mesmo bloco, ou seja, sabemos apenas a ordem de preferência dada pelo entrevistado dentro de um mesmo bloco. Denotamos por y_{ilk}^* os postos atribuídos. Suponha que o indivíduo i tenha recebido 2 blocos de 4 estímulos cada, e que os postos por ele atribuídos sejam dados pela Tabela 3.6.

Tabela 3.6: Exemplo de apresentação de estímulos em blocos com avaliação feita por ordenamento

Bloco	Estímulo	Posto	Bloco	Estímulo	Posto
1	<i>Gol verde</i>	$y_{i13}^* = 3$	2	<i>Palio verde</i>	$y_{i21}^* = 1$
1	<i>Corsa azul</i>	$y_{i14}^* = 4$	2	<i>Corsa vermelho</i>	$y_{i23}^* = 3$
1	<i>Palio vermelho</i>	$y_{i11}^* = 1$	2	<i>Celta azul</i>	$y_{i22}^* = 2$
1	<i>Celta verde</i>	$y_{i12}^* = 2$	2	<i>Gol azul</i>	$y_{i24}^* = 4$

Podemos entender que as medidas de preferência, y_{ijk} , que não observamos diretamente, devem obedecer, em cada bloco, essa relação definida pelos postos observados (y_{ijk}^*).

O Modelo Bayesiano e o amostrador de Gibbs podem ser utilizados pois, dados y_{ijk} , as distribuições condicionais dos parâmetros, (2.6), não utilizam y_{ijk}^* diretamente; logo, basta obtermos, em

cada iteração, os valores y_{ijk} condicionados aos valores observados y_{ijk}^* .

Para gerar valores de y_{ijk} dada a informação y_{ijk}^* , primeiro geramos, para cada bloco, o valor y_{ijr} a partir de (2.2), que corresponde ao estímulo de maior posto; assumimos, então, esse valor gerado como sendo a menor nota presente neste bloco. Depois, geramos, em cada bloco, realizações para o segundo menor valor, $y_{ij(r-1)}$, também através de (2.2), mas restritas à condição de serem maior que o valor anterior. Repete-se esse processo até que os r valores de cada bloco sejam simulados. A restrição imposta pode ser atendida rejeitando-se os valores que não a obedecem.

No processo de simulação, a cada iteração t , geram-se os valores de y_{ijk} , que chamamos de $y_{ijk}^{(t)}$ e os demais valores das distribuições condicionais dos parâmetros do modelo.

Por exemplo, os valores que constituirão a $(t + 1)$ -ésima iteração são gerados por

$$\begin{aligned} & \left[y_i | y_i^*, u_i, \mu_i^{(t)}, \sigma^{2(t)}, \Omega^{(t)} \right] \\ & \left[\mu_i | y_i^{(t)}, y^*, u_i, \sigma^{2(t)}, \Omega^{(t)} \right] \\ & \left[\sigma^2 | y^{(t)}, y^*, u, \mu^{(t)}, \Omega^{(t)} \right] \\ & \left[\Omega | y^{(t)}, y^*, u, \mu^{(t)}, \sigma^{2(t)} \right] \end{aligned} \quad (3.6)$$

sendo que, por indivíduo e bloco, a simulação da primeira linha de (3.6) consiste nos seguintes passos:

(1) Gerar o valor mínimo $y_{ijr}^{(t+1)}$, utilizando a distribuição *a posteriori* definida em (2.2) atualizada com os valores gerados na iteração t ,

$$\left[y_{ijr} | \mu_i^{(t)}, \sigma^{2(t)} \right] \sim N \left(X_{ij,r} \mu_i^{(t)}; \sigma^{2(t)} \right) \quad (3.7)$$

sendo que, $X_{ij,r}$ corresponde à linha da matriz X_i relativa ao estímulo que recebeu maior posto, segundo o indivíduo i . Note que, por indivíduo e bloco, os valores $y_{ijk}^{(t+1)}$ serão gerados separadamente. Isso é possível pois assumimos, no Modelo Bayesiano, que a matriz de variâncias-covariâncias da distribuição Normal multivariada de y_i é diagonal, logo, temos que os elementos de y_i são independentes entre si.

(2) Gerar o segundo menor valor $y_{ij(r-1)}^{(t+1)}$, utilizando, a distribuição *a posteriori* definida em (2.2), também atualizada com os valores da iteração t , (3.7), porém este valor deve ser maior que o último valor gerado para este bloco (neste caso o valor mínimo), pois sabemos que este estímulo deve receber

uma nota maior.

$$\left[y_{ij(r-1)} | \mu_i^{(t)}, \sigma^{2(t)} \right] \sim N \left(X_{i,j,(r-1)} \mu_i^{(t)}; \sigma^{2(t)} \right), \text{ tal que, } y_{ij(r-1)} > y_{ijr}^{(t+1)}$$

na qual, $X_{i,j,(r-1)}$ corresponde à linha da matriz X_i relativa ao estímulo que recebeu o segundo maior posto no bloco j , dado pelo indivíduo i . Se o valor gerado não satisfizer a condição desejada, deve ser descartado, e o processo repetido até que um valor que satisfaça seja gerado (uma outra maneira é utilizando a distribuição de uma Normal truncada, veja Allenby, Arora e Ginter, 1995).

(3) Gerar os demais valores de $y_{ijk}^{(t+1)}$ até chegar ao $y_{ij1}^{(t+1)}$ de cada bloco e indivíduo, utilizando sempre a distribuição *a posteriori* definida em (3.7) truncada inferiormente pelo valor gerado para o estímulo anterior. Note que os valores $y_{ij1}^{(t+1)}$ devem ser os maiores, dentro de cada bloco e para cada indivíduo.

No exemplo da Tabela 3.6, o indivíduo i atribuiu posto 4 ao estímulo “Corsa azul” do bloco 1 e ao estímulo “Gol azul” do bloco 2. Suponha que já tenhamos gerado os valores da t -ésima iteração do processo de estimação dos parâmetros e desejamos gerar os valores da $(t+1)$ -ésima iteração. As linhas da matriz X_i que correspondem a esses dois estímulos são respectivamente

$$X_{i1,4} = \begin{bmatrix} 1 & 0 & 1 & 0 & -1 & -1 \end{bmatrix} \text{ e } X_{i2,4} = \begin{bmatrix} 1 & 1 & 0 & 0 & -1 & -1 \end{bmatrix}$$

Neste caso, devemos, então, gerar valores para $y_{i14}^{(t+1)}$ e $y_{i24}^{(t+1)}$ a partir das seguintes distribuições condicionais

$$\left[y_{i14} | \mu_i^{(t)}, \sigma^{2(t)} \right] \sim N \left(X_{i1,4} \mu_i^{(t)}; \sigma^{2(t)} \right)$$

e

$$\left[y_{i24} | \mu_i^{(t)}, \sigma^{2(t)} \right] \sim N \left(X_{i2,4} \mu_i^{(t)}; \sigma^{2(t)} \right).$$

Sejam estes os números gerados

$$y_{i14}^{(t+1)} = 25 \text{ e } y_{i24}^{(t+1)} = 10.$$

Com isso, o valor $y_{i13}^{(t+1)}$, que se refere ao estímulo “Gol verde”, deve ser gerado a partir da distribuição condicional que segue, truncada no valor $y_{i14}^{(t+1)}$, ou seja, deve ser maior que 25

$$\left[y_{i13} | \mu_i^{(t)}, \sigma^{2(t)} \right] \sim N \left(X_{i1,3} \mu_i^{(t)}; \sigma^{2(t)} \right)$$

sendo

$$X_{i1,3} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Para atender à restrição de ser maior que um certo valor, no caso $y_{i14}^{(t+1)}$, podemos simplesmente gerar valores até que obedeçam a essa condição.

De maneira análoga, geramos o valor de $y_{i23}^{(t+1)}$ a partir de sua distribuição condicional truncada no valor $y_{i24}^{(t+1)}$, ou seja, tal que seja maior que 10.

Gerados os valores $y_{ijk}^{(t+1)}$, o processo continua o mesmo para obtenção dos demais parâmetros.

3.4 Avaliações da fase de decomposição feitas por escolha discreta

Assuma que os m estímulos sejam apresentados em b blocos de r estímulos cada, e suponha que o entrevistado tenha escolhido, em cada bloco, o estímulo de sua maior preferência. Entendemos que essa avaliação nos informa qual estímulo, dentro de cada bloco, recebe a maior medida de preferência (maior nota, por exemplo). Essa informação é denotada por y_{ijk}^* e definida da seguinte maneira

$$y_{ijk}^* = \begin{cases} 1, & \text{se o estímulo } k \text{ é o preferido no bloco } j \\ 0, & \text{caso contrário} \end{cases}, \text{ para } i = 1, 2, \dots, n; j = 1, \dots, b \text{ e } k = 1, \dots, r.$$

Ou seja, o vetor de valores de cada bloco segue distribuição multinomial. A probabilidade de o k -ésimo valor ser igual a um, $P(y_{ijk}^* = 1)$, é modelada pelos autores segundo a abordagem logística de Allenby et al., (1995), que é dada por

$$P(y_{ijk}^* = 1) = \frac{e^{y_{ijk}}}{\sum_{k=1}^r e^{y_{ijk}}}, \quad i = 1, 2, \dots, n; j = 1, \dots, b \text{ e } k = 1, \dots, r \quad (3.8)$$

na qual, y_{ijk} são as medidas de preferência, que não observamos diretamente, mas seguem as especificações dadas pelo Modelo Bayesiano. Este modelo confere, então, ao estímulo com maior nota (maior y_{ijk}) a maior probabilidade de ser o escolhido dentre os estímulos pertencentes ao mesmo bloco, o que é razoável, uma vez que altas notas implicam em maior preferência do entrevistado a este produto hipotético.

Mais uma vez, dados os valores y_{ijk} , podemos utilizar o Modelo Bayesiano, pois as distribuições condicionais dos parâmetros, (2.6), não dependem de y_{ijk}^* .

No processo de estimação de Gibbs, os valores que constituirão a $(t + 1)$ -ésima iteração são gerados por

$$\begin{aligned} & \left[y|y^*, u_i, \mu_i^{(t)}, \sigma^{2(t)}, \Omega^{(t)} \right] \\ & \left[\mu_i|y^{(t)}, y^*, u_i, \sigma^{2(t)}, \Omega^{(t)} \right] \\ & \left[\sigma^2|y^{(t)}, y^*, u_i, \mu_i^{(t)}, \Omega^{(t)} \right] \\ & \left[\Omega|y^{(t)}, y^*, u_i, \mu_i^{(t)}, \sigma^{2(t)} \right] \end{aligned} \quad (3.9)$$

sendo que a simulação de valores da primeira linha de (3.9) consiste nas etapas:

- (1) Gerar, por bloco, a partir da distribuição *a posteriori* de y_i (2.2) atualizada com os valores da iteração t , o valor $y_{ijl}^{(t+1)}$ caso $y_{ijl}^* = 1$;
- (2) Gerar os demais valores de y_{ijk} , $k \neq l$, de cada bloco, também a partir da distribuição *a posteriori* definida pelo Modelo Bayesiano, porém, tais que sejam menores que o valor gerado no passo anterior, pois sabemos que todos os estímulos do bloco que apresentam valor 0 em y_{ijk}^* devem receber nota inferior a y_{ijl} .

Uma alternativa para a primeira linha de (3.9) é

- (1') Gerar os valores correspondentes aos blocos, $y_{ij}^{(t+1)}$, e aceitar apenas os que respeitam a condição dada pelas informações y_{ij}^* correspondentes, ou seja, se no bloco j tivermos $y_{ijl}^* = 1$, então o valor y_{ijl} desse bloco deve ser a maior nota.

Como exemplo, suponha que a informação obtida na coleta de dados da fase de decomposição seja esta da Tabela 3.7.

Suponha que na iteração $(t + 1)$ tenhamos gerado os valores

$$y_{i12}^{(t+1)} = 85 \text{ e } y_{i24}^{(t+1)} = 90$$

a partir de

$$\left[y_{i12}|\mu_i^{(t)}, \sigma^{2(t)} \right] \sim N \left(X_{i1,2}\mu_i^{(t)}; \sigma^{2(t)} \right)$$

e

$$\left[y_{i24}|\mu_i^{(t)}, \sigma^{2(t)} \right] \sim N \left(X_{i2,4}\mu_i^{(t)}; \sigma^{2(t)} \right)$$

Tabela 3.7: Exemplo de apresentação de estímulos em blocos com avaliação feita por escolha discreta

Bloco	Estímulo	Estímulo preferido	Informação
1	<i>Gol verde</i>		$y_{i11}^* = 0$
1	<i>Corsa azul</i>	X	$y_{i12}^* = 1$
1	<i>Palio vermelho</i>		$y_{i13}^* = 0$
1	<i>Celta verde</i>		$y_{i14}^* = 0$
Bloco	Estímulo	Estímulo preferido	Informação
2	<i>Palio verde</i>		$y_{i21}^* = 0$
2	<i>Corsa vermelho</i>		$y_{i22}^* = 0$
2	<i>Celta azul</i>		$y_{i23}^* = 0$
2	<i>Gol azul</i>	X	$y_{i24}^* = 1$

sendo que as linhas da matriz de planejamento que se referem aos estímulos “Corsa azul” e “Gol azul” são, respectivamente

$$X_{i1,2} = \begin{bmatrix} 1 & 0 & 1 & 0 & -1 & -1 \end{bmatrix} \text{ e } X_{i2,4} = \begin{bmatrix} 1 & 1 & 0 & 0 & -1 & -1 \end{bmatrix}.$$

Os valores $y_{ijk}^* = 1$ definem as notas máximas de cada bloco; sendo assim, os demais valores devem ser gerados de forma que sejam menores que os máximos gerados. Neste caso, devem ser menores que 85 e 90, respectivamente. Essa restrição pode ser obedecida descartando-se os valores que não a satisfazem e continuando o processo até que sejam gerados valores adequados.

3.5 Comentários

Desta maneira, mostramos como o modelo de Marshall e Bradlow agrega informações provenientes dos métodos de composição e de decomposição, atribuindo maior peso às informações de maior precisão. Mostramos, também, que unifica os modelos de cada um dos métodos, pois, analiticamente, mostramos que os modelos separados são casos particulares deste. Vimos que o Modelo Bayesiano apresenta matriz de variâncias-covariâncias menor que os modelos separados.

Por fim, vimos que a informação *a priori* do vetor de utilidades pode ter diversas origens, pois é possível introduzi-la no modelo de forma adequada através de adaptações no método de estimação. Com isso, concluímos a apresentação teórica do Modelo Bayesiano de Marshall e Bradlow.

Para completar a presente dissertação, conduzimos uma coleta de dados e aplicamos o Modelo Bayesiano de Marshall e Bradlow. Essa parte prática é apresentada no próximo capítulo.

Capítulo 4

Comparação dos modelos

Nesta seção apresentamos e comparamos alguns resultados obtidos no ajuste dos modelos de decomposição, de composição e variações do Modelo Bayesiano a um determinado conjunto de dados, que foi coletado exclusivamente para a presente dissertação.

Descrevemos aqui as etapas da aplicação, incluindo a elaboração do questionário, o ajuste e a comparação dos modelos. Essa aplicação tem dois objetivos:

1. Verificar o efeito da redução do número de estímulos sobre os resultados finais;
2. Apresentar uma aplicação da técnica a dados reais (parte desenvolvida no Capítulo 5).

O questionário desenvolvido foi aplicado a alunos de primeiro ano de pós-graduação de uma faculdade particular da cidade de São Paulo (SP).

4.1 Definição do produto e planejamento do questionário

O “produto” escolhido para essa pesquisa não é do tipo físico, mas um conceito: *Instituição de ensino superior para realização de um curso de pós-graduação*. Além de comparar o ajuste de diferentes modelos, queremos entender que motivos levam um aluno a escolher uma determinada instituição de ensino, e não outra, para realizar sua pós-graduação.

No modelo Bayesiano de Marshall e Bradlow não há restrições quanto ao número de atributos ou número de níveis a serem adotados. Nesta aplicação definimos 5 atributos, com 2 níveis cada, para

compor esse produto, sendo que o atributo referente ao corpo docente foi inspirado na aplicação feita por Friedmann (1998). São eles

(A) *Tradição do Instituto e Qualidade do curso*

(A1) *Instituição tradicional (reconhecida) com razoável qualificação para o curso desejado.*

(A2) *Instituição nova e com alta qualificação para o curso desejado.*

(B) *Dias das aulas*

(B1) *Todas as aulas em dias de semana.*

(B2) *Menos aulas em dias de semana e algumas aulas aos sábados de manhã.*

(C) *Quantidade de disciplinas*

(C1) *Disciplinas com maior carga horária, mas em menor quantidade.*

(C2) *Disciplinas com carga horária menor, mas em maior quantidade.*

(D) *Corpo docente*

(D1) *Corpo docente com formação acadêmica razoável, mas com grande experiência de mercado.*

(D2) *Corpo docente com excelente formação acadêmica, mas sem grande experiência de mercado.*

(E) *Grade horária das disciplinas*

(E1) *Ter apenas 1 disciplina por dia, cada disciplina com 1 dia de aula por semana.*

(E2) *Ter 2 disciplinas por dia, cada disciplina com 2 dias de aula por semana.*

A partir de 5 atributos com 2 níveis cada, obtemos um planejamento fatorial completo com 32 estímulos no total (Tabela 4.1). Porém, através de um planejamento ortogonal fracionário (1/4) (Neter et al., 1996), chegamos a um total de 8 estímulos - número mínimo necessário para se ajustar aos dados um modelo de regressão linear aditivo. Para aplicar o Modelo Bayesiano, não precisamos da avaliação destes 8 estímulos (este número pode ser menor), porém coletamos as 8 avaliações para ajustar um modelo de decomposição e, com isso, comparamos os resultados obtidos.

Seguem na Tabela 4.2 os 8 estímulos que resultam do planejamento ortogonal fracionário feito pela versão 13.0 do pacote estatístico Minitab.

O questionário respondido por cada aluno, apresentado no Apêndice B, contém três partes:

- A primeira (Figura B.1) é composta de um conjunto de perguntas de perfil sócio-econômico, que coletamos a fim de caracterizar os alunos que responderam ao questionário.

- A segunda parte foi estruturada para obter dados via método de decomposição (Figura B.2), nela apresentamos os 8 estímulos ao aluno e solicitamos que atribuisse notas de 0 a 100 para cada estímulo segundo sua preferência por cada um, sendo que notas altas indicam maior preferência.

Tabela 4.1: Estímulos formados pela combinação de 5 atributos com 2 níveis cada ($2^5=32$)

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
A1	B1	C1	D1	E1	A2	B1	C1	D1	E1
A1	B1	C1	D1	E2	A2	B1	C1	D1	E2
A1	B1	C1	D2	E1	A2	B1	C1	D2	E1
A1	B1	C1	D2	E2	A2	B1	C1	D2	E2
A1	B1	C2	D1	E1	A2	B1	C2	D1	E1
A1	B1	C2	D1	E2	A2	B1	C2	D1	E2
A1	B1	C2	D2	E1	A2	B1	C2	D2	E1
A1	B1	C2	D2	E2	A2	B1	C2	D2	E2
A1	B2	C1	D1	E1	A2	B2	C1	D1	E1
A1	B2	C1	D1	E2	A2	B2	C1	D1	E2
A1	B2	C1	D2	E1	A2	B2	C1	D2	E1
A1	B2	C1	D2	E2	A2	B2	C1	D2	E2
A1	B2	C2	D1	E1	A2	B2	C2	D1	E1
A1	B2	C2	D1	E2	A2	B2	C2	D1	E2
A1	B2	C2	D2	E1	A2	B2	C2	D2	E1
A1	B2	C2	D2	E2	A2	B2	C2	D2	E2

Tabela 4.2: Estímulos escolhidos pelo planejamento ortogonal fracionário

Estímulo	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
I	A2	B1	C1	D1	E2
II	A2	B1	C2	D1	E1
III	A2	B2	C1	D2	E1
IV	A1	B2	C1	D1	E1
V	A1	B1	C1	D2	E2
VI	A1	B2	C2	D1	E2
VII	A2	B2	C2	D2	E2
VIII	A1	B1	C2	D2	E1

- A terceira parte (Figura B.3) foi formulada para obter dados via método de composição, nesta fase solicitamos ao entrevistado que atribuisse notas de 0 a 10 a cada um dos níveis dos atributos conforme a preferência; em seguida, que distribuisse um total de 100 pontos dentre os 5 atributos, representando a importância de cada um deles na sua concepção.

A ordem de apresentação das duas últimas partes foi aleatorizada e todos os indivíduos receberam os mesmos 8 estímulos para avaliar. Ao todo, 29 alunos responderam aos questionários.

4.2 Definição dos Modelos ajustados

A partir dos dados coletados, foi possível aplicar o Modelo Bayesiano de Marshall e Bradlow. Tomamos o cuidado de obter as avaliações sob escalas distintas, ou seja, as avaliações de estímulos não estão na mesma escala das avaliações obtidas pelo método de composição. Isso nos permite aplicar o Modelo Bayesiano adaptado, como apresentamos na Seção 3.1. Depois, igualamos a escala das avaliações manualmente, para também aplicar o Modelo Bayesiano original, Seção 2.1.

Finalmente, uma vez que coletamos número suficiente de estímulos tais que garantissem que a matriz de planejamento tivesse posto completo, ajustamos o Modelo Bayesiano (com escalas distintas e com escalas iguais) a todas as avaliações de estímulos coletadas de cada indivíduo e ajustamos também a um número reduzido dessas avaliações feitas a estímulos. Verificamos assim, a performance do Modelo Bayesiano aplicado a um número reduzido de estímulos e a performance obtida com todas as avaliações coletadas.

Listamos aqui os modelos que foram ajustados e que serão detalhados nas seções seguintes:

- *Decomposição*: Modelo utilidade usual do método de decomposição.
- *Composição*: Modelo de composição.
- *Bayesiano I*: Modelo Bayesiano com verossimilhança gerada por 8 estímulos.
- *Bayesiano II*: Modelo Bayesiano com verossimilhança gerada por apenas 5 estímulos.
- *Bayesiano III*: Modelo Bayesiano com verossimilhança gerada por 8 estímulos e utilizando a mesma escala nas duas partes.
- *Bayesiano IV*: Modelo Bayesiano com verossimilhança gerada por apenas 5 estímulos e utilizando a mesma escala nas duas partes.

Importante lembrar que, assumimos para todos os modelos, a não existência de interação entre os atributos.

4.2.1 Modelo utilidade do método de decomposição

O modelo utilidade do método de decomposição possui estrutura análoga ao modelo de regressão linear, como comentamos na Seção 1.2.4. Para ajustar este modelo, utilizamos apenas as respostas coletadas na parte de decomposição - que podem ser conferidas na Tabela B.1 do Apêndice B. Ajustamos um modelo de regressão linear aditivo, equivalente ao modelo utilidade do exemplo (1.1) e (1.2).

O total de questionários respondidos é $n = 29$ e o número de coeficientes a serem estimados é $p = (2 - 1) \times 5 + 1 = 6$. A matriz de planejamento, (4.1), nesta aplicação, é igual para todos os indivíduos e define os $m = 8$ estímulos apresentados obedecendo restrições de identificabilidade do tipo casela de referência (Neter et al., 1996). O tratamento de referência é formado pelo primeiro nível de cada atributo. Lembrando que, neste caso, todos os atributos apresentam apenas 2 níveis cada.

$$X = X_i = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}, i = 1, 2, \dots, 29. \quad (4.1)$$

O modelo utilidade, correspondente ao indivíduo i , deste caso, pode ser escrito matricialmente da seguinte maneira

$$y_i = X\mu_i + \epsilon_i \quad (4.2)$$

no qual,

y_i são vetores (8×1) com as notas dadas pelo indivíduo i a cada um dos estímulos;

μ_i é um vetor (6×1) com as utilidades dos níveis de atributos, referentes ao indivíduo i ;

ϵ_i é um vetor (8×1) com os erros aleatórios atribuídos para cada y_{ijk} , tal que, $\epsilon_i \sim N_8(0; \sigma^2 I)$ independentes;

σ^2 é a variância desconhecida dos erros ϵ_i e

I é uma matriz identidade de dimensão 8.

O estimador utilizado foi o estimador de mínimos quadrados (1.5) e o pacote estatístico foi o Minitab (versão 13.0). A partir das utilidades estimadas, $\hat{\mu}_i^d$, que estão dispostas nas tabelas E.1, E.2, E.3 e E.4 do

Apêndice E, obtivemos as estimativas das notas correspondentes aos estímulos que foram avaliados (Tabela 4.2), através de

$$\hat{y}_i^d = X_i \hat{\mu}_i^d. \quad (4.3)$$

As estimativas das avaliações dos 8 estímulos, para cada um dos indivíduos, estão dispostas no Apêndice E, Tabela E.5. Para exemplificar, apresentamos abaixo o cálculo dessas mesmas estimativas referentes ao primeiro indivíduo. As notas dadas na fase de decomposição pelo indivíduo 1 são,

$$y_1 = \begin{bmatrix} 60 \\ 50 \\ 55 \\ 45 \\ 55 \\ 52 \\ 58 \\ 50 \end{bmatrix}.$$

A estimativa $\hat{\mu}_1^d$ foi obtida a partir de (1.5) aplicado sobre o vetor de avaliações observadas y_1 e a estimativa $\hat{y}_1^d$ foi obtida através de (4.3); são elas

$$\hat{\mu}_1^d = \begin{bmatrix} 47,3 \\ 5,3 \\ -1,3 \\ -1,3 \\ 2,8 \\ 6,3 \end{bmatrix} \quad e \quad \hat{y}_1^d = \begin{bmatrix} 58,8 \\ 51,3 \\ 54,0 \\ 46,0 \\ 56,3 \\ 51,0 \\ 59,0 \\ 48,8 \end{bmatrix}.$$

Esse mesmo procedimento foi realizado para a obtenção das estimativas dos coeficientes e das avaliações dos 28 indivíduos restantes.

4.2.2 Modelo de composição

Neste caso utilizamos as respostas coletadas na parte de composição e aplicamos o modelo descrito na Seção 1.2.4. Todas as respostas dadas na fase de composição estão expostas na Tabela B.2 do Apêndice B.

Seguindo o exemplo da Tabela 1.9, compomos as avaliações correspondentes aos 8 estímulos que foram avaliados na fase de decomposição. Chamamos os vetores com essas estimativas “compostas” de $\hat{y}_i^c$. O indivíduo 1, na parte de composição, atribuiu aos atributos e níveis definidos na Seção 4.1 os valores da Tabela 4.3.

Tabela 4.3: Avaliações feitas pelo indivíduo 1 aos níveis e atributos na fase auto-explicativa

<i>Atribuir notas de 0 a 10</i>			
Nível	Nota		
A1	6		
A2	7		
B1	5		
B2	6		
C1	8		
C2	4		
D1	6		
D2	5		
E1	5		
E2	6		

<i>Distribuir 100 pontos</i>	
Atributo	Pontos
A	25
B	5
C	25
D	35
E	10
Total de pontos	100

Por exemplo, como o primeiro estímulo avaliado pelos indivíduos é dado pela combinação de níveis (A2, B1, C1, D1, E2) (Tabela 4.2), então a estimativa da avaliação deste estímulo, referente ao indivíduo 1, fica

$$\begin{aligned}\hat{y}_{11}^c &= (A \times A2) + (B \times B1) + (C \times C1) + (D \times D1) + (E \times E2) = \\ &= (25 \times 7) + (5 \times 5) + (25 \times 8) + (35 \times 6) + (10 \times 6) = 670.\end{aligned}$$

As estimativas obtidas, pelo modelo de composição, para as avaliações y_i , referentes ao indivíduo 1 são

$$\hat{y}_1^c = \begin{bmatrix} 670,0 \\ 560,0 \\ 630,0 \\ 627,5 \\ 597,5 \\ 537,5 \\ 540,0 \\ 487,5 \end{bmatrix}.$$

No Apêndice E, Tabela E.7, apresentamos as estimativas das avaliações dos estímulos obtidas pelo método de composição.

Note que as avaliações compostas variam de 0 a 1.000, pois foram compostas a partir de escalas que variam de 0 a 100 e de 0 a 10.

4.2.3 Modelos bayesianos

O Modelo Bayesiano de Marshall e Bradlow utiliza a informação coletada no método de composição como informação *a priori* para o vetor de utilidades μ_i , e a informação proveniente da parte de decomposição do questionário na composição da função de verossimilhança. Logo, utilizamos, no ajuste de cada um dos Modelos Bayesianos definidos na Seção 4.2, as respostas provenientes das duas partes do questionário (tabelas B.1 e B.2 do Apêndice B).

$$\hat{y}_1^c = \begin{bmatrix} 670,0 \\ 560,0 \\ 630,0 \\ 627,5 \\ 597,5 \\ 537,5 \\ 540,0 \\ 487,5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \\ u_{14} \\ u_{15} \\ u_{16} \end{bmatrix} \quad (4.4)$$

Utilizando como exemplo a equação (2.1.1), os vetores de utilidades u_i , que utilizaremos como informação *a priori* dos vetores μ_i neste caso, podem ser obtidos compondo-se as 8 avaliações ($\hat{y}_i^c$) referentes aos estímulos da Tabela 4.2 e utilizando a matriz de planejamento (4.1). Por exemplo, o vetor u_1 , referente ao indivíduo 1, foi obtido resolvendo-se a equação (4.4) e resulta em

$$\begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \\ u_{14} \\ u_{15} \\ u_{16} \end{bmatrix} = \begin{bmatrix} 622,5 \\ 37,5 \\ 5,0 \\ -100,0 \\ -35,0 \\ 10,0 \end{bmatrix}$$

Os vetores u_i , para $i = 1, 2, \dots, 29$, estão dispostos na Tabela E.6 do Apêndice E.

Adotamos, para os quatro Modelos Bayesianos que iremos comparar, a mesma estrutura do modelo utilidade que foi definida em (4.2).

Apesar da estrutura ser a mesma, definimos algumas variações na aplicação a fim de comparar os resultados obtidos. Essas variações são detalhadas nas seções seguintes.

Modelo Bayesiano I

A matriz de planejamento (4.1), que define os $m = 8$ estímulos avaliados, será adotada no ajuste deste Modelo Bayesiano pois, nesta primeira variação, os 8 estímulos avaliados compõem a função de verossimilhança de cada indivíduo.

Como os vetores u_i foram gerados a partir de avaliações com escala de 0 a 1.000, e os vetores μ_i serão estimados a partir das avaliações que variam de 0 a 100, introduzimos nesse modelo um vetor α (2×1) que corrige essa diferença; garantindo, então, que a informação *a priori* de μ_i , u_i , fique na mesma escala de μ_i - veja Seção 3.1.

Note, através das informações do indivíduo 1, que as avaliações estão em escalas distintas

$$\hat{y}_1^c = \begin{bmatrix} 670,0 \\ 560,0 \\ 630,0 \\ 627,5 \\ 597,5 \\ 537,5 \\ 540,0 \\ 487,5 \end{bmatrix} \quad e \quad y_1 = \begin{bmatrix} 60 \\ 50 \\ 55 \\ 45 \\ 55 \\ 52 \\ 58 \\ 50 \end{bmatrix},$$

logo, geram vetores utilidades em escalas distintas

$$u_1 = \begin{bmatrix} 622,5 \\ 37,5 \\ 5,0 \\ -100,0 \\ -35,0 \\ 10,0 \end{bmatrix} \quad e \quad \hat{\mu}_1^d = \begin{bmatrix} 47,3 \\ 5,3 \\ -1,3 \\ -1,3 \\ 2,8 \\ 6,3 \end{bmatrix}.$$

O vetor α deve também ser estimado; logo, será introduzido no processo de estimação dos parâmetros.

Modelo Bayesiano II

No Modelo Bayesiano II utilizamos apenas $m = 5$ avaliações de estímulos para gerar a função de verossimilhança. Queremos, com isso, verificar o desempenho do modelo quando um número menor de estímulos é utilizado, ou seja, quando a matriz de planejamento não apresenta posto completo.

Dessa maneira, a matriz de planejamento (4.1) foi trocada por uma matriz “reduzida” X_i^r , também igual para todos os indivíduos, dada por

$$X_i^r = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}, i = 1, 2, \dots, 29. \quad (4.5)$$

Os 5 estímulos que compõem a matriz (4.1) foram escolhidos de maneira que não houvesse confusão entre os efeitos dos atributos. As combinações de níveis correspondentes estão na Tabela 4.4.

Tabela 4.4: Estímulos escolhidos para compor a matriz reduzida de planejamento

Estímulo	A	B	C	D	E
I	A2	B1	C1	D1	E2
II	A2	B1	C2	D1	E1
III	A2	B2	C1	D2	E1
V	A1	B1	C1	D2	E2
VI	A1	B2	C2	D1	E2

Os vetores y_i também são reduzidos, pois devem se referir aos 5 estímulos escolhidos. No caso do indivíduo 1, temos

$$y_1^r = \begin{bmatrix} 60 \\ 50 \\ 55 \\ 55 \\ 52 \end{bmatrix}.$$

A estrutura desta versão é a mesma do Modelo Bayesiano I, apenas os vetores y_i e a matriz de planejamento sofrem alterações.

Modelos bayesianos III e IV

Os Modelos Bayesianos III e IV foram preparados para avaliar o desempenho dos modelos Bayesianos I e II sem a necessidade de introduzir um parâmetro de correção de escala. Para isso, transformamos manualmente as avaliações dos estímulos antes de aplicar o Modelo Bayesiano, ou seja, multiplicamos por 10 as avaliações dadas na parte de decomposição, obtendo assim notas entre 0 e 1.000, escala igual à obtida na parte de composição. Com isso, μ_i e u_i ficam sob mesma escala.

No ajuste do Modelo Bayesiano III, a verossimilhança foi gerada a partir das avaliações dos $m = 8$ estímulos com a escala das avaliações de decomposição alterada. A matriz de planejamento utilizada é a mesma do Modelo Bayesiano I.

O Modelo Bayesiano IV, é análogo ao Modelo Bayesiano II, ou seja, apresenta função de verossimilhança gerada com apenas $m = 5$ avaliações (referentes aos mesmos 5 estímulos), a diferença está apenas na escala das avaliações do método de decomposição, que foi igualada à das avaliações obtidas no método de composição.

4.2.4 Estimação dos coeficientes dos Modelos Bayesianos

Para estimar os vetores de utilidades, μ_i , dos quatro Modelos Bayesianos desta análise, utilizamos a interface 0.2-3 do OpenBUGS para o software R versão 2.1.1, denominada *BRugs*¹. Este pacote implementa o amostrador de Gibbs e nos permite utilizar processos com múltiplas e independentes cadeias. É possível, também, utilizar uma função que gera automaticamente valores iniciais para os todos os parâmetros e construir gráficos para o diagnóstico de convergência das cadeias.

A implementação desenvolvida e os passos utilizados nessa estimação estão presentes no Apêndice G.

Na estimação dos parâmetros dos Modelos Bayesianos, utilizamos múltiplas cadeias de Markov e a função que gera valores iniciais automaticamente para dar início às iterações. Realizamos, para cada cadeia, um primeiro aquecimento (*burn in*) de 50.000 iterações, sendo que, apenas 5.000 valores de cada uma foram considerados - armazenamos 1 valor a cada 10 iterações.

Utilizamos, na estimação de todos os modelos, distribuições *a priori* não-informativas para os parâmetros σ^2 e Ω . Para os modelos bayesianos I e II, assumimos na distribuição *a priori* de α , (3.2), os seguintes valores:

$$\alpha_0 = \begin{bmatrix} 0 \\ 0,1 \end{bmatrix} \quad e \quad \Sigma_0 = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}.$$

O diagnóstico de convergência foi feito através da estatística $\hat{R}$ proposta por Gelman e Rubin (1992) (corrigida por Brooks and Gelman, 1998), dos gráficos de densidade das estimativas e dos históricos das iterações.

¹<http://mathstat.helsinki.fi/openbugs/>

Os gráficos de densidade foram construídos com os valores das iterações realizadas após o período de aquecimento que, por sua vez, foi determinado através da estatística de Gelman e Rubin. No Apêndice D estão presentes os gráficos com a evolução dessas estatísticas de convergências, $\hat{R}$, e os gráficos de densidade de cada um dos parâmetros, para cada um dos indivíduos, sob cada um dos quatro Modelos Bayesianos.

Após este primeiro aquecimento, realizamos 25.000 iterações (*update*) para cada uma das 5 cadeias; totalizando 30.000 valores para cada. No caso dos Modelos Bayesianos I e II, a convergência das cadeias foi observada após as primeiras 15.000 iterações; logo, as estimativas foram compostas pelos 15.000 valores finais de cada cadeia. Para os modelos III e IV, realizamos mais 25.000 iterações, pois as cadeias dos parâmetros σ^2 e Ω não haviam convergido. Porém, mesmo depois de dobrar o número de iterações de cada cadeia, a convergência dos parâmetros σ^2 e Ω dos modelos III e IV não foi alcançada, indicando que estes modelos, provavelmente, sejam inadequados para este conjunto de dados.

A Figura 4.1 exemplifica um diagnóstico de convergência das cadeias para o primeiro valor do vetor de utilidades do indivíduo 1, sob o Modelo Bayesiano I.

Figura 4.1: Gráficos e tabela referentes às simulações feitas sob o Modelo Bayesiano I - Indivíduo 1

Note que a estatística $\hat{R}$ (apresentada em vermelho no primeiro gráfico da Figura 4.1) desde o início

das iterações já apresenta valor próximo de 1, o que indica que o período de aquecimento foi concluído, pois já há indício de que as cadeias tenham convergido. Observando essa mesma estatística para cada um dos escalares estimados, percebemos que após as primeiras 15.000 iterações, todos já apresentam indicação de que o período de aquecimento pode ser interrompido.

O gráfico de densidade foi construído com os valores das iterações sem considerar o período de aquecimento. Nesta aplicação estimamos valores que seguem distribuições com apenas uma moda, que são a distribuição Normal e a distribuição Qui-quadrado, com isso, esperamos que os gráficos de densidade apresentem perfil semelhante. Na Figura 4.1, percebe-se que o gráfico de densidade se apresenta semelhante a uma densidade de distribuição Normal.

O histórico, apresentado no terceiro gráfico da Figura 4.1, indica que as 5 cadeias constituídas oscilam em torno do mesmo valor. O que indica que convergiram. Por fim, observamos que a estimativa do erro padrão de Monte Carlo é inferior a 5% da estimativa do desvio padrão amostral, isso indica que não é mais necessário continuar a simulação de valores (veja tutorial do Brugs).

Como todos os escalares estimados para o Modelo Bayesiano I apresentaram mesmos diagnósticos, acreditamos que a convergência tenha sido alcançada.

Na Figura 4.2 apresentamos um caso em que a convergência não foi diagnosticada. Trata-se da estimativa de σ^2 do Modelo Bayesiano IV. Note no primeiro gráfico, por exemplo, que a estatística de Gelman e Rubin não se aproxima de 1. Logo, não há indicação de que o período de aquecimento tenha chegado ao fim.

Figura 4.2: Gráficos e tabela referentes às simulações feitas sob o Modelo Bayesiano IV - Sigma

No gráfico de densidade da Figura 4.2 percebe-se mais de uma moda, o que não é esperado para um gráfico de densidade da distribuição Qui-quadrado Invertida. No histórico das simulações das cadeias, terceiro gráfico, notamos que estas não variam em torno do mesmo valor, não convergem, portanto, para o mesmo ponto. Como não houve convergência de todos os parâmetros estimados do modelo, as demais estimativas obtidas sob ele não são confiáveis, pois não há evidências de que o processo de simulação tenha convergido.

Como não houve convergência dos parâmetros σ^2 e Ω para os Modelos Bayesianos III e IV, como pode ser visto nas figuras D.25, D.26, D.27 e D.28, entendemos que as estimativas obtidas nas suas simulações não são confiáveis.

Quando transformamos a escala das avaliações dos estímulos, a fim de igualar com a escala das avaliações obtidas com o método de composição, acreditávamos que os modelos (modelos III e IV) apresentariam melhor desempenho na simulação, pois haveria um vetor de parâmetros (α) a menos para ser estimado. Porém, notamos que, apesar de termos multiplicado por 10 as avaliações de perfil-completo (o que as faria variar na mesma escala da informação *a priori*, de 0 a 1.000), a amplitude não se igualou. Concluimos, então, que o fato de termos retirado o vetor de transformação de escala pode ter impedido a convergência.

Note, pelas tabelas E.8 e E.14, que os valores estimados para o vetor de transformação de escala, α , sob o Modelo Bayesiano I e Modelo Bayesiano II são, respectivamente,

$$\hat{\alpha}^I = \begin{bmatrix} 0,05 \\ 0,09 \end{bmatrix} \text{ e } \hat{\alpha}^{II} = \begin{bmatrix} 0,03 \\ 0,09 \end{bmatrix}.$$

Ou seja, para transformar as notas precisaríamos multiplicá-las, neste conjunto de dados, por $(1/0,09) \cong 11,11$ e ainda somar uma constante, o que não foi feito. Isso pode justificar a não convergência dos Modelos Bayesianos III e IV.

Na prática, não faz sentido ajustar primeiro o modelo com o vetor α , apenas para encontrar estes valores de transformação e, então, refazer a estimação com a escala modificada manualmente; logo, introduzir o vetor de transformação parece ser a melhor opção. Com isso, podemos ver que, mesmo coletando dados na mesma escala (decomposição e composição), pode ocorrer de a convergência das cadeias só ser alcançada com a introdução do vetor de transformação α .

As utilidades estimadas em cada um dos Modelos Bayesianos que convergiram, denominamos de $\hat{\mu}_i^I$ e $\hat{\mu}_i^{II}$, para $i = 1, \dots, 29$. Através delas estimamos as notas referentes aos 8 estímulos que foram avaliados pelos alunos. Estas estimativas foram denominadas, respectivamente, por $\hat{y}_i^I$ e $\hat{y}_i^{II}$, para $i = 1, \dots, 29$.

Estão presentes na tabela E.8 as estimativas dos parâmetros σ^2 e Ω ; nas tabelas E.9, E.10, E.11 e E.12 as estimativas das utilidades e na tabela E.13 as estimativas das avaliações; todas obtidas sob o Modelo Bayesiano I. Nas tabelas E.14, E.15, E.16, E.17, E.18 e E.19, apresentamos as estimativas referentes aos mesmos parâmetros obtidas sob o Modelo Bayesiano II.

4.3 Resultados da comparação

Comparamos as estimativas dos modelos de composição, decomposição e dos modelos bayesianos que apresentaram convergência, através de algumas estatísticas também utilizadas em Marshall e Bradlow (2002), que são: a Correlação de Pearson, a Correlação de Spearman, a Raiz do Erro Quadrático Médio ($\sqrt{EQM}$) e o Desvio Médio Absoluto (MAD) (veja as respectivas fórmulas no Apêndice F). Além dessas estatísticas, incluímos a porcentagem de “acertos de primeira escolha” - na qual observamos, para cada modelo, quantas vezes o estímulo que recebera a maior avaliação no questionário coincidiu com o estímulo que apresentou maior avaliação estimada.

Tabela 4.5: Estatísticas de comparação dos modelos com avaliações originais (escalas distintas)

Estatísticas para comparação	Decomposição	Modelo			Composição
		Bayesiano I (m=8)	Bayesiano II (m=5)	Bayesiano II (m=5)*	
Correlação de Pearson(y_{obs}, y_{est})	0,98	0,84	0,81	0,89	0,48
Correlação de Spearman(p_{obs}, p_{est})	0,95	0,68	0,66	0,78	0,55
$\sqrt{EQM}(y_{obs}, y_{est})$	4,00	10,03	10,80	8,78	19,37
$\sqrt{EQM}(p_{obs}, p_{est})$	0,69	1,80	1,85	0,92	2,12
MAD(y_{obs}, y_{est})	2,88	7,69	8,17	6,76	15,11
MAD(p_{obs}, p_{est})	0,47	1,33	1,38	0,63	1,53
Acerto de primeira escolha	0,86	0,69	0,69	0,76	0,66
Total de indivíduos	29	29	29	29	29
Cadeias de Markov	—	5	5	5	—
Total de iterações por cadeia	—	75000	75000	75000	—
Iterações de aquecimento por cadeia	—	15000	15000	15000	—
Total de simulações das estimativas	—	75000	75000	75000	—

* Estatísticas calculadas apenas com os estímulos utilizados na composição da verossimilhança.

A notação (y_{obs}, y_{est}) indica que a estatística compara, para cada um dos modelos, os valores por eles previstos, y_{est} , com os valores de fato observados, y_{obs} . Já a notação (p_{obs}, p_{est}) indica que a medida compara, também para cada um dos modelos, os postos dos estímulos por eles previstos, p_{est} , com os correspondentes postos observados, p_{obs} .

A Tabela 4.5 mostra as estatísticas calculadas para os modelos que convergiram, ou seja, Modelos Bayesianos I e II. Podemos notar que a performance do Modelo Bayesiano II é semelhante à do modelo I. Esse fato é de grande importância, pois concretiza uma das propostas do Modelo Bayesiano, que é a de reduzir o trabalho do entrevistado; note que obtemos praticamente o mesmo desempenho de um modelo com avaliações

de 8 estímulos coletando apenas 5 estímulos.

Observamos que, dentre estes quatro modelos da Tabela 4.5, o que se aproxima mais das avaliações do método de decomposição é exatamente o Modelo de decomposição (modelo utilidade). Este fato indica que as avaliações provenientes do método de composição são informativas, pois participaram com peso na composição da média *a posteriori*.

Com estes resultados, concluímos que, se houvesse necessidade de redução do número de estímulos a serem avaliados no método de decomposição, o Modelo Bayesiano seria uma alternativa razoável, pois vimos que apresenta praticamente o mesmo desempenho quando ajustado com um número reduzido de estímulos.

A Tabela 4.6 apresenta as mesmas estatísticas utilizadas na comparação dos modelos com os valores observados, porém aplicadas aos valores de y_{ijk} previstos sob os Modelos Bayesianos com e sem redução do número de estímulos. A notação $\sqrt{EQM}_{postos}$ e MAD_{postos} indica que essas medidas comparam os postos previstos pelos dois modelos.

Tabela 4.6: Estatísticas de comparação dos Modelos Bayesianos

Estatísticas para comparação	I versus II
Correlação de Pearson	0,98
Correlação de Spearman	0,76
$\sqrt{EQM}$	2,95
$\sqrt{EQM}_{postos}$	1,54
MAD	2,34
MAD_{postos}	1,00
Coincidência na primeira escolha	0,69
Total de indivíduos	29
Número de Cadeias de Markov	5
Total de iterações por cadeia	75000
Iterações de aquecimento por cadeia	15000
Total de simulações das estimativas	75000

Note, pela Tabela 4.6, que as estimativas das avaliações apresentam correlação de Pearson igual a 98% e correlação de Spearman igual a 76%, com isso, vemos que e as estimativas obtidas com estes modelos são realmente semelhantes.

Capítulo 5

Aplicação da Análise de Preferência Conjunta

Desenvolvemos, com intuito ilustrativo, uma análise de preferência conjunta utilizando as estimativas obtidas sob os Modelos Bayesianos I e II.

Através dessa análise, pretendemos inferir quais atributos foram os mais importantes para os alunos de primeiro ano pós-graduação de uma faculdade particular da cidade de São Paulo (SP), na escolha da *Instituição de ensino* para realizar o curso de pós-graduação desejado.

Como a mesma análise foi coduzida para os dois modelos, continuamos comparando os resultados obtidos através deles.

5.1 Cálculo das Importâncias dos atributos

Calculamos para cada indivíduo (sob cada um dos modelos) as importâncias referentes a cada um dos atributos que compunham nosso produto. O indivíduo 1, por exemplo, apresentou o seguinte vetor de estimativas sob o Modelo Bayesiano I:

$$\hat{\mu}_1^I = \begin{bmatrix} 54,94 \\ 3,60 \\ 0,96 \\ -8,63 \\ -2,01 \\ 1,79 \end{bmatrix}.$$

Para estimar a importância de um atributo, Tabela 5.1, calcula-se primeiro a amplitude existente entre as utilidades estimadas para cada um dos atributos, por exemplo, a amplitude referente ao atributo *A* é de 3,60.

A importância atribuída a um atributo é, então, dada pelo percentual que a sua amplitude representa na soma de todas as amplitudes. O atributo *A* corresponde a uma importância de 21% na preferência do indivíduo 1.

Tabela 5.1: Cálculo das Importâncias dos atributos para o indivíduo 1

Utilidades estimadas				Amplitude	Percentual
A1	0	A2	3,60	3,60	21%
B1	0	B2	0,96	0,96	6%
C1	0	C2	-8,63	8,63	51%
D1	0	D2	-2,01	2,01	12%
E1	0	E2	1,79	1,79	11%
Soma das amplitudes				16,99	100%

Calculamos as importâncias atribuídas aos atributos para cada indivíduo, através das utilidades estimadas por cada um dos Modelos Bayesianos. Na Figura 5.1, apresentamos os gráficos de barras construídos com as importâncias estimadas, para cada um dos 29 indivíduos. Note que os modelos não nos levam a conclusões distintas sobre a distribuição da importância dada pelos indivíduos aos atributos.

Figura 5.1: Gráfico de barras com as importâncias estimadas dos atributos por indivíduo segundo os Modelos Bayesianos I e II

5.2 Análise de Agrupamento

A partir das importâncias estimadas (calculadas), realizamos uma análise de agrupamento dos indivíduos (Johnson, 1998). Pelo dendrograma obtido com uma técnica de agrupamento hierárquica (utilizando a distância de Ward) notamos que, sob qualquer um dos Modelos Bayesianos considerados, era possível identificar 3 grupos de indivíduos.

Com esse resultado, realizamos um agrupamento através do método de *k-médias*, solicitando, então, que fossem formados 3 grupos de indivíduos a partir das importâncias estimadas. O agrupamento foi feito com as estimativas de cada um dos modelos separadamente.

Calculamos, para cada grupo, as importâncias médias correspondentes aos atributos, como pode ser visto na Figura 5.2. Note que os dois Modelos Bayesianos formaram grupos cujas importâncias médias se assemelham.

Figura 5.2: Gráfico de barras com as importâncias médias estimadas dos atributos segundo os grupos obtidos pelos Modelos Bayesianos I e II

Análise descritiva sócio-econômica

Realizamos ainda, para cada grupo, uma análise descritiva das informações que coletamos na parte sócio-econômica do questionário (Figura B.1 do Apêndice B). Veja as figuras 5.3 e 5.4.

Figura 5.3: Tabelas com variáveis sociais segundo os grupos obtidos pelos Modelos Bayesianos I e II

Percentual de indivíduos que recebem bolsa				
	Grupo			
Estimativas utilizadas para formar os grupos	1	2	3	
Modelo Bayesiano I	75%	82%	79%	
Modelo Bayesiano II	100%	82%	73%	
Percentual de indivíduos do sexo masculino				
	Grupo			
Estimativas utilizadas para formar os grupos	1	2	3	
Modelo Bayesiano I	75%	82%	86%	
Modelo Bayesiano II	100%	91%	73%	
Percentual de indivíduos com estado civil solteiro				
	Grupo			
Estimativas utilizadas para formar os grupos	1	2	3	
Modelo Bayesiano I	50%	45%	71%	
Modelo Bayesiano II	67%	45%	67%	
Percentual de indivíduos que moram na cidade de São Paulo				
	Grupo			
Estimativas utilizadas para formar os grupos	1	2	3	
Modelo Bayesiano I	75%	82%	86%	
Modelo Bayesiano II	67%	82%	87%	

Figura 5.4: Gráfico de barras com os hábitos dos indivíduos segundo os grupos obtidos pelos Modelos Bayesianos I e II

Perfil dos grupos formados

Os três grupos, formados de maneira semelhante pelos Modelos Bayesianos I e II, podem ser caracterizados, através das figuras 5.2, 5.3 e 5.4, como segue abaixo. Informamos, entre parênteses, o número de pessoas que forma o grupo segundo os modelos I e II, respectivamente.

-Grupo 1 (4 e 3 indivíduos):

Dentre os três grupos, este é o que apresenta maior importância média atribuída ao atributo “Quantidade de disciplinas” e o que menos dá importância ao atributo “Tradição da Instituição e qualidade do curso”. Este grupo é formado por pessoas que têm o hábito de ir a shows e ao teatro, mas que não apreciam viajar; é o grupo com menor porcentagem de pessoas que moram em São Paulo e menor percentual de solteiros.

-Grupo 2 (11 e 11 indivíduos):

Este grupo é o que apresenta maior importância média atribuída ao atributo “Tradição da Instituição e qualidade do curso” e menor importância média associada ao atributo “Grade horária das disciplinas”. É formado basicamente por bolsistas, casados e que moram na cidade de São Paulo; a atividade que aparece com maior destaque neste grupo é a leitura de livros e, dentre os três grupos, é o que menos aprecia ir a shows.

-*Grupo 3* (14 e 15 indivíduos):

Para este grupo, o atributo com maior importância média é o “Dias de aula” e dentre os três, é o que atribui maior importância ao atributo “Corpo docente”. É caracterizado pelos hábitos de viajar, praticar esportes e ir ao cinema, além disso apresenta elevada porcentagem do estado civil solteiro e de moradia em São Paulo.

5.3 Comentários

Note que a redução do número de avaliações, que fizemos no Modelo Bayesiano II, não resultou em grupos muito distintos dos que foram formados pelo Modelo Bayesiano I, ou seja, o desempenho não é prejudicado. Além disso, ressaltamos que neste exemplo, os grupos foram formados a partir das importâncias dadas aos atributos e não pelo nível preferido em cada um.

Capítulo 6

Comentários Finais

No desenvolvimento do Modelo Bayesiano verificamos que a proposta de Marshall e Bradlow é atraente, uma vez que agrega as informações dos dois principais métodos de coletas de dados utilizados na Análise de Preferência Conjunta. Além de apresentar como casos particulares os modelos usuais desses dois métodos.

As propriedades decorrentes dessa abordagem também são interessantes, dentre elas vimos que a matriz de planejamento não precisa ter posto completo, logo, não é necessário coletar um número mínimo de avaliações para se obter uma matriz ortogonal, o que reduz o trabalho do entrevistado se o número de atributos e níveis for elevado; e vimos que a matriz de variâncias-covariâncias resultante é “menor” que as matrizes dos modelos agregados, o que nos remete à uma maior precisão do estimador no sentido de apresentar menor matriz de variâncias-covariâncias.

Na aplicação verificamos que o Modelo Bayesiano ajustado com número reduzido de avaliações de estímulos (5 avaliações) apresentou performance semelhante ao Modelo Bayesiano ajustado com 8 avaliações (número mínimo necessário para se obter um planejamento ortogonal).

Na análise de agrupamentos, feita com as estimativas obtidas através dos Modelos Bayesianos com 8 e 5 avaliações cada, obtivemos a formação de grupos semelhantes. Percebemos, novamente, que a redução do número de avaliações não prejudicam o desempenho do Modelo Bayesiano.

6.1 Futuros estudos

O efeito da redução do número de estímulos a serem avaliados nas estimativas dos parâmetros do modelo foi estudado através de uma aplicação, portanto, merece ser corroborado por estudos adicionais. De qualquer modo, aponta na direção de uma propriedade com forte apelo prático, pois facilita a tarefa do entrevistado.

Associado à essa questão, sugerimos algum estudo que facilite ou contribua para a escolha deste número reduzido de estímulos, como por exemplo, técnicas relacionadas à Teoria da Decisão, que também se trata de

uma abordagem bayesiana.

Outro ponto que poderia ser estudado é a incorporação de efeitos de interação no modelo bayesiano aditivo, pois essa suposição é delicada e bastante forte. Além das alterações que serão necessárias na formulação do modelo, deve-se focar no planejamento mínimo necessário para a estimação do modelo, ou seja, no número mínimo de estímulos a serem avaliados.

Marshall e Bradlow (2002) desenvolveram uma abordagem bayesiana para a estimação de um modelo aditivo. Seria bem-vinda a extensão dessa abordagem para a estimação de modelos ponto-ideal.

Apêndice A

Demonstrações de resultados

Neste apêndice demonstramos resultados que fizeram parte do desenvolvimento do Modelo Bayesiano dessa dissertação.

A.1 Distribuição *a posteriori* de μ_i , $f(\mu_i|y_i)$

Nesta seção apresentamos uma maneira de se obter a distribuição *a posteriori* de μ_i , equação (2.3) da Seção 2.1.3, supondo conhecidos os parâmetros σ^2 e Ω . Como os modelos são individuais considere que $i = 1, \dots, n$, sendo i o indexador desses indivíduos. Tomamos como base a demonstração feita em Artes (1991) para a obtenção do estimador James-Stein.

Sabemos que $y_i|\mu_i$ segue distribuição Normal m-variada (Seção 2.1.2) com a seguinte função densidade de probabilidade

$$g(y_i|\mu_i) \propto \exp \left\{ -\frac{1}{2\sigma^2} (y_i - X_i\mu_i)' (y_i - X_i\mu_i) \right\}.$$

Se tomarmos o estimador de mínimos quadrados dado por

$$b_i = (X_i'X_i)^{-1} X_i'y_i$$

teremos que

$$(y_i - X_i\mu_i)' (y_i - X_i\mu_i) =$$

$$(y_i - X_i b_i + X_i b_i - X_i \mu_i)'(y_i - X_i b_i + X_i b_i - X_i \mu_i) =$$

$$(y_i - X_i b_i)'(y_i - X_i b_i) + 2(y_i - X_i b_i)'X_i(b_i - \mu_i) + (\mu_i - b_i)'X_i'X_i(\mu_i - b_i)$$

porém, o segundo termo da última igualdade é nulo, como mostramos abaixo

$$(y_i - X_i b_i)'X_i(b_i - \mu_i) = (y_i - X_i(X_i'X_i)^{-1}X_i'y_i)'X_i(b_i - \mu_i) =$$

$$y_i'(I - X_i(X_i'X_i)^{-1}X_i')X_i(b_i - \mu_i) = y_i'(X_i - X_i)(b_i - \mu_i) = 0.$$

Vale então a igualdade

$$(y_i - X_i \mu_i)'(y_i - X_i \mu_i) = (y_i - X_i b_i)'(y_i - X_i b_i) + (\mu_i - b_i)'X_i'X_i(\mu_i - b_i).$$

Finalmente, chegamos à seguinte expressão

$$g(y_i|\mu_i) \propto \exp \left\{ -\frac{1}{2\sigma^2} [(y_i - X_i b_i)'(y_i - X_i b_i) + (\mu_i - b_i)'X_i'X_i(\mu_i - b_i)] \right\}.$$

Provado este resultado, partimos para a distribuição de $\mu_i|y_i$. Pelo Teorema de Bayes temos

$$f(\mu_i|y_i) \propto h(\mu_i)g(y_i|\mu_i).$$

Sabendo que a distribuição *a priori* de μ_i é Normal p-variada dada por (2.1), sua função densidade de probabilidade é

$$h(\mu_i) \propto \exp \left\{ -\frac{1}{2}(\mu_i - u_i)'\Omega^{-1}(\mu_i - u_i) \right\}.$$

A partir das funções de densidade $h(\mu_i)$ e $g(y_i|\mu_i)$ chegamos a

$$\begin{aligned} f(\mu_i|y_i) &\propto \exp \left\{ -\frac{1}{2} \left[(\mu_i - u_i)'\Omega^{-1}(\mu_i - u_i) + \frac{1}{\sigma^2}(\mu_i - b_i)'X_i'X_i(\mu_i - b_i) \right] \right\} = \\ \exp \left\{ -\frac{1}{2} \left[\mu_i'\Omega^{-1}\mu_i - \underbrace{u_i'\Omega^{-1}\mu_i - \mu_i'\Omega^{-1}u_i}_{2\mu_i'\Omega^{-1}u_i} + \mu_i'\frac{X_i'X_i}{\sigma^2}\mu_i - \underbrace{b_i'\frac{X_i'X_i}{\sigma^2}\mu_i - \mu_i'\frac{X_i'X_i}{\sigma^2}b_i}_{2\mu_i'\frac{X_i'X_i}{\sigma^2}b_i} + b_i'\frac{X_i'X_i}{\sigma^2}b_i \right] \right\} = \\ \exp \left\{ -\frac{1}{2} \left[\mu_i'\Omega^{-1}\mu_i - 2\mu_i'\Omega^{-1}u_i + u_i'\Omega^{-1}u_i + \mu_i'\frac{X_i'X_i}{\sigma^2}\mu_i - 2\mu_i'\frac{X_i'X_i}{\sigma^2}b_i + b_i'\frac{X_i'X_i}{\sigma^2}b_i \right] \right\} = \\ \exp \left\{ -\frac{1}{2} \left[\mu_i' \left(\Omega^{-1} + \frac{X_i'X_i}{\sigma^2} \right) \mu_i - 2\mu_i' \left(\Omega^{-1}u_i + \frac{X_i'X_i}{\sigma^2}b_i \right) + u_i'\Omega^{-1}u_i + b_i'\frac{X_i'X_i}{\sigma^2}b_i \right] \right\} = (*) \end{aligned}$$

tomando

$$A_i = \left(\Omega^{-1} + \frac{X_i' X_i}{\sigma^2} \right)$$

e voltando à equação (*) teremos

$$(*) = \exp \left\{ -\frac{1}{2} \left[\mu_i' A_i \mu_i - 2 \mu_i' A_i A_i^{-1} \left(\Omega^{-1} u_i + \frac{X_i' X_i}{\sigma^2} b_i \right) + u_i' \Omega^{-1} u_i + b_i' \frac{X_i' X_i}{\sigma^2} b_i \right] \right\} = (**)$$

se tomarmos ainda

$$\mu_i^* = A_i^{-1} \left(\Omega^{-1} u_i + \frac{X_i' X_i}{\sigma^2} b_i \right)$$

e substituirmos este valor na equação anterior, obteremos

$$\begin{aligned} (**) &= \exp \left\{ -\frac{1}{2} \left[\mu_i' A_i^{-1} \mu_i - 2 \mu_i' A_i \mu_i^* + u_i' \Omega^{-1} u_i + b_i' \frac{X_i' X_i}{\sigma^2} b_i \right] \right\} = \\ &\exp \left\{ -\frac{1}{2} \left[\underbrace{\mu_i' A_i \mu_i - \mu_i' A_i \mu_i^* - \mu_i^{*'} A_i \mu_i + \mu_i^{*'} A_i \mu_i^*}_{=0} - \mu_i^{*'} A_i \mu_i^* + u_i' \Omega^{-1} u_i + b_i' \frac{X_i' X_i}{\sigma^2} b_i \right] \right\} = \\ &\exp \left\{ -\frac{1}{2} \left[(\mu_i - \mu_i^*)' A_i (\mu_i - \mu_i^*) - \mu_i^{*'} A_i \mu_i^* + u_i' \Omega^{-1} u_i + b_i' \frac{X_i' X_i}{\sigma^2} b_i \right] \right\} \end{aligned}$$

logo,

$$f(\mu_i | y_i) \propto \exp \left\{ -\frac{1}{2} [(\mu_i - \mu_i^*)' A_i (\mu_i - \mu_i^*)] \right\}.$$

Note que essa função densidade nos remete a uma distribuição Normal cujos parâmetros, média e matriz de variâncias-covariâncias, são respectivamente

$$\mu_i^* = A_i^{-1} \left(\Omega^{-1} u_i + \frac{X_i' X_i}{\sigma^2} (X_i' X_i)^{-1} X_i' y_i \right) = A_i^{-1} \left(\Omega^{-1} u_i + \frac{X_i' y_i}{\sigma^2} \right)$$

e

$$A_i^{-1} = \left(\Omega^{-1} + \frac{X_i' X_i}{\sigma^2} \right)^{-1}.$$

Deduzimos, então, a distribuição *a posteriori* $\mu_i | y_i$.

A.2 Distribuições condicionais completas dos parâmetros σ^2 e Ω

Demostramos, aqui, como obter as distribuições condicionais completas de σ^2 e Ω , dadas respectivamente pelos resultados (2.7) e (2.8) da Seção 2.5. Lembrando que no desenvolvimento do Modelo Bayesiano, assumimos que σ^2 e Ω são parâmetros independentes entre si e iguais para todos os indivíduos.

Admita, como notação, que $y = (y'_1, \dots, y'_n)'$, ou seja, que y represente todos os vetores observados y_i e que $\mu = (\mu'_1, \dots, \mu'_n)'$.

A distribuição Normal m-variada dos vetores y_i , dada em (2.2), nos fornece a seguinte função densidade de probabilidade para $(y_i|\sigma^2, \mu_i)$

$$P(y_i|\sigma^2, \mu_i) = \left[\frac{|I\sigma^2|^{-\frac{1}{2}}}{2\pi^{\frac{m}{2}}} \exp \left\{ -\frac{1}{2} (y_i - X_i\mu_i)' (\sigma^2 I)^{-1} (y_i - X_i\mu_i) \right\} \right], \quad (\text{A.1})$$

$$-\infty < y_{ij} < \infty, \quad i = 1, 2, \dots, n \text{ e } j = 1, 2, \dots, m$$

tal que, $\sigma^2 > 0$.

Os vetores μ_i , apresentam distribuição *a priori* dada pela Normal p-variada (2.1), logo, a função densidade de probabilidade fica dada por

$$P(\mu_i|\Omega) = \left[\frac{|\Omega|^{-\frac{1}{2}}}{2\pi^{\frac{p}{2}}} \exp \left\{ -\frac{1}{2} (\mu_i - u_i)' \Omega^{-1} (\mu_i - u_i) \right\} \right], \quad (\text{A.2})$$

$$-\infty < \mu_{ij} < \infty, \quad i = 1, 2, \dots, n \text{ e } j = 1, 2, \dots, p$$

sendo, $-\infty < u_{ij} < \infty$ e Ω uma matriz $(p \times p)$ simétrica e positiva definida.

Sabemos que a distribuição *a priori* de σ^2 é dada por (2.4), assim, temos a seguinte função densidade de probabilidade

$$P(\sigma^2) = \left(\frac{1}{2} \right)^{\frac{g_0}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \right\} \frac{(\sigma^2)^{-\frac{g_0}{2}-1}}{\Gamma(\frac{g_0}{2})}, \quad 0 < \sigma^2 < \infty, \quad (\text{A.3})$$

sendo $\Gamma()$ a função Gamma e $0 < g_o < \infty$ os graus de liberdade da distribuição Qui-quadrado Invertida.

Sabemos também que a distribuição *a priori* de Ω é dada por (2.5), o que nos fornece a função densidade de probabilidade abaixo

$$P(\Omega) \propto |D_o|^{\frac{d_o}{2}} |\Omega|^{-\frac{1}{2}(d_o+p+1)} \exp \left\{ -\frac{1}{2} \text{tr} (D_o \Omega^{-1}) \right\}, \quad (\text{A.4})$$

na qual, $|D_o|$ é uma matriz $(p \times p)$ positiva definida e d_o são os graus de liberdade dessa distribuição de Wishart Invertida.

Distribuição condicional completa de σ^2

Queremos mostrar que a distribuição condicional completa *a posteriori* de σ^2 é dada por

$$\begin{aligned} [\sigma^2 | y, \mu, \Omega] &\sim IW_1(g, G), \\ g &= g_0 + nm \quad e \quad G = G_0 + \sum_{i=1}^n (y_i - X_i \mu_i)' (y_i - X_i \mu_i). \end{aligned}$$

Antes disso, note que

$$P(\sigma^2 | y, \mu, \Omega) = P(\sigma^2 | y, \mu)$$

devido à estrutura do modelo e à suposição de que σ^2 e Ω sejam independentes. Faremos, então, a demonstração a partir de $(\sigma^2 | y, \mu)$, pois através do teorema de Bayes, chegamos à seguinte relação

$$P(\sigma^2 | y, \mu) \propto P(\sigma^2) P(y | \sigma^2, \mu). \quad (\text{A.5})$$

Como as distribuições de $(y_i | \sigma^2, \mu_i)$ são independentes para diferentes i 's, substituímos em (A.5), as funções de densidade $P(\sigma^2)$ e $P(y | \sigma^2, \mu)$, dadas por (A.3) e (A.1) e chegamos a

$$P(\sigma^2 | y, \mu, \Omega) \propto P(\sigma^2) \prod_{i=1}^n P(y_i | \sigma^2, \mu_i) =$$

$$\begin{aligned}
&= \left(\frac{1}{2}\right)^{\frac{g_0}{2}} \exp\left\{-\frac{1}{2\sigma^2}\right\} \frac{(\sigma^2)^{-\frac{g_0}{2}-1}}{\Gamma\left(\frac{g_0}{2}\right)} \prod_{i=1}^n \left[\frac{|I\sigma^2|^{-\frac{1}{2}}}{(2\pi)^{\frac{m}{2}}} \exp\left\{-\frac{1}{2}(y_i - X_i\mu_i)' (\sigma^2 I)^{-1} (y_i - X_i\mu_i)\right\} \right] \\
&= \frac{(\sigma^2)^{-\frac{g_0}{2}-1}}{2^{\frac{g_0}{2}} \Gamma\left(\frac{g_0}{2}\right)} \exp\left\{-\frac{1}{2\sigma^2}\right\} \frac{(\sigma^2)^{-\frac{nm}{2}}}{(2\pi)^{\frac{nm}{2}}} \exp\left\{-\frac{1}{\sigma^2} \sum_{i=1}^n (y_i - X_i\mu_i)' (y_i - X_i\mu_i)\right\} \\
&= \frac{(\sigma^2)^{-\frac{1}{2}(g_0+nm+2)}}{\Gamma\left(\frac{g_0}{2}\right) 2^{\frac{g_0+nm}{2}}} \exp\left\{-\frac{1}{2} \text{tr} \left[\left(\sum_{i=1}^n (y_i - X_i\mu_i)' (y_i - X_i\mu_i) + 1 \right) (\sigma^2)^{-1} \right] \right\}
\end{aligned}$$

sendo tr a função *traço* de uma matriz.

Chegamos, então à seguinte relação

$$P(\sigma^2|y, \mu, \Omega) \propto (\sigma^2)^{-\frac{1}{2}(g_0+nm+1+1)} \exp\left\{-\frac{1}{2} \text{tr} \left[\left(\sum_{i=1}^n (y_i - X_i\mu_i)' (y_i - X_i\mu_i) + 1 \right) (\sigma^2)^{-1} \right] \right\}$$

que, segundo Anderson (1984), é suficiente pra dizer que

$$[\sigma^2|y, \mu, \Omega] \sim IW_1 \left((g_0 + nm), \left(G_0 + \sum_{i=1}^n (y_i - X_i\mu_i)' (y_i - X_i\mu_i) \right) \right)$$

sendo $G_0 = 1$. Com isso mostramos a distribuição condicional completa de σ^2 do Modelo Bayesiano.

Distribuição condicional completa de Ω

Agora mostraremos que a distribuição condicional completa de Ω é dada por

$$[\Omega|y, \mu, \sigma^2] \sim IW_p(d, D),$$

$$d = d_0 + n \quad e \quad D = D_0 + \sum_{i=1}^n (\mu_i - u_i) (\mu_i - u_i)'.$$

Devido à independência assumida entre σ^2 e Ω ; e percebendo que, pela estrutura construída no Modelo Bayesiano, Ω e y também são independentes, temos a igualdade

$$P(\Omega|y, \mu, \sigma^2) = P(\Omega|\mu).$$

Trabalhamos, então, com a seguinte relação dada pelo teorema de Bayes

$$P(\Omega|\mu) \propto P(\Omega) P(\mu|\Omega). \tag{A.6}$$

Sabemos que a distribuição *a priori* de Ω , $P(\Omega)$, é dada por (A.4), além disso, temos que as distribuições *a priori*, $\mu_i|\Omega$, são independentes para diferentes i 's e dadas por (A.2). Substituindo estas equações na relação (A.6), temos

$$\begin{aligned} P(\Omega|y, \mu, \sigma^2) &\propto P(\Omega) \prod_{i=1}^n P(\mu_i|\Omega) \\ &= |D_o|^{\frac{d_o}{2}} |\Omega|^{-\frac{1}{2}(d_o+p+1)} \exp\left\{-\frac{1}{2}\text{tr}(D_o\Omega^{-1})\right\} \prod_{i=1}^n \left[\frac{|\Omega|^{-\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \exp\left\{-\frac{1}{2}(\mu_i - u_i)' \Omega^{-1} (\mu_i - u_i)\right\} \right] \\ &= \frac{|D_o|^{\frac{d_o}{2}}}{(2\pi)^{\frac{np}{2}}} |\Omega|^{-\frac{1}{2}(d_o+n+p+1)} \exp\left\{-\frac{1}{2}\left[\text{tr}(D_o\Omega^{-1}) + \sum_{i=1}^n (\mu_i - u_i)' \Omega^{-1} (\mu_i - u_i)\right]\right\} \end{aligned}$$

note que $(\mu_i - u_i)' \Omega^{-1} (\mu_i - u_i)$ resulta num escalar, logo, podemos tomar a igualdade

$$(\mu_i - u_i)' \Omega^{-1} (\mu_i - u_i) = \text{tr}[(\mu_i - u_i)' \Omega^{-1} (\mu_i - u_i)].$$

Além disso, aplicando as propriedades da função *traço* de matrizes, temos essa outra igualdade

$$\text{tr}(\mu_i - u_i)' \Omega^{-1} (\mu_i - u_i) = \text{tr}((\mu_i - u_i) (\mu_i - u_i)' \Omega^{-1}).$$

De volta à equação, temos

$$\begin{aligned} P(\Omega|y, \mu, \sigma^2) &\propto \\ &\frac{|D_o|^{\frac{d_o}{2}}}{(2\pi)^{\frac{np}{2}}} |\Omega|^{-\frac{1}{2}(d_o+n+p+1)} \exp\left\{-\frac{1}{2}\left[\text{tr}(D_o\Omega^{-1}) + \sum_{i=1}^n \text{tr}[(\mu_i - u_i) (\mu_i - u_i)' \Omega^{-1}]\right]\right\} \\ &= \frac{|D_o|^{\frac{d_o}{2}}}{(2\pi)^{\frac{np}{2}}} |\Omega|^{-\frac{1}{2}(d_o+n+p+1)} \exp\left\{-\frac{1}{2}\text{tr}\left[D_o\Omega^{-1} + \sum_{i=1}^n [(\mu_i - u_i) (\mu_i - u_i)' \Omega^{-1}]\right]\right\} \\ &= \frac{|D_o|^{\frac{d_o}{2}}}{(2\pi)^{\frac{np}{2}}} |\Omega|^{-\frac{1}{2}(d_o+n+p+1)} \exp\left\{-\frac{1}{2}\text{tr}\left[\left(D_o + \sum_{i=1}^n (\mu_i - u_i) (\mu_i - u_i)'\right) \Omega^{-1}\right]\right\}, \end{aligned}$$

essa função densidade nos leva à seguinte distribuição (ver Anderson, 1984)

$$[\Omega|y, \mu, \sigma^2] \sim IW_p\left((d_o + n), \left(D_o + \sum_{i=1}^n (\mu_i - u_i) (\mu_i - u_i)'\right)\right).$$

Chegamos, então, à distribuição condicional completa de Ω .

A.3 Alguns resultados matriciais

Nessa seção citamos alguns resultados sobre matrizes positivas definidas retirados do livro de Graybill (1983) que são utilizados na seção seguinte. A fim de facilitar uma possível consulta ao livro de Graybill, mantivemos a numeração original dos teoremas e corolários.

Teorema 12.2.1

Se $A = a_{ij}$ é uma matriz $(k \times k)$ positiva definida, então valem os seguintes resultados

(1b) O posto de A é igual a k .

(2b) $a_{ii} > 0$ para todo $i=1,2,\dots,k$.

(3b) Para qualquer matriz P $(k \times k)$ de posto completo, $P'AP$ é uma matriz positiva definida

(Em particular, A^{-1} é positiva definida).

Corolário 12.2.2

Se B é uma matriz $(p \times k)$ de posto r , então

(1) $B'B$ e BB' são matrizes não-negativas.

(2) $B'B$ é uma matriz positiva semidefinida se $r < k$.

(3) $B'B$ é uma matriz positiva definida se $r = k$.

Teorema 12.2.14

Sejam A , B e C matrizes simétricas $(k \times k)$.

(2) Se A e B são matrizes não singulares, ou seja, têm posto completo, e $(A - B)$ é positiva definida, então $(B^{-1} - A^{-1})$ é positiva definida.

Teorema 12.2.18

Seja $\{A_1, A_2, \dots, A_n\}$ uma coleção de matrizes $k \times k$ positivas definidas e seja $\{a_1, a_2, \dots, a_n\}$ um conjunto de escalares positivos. A matriz B é também positiva definida, no qual

$$B = \sum_{i=1}^n a_i A_i$$

A.4 Matriz de variâncias-covariâncias do Modelo Bayesiano

Na Seção 2.4 dissemos que a matriz de variâncias-covariâncias do estimador de μ_i referente ao Modelo Bayesiano de Marshall e Bradlow é “menor” que as matrizes de variâncias-covariâncias dos estimadores de μ_i quando estes são obtidos pelos métodos de composição ou decomposição. Provamos, nesta seção, este fato.

Dizemos que uma matriz A é menor que uma a outra matriz B , se a diferença entre essas matrizes, $(B - A)$, resulta em matriz positiva definida.

Modelo Bayesiano x Modelo de Decomposição

Primeiro, mostramos que a matriz de variância-covariâncias do Modelo Bayesiano é menor ou igual que a de um modelo de perfil-completo.

Assumindo, que o método adotado é o de perfil-completo e que um modelo de regressão linear foi utilizado, temos que a matriz de variâncias-covariâncias do estimador de μ_i é

$$(X'X)^{-1}\sigma^2, \text{ de dimensão } (p \times p).$$

Neste caso, X deve ter posto completo.

Pelo Corolário 12.2.2 e Teorema 12.2.1(3), temos que $(X_i'X_i)^{-1}$ é positiva definida e pelo Teorema 12.2.18 concluímos que $(X_i'X_i)^{-1}\sigma^2$ também é uma matriz positiva definida, pois σ^2 é um escalar positivo. Além disso, pelo Teorema 12.2.1(1b), temos que $(X_i'X_i)^{-1}\sigma^2$ tem posto completo.

Ou seja, a matriz de variâncias-covariâncias do estimador obtido pelo modelo de decomposição, $(X_i'X_i)^{-1}\sigma^2$, é uma matriz positiva definida e tem posto completo.

Por outro lado, a matriz de variâncias-covariâncias do estimador do Modelo Bayesiano é

$$A_i^{-1} = \left[\Omega^{-1} + \frac{X_i'X_i}{\sigma^2} \right]^{-1}, \text{ de dimensão } (p \times p)$$

Como a matriz Ω é, por definição, positiva definida, temos pelos Teoremas 12.2.18 e 12.2.1(3b) que A_i^{-1} é também positiva definida e tem posto completo.

Temos, então, que A_i e $\left[\frac{X_i'X_i}{\sigma^2} \right]$, são matrizes com dimensão $(p \times p)$ e com posto completo.

Pelo Teorema 12.2.14 podemos mostrar que a matriz de variâncias-covariâncias do estimador do Modelo Bayesiano, $A_i^{-1} = A^{-1}$, é menor que a matriz de variâncias-covariâncias do estimador do modelo de decomposição, $(X_i'X_i)^{-1}\sigma^2 = B^{-1}$, pois temos que

$$(A - B) = \left[\Omega^{-1} + \frac{X_i'X_i}{\sigma^2} \right] - \left[\frac{X_i'X_i}{\sigma^2} \right] = \Omega^{-1} \text{ é positiva definida}$$

logo,

$$(B^{-1} - A^{-1}) = (X_i'X_i)^{-1}\sigma^2 - A_i^{-1} \text{ também é uma matriz positiva definida.}$$

Dessa maneira, as estimativas de μ_i provenientes do Modelo Bayesiano são mais precisas que as provenientes do modelo referente ao método de perfil-completo no sentido de apresentarem matriz de variâncias-covariâncias menor.

Modelo Bayesiano x Modelo de Composição

Se o método adotado é o de composição assumimos, como matriz de variâncias-covariâncias do estimador de μ_i , a matriz positiva definida Ω com dimensão $(p \times p)$ e com posto completo.

Dessa maneira, temos que A_i e Ω^{-1} são matrizes com dimensão $(p \times p)$ e com posto completo.

Com o mesmo raciocínio, pelo Teorema 12.2.14 podemos mostrar que a matriz de variâncias-covariâncias do estimador do Modelo bayesiano, $A_i^{-1} = A^{-1}$, é menor que a matriz de variâncias-covariâncias do estimador do modelo de composição, $\Omega = B^{-1}$, pois temos que

$$(A - B) = \left[\frac{X_i'X_i}{\sigma^2} \right] \text{ é positiva definida}$$

logo,

$$(B^{-1} - A^{-1}) = \Omega - [A_i^{-1}] \text{ é positiva definida.}$$

Mostramos assim, que as estimativas de μ_i provenientes do Modelo Bayesiano são mais precisas que as estimativas obtidas pelo método de composição no sentido de apresentarem matriz de variâncias-covariâncias menor.

Apêndice B

Questionário aplicado e respostas dadas pelos indivíduos

Figura B.1: Parte sócio-econômica do Questionário

Questionário desenvolvido para avaliar a preferência dos alunos na escolha da Instituição (Faculdade) para realizar um curso de pós-graduação.

Toda informação coletada neste questionário terá finalidade exclusivamente acadêmica.

Informações Pessoais

Sexo: _____ Idade: _____ Estado civil: _____

Cidade onde reside _____

Estudos e Trabalho

Possui bolsa de estudo? Sim () Não () Percentual de desconto: _____

Curso(s) de graduação concluído(s) _____

Área de atuação profissional (Finança, Economia... _____

Tempo médio gasto para se deslocar de seu trabalho (ou residência, se for o caso) até a Instituição na qual realiza o curso de pós-graduação, num dia de semana:

Meio(s) de locomoção utilizado(s) para esse trajeto Carro () Ônibus () Metrô () Outros ()

Lazer

Atividade(s) que lhe agradam e que você pratica com frequência:

Ler livros () Viajar () Ir a shows () Ir ao cinema () Ir ao teatro () Praticar esportes ()

Figura B.2: Parte de decomposição do Questionário

Suponha que você precise escolher uma Instituição (Faculdade) para realizar um curso de pós-graduação.
Listamos abaixo alguns perfis de Instituições (Faculdades) que oferecem cursos de pós-graduação.
Dê notas de 0 a 100 para cada uma dessas Instituições, conforme sua preferência por cada uma delas.

Preferência: 0 100

		Nota
I	Instituição nova e com alta qualificação para o curso desejado. Todas as aulas em dias de semana. Disciplinas com maior carga horária, mas em menor quantidade. Corpo docente com formação acadêmica razoável, mas com grande experiência de mercado. Ter 2 disciplinas por dia, cada disciplina com 2 dias de aula por semana.	
II	Instituição nova e com alta qualificação para o curso desejado. Todas as aulas em dias de semana. Disciplinas com carga horária menor, mas em maior quantidade. Corpo docente com formação acadêmica razoável, mas com grande experiência de mercado. Ter apenas 1 disciplina por dia, cada disciplina com 1 dia de aula por semana.	
III	Instituição nova e com alta qualificação para o curso desejado. Menos aulas em dias de semana e algumas aulas aos sábados de manhã. Disciplinas com maior carga horária, mas em menor quantidade. Corpo docente com excelente formação acadêmica, mas sem grande experiência de mercado. Ter apenas 1 disciplina por dia, cada disciplina com 1 dia de aula por semana.	
IV	Instituição tradicional (reconhecida) com razoável qualificação para o curso desejado. Menos aulas em dias de semana e algumas aulas aos sábados de manhã. Disciplinas com maior carga horária, mas em menor quantidade. Corpo docente com formação acadêmica razoável, mas com grande experiência de mercado. Ter apenas 1 disciplina por dia, cada disciplina com 1 dia de aula por semana.	
V	Instituição tradicional (reconhecida) com razoável qualificação para o curso desejado. Todas as aulas em dias de semana. Disciplinas com maior carga horária, mas em menor quantidade. Corpo docente com excelente formação acadêmica, mas sem grande experiência de mercado. Ter 2 disciplinas por dia, cada disciplina com 2 dias de aula por semana.	
VI	Instituição tradicional (reconhecida) com razoável qualificação para o curso desejado. Menos aulas em dias de semana e algumas aulas aos sábados de manhã. Disciplinas com carga horária menor, mas em maior quantidade. Corpo docente com formação acadêmica razoável, mas com grande experiência de mercado. Ter 2 disciplinas por dia, cada disciplina com 2 dias de aula por semana.	
VII	Instituição nova e com alta qualificação para o curso desejado. Menos aulas em dias de semana e algumas aulas aos sábados de manhã. Disciplinas com carga horária menor, mas em maior quantidade. Corpo docente com excelente formação acadêmica, mas sem grande experiência de mercado. Ter 2 disciplinas por dia, cada disciplina com 2 dias de aula por semana.	
VIII	Instituição tradicional (reconhecida) com razoável qualificação para o curso desejado. Todas as aulas em dias de semana. Disciplinas com carga horária menor, mas em maior quantidade. Corpo docente com excelente formação acadêmica, mas sem grande experiência de mercado. Ter apenas 1 disciplina por dia, cada disciplina com 1 dia de aula por semana.	

Figura B.3: Parte de composição do Questionário

Suponha que você precise escolher uma Instituição (Faculdade) para realizar um curso de pós-graduação.
Listamos nestes 8 blocos alguns fatores importantes na escolha de uma Instituição para se realizar um curso de pós-graduação.

Dê notas de 0 a 10 para cada uma dessas frases, conforme sua preferência por cada uma delas.

Preferência: 0 10

		Nota
Tradição do Instituto e Qualidade do curso	Instituição tradicional (reconhecida) com razoável qualificação para o curso desejado.	
	Instituição nova e com alta qualificação para o curso desejado.	
Dias das aulas	Todas as aulas em dias de semana.	
	Menos aulas em dias de semana e algumas aulas aos sábados de manhã.	
Quantidade de disciplinas	Disciplinas com maior carga horária, mas em menor quantidade.	
	Disciplinas com carga horária menor, mas em maior quantidade.	
Corpo docente	Corpo docente com formação acadêmica razoável, mas com grande experiência de mercado.	
	Corpo docente com excelente formação acadêmica, mas sem grande experiência de mercado.	
Grade horária das disciplinas	Ter apenas 1 disciplina por dia, cada disciplina com 1 dia de aula por semana.	
	Ter 2 disciplinas por dia, cada disciplina com 2 dias de aula por semana.	

Agora distribua 100 pontos dentre estes fatores, conforme a importância de cada um deles na sua concepção.

Importância: menos pontos mais pontos

PONTOS	
Tradição do Instituto e Qualidade do curso	
Dias das aulas	
Quantidade de disciplinas	
Corpo docente	
Grade horária das disciplinas	
Soma total dos PONTOS	100

Tabela B.1: Avaliações dadas pelos indivíduos aos 8 estímulos apresentados - Fase de decomposição

indivíduo	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>	<i>VII</i>	<i>VIII</i>
1	60	50	55	45	55	52	58	50
2	95	90	80	75	60	60	65	75
3	70	90	50	30	10	30	40	50
4	70	70	80	70	60	50	70	80
5	65	80	70	60	55	50	40	45
6	80	90	60	70	70	40	50	50
7	70	60	50	40	60	50	60	50
8	70	75	65	55	57	56	63	57
9	40	90	20	30	50	15	10	80
10	80	80	70	70	80	70	70	80
11	90	85	75	40	45	40	75	40
12	70	80	60	60	70	80	85	75
13	50	50	40	50	80	60	50	70
14	70	60	50	40	60	50	50	60
15	70	70	30	30	30	30	30	40
16	70	60	100	60	65	50	50	50
17	70	80	70	50	50	50	50	90
18	30	40	70	80	90	80	60	80
19	30	40	10	20	60	50	10	30
20	100	100	80	60	70	50	90	50
21	100	90	70	80	90	80	70	80
22	40	50	30	60	60	70	30	70
23	70	65	60	65	60	60	50	55
24	50	40	30	60	80	90	20	70
25	70	70	65	60	55	50	50	50
26	85	80	70	60	55	70	70	55
27	60	60	80	70	40	70	80	60
28	65	50	50	75	85	75	60	85
29	50	60	70	40	50	50	60	80

Tabela B.2: Avaliações dadas pelos indivíduos aos atributos e níveis - Fase de composição

indivíduo	Atributos					Níveis de atributos									
	A	B	C	D	E	A1	A2	B1	B2	C1	C2	D1	D2	E1	E2
1	25	5	25	35	10	5,5	7	5	6	8	4	6	5	5	6
2	25	15	10	45	5	7	9	10	6	9	7	8	7	10	7
3	80	4	4	10	2	8	9	6	4	2	4	6	4	6	4
4	45	10	0	30	15	3	8	3	8	8	8	9	3	10	0
5	50	5	15	20	10	8	6	4	7	5	8	8	7	8	7
6	70	16	2	2	10	8	2	9	1	7	3	8	2	9	1
7	30	20	10	30	10	6	7	8	6	7	6	6	7	6	7
8	35	15	5	30	15	3	8	8	5	8	9	8	8	9	7
9	30	15	15	30	10	5	5	7	2	7	4	7	4	8	2
10	23	23	19	23	12	5	5	10	5	10	10	5	5	10	10
11	60	0	5	25	10	0	10	10	0	10	0	6	8	5	8
12	30	10	20	30	10	8	9	8	2	8	3	7	8	7	4
13	60	20	5	10	5	10	5	10	5	8	6	9	6	6	9
14	40	10	10	30	10	7	8	10	5	9	7	7	7	5	5
15	25	25	25	25	0	4	8	10	3	7	5	6	6	8	2
16	50	5	5	20	20	5	10	5	10	10	3	7	7	10	5
17	30	20	10	30	10	9	8	8	5	8	6	9	7	3	9
18	54	9	5	9	23	9	8	7	8	9	8	9	8	9	8
19	30	30	5	30	5	6	6	10	0	8	4	7	6	10	0
20	30	10	10	30	20	6	10	10	0	10	5	10	7	10	7
21	30	15	15	30	10	10	10	10	8	10	8	8	8	10	8
22	60	5	10	15	10	7	7	10	5	10	8	7	7	10	6
23	30	20	15	30	5	9	8	9	6	9	7	8	7	6	9
24	50	15	0	30	5	8	5	8	4	9	3	8	6	9	3
25	50	5	10	30	5	6	8	8	3	5	5	7	5	7	5
26	30	15	10	35	10	6	10	6	7	8	4	8	6	9	6
27	30	10	10	30	20	7	9	7	9	7	8	7	9	10	6
28	50	10	10	20	10	10	5	10	0	8	5	7	7	5	5
29	20	20	30	25	5	8	9	7	8	6	8	7	8	9	7

Apêndice C

Métodos de estimação

No processo de estimação dos modelos citamos o método de amostragem de Gibbs (Geman e Geman, 1984) para simular a distribuição condicional *a posteriori* (2.3) dos vetores utilidade sob o Modelo Bayesiano.

O método de amostragem de Gibbs pertence aos métodos denominados “MCMC”, que são uma combinação dos processos de simulação de distribuições via Cadeias de Markov com o método de aproximação de integrais de Monte Carlo (Hastings, 1970).

Quando há a necessidade de se obter a função de distribuição de probabilidade de uma variável aleatória, em que seu cálculo é analiticamente inviável, ou então, resulta numa função de distribuição da qual não é fácil de se obter realizações (gerar valores) recorre-se, por exemplo, a processos de simulação de distribuições.

Estes processos são algoritmos que permitem simular uma distribuição sem precisar de sua função densidade de probabilidade; utilizam para isso, processos estocásticos temporais fáceis de simular e que convergem para a distribuição desejada. Geralmente o processo estocástico envolvido é uma Cadeia de Markov homogênea que tenha como distribuição de equilíbrio exatamente a distribuição de interesse. Para construir Cadeias de Markov com distribuição de equilíbrio determinadas, existe o algoritmo de Metropolis-Hastings, que foi desenvolvido para isso, veja o artigo de Smith e Roberts (1993).

Neste apêndice fazemos uma introdução do método de amostragem de Gibbs, que, na realidade, é um caso particular do algoritmo de Metropolis-Hastings (), tomando como base o livro de Paulino, Turkman e Murteira (2003). Antes disso, abordamos, na primeira seção, o método de Monte Carlo.

C.1 Método Monte Carlo

Considere que X seja uma variável contínua aleatória e $f(x)$ sua função densidade de probabilidade. Seja $g()$ uma função integrável e $E[g(x)]$ o valor esperado de $g(x)$, dado pela seguinte integral

$$E[g(x)] = \int_b^a g(x)f(x)dx, \quad (C.1)$$

na qual x e $g(x)$ podem ser vetores, e a e b são seus limites de integração, ou seja $a < g(x) < b$.

Suponha $x_1, x_2, \dots, x_n$ uma amostra aleatória gerada a partir da distribuição $f(x)$. O método de Monte Carlo aproxima o valor de (C.1) pela seguinte média empírica

$$\hat{E}[g(x)] = \frac{1}{n} \sum_{i=1}^n g(x_i)$$

que, pela Lei Forte dos Grandes Números, converge quase certamente para $E[g(x)]$.

O erro padrão associado à estimativa de Monte Carlo pode ser dado por

$$EP = \frac{1}{\sqrt{n(n-1)}} \left\{ \sum_{i=1}^n \left[g(\theta_i) - \frac{1}{n} \sum_{i=1}^n g(\theta_i) \right]^2 \right\}^{\frac{1}{2}}.$$

C.2 Amostrador Gibbs

Suponha que desejamos estimar o valor esperado de $g(x)$ dado por (C.1), através do método de Monte Carlo e que, por algum motivo, não é possível gerar a amostra $x_1, x_2, \dots, x_n$ de $f(x)$.

Este problema pode ser resolvido utilizando o método de Metrópolis-Hastings para simular a distribuição $f(x)$ e, então, chegar à amostra desejada. O amostrador de Gibbs é um caso particular desse método de simulação de distribuição, para casos nos quais a distribuição de interesse é multivariada (Geman e Geman, 1984). Gelfand e Smith (1990) mostram como aplicar esse algoritmo para simular distribuições *a posteriori*, resolvendo assim problemas de estimação em modelos bayesianos.

Seja $f(x_1, x_2, \dots, x_k)$ a função densidade de probabilidade conjunta das variáveis aleatórias $X_1, \dots, X_k$ e seja Θ_i o suporte da distribuição de x_i .

O amostrador de Gibbs se baseia no resultado de Besag (1974), que diz que se a distribuição conjunta de $(X_1, X_2, \dots, X_k)$ for positiva em $\Theta_1 \times \dots \times \Theta_k$ então ela é unicamente determinada pelas distribuições condicionais completas dadas por

$$\begin{aligned} &f(x_1|x_2, x_3, \dots, x_k) \\ &f(x_2|x_1, x_3, \dots, x_k) \\ &f(x_3|x_1, x_2, x_4, \dots, x_k) \\ &\dots \\ &h(x_k|x_1, x_2, x_3, \dots, x_{k-1}) \end{aligned}$$

O método de amostragem de Gibbs utiliza estas distribuições para compor a matriz de transição da Cadeia de Markov e obter como distribuição de equilíbrio (estacionária) a distribuição conjunta $f(\mathbf{x}) = f(x_1, x_2, \dots, x_k)$.

Suponha, por exemplo, que queiramos obter o valor de $E[X_1]$. Para isso, precisaríamos da função densidade de probabilidade da distribuição marginal de X_1 , que é dada por

$$f_{X_1}(x_1) = \int_{\Theta_k} \dots \int_{\Theta_2} f(x_1, x_2, \dots, x_k) dx_2 \dots dx_k \quad (\text{C.2})$$

porém, o cálculo de (C.2) pode ser difícil ou até impossível. Neste caso, seria interessante obter uma amostra da distribuição marginal de X_1 , para poder aplicar o método de Monte Carlo. O método de amostragem de Gibbs permite obter essa amostra sem precisar calcular (C.2).

Para iniciar o processo de simulação da distribuição será necessário tomar valores iniciais para os elementos do vetor $\mathbf{x}$. Algoritmo:

Seja, então, $\mathbf{x}^0 = (x_1^{(0)}, \dots, x_k^{(0)})$ um valor arbitrário inicial para o vetor $\mathbf{x}$.

Passo 1

Gera-se um valor aleatório, $x_1^{(1)}$, a partir da distribuição de $f_{X_1}(x_1|x_2^{(0)}, \dots, x_k^{(0)})$,

Gera-se um valor aleatório, $x_2^{(1)}$, a partir da distribuição de $f_{X_2}(x_2|x_1^{(1)}, x_3^{(0)}, \dots, x_k^{(0)})$,

Gera-se um valor aleatório, $x_3^{(1)}$, a partir da distribuição de $f_{X_3}(x_3|x_1^{(1)}, x_2^{(1)}, x_4^{(0)}, \dots, x_k^{(0)})$,

...

Gera-se um valor aleatório, $x_k^{(1)}$, a partir da distribuição de $f_{X_k}(x_k|x_1^{(1)}, \dots, x_{k-1}^{(1)})$.

Após gerar os k elementos, a cadeia transita de $\mathbf{x}^0 = (x_1^{(0)}, \dots, x_k^{(0)})$ para $\mathbf{x}^1 = (x_1^{(1)}, \dots, x_k^{(1)})$, completando a primeira *iteração*.

Passo 2

Seja t o indexador das iterações. Repete-se o passo 1, tomando $\mathbf{x}^{t-1}$ como valor inicial para obter o novo vetor $\mathbf{x}^t$. Completando a iteração de número t .

Passo 3

Repete-se o passo 2, até se obter T iterações.

Os vetores $\mathbf{x}^1, \dots, \mathbf{x}^t, \dots, \mathbf{x}^T$ são realizações de uma Cadeia de Markov pois a iteração $t + 1$ depende apenas da iteração t (Paulino, Turkman e Murteira, 2003).

Temos que, quando $t \rightarrow \infty$, o conjunto $x_1^t, \dots, x_k^t$ tende em distribuição para um vetor aleatório cuja função densidade de probabilidade conjunta é $f(\mathbf{x})$.

Uma das questões que se levanta na utilização desse método de amostragem, está relacionada com a correlação que existe entre os valores gerados para $\mathbf{x}$. Para obtermos, então, uma amostra independente de $\mathbf{x}$, podemos, por exemplo, tomar valores suficientemente espaçados, pois isso irá reduzir a autocorrelação entre as iterações.

Uma outra alternativa é a geração de múltiplas Cadeias de Markov, que podem partir do mesmo vetor inicial (Gelfand e Smith, 1990) ou de vetores iniciais distintos (Gelman e Rubin, 1992). Em qualquer caso, deve-se utilizar os valores gerados após a cadeia ter atingido estado de equilíbrio, descartando-se o período que antecede esse fato, que é chamado de “aquecimento”.

Existem métodos que indicam se a cadeia já atingiu o estado de equilíbrio ou não, determinando as iterações de aquecimento, veja Paulino, Turkman e Murteira (2003). Ao final desta seção, descreveremos o método de diagnóstico de convergência, para múltiplas cadeias, desenvolvido por Gelman e Rubin em 1992.

Uma vez obtida a amostra de Gibbs de, por exemplo, m vetores $(1 \times k)$, $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)}$, as inferências podem ser feitas a partir do método de Monte Carlo. Suponha que desejamos obter uma estimativa da componente X_j do vetor $\mathbf{x}$, então toma-se

$$\hat{E}[X_j] = \frac{1}{m} \sum_{r=1}^m x_j^{(r)}$$

no qual, $x_j^{(r)}$ é a j -ésima componente de $\mathbf{x}^{(r)}$.

ou

$$\hat{E}[X_j] = \frac{1}{m} \sum_{r=1}^m E[X_j | x_l^{(r)}, l = 1, 2, \dots, k \text{ tal que } l \neq j]$$

Como foi dito, devemos analisar se a(s) cadeia(s) está(ão) no seu estado de equilíbrio ou não. Para isso, existem diversos métodos que diagnosticam a convergência de uma cadeia, dentre estes métodos, vamos descrever, brevemente, o sugerido por Gelman e Rubin (1992).

C.3 Método de diagnóstico de convergência para múltiplas cadeias

O método de Gelman e Rubin (1992) foi desenvolvido para averiguar a convergência do método de Gibbs, quando obtido através de múltiplas e independentes cadeias, simuladas a partir de pontos iniciais dispersos. Seu processo é descrito pelos autores em algumas etapas, que são:

1) Simula-se $m \geq 2$ sequências independentes e de tamanho $2n$, com valores iniciais gerados de uma distribuição sobredispersa relativamente à distribuição de equilíbrio. E, para amenizar o efeito causado pelos pontos iniciais, descarta-se as n primeiras iterações de cada uma.

2) Seja x um dos parâmetros escalares de interesse, para cada um deles, calcular as seguintes quantidades:

média dos n valores de x referentes à i -ésima cadeia:

$$\bar{x}_{i.} = \frac{1}{n} \sum_{j=1}^n x_{ij}$$

média de todos os mn valores de x gerados:

$$\bar{x}_{..} = \frac{1}{m} \sum_{i=1}^m \bar{x}_{i.}$$

variância entre as m médias, $\bar{x}_{i.}$:

$$\frac{B}{n} = \frac{1}{m-1} \sum_{i=1}^m (\bar{x}_{i.} - \bar{x}_{..})^2$$

variância dos n valores de x referentes à i -ésima cadeia:

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_{i.})^2$$

média das m variâncias de cada cadeia:

$$W = \frac{1}{m} \sum_{i=1}^m s_i^2$$

3) Estimar a média de X como a média dos nm valores simulados, x_{ij} , o já calculado $\bar{x}_{..}$.

$$\hat{\mu} = \bar{x}_{..} = \frac{1}{m} \sum_{i=1}^m \bar{x}_{i.}$$

4) Estimar a variância de X como uma média ponderada de B e W .

$$\hat{\sigma}^2 = \left(\frac{n-1}{n}\right)W + \left(\frac{1}{n}\right)B$$

essa quantidade sobrestima (superestima) σ^2 se a distribuição inicial for sobredispersa, mas é não viesada sob estacionariedade ou no limite $n \rightarrow \infty$.

5) O resultado é uma distribuição aproximada t-Student para x centrada em $\hat{\mu}$, com escala

$$\sqrt{\hat{V}} = \sqrt{\hat{\sigma}^2 + \frac{B}{mn}}$$

e graus de liberdade

$$df = \frac{2\hat{V}^2}{\text{var}(\hat{V})}$$

na qual,

$$\begin{aligned} \text{var}(\hat{V}) = & \left(\frac{n-1}{n}\right)^2 \frac{1}{m} \text{var}(s_i^2) + \left(\frac{m-1}{mn}\right)^2 \frac{2}{m-1} B^2 + 2 \frac{(m+1)(n-1)}{mn^2} \\ & \times \frac{n}{m} [\text{cov}(s_i^2, \bar{x}_{i.}) - 2\bar{x}_{..} \text{cov}(s_i^2, \bar{x}_{i.})] \end{aligned}$$

as variâncias e covariâncias são obtidas dos m valores amostrais de $\bar{x}_{i.}$ e s_i^2 . Note que $df \rightarrow \infty$ conforme $n \rightarrow \infty$.

6) Monitorar a convergência das simulações iterativas estimando o *fator de redução da escala*, que é o fator de redução que o parâmetro de escala da distribuição t-Student para X sofreria, se as simulações continuassem. Esse potencial de redução de escala é estimado por

$$\sqrt{\hat{R}} = \sqrt{\left(\frac{\hat{V}}{\bar{W}}\right) \frac{df}{(df - 2)}}$$

Valores altos de $\hat{R}$ nos levam a acreditar que mais iterações poderiam melhorar a simulação da distribuição. Valores de $\hat{R} \approx 1$, para todas as estimativas de escalares de interesse, sugerem o final do período de aquecimento da cadeia.

Esse método de Gelman e Rubin (1992) assume que os valores iniciais das cadeias sejam sobredispersos. Para isso, cria uma distribuição para gerar esses valores iniciais, que é centrada na moda da distribuição desejada, porém tem maior variabilidade. Esta estatística de Gelman e Rubin é calculada, por exemplo, nos pacotes CODA e Brugs do software R.

Apêndice D

Gráficos

Neste apêndice apresentamos os gráficos de diagnóstico de convergência de Gelman e Rubin (1992) e os gráficos de densidade dos parâmetros estimados pelos modelos da Seção 4.2.

D.1 Modelo Bayesiano I

O Modelo Bayesiano I possui a função de verossimilhança composta pelas avaliações feitas aos 8 estímulos apresentados a cada um dos entrevistados. Apresenta o vetor α para correção das escalas das avaliações.

Figura D.1: Gráficos da estatística R dos parâmetros alpha e sigma sob o Modelo Bayesiano I

Figura D.2: Gráficos de densidade dos parâmetros alpha e sigma sob o Modelo Bayesiano I

Figura D.3: Gráficos da estatística R do parâmetro omega sob o Modelo Bayesiano I

Figura D.4: Gráficos de densidade do parâmetro ω sob o Modelo Bayesiano I

Figura D.5: Gráficos da estatística R das utilidades sob o Modelo Bayesiano I

Figura D.6: Gráficos da estatística R das utilidades sob o Modelo Bayesiano I

Figura D.7: Gráficos da estatística R das utilidades sob o Modelo Bayesiano I

Figura D.8: Gráficos da estatística R das utilidades sob o Modelo Bayesiano I

Figura D.9: Gráficos de densidade das utilidades sob o Modelo Bayesiano I

Figura D.10: Gráficos de densidade das utilidades sob o Modelo Bayesiano I

Figura D.11: Gráficos de densidade das utilidades sob o Modelo Bayesiano I

Figura D.12: Gráficos de densidade das utilidades sob o Modelo Bayesiano I

D.2 Modelo Bayesiano II

O Modelo Bayesiano II possui a função de verossimilhança composta por apenas 5 das avaliações feitas aos estímulos apresentados. Apresenta o vetor α para correção das escalas das avaliações.

Figura D.13: Gráficos da estatística R dos parâmetros alpha e sigma sob o Modelo Bayesiano II

Figura D.14: Gráficos de densidade dos parâmetros alpha e sigma sob o Modelo Bayesiano II

Figura D.15: Gráficos da estatística R do parâmetro ω sob o Modelo Bayesiano II

Figura D.16: Gráficos de densidade do parâmetro ω sob o Modelo Bayesiano II

Figura D.17: Gráficos da estatística R das utilidades sob o Modelo Bayesiano II

Figura D.18: Gráficos da estatística R das utilidades sob o Modelo Bayesiano II

Figura D.19: Gráficos da estatística R das utilidades sob o Modelo Bayesiano II

Figura D.20: Gráficos da estatística R das utilidades sob o Modelo Bayesiano II

Figura D.21: Gráficos de densidade das utilidades sob o Modelo Bayesiano II

Figura D.22: Gráficos de densidade das utilidades sob o Modelo Bayesiano II

Figura D.23: Gráficos de densidade das utilidades sob o Modelo Bayesiano II

Figura D.24: Gráficos de densidade das utilidades sob o Modelo Bayesiano II

D.3 Modelo Bayesiano III

A função de verossimilhança do Modelo Bayesiano III é composta pelas 8 avaliações feitas aos estímulos apresentados a cada indivíduo. A escala dessas avaliações foi alterada previamente, para que coincidissem com a escala das avaliações da parte de composição. Com essa medida, não há, teoricamente, necessidade de introduzir o vetor de correção de escala α .

Figura D.25: Gráfico da estatística R do parâmetro sigma sob o Modelo Bayesiano III

Figura D.26: Gráfico de densidade do parâmetro sigma sob o Modelo Bayesiano III

Figura D.27: Gráficos da estatística R do parâmetro ω sob o Modelo Bayesiano III

Figura D.28: Gráficos de densidade do parâmetro ω sob o Modelo Bayesiano III

Figura D.29: Gráficos da estatística R das utilidades sob o Modelo Bayesiano III

Figura D.30: Gráficos da estatística R das utilidades sob o Modelo Bayesiano III

Figura D.31: Gráficos da estatística R das utilidades sob o Modelo Bayesiano III

Figura D.32: Gráficos da estatística R das utilidades sob o Modelo Bayesiano III

Figura D.33: Gráficos de densidade das utilidades sob o Modelo Bayesiano III

Figura D.34: Gráficos de densidade das utilidades sob o Modelo Bayesiano III

Figura D.35: Gráficos de densidade das utilidades sob o Modelo Bayesiano III

Figura D.36: Gráficos de densidade das utilidades sob o Modelo Bayesiano III

D.4 Modelo Bayesiano IV

A função de verossimilhança do Modelo Bayesiano IV é composta pelas 8 avaliações feitas aos estímulos apresentados a cada indivíduo. A escala dessas avaliações foi alterada previamente, para que coincidissem com a escala das avaliações da parte de composição. Com essa medida, não há, teoricamente, necessidade de introduzir o vetor de correção de escala α .

Figura D.37: Gráficos da estatística R dos parâmetros alpha e sigma sob o Modelo Bayesiano IV

Figura D.38: Gráficos de densidade dos parâmetros alpha e sigma sob o Modelo Bayesiano IV

Figura D.39: Gráficos da estatística R do parâmetro ω sob o Modelo Bayesiano IV

Figura D.40: Gficos de densidade do parmetro omega sob o Modelo Bayesiano IV

Figura D.41: Gráficos da estatística R das utilidades sob o Modelo Bayesiano IV

Figura D.42: Gráficos da estatística R das utilidades sob o Modelo Bayesiano IV

Figura D.43: Gráficos da estatística R das utilidades sob o Modelo Bayesiano IV

Figura D.44: Gráficos da estatística R das utilidades sob o Modelo Bayesiano IV

Figura D.45: Gráficos de densidade das utilidades sob o Modelo Bayesiano IV

Figura D.46: Gráficos de densidade das utilidades sob o Modelo Bayesiano IV

Figura D.47: Gráficos de densidade das utilidades sob o Modelo Bayesiano IV

Figura D.48: Gráficos de densidade das utilidades sob o Modelo Bayesiano IV

Apêndice E

Estimativas obtidas no ajuste dos modelos

Anexamos neste apêndice as estimativas obtidas pelos ajustes dos modelos da Seção 4.2.

E.1 Modelo de Decomposição

Estimativas das utilidades (coeficientes dos modelos de regressão linear) dos níveis referentes a cada indivíduo. Os dados utilizados na obtenção dessas estimativas são as notas (0 a 100) dadas aos 8 estímulos no método de decomposição.

As tabelas contém a estimativa do parâmetro, dada pelo método de Mínimos Quadrados (Estimativa), o Erro Padrão a ela associado e os Limites Inferior (LI) e Superior (LS) do Intervalo de Confiança de 95%.

Tabela E.1: Estimativas das utilidades sob o modelo de decomposição

Indivíduo	Parâmetro	Estimativa	Erro Padrão	$LI_{(IC95\%)}$	$LS_{(IC95\%)}$
1	<i>cte</i>	47,25	1,96	38,81	55,69
1	a_1	5,25	1,6	-1,64	12,14
1	b_1	-1,25	1,6	-8,14	5,64
1	c_1	-1,25	1,6	-8,14	5,64
1	d_1	2,75	1,6	-4,14	9,64
1	e_1	6,25	1,6	-0,64	13,14
2	<i>cte</i>	85	6,12	58,65	111,35
2	a_1	15	5	-6,51	36,51
2	b_1	-10	5	-31,51	11,51
2	c_1	-5	5	-26,51	16,51
2	d_1	-10	5	-31,51	11,51
2	e_1	-10	5	-31,51	11,51
3	<i>cte</i>	50	6,85	20,54	79,46
3	a_1	32,5	5,59	8,45	56,55
3	b_1	-17,5	5,59	-41,55	6,55
3	c_1	12,5	5,59	-11,55	36,55
3	d_1	-17,5	5,59	-41,55	6,55
3	e_1	-17,5	5,59	-41,55	6,55
4	<i>cte</i>	70	6,85	40,54	99,46
4	a_1	7,5	5,59	-16,55	31,55
4	b_1	-2,5	5,59	-26,55	21,55
4	c_1	-2,5	5,59	-26,55	21,55
4	d_1	7,5	5,59	-16,55	31,55
4	e_1	-12,5	5,59	-36,55	11,55
5	<i>cte</i>	71,25	9,8	29,07	113,43
5	a_1	11,25	8	-23,19	45,69
5	b_1	-6,25	8	-40,69	28,19
5	c_1	-8,75	8	-43,19	25,69
5	d_1	-11,25	8	-45,69	23,19
5	e_1	-11,25	8	-45,69	23,19
6	<i>cte</i>	82,5	11,04	35	130
6	a_1	12,5	9,01	-26,28	51,28
6	b_1	-17,5	9,01	-56,28	21,28
6	c_1	-12,5	9,01	-51,28	26,28
6	d_1	-12,5	9,01	-51,28	26,28
6	e_1	-7,5	9,01	-46,28	31,28
7	<i>cte</i>	50	0	50	50
7	a_1	10	0	10	10
7	b_1	-10	0	-10	-10
7	c_1	0	0	0	0
7	d_1	0	0	0	0
7	e_1	10	0	10	10
8	<i>cte</i>	60,75	1,79	53,07	68,43
8	a_1	12	1,46	5,73	18,27

Tabela E.2: Estimativas das utilidades sob o modelo de decomposição

Indivíduo	Parâmetro	Estimativa	Erro Padrão	$LI_{(IC95\%)}$	$LS_{(IC95\%)}$
8	b_1	-5	1,46	-11,27	1,27
8	c_1	1	1,46	-5,27	7,27
8	d_1	-3,5	1,46	-9,77	2,77
8	e_1	-1,5	1,46	-7,77	4,77
9	cte	75	6,31	47,84	102,16
9	a_1	-3,75	5,15	-25,93	18,43
9	b_1	-46,25	5,15	-68,43	-24,07
9	c_1	13,75	5,15	-8,43	35,93
9	d_1	-3,75	5,15	-25,93	18,43
9	e_1	-26,25	5,15	-48,43	-4,07
10	cte	80	0	80	80
10	a_1	0	0	0	0
10	b_1	-10	0	-10	-10
10	c_1	0	0	0	0
10	d_1	0	0	0	0
10	e_1	0	0	0	0
11	cte	47,5	0	47,5	47,5
11	a_1	40	0	40	40
11	b_1	-7,5	0	-7,5	-7,5
11	c_1	-2,5	0	-2,5	-2,5
11	d_1	-5	0	-5	-5
11	e_1	2,5	0	2,5	2,5
12	cte	61,25	2,17	51,93	70,57
12	a_1	2,5	1,77	-5,11	10,11
12	b_1	-2,5	1,77	-10,11	5,11
12	c_1	15	1,77	7,39	22,61
12	d_1	0	1,77	-7,61	7,61
12	e_1	7,5	1,77	-0,11	15,11
13	cte	62,5	3,06	49,33	75,67
13	a_1	-17,5	2,5	-28,26	-6,74
13	b_1	-12,5	2,5	-23,26	-1,74
13	c_1	2,5	2,5	-8,26	13,26
13	d_1	7,5	2,5	-3,26	18,26
13	e_1	7,5	2,5	-3,26	18,26
14	cte	57,5	4,33	38,87	76,13
14	a_1	5	3,54	-10,21	20,21
14	b_1	-15	3,54	-30,21	0,21
14	c_1	0	3,54	-15,21	15,21
14	d_1	0	3,54	-15,21	15,21
14	e_1	5	3,54	-10,21	20,21
15	cte	52,5	3,06	39,33	65,67
15	a_1	17,5	2,5	6,74	28,26
15	b_1	-22,5	2,5	-33,26	-11,74
15	c_1	2,5	2,5	-8,26	13,26

Tabela E.3: Estimativas das utilidades sob o modelo de decomposição

Indivíduo	Parâmetro	Estimativa	Erro Padrão	$LI_{(IC95\%)}$	$LS_{(IC95\%)}$
15	d_1	-17,5	2,5	-28,26	-6,74
15	e_1	-2,5	2,5	-13,26	8,26
16	cte	66,25	12,34	13,14	119,36
16	a_1	13,75	10,08	-29,61	57,11
16	b_1	3,75	10,08	-39,61	47,11
16	c_1	-21,25	10,08	-64,61	22,11
16	d_1	6,25	10,08	-37,11	49,61
16	e_1	-8,75	10,08	-52,11	34,61
17	cte	72,5	11,04	25	120
17	a_1	7,5	9,01	-31,28	46,28
17	b_1	-17,5	9,01	-56,28	21,28
17	c_1	7,5	9,01	-31,28	46,28
17	d_1	2,5	9,01	-36,28	41,28
17	e_1	-17,5	9,01	-56,28	21,28
18	cte	70	6,85	40,54	99,46
18	a_1	-32,5	5,59	-56,55	-8,45
18	b_1	12,5	5,59	-11,55	36,55
18	c_1	-2,5	5,59	-26,55	21,55
18	d_1	17,5	5,59	-6,55	41,55
18	e_1	-2,5	5,59	-26,55	21,55
19	cte	45	15,31	-20,87	110,87
19	a_1	-17,5	12,5	-71,28	36,28
19	b_1	-17,5	12,5	-71,28	36,28
19	c_1	2,5	12,5	-51,28	56,28
19	d_1	-7,5	12,5	-61,28	46,28
19	e_1	12,5	12,5	-41,28	66,28
20	cte	65	8,66	27,74	102,26
20	a_1	35	7,07	4,58	65,42
20	b_1	-10	7,07	-40,42	20,42
20	c_1	-5	7,07	-35,42	25,42
20	d_1	-5	7,07	-35,42	25,42
20	e_1	5	7,07	-25,42	35,42
21	cte	95	0	95	95
21	a_1	0	0	0	0
21	b_1	-15	0	-15	-15
21	c_1	-5	0	-5	-5
21	d_1	-10	0	-10	-10
21	e_1	5	0	5	5
22	cte	70	3,06	56,83	83,17
22	a_1	-27,5	2,5	-38,26	-16,74
22	b_1	-7,5	2,5	-18,26	3,26
22	c_1	7,5	2,5	-3,26	18,26
22	d_1	-7,5	2,5	-18,26	3,26

Tabela E.4: Estimativas das utilidades sob o modelo de decomposição

Indivíduo	Parâmetro	Estimativa	Erro Padrão	$LI_{(IC95\%)}$	$LS_{(IC95\%)}$
22	e_1	-2,5	2,5	-13,26	8,26
23	cte	70	1,53	63,41	76,59
23	a_1	1,25	1,25	-4,13	6,63
23	b_1	-3,75	1,25	-9,13	1,63
23	c_1	-6,25	1,25	-11,63	-0,87
23	d_1	-8,75	1,25	-14,13	-3,37
23	e_1	-1,25	1,25	-6,63	4,13
24	cte	80	12,25	27,3	132,7
24	a_1	-40	10	-83,03	3,03
24	b_1	-10	10	-53,03	33,03
24	c_1	0	10	-43,03	43,03
24	d_1	-10	10	-53,03	33,03
24	e_1	10	10	-33,03	53,03
25	cte	66,25	2,17	56,93	75,57
25	a_1	10	1,77	2,39	17,61
25	b_1	-5	1,77	-12,61	2,61
25	c_1	-7,5	1,77	-15,11	0,11
25	d_1	-7,5	1,77	-15,11	0,11
25	e_1	-5	1,77	-12,61	2,61
26	cte	63,75	3,42	49,02	78,48
26	a_1	16,25	2,8	4,22	28,28
26	b_1	-1,25	2,8	-13,28	10,78
26	c_1	1,25	2,8	-10,78	13,28
26	d_1	-11,25	2,8	-23,28	0,78
26	e_1	3,75	2,8	-8,28	15,78
27	cte	50	6,12	23,65	76,35
27	a_1	10	5	-11,51	31,51
27	b_1	20	5	-1,51	41,51
27	c_1	5	5	-16,51	26,51
27	d_1	0	5	-21,51	21,51
27	e_1	-5	5	-26,51	16,51
28	cte	78,75	5,52	55	102,5
28	a_1	-23,75	4,51	-43,14	-4,36
28	b_1	-6,25	4,51	-25,64	13,14
28	c_1	-1,25	4,51	-20,64	18,14
28	d_1	3,75	4,51	-15,64	23,14
28	e_1	6,25	4,51	-13,14	25,64
29	cte	50	8,66	12,74	87,26
29	a_1	5	7,07	-25,42	35,42
29	b_1	-5	7,07	-35,42	25,42
29	c_1	10	7,07	-20,42	40,42
29	d_1	15	7,07	-15,42	45,42
29	e_1	-10	7,07	-40,42	20,42

Tabela E.5: Estimativas das 8 avaliações de perfil-completo pelo modelo de decomposição

indivíduo	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>	<i>VII</i>	<i>VIII</i>
1	59	51	54	46	56	51	59	49
2	90	95	80	75	65	60	65	70
3	65	95	48	33	15	28	43	45
4	65	75	83	68	65	53	68	75
5	71	74	65	65	49	45	45	51
6	88	83	65	65	63	45	45	58
7	70	60	50	40	60	50	60	50
8	71	74	64	56	56	55	64	58
9	45	85	21	29	45	16	9	85
10	80	80	70	70	80	70	70	80
11	90	85	75	40	45	40	75	40
12	71	79	61	59	69	81	84	76
13	53	48	40	50	78	60	50	73
14	68	63	48	43	63	48	53	58
15	68	73	30	30	33	30	30	38
16	71	59	90	70	64	40	60	51
17	63	88	65	55	58	45	55	83
18	35	35	68	83	85	78	63	85
19	40	30	3	28	50	43	18	40
20	105	95	85	55	65	55	85	55
21	100	90	70	80	90	80	70	80
22	40	50	28	63	60	68	33	70
23	70	65	59	66	60	59	51	55
24	50	40	20	70	80	80	30	70
25	71	69	64	61	54	49	51	51
26	84	81	68	63	56	68	73	54
27	55	65	80	70	45	70	80	55
28	61	54	53	73	89	78	58	81
29	45	65	65	45	55	45	65	75

E.2 Modelo de Composição

As utilidades e avaliações compostas a partir dos pontos distribuídos e notas dadas na parte de composição, presentes nos questionários, são apresentadas nas tabelas que seguem.

Tabela E.6: Estimativas das utilidades pelo modelo de composição

indivíduo	<i>cte</i>	<i>a1</i>	<i>b1</i>	<i>c1</i>	<i>d1</i>	<i>e1</i>
1	622,5	37,5	5	-100	-35	10
2	825	50	-60	-20	-45	-15
3	744	80	-8	8	-20	-4
4	585	225	50	0	-180	-150
5	735	-100	15	45	-20	-10
6	824	-420	-128	-8	-12	-80
7	650	30	-40	-10	30	10
8	640	175	-45	5	0	-30
9	650	0	-75	-45	-90	-60
10	770	0	-115	0	0	0
11	250	600	0	-50	50	30
12	760	30	-60	-100	30	-30
13	960	-300	-100	-10	-30	15
14	730	40	-50	-20	0	0
15	675	100	-175	-50	0	0
16	665	250	25	-35	0	-100
17	810	-30	-60	-20	-60	60
18	881,8	-54,6	9,1	-4,6	-9,1	-22,7
19	780	0	-300	-20	-30	-50
20	880	120	-100	-50	-90	-60
21	940	0	-30	-30	0	-20
22	775	0	-25	-20	0	-40
23	855	-30	-60	-30	-30	15
24	805	-150	-60	0	-60	-30
25	635	100	-25	0	-60	-10
26	720	120	15	-40	-70	-30
27	760	60	20	10	60	-80
28	870	-250	-100	-30	0	0
29	700	20	20	60	25	-10

Tabela E.7: Estimativas das 8 avaliações de perfil-completo pelo modelo de composição

indivíduo	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>	<i>VII</i>	<i>VIII</i>
1	670	560	630	628	598	538	540	488
2	860	855	770	765	765	730	735	760
3	820	832	796	736	720	740	800	732
4	660	810	680	635	255	485	530	405
5	625	680	630	750	705	785	665	760
6	324	396	264	696	732	608	176	804
7	690	670	670	610	690	610	670	670
8	785	820	770	595	610	570	745	645
9	590	605	485	575	500	470	380	515
10	770	770	655	655	770	655	655	770
11	880	800	900	250	330	230	880	250
12	760	690	760	700	760	570	630	690
13	675	650	530	860	945	865	535	920
14	770	750	720	680	730	660	700	710
15	775	725	600	500	675	450	550	625
16	815	880	940	690	565	555	805	630
17	840	760	660	750	810	790	700	730
18	805	823	827	891	850	864	800	868
19	730	760	450	480	700	410	380	730
20	940	950	810	780	730	670	700	740
21	920	910	910	910	920	860	860	910
22	735	755	750	750	735	690	690	755
23	840	795	735	795	840	780	720	795
24	625	655	535	745	715	715	505	745
25	725	735	650	610	565	600	640	575
26	810	800	785	735	620	665	715	610
27	740	830	900	780	740	710	830	830
28	620	590	520	770	870	740	490	840
29	710	780	765	720	715	770	815	785

E.3 Modelo Bayesiano I

Seguem abaixo as estimativas de σ^2 , de Ω , do vetor α , das utilidades e das avaliações referentes aos estímulos avaliados. Os dados utilizados nestas estimativas provêm das avaliações feitas na parte de decomposição e das avaliações compostas pelas respostas dadas na parte de composição. As estimativas desta seção foram obtidas pelo Modelo Bayesiano I.

As tabelas contém a estimativa do parâmetro, dada pelo método de Monte Carlo (Média), o Desvio Padrão amostral (DP), o Erro Padrão da estimativa de Monte Carlo (EP_{MV}), os percentis de 2,5%, 5% (Mediana) e 97,5%.

Tabela E.8: Estimativa de σ^2 e Ω sob o Modelo Bayesiano I

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
—	σ^2	133,6	15,72	0,19	106,7	132,4	167,8
—	α_1	0,05	0,29	0,01	-0,54	0,06	0,61
—	α_2	0,09	0	0	0,08	0,09	0,09
—	Ω_{11}	35,34	22,84	0,7	6,75	30,28	93,03
—	Ω_{12}	18,5	31,56	0,96	-49,55	19,98	78,69
—	Ω_{13}	2,14	14,4	0,46	-31,5	3,21	27,6
—	Ω_{14}	-6,88	10,62	0,35	-32,3	-5,31	10,17
—	Ω_{15}	8,76	12,16	0,36	-15,47	8,36	34,31
—	Ω_{16}	6,2	12,81	0,38	-21,04	6,42	31,43
—	Ω_{21}	18,5	31,56	0,96	-49,55	19,98	78,69
—	Ω_{22}	162,9	64,98	0,98	67,56	152,6	319,9
—	Ω_{23}	-44,31	29	0,79	-110,3	-41,33	5,28
—	Ω_{24}	-7,25	22,09	0,73	-54,31	-6	33,83
—	Ω_{25}	-32,57	27,02	0,76	-94,34	-29,41	11,74
—	Ω_{26}	-41,2	27,53	0,71	-104,5	-38,21	4,77
—	Ω_{31}	2,14	14,4	0,46	-31,5	3,21	27,6
—	Ω_{32}	-44,31	29	0,79	-110,3	-41,33	5,28
—	Ω_{33}	23,7	17,81	0,55	2,05	19,63	68,78
—	Ω_{34}	-0,88	8,61	0,28	-19,86	-0,45	15,81
—	Ω_{35}	15,34	12,1	0,36	-2,3	13,4	44,39
—	Ω_{36}	18,4	13,6	0,41	-1,03	16,16	51,29
—	Ω_{41}	-6,88	10,62	0,35	-32,3	-5,31	10,17
—	Ω_{42}	-7,25	22,09	0,73	-54,31	-6	33,83
—	Ω_{43}	-0,88	8,61	0,28	-19,86	-0,45	15,81
—	Ω_{44}	6,54	6,57	0,2	0,38	4,54	24,4
—	Ω_{45}	-1,85	8,23	0,27	-20	-1,22	13,78
—	Ω_{46}	-1,12	8,61	0,28	-20,02	-0,88	16,18
—	Ω_{51}	8,76	12,16	0,36	-15,47	8,36	34,31
—	Ω_{52}	-32,57	27,02	0,76	-94,34	-29,41	11,74
—	Ω_{53}	15,34	12,1	0,36	-2,3	13,4	44,39
—	Ω_{54}	-1,85	8,23	0,27	-20	-1,22	13,78
—	Ω_{55}	20,43	15,49	0,47	1,64	16,94	59,37
—	Ω_{56}	16,24	11,39	0,32	-0,11	14,28	43,72
—	Ω_{61}	6,2	12,81	0,38	-21,04	6,42	31,43
—	Ω_{62}	-41,2	27,53	0,71	-104,5	-38,21	4,77
—	Ω_{63}	18,4	13,6	0,41	-1,03	16,16	51,29
—	Ω_{64}	-1,12	8,61	0,28	-20,02	-0,88	16,18
—	Ω_{65}	16,24	11,39	0,32	-0,11	14,28	43,72
—	Ω_{66}	23,91	17,44	0,53	2,48	19,89	68,19

Tabela E.9: Estimativas das utilidades sob o Modelo Bayesiano I

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
1	cte	54,94	2,62	0,04	49,83	54,87	60,31
1	a_1	3,6	6,07	0,04	-8,44	3,6	15,48
1	b_1	0,96	2,48	0,02	-4,02	0,92	6,09
1	c_1	-8,63	1,43	0,02	-11,62	-8,61	-5,8
1	d_1	-2,01	2,31	0,03	-6,41	-2,15	3,03
1	e_1	1,79	2,51	0,02	-3,08	1,67	7,11
2	cte	74,32	3,02	0,07	68,59	74,21	80,65
2	a_1	14,94	6,11	0,04	3,01	14,9	26,96
2	b_1	-7,33	2,92	0,06	-13,79	-7,08	-2,15
2	c_1	-2,46	1,94	0,05	-6,6	-2,34	1,21
2	d_1	-4,87	2,68	0,06	-10,73	-4,67	0,09
2	e_1	-3,04	2,88	0,06	-9,38	-2,81	2,11
3	cte	54,32	5,44	0,18	42,98	54,61	64,07
3	a_1	23,48	7,14	0,11	9,49	23,44	37,54
3	b_1	-10,64	5,12	0,16	-21,29	-10,42	-1,16
3	c_1	2,3	4,56	0,16	-6,24	2,09	11,81
3	d_1	-11,92	4,78	0,15	-21,75	-11,78	-3,16
3	e_1	-11,26	4,84	0,15	-21,57	-10,93	-2,69
4	cte	59,86	4,33	0,13	52,32	59,53	69,17
4	a_1	15,73	6,83	0,11	2,14	15,8	28,85
4	b_1	9,26	3,89	0,11	1,46	9,25	17,06
4	c_1	-1,48	3,36	0,11	-8,33	-1,38	4,93
4	d_1	-8,55	3,91	0,12	-15,21	-8,88	0
4	e_1	-6,56	3,6	0,09	-13,06	-6,8	1,12
5	cte	62,3	3,49	0,09	55,04	62,45	68,91
5	a_1	6,03	6,4	0,07	-6,34	5,95	18,8
5	b_1	-3,99	3,3	0,08	-11,29	-3,68	1,54
5	c_1	3,53	2,47	0,08	-1,71	3,65	8,26
5	d_1	-6,25	3,14	0,08	-13,25	-5,93	-1,08
5	e_1	-6,18	3,28	0,08	-13,5	-5,88	-0,72
6	cte	80,84	5,52	0,19	69,46	80,86	91,66
6	a_1	-2,31	7,6	0,13	-17,14	-2,31	12,65
6	b_1	-17,65	5,37	0,18	-28,56	-17,56	-7
6	c_1	-3,28	4,63	0,16	-12,74	-3,14	5,51
6	d_1	-4,48	5,1	0,17	-15,19	-4,2	4,79
6	e_1	-12,09	4,88	0,15	-21,81	-12,12	-2,5
7	cte	55,17	2,67	0,04	49,58	55,28	60,22
7	a_1	6,19	6,11	0,05	-5,63	6,18	18,35
7	b_1	-4,94	2,53	0,03	-10,46	-4,74	-0,27
7	c_1	-0,76	1,45	0,03	-3,74	-0,78	2,28
7	d_1	1,3	2,33	0,03	-3,84	1,5	5,61
7	e_1	-0,34	2,47	0,02	-5,47	-0,25	4,53
8	cte	56	2,62	0,04	50,99	55,93	61,43
8	a_1	14,47	6,03	0,03	2,47	14,51	26,17

Tabela E.10: Estimativas das utilidades sob o Modelo Bayesiano I

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
8	b_1	-3,3	2,45	0,02	-8,21	-3,38	1,82
8	c_1	0,38	1,36	0,02	-2,41	0,4	3,13
8	d_1	0,7	2,24	0,02	-3,78	0,62	5,4
8	e_1	-1,79	2,47	0,02	-6,63	-1,86	3,4
9	cte	53,74	3,16	0,07	47,35	53,8	59,91
9	a_1	6,44	7,1	0,14	-7,98	6,6	19,92
9	b_1	-10,77	3,66	0,1	-19,44	-10,16	-5,39
9	c_1	-2,96	2,14	0,06	-6,73	-3,19	1,89
9	d_1	-10,84	2,85	0,07	-17,06	-10,59	-5,88
9	e_1	-9,22	3,46	0,09	-17,23	-8,72	-3,83
10	cte	72,25	3,19	0,08	66,65	72,02	79,1
10	a_1	3,15	6,15	0,05	-9,06	3,19	15,16
10	b_1	-8,3	3	0,07	-14,24	-8,37	-2,16
10	c_1	-0,99	2,16	0,06	-5,51	-0,89	3,14
10	d_1	2,68	2,73	0,06	-2,4	2,52	8,48
10	e_1	2,34	2,85	0,05	-3,06	2,22	8,32
11	cte	26,72	3,19	0,08	21,56	26,34	33,97
11	a_1	50,51	6,37	0,08	37,58	50,67	62,67
11	b_1	2,07	2,88	0,06	-3,72	2,02	7,94
11	c_1	-4,88	2	0,06	-9,07	-4,8	-0,93
11	d_1	7,1	2,59	0,05	2,29	6,94	12,61
11	e_1	5,47	2,83	0,05	0,23	5,3	11,58
12	cte	70,66	3,12	0,07	65,1	70,47	77,29
12	a_1	3,32	6,42	0,08	-9,34	3,38	15,85
12	b_1	-3	3,03	0,07	-8,9	-3,1	3,26
12	c_1	-8,93	2,11	0,06	-13,03	-8,97	-4,46
12	d_1	5,48	2,7	0,06	0,43	5,32	11,24
12	e_1	0,47	3,01	0,06	-4,9	0,23	7,17
13	cte	79,1	3,24	0,08	72,14	79,32	84,83
13	a_1	-24,7	6,18	0,06	-36,59	-24,75	-12,35
13	b_1	-10,18	2,72	0,05	-16	-10,04	-5,02
13	c_1	-0,3	1,67	0,04	-3,47	-0,41	3,4
13	d_1	-4,24	2,41	0,04	-9,27	-4,11	0,42
13	e_1	-0,44	2,61	0,04	-6,02	-0,33	4,59
14	cte	59,71	2,86	0,06	53,59	59,88	64,85
14	a_1	3,81	6,09	0,04	-8,2	3,83	15,72
14	b_1	-5,88	2,65	0,04	-11,67	-5,68	-1,02
14	c_1	-1,12	1,63	0,04	-4,16	-1,24	2,5
14	d_1	-1,69	2,39	0,03	-6,83	-1,53	2,8
14	e_1	-1,56	2,57	0,03	-6,98	-1,46	3,47
15	cte	52,32	3,54	0,1	44,74	52,56	58,56
15	a_1	14,68	6,34	0,07	2,28	14,65	27,25
15	b_1	-19,56	3,35	0,09	-26,86	-19,28	-13,71
15	c_1	-3,21	2,6	0,08	-8,07	-3,36	2,4

Tabela E.11: Estimativas das utilidades sob o Modelo Bayesiano I

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
15	d_1	-5,16	3,19	0,08	-12,33	-4,81	0
15	e_1	-5,04	3,11	0,07	-11,86	-4,78	0,33
16	cte	57,35	2,8	0,05	51,95	57,29	63,13
16	a_1	16,62	6,17	0,06	4,34	16,67	28,68
16	b_1	3,87	2,61	0,04	-0,81	3,62	9,65
16	c_1	-3,24	1,52	0,03	-6,65	-3,16	-0,37
16	d_1	1,66	2,46	0,04	-2,63	1,38	7,19
16	e_1	-6,81	2,65	0,04	-11,7	-7,02	-1,01
17	cte	68,72	3,11	0,07	62,32	68,8	74,74
17	a_1	7,11	6,37	0,08	-5,35	7,09	19,66
17	b_1	-9,1	3,14	0,07	-16,27	-8,69	-4,16
17	c_1	-1,45	2	0,06	-5,36	-1,51	2,78
17	d_1	-8,08	2,62	0,05	-13,79	-7,85	-3,5
17	e_1	0,99	3,22	0,08	-6,51	1,43	5,94
18	cte	72,85	4,34	0,14	64,39	72,77	81,84
18	a_1	-25,03	6,68	0,09	-38,37	-24,96	-12,08
18	b_1	6,56	3,93	0,12	-0,55	6,33	14,95
18	c_1	0,47	2,88	0,09	-5,25	0,41	6,31
18	d_1	3,68	3,76	0,11	-2,68	3,32	12,02
18	e_1	3,25	3,55	0,09	-3,19	3,07	10,66
19	cte	58,28	4,45	0,14	48,87	58,54	66,42
19	a_1	-16,89	6,47	0,07	-29,6	-16,88	-4,27
19	b_1	-23,44	4,2	0,13	-31,43	-23,58	-14,58
19	c_1	0,16	3,04	0,1	-5,57	0,03	6,48
19	d_1	-2,77	3,45	0,1	-9,45	-2,84	4,28
19	e_1	-2,86	3,74	0,1	-9,68	-3,06	5,07
20	cte	79,41	3,25	0,08	72,9	79,43	85,81
20	a_1	23,52	6,43	0,08	11,29	23,39	36,6
20	b_1	-11,09	3,06	0,07	-17,7	-10,88	-5,51
20	c_1	-5,25	2,2	0,06	-9,94	-5,14	-1,09
20	d_1	-8,8	2,87	0,07	-14,92	-8,64	-3,41
20	e_1	-6,85	2,92	0,06	-12,89	-6,75	-1,25
21	cte	82,8	2,83	0,06	77,42	82,72	88,63
21	a_1	3,5	6,17	0,06	-8,73	3,54	15,57
21	b_1	-3,18	2,6	0,04	-8,89	-3,02	1,73
21	c_1	-2,88	1,47	0,03	-6,06	-2,81	-0,06
21	d_1	0,02	2,38	0,03	-5,04	0,11	4,67
21	e_1	-1,6	2,59	0,03	-6,76	-1,62	3,76
22	cte	60,28	3,92	0,12	52,41	60,32	68,31
22	a_1	-15,79	6,64	0,09	-29,16	-15,67	-3,1
22	b_1	0,12	3,48	0,09	-6,72	0,06	7,21
22	c_1	0,01	2,7	0,09	-4,92	-0,18	5,83
22	d_1	0,54	3,17	0,08	-5,53	0,46	7,07

Tabela E.12: Estimativas das utilidades sob o Modelo Bayesiano I

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
22	e_1	-1,74	3,37	0,08	-8,02	-1,87	5,31
23	cte	70,11	3,02	0,07	63,67	70,3	75,48
23	a_1	-0,99	6,08	0,04	-12,82	-1,03	11,2
23	b_1	-6,95	2,65	0,04	-12,57	-6,81	-1,93
23	c_1	-2,07	1,72	0,04	-5,42	-2,15	1,61
23	d_1	-4,85	2,49	0,04	-10,37	-4,61	-0,46
23	e_1	-0,91	2,63	0,04	-6,62	-0,73	3,94
24	cte	67,94	3,86	0,12	60,74	67,76	76,24
24	a_1	-28,23	6,64	0,09	-41,55	-28,11	-15,55
24	b_1	-0,76	3,22	0,08	-6,61	-0,95	6,06
24	c_1	0,65	2,36	0,07	-3,99	0,55	5,64
24	d_1	-1,72	3	0,07	-6,99	-1,96	4,75
24	e_1	2,05	3,36	0,08	-3,51	1,71	9,53
25	cte	56,27	2,64	0,04	51,25	56,17	61,79
25	a_1	11,81	5,99	0,03	0,01	11,81	23,59
25	b_1	-2,57	2,48	0,03	-7,83	-2,47	2,29
25	c_1	-0,4	1,43	0,02	-3,55	-0,29	2,26
25	d_1	-5,04	2,31	0,03	-9,85	-5,01	-0,4
25	e_1	-1,05	2,5	0,02	-6,36	-0,96	3,82
26	cte	64,71	2,72	0,05	59,58	64,61	70,38
26	a_1	15,03	6,09	0,04	3,12	15,01	27,02
26	b_1	0,92	2,62	0,04	-4,65	1,02	5,97
26	c_1	-3,83	1,57	0,03	-7,19	-3,76	-0,78
26	d_1	-5,68	2,42	0,04	-10,82	-5,63	-0,9
26	e_1	-2,32	2,59	0,03	-7,49	-2,33	3,02
27	cte	63,77	2,78	0,05	57,91	63,93	68,91
27	a_1	4,42	6,2	0,07	-7,56	4,34	16,76
27	b_1	1,93	2,52	0,03	-2,78	1,77	7,5
27	c_1	1,15	1,39	0,02	-1,55	1,09	4,17
27	d_1	4,65	2,31	0,03	-0,3	4,73	9,18
27	e_1	-7,05	2,6	0,04	-12,22	-7,08	-1,67
28	cte	79,08	3,13	0,08	73,33	78,91	85,72
28	a_1	-22,29	6,14	0,05	-34,52	-22,26	-10,37
28	b_1	-6,27	2,85	0,06	-11,52	-6,46	-0,15
28	c_1	-3,22	1,92	0,05	-7,29	-3,12	0,49
28	d_1	2,95	2,62	0,05	-1,56	2,69	8,78
28	e_1	2,95	2,78	0,05	-1,99	2,71	9,09
29	cte	56,69	2,91	0,06	50,4	56,88	61,86
29	a_1	2,62	6,18	0,06	-9,51	2,61	14,8
29	b_1	-0,2	2,72	0,05	-6,16	0,02	4,78
29	c_1	5,83	1,76	0,04	2,62	5,68	9,78
29	d_1	0,33	2,39	0,03	-4,63	0,43	5
29	e_1	-3,13	2,76	0,05	-9,3	-2,88	1,76

Tabela E.13: Estimativas das 8 avaliações de perfil-completo sob o Modelo Bayesiano I

indivíduo	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>	<i>VII</i>	<i>VIII</i>
1	60	50	57	56	55	49	51	44
2	86	87	77	67	66	61	72	67
3	67	80	55	44	31	35	46	45
4	69	74	76	69	45	61	68	50
5	62	72	58	58	50	56	55	60
6	66	75	56	63	64	48	41	73
7	61	61	58	50	56	49	57	56
8	69	71	68	53	55	51	66	57
9	51	57	39	43	34	31	26	40
10	78	74	70	64	77	65	71	74
11	83	72	86	29	39	29	87	29
12	74	65	76	68	77	59	68	67
13	54	54	40	69	74	68	39	75
14	62	62	56	54	56	51	53	57
15	62	64	42	33	42	25	34	44
16	67	71	79	61	52	51	69	56
17	77	74	59	60	62	59	58	59
18	51	48	58	79	80	83	62	77
19	39	42	15	35	53	32	12	56
20	96	98	83	68	64	56	71	65
21	85	83	83	80	81	75	79	80
22	43	45	45	60	59	59	43	61
23	68	67	57	63	64	60	54	63
24	42	40	37	67	68	70	40	67
25	67	68	60	54	50	52	59	51
26	77	76	75	66	57	59	69	55
27	61	69	75	66	61	60	69	70
28	60	54	53	73	85	73	53	79
29	56	65	59	56	54	59	62	63

E.4 Modelo Bayesiano II

Nesta seção apresentamos as estimativas das utilidades, de σ^2 , de Ω , do vetor α e das avaliações referentes aos estímulos avaliados, obtidas pelo Modelo Bayesiano II. Os dados utilizados nestas estimativas também provêm das avaliações feitas na parte de decomposição e das avaliações compostas pelas respostas dadas na parte de composição.

As tabelas contém a estimativa do parâmetro, dada pelo método de Monte Carlo (Média), o Desvio Padrão amostral (DP), o Erro Padrão da estimativa de Monte Carlo (EP_{MV}), os percentis de 2,5%, 5% (Mediana) e 97,5%.

Tabela E.14: Estimativa de σ^2 e Ω sob o Modelo Bayesiano II

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
—	σ^2	121,5	31,14	0,93	82,36	114,5	208,8
—	α_1	0,03	0,31	0,01	-0,58	0,03	0,63
—	α_2	0,09	0	0	0,08	0,09	0,09
—	Ω_{11}	54,54	57,75	2,04	3,39	34,36	214,5
—	Ω_{12}	13,69	71,17	2,63	-155,6	25,85	127,8
—	Ω_{13}	-15,41	36,74	1,34	-107,8	-9,59	42,57
—	Ω_{14}	-15,94	26,21	0,93	-87,29	-8,4	13,56
—	Ω_{15}	-9,22	38,11	1,38	-117,4	1,47	34,33
—	Ω_{16}	-12,31	39,73	1,42	-119,3	-0,59	29,5
—	Ω_{21}	13,69	71,17	2,63	-155,6	25,85	127,8
—	Ω_{22}	169,9	105,8	3,3	26,94	148,5	432,5
—	Ω_{23}	-63,13	38,26	1,11	-153,4	-58,14	-1,15
—	Ω_{24}	-2,33	40,15	1,47	-79,13	-2,86	84,85
—	Ω_{25}	-27,43	40,75	1,36	-112,1	-27,05	52,03
—	Ω_{26}	-33,36	39,3	1,27	-111	-33,57	41,47
—	Ω_{31}	-15,41	36,74	1,34	-107,8	-9,59	42,57
—	Ω_{32}	-63,13	38,26	1,11	-153,4	-58,14	-1,15
—	Ω_{33}	36,34	28,11	0,9	2,05	30,21	107
—	Ω_{34}	0,85	19,14	0,69	-37,79	0,01	42,2
—	Ω_{35}	16,27	22,61	0,77	-21,88	13,4	68,6
—	Ω_{36}	22,67	24,5	0,81	-10,87	18,27	82,47
—	Ω_{41}	-15,94	26,21	0,93	-87,29	-8,4	13,56
—	Ω_{42}	-2,33	40,15	1,47	-79,13	-2,86	84,85
—	Ω_{43}	0,85	19,14	0,69	-37,79	0,01	42,2
—	Ω_{44}	13,46	15,61	0,52	0,47	8,52	55,46
—	Ω_{45}	7,84	20,68	0,73	-17,48	2,38	65,16
—	Ω_{46}	4,83	20,26	0,72	-23,77	0,47	58,64
—	Ω_{51}	-9,22	38,11	1,38	-117,4	1,47	34,33
—	Ω_{52}	-27,43	40,75	1,36	-112,1	-27,05	52,03
—	Ω_{53}	16,27	22,61	0,77	-21,88	13,4	68,6
—	Ω_{54}	7,84	20,68	0,73	-17,48	2,38	65,16
—	Ω_{55}	33,22	32,62	1,09	1,77	23,59	121
—	Ω_{56}	27,92	30,24	1,03	-4,73	19,91	109,6
—	Ω_{61}	-12,31	39,73	1,42	-119,3	-0,59	29,5
—	Ω_{62}	-33,36	39,3	1,27	-111	-33,57	41,47
—	Ω_{63}	22,67	24,5	0,81	-10,87	18,27	82,47
—	Ω_{64}	4,83	20,26	0,72	-23,77	0,47	58,64
—	Ω_{65}	27,92	30,24	1,03	-4,73	19,91	109,6
—	Ω_{66}	38,98	37,25	1,24	1,49	28,87	135,7

Tabela E.15: Estimativas das utilidades sob o Modelo Bayesiano II

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
1	cte	54,93	3,62	0,06	47,58	54,9	62,4
1	a_1	2,7	5,77	0,04	-9,19	2,81	13,99
1	b_1	0,83	2,87	0,02	-4,92	0,76	6,84
1	c_1	-8,7	1,93	0,03	-12,64	-8,7	-4,67
1	d_1	-2,33	2,93	0,03	-8,15	-2,45	3,96
1	e_1	1,76	3,18	0,02	-4,51	1,59	8,63
2	cte	75,17	6,1	0,2	63,23	75	88,32
2	a_1	16,1	6,74	0,14	4,37	15,58	30,8
2	b_1	-9,84	3,55	0,08	-17,56	-9,52	-3,66
2	c_1	-2,37	3,35	0,11	-9,76	-2,09	3,9
2	d_1	-5,71	4,21	0,12	-15,21	-5,32	1,7
2	e_1	-3,56	4,21	0,11	-13,03	-3,1	3,68
3	cte	59,84	10,21	0,38	39,24	60,18	81,42
3	a_1	21,6	9,36	0,27	3,2	21,52	40,62
3	b_1	-8,75	5,86	0,18	-20,69	-8,6	2,27
3	c_1	1,17	6,3	0,23	-12,07	1,32	13,08
3	d_1	-13,18	7,45	0,25	-29,51	-12,63	-0,7
3	e_1	-13,18	7,7	0,26	-29,42	-12,92	0,25
4	cte	56,47	5,64	0,17	45,55	56,11	68,71
4	a_1	18,74	6,62	0,12	5,04	18,95	31,32
4	b_1	5,97	4,1	0,11	-2,27	5,91	14,34
4	c_1	-0,79	3,54	0,11	-8,05	-0,72	6,45
4	d_1	-9,73	4,65	0,13	-17,49	-10,21	0,69
4	e_1	-6,79	4,74	0,13	-15,07	-7,11	3,55
5	cte	66,56	7	0,25	53,09	66,36	81,66
5	a_1	5,25	7,17	0,16	-7,68	4,8	20,65
5	b_1	-4,62	3,88	0,1	-13,09	-4,31	2,05
5	c_1	3,36	3,87	0,13	-5,21	3,66	10,62
5	d_1	-5,03	4,62	0,14	-15,36	-4,59	3,02
5	e_1	-5,24	4,81	0,14	-16,15	-4,63	2,36
6	cte	82,95	14,43	0,56	55,36	82,9	112,2
6	a_1	-3,58	10,9	0,35	-24,18	-3,97	18,86
6	b_1	-23,99	7,03	0,24	-37,86	-23,84	-10,55
6	c_1	-2,66	8,16	0,31	-19,24	-1,98	12,52
6	d_1	-3,77	8,95	0,33	-22,92	-3,47	12,72
6	e_1	-11,22	8,49	0,31	-28,11	-11,18	4,34
7	cte	56,97	3,73	0,06	49,6	56,9	64,83
7	a_1	4,76	5,84	0,05	-6,61	4,59	16,81
7	b_1	-4,58	2,92	0,03	-10,96	-4,38	0,95
7	c_1	-0,93	1,96	0,03	-5,15	-0,88	3
7	d_1	1,75	2,99	0,03	-4,92	1,98	7,44
7	e_1	0,12	3,22	0,03	-6,79	0,29	6,49
8	cte	56,59	3,64	0,06	49,3	56,54	64,23
8	a_1	14,75	5,79	0,04	2,85	14,87	26,12

Tabela E.16: Estimativas das utilidades sob o Modelo Bayesiano II

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
8	b_1	-3,62	2,88	0,02	-9,51	-3,64	2,26
8	c_1	0,48	1,9	0,02	-3,38	0,48	4,46
8	d_1	0,73	2,94	0,02	-5,17	0,61	7,03
8	e_1	-1,71	3,22	0,03	-8,06	-1,88	5,3
9	cte	56,48	6,05	0,2	45,14	56,18	70,16
9	a_1	9,04	6,78	0,13	-3,66	8,85	23,05
9	b_1	-11,8	4,22	0,12	-21,49	-11,23	-5,3
9	c_1	-3,6	3,43	0,11	-10,79	-3,59	3,37
9	d_1	-12,41	4,33	0,12	-22,08	-11,92	-5,26
9	e_1	-10,95	4,8	0,14	-21,95	-10,33	-3,76
10	cte	71,5	4,55	0,12	62,92	71,28	81,44
10	a_1	4,21	6,05	0,07	-7,7	4,1	16,59
10	b_1	-10,99	3,39	0,07	-18,05	-10,88	-4,26
10	c_1	-0,47	2,64	0,07	-5,96	-0,41	4,94
10	d_1	2,39	3,59	0,08	-4,67	2,24	9,89
10	e_1	2,49	3,73	0,07	-4,65	2,32	10,45
11	cte	25,67	4,27	0,1	17,94	25,26	35,4
11	a_1	53,67	6,12	0,08	41,19	53,78	65,69
11	b_1	-0,02	3,37	0,07	-7,15	0,07	6,63
11	c_1	-4,55	2,56	0,07	-9,66	-4,59	0,88
11	d_1	7,1	3,5	0,07	0,41	6,88	14,63
11	e_1	5,86	3,76	0,07	-0,91	5,52	14,25
12	cte	68,95	4,16	0,09	61,17	68,74	78,03
12	a_1	2,62	6,22	0,09	-10,43	2,84	14,64
12	b_1	-4,81	3,22	0,05	-11,37	-4,83	1,81
12	c_1	-8,5	2,32	0,05	-12,88	-8,63	-3,32
12	d_1	4,57	3,25	0,05	-1,73	4,37	11,61
12	e_1	-0,12	3,61	0,06	-6,72	-0,45	7,94
13	cte	81,39	4,34	0,11	72,25	81,5	89,77
13	a_1	-27,78	5,98	0,07	-39,99	-27,69	-16
13	b_1	-8,65	3,08	0,05	-14,96	-8,62	-2,44
13	c_1	-0,6	2,2	0,05	-5,17	-0,62	4,01
13	d_1	-4,22	3,24	0,06	-11,15	-4,04	2,07
13	e_1	-0,55	3,45	0,05	-7,99	-0,29	6,02
14	cte	62,14	3,87	0,07	54,06	62,22	69,69
14	a_1	2,74	5,84	0,05	-9,04	2,78	14,37
14	b_1	-4,52	2,95	0,03	-10,65	-4,49	1,45
14	c_1	-1,57	2,05	0,04	-5,87	-1,57	2,67
14	d_1	-1,42	3,1	0,04	-8,21	-1,18	4,46
14	e_1	-1,29	3,3	0,04	-8,27	-1,13	5,21
15	cte	55,7	5,53	0,17	44,42	55,79	67,03
15	a_1	13,56	6,61	0,12	1,33	13,28	27,58
15	b_1	-18,4	3,85	0,1	-26,65	-18,15	-11,31
15	c_1	-3,87	3,36	0,11	-10,95	-3,84	2,88

Tabela E.17: Estimativas das utilidades sob o Modelo Bayesiano II

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
15	d_1	-5,63	4,46	0,13	-15,88	-5,05	1,25
15	e_1	-5,87	4,5	0,12	-15,89	-5,45	1,51
16	cte	59,18	4,87	0,14	48,68	59,2	69,02
16	a_1	16,78	6,24	0,09	3,74	17,03	28,49
16	b_1	5,05	3,41	0,07	-0,9	4,71	12,68
16	c_1	-3,36	2,69	0,08	-9	-3,32	2,22
16	d_1	3,64	3,71	0,09	-2,14	3,07	12,31
16	e_1	-4,84	3,89	0,09	-11,07	-5,32	4,08
17	cte	70,22	6,28	0,21	57,77	70,13	83,68
17	a_1	7,87	7,02	0,15	-4,61	7,43	22,94
17	b_1	-10,05	3,67	0,09	-17,98	-9,78	-3,62
17	c_1	-1,86	3,54	0,12	-9,65	-1,67	4,86
17	d_1	-9,87	4,41	0,13	-19,94	-9,36	-2,5
17	e_1	-0,54	4,75	0,13	-11,42	0,1	6,56
18	cte	72,34	10,06	0,38	51,16	72,45	91,88
18	a_1	-25,56	8,79	0,25	-44,08	-25,1	-9,15
18	b_1	9,43	4,85	0,15	0,68	9,15	19,68
18	c_1	0,54	5,39	0,2	-9,71	0,12	11,94
18	d_1	3,55	6,16	0,21	-6,84	2,95	17,28
18	e_1	2,95	5,97	0,2	-7,46	2,55	15,78
19	cte	61,61	9,81	0,37	41,27	61,64	81,15
19	a_1	-20,75	8,55	0,24	-39,09	-20,16	-5,53
19	b_1	-18,19	4,91	0,15	-27,2	-18,4	-8,1
19	c_1	-0,31	5,34	0,2	-10,27	-0,77	10,95
19	d_1	-0,72	6,06	0,21	-11,63	-1,09	12,66
19	e_1	-1,42	5,95	0,2	-12,29	-1,71	11,05
20	cte	80,45	5,79	0,19	69,16	80,26	92,99
20	a_1	21,09	6,45	0,11	9,68	20,64	35,09
20	b_1	-12,96	3,58	0,08	-20,81	-12,61	-6,86
20	c_1	-5,02	3,16	0,1	-12	-4,79	0,99
20	d_1	-8,73	4,04	0,11	-17,67	-8,47	-1,3
20	e_1	-6,61	4,07	0,1	-15,52	-6,26	0,75
21	cte	84,7	4,33	0,11	76,48	84,55	94,02
21	a_1	4,12	5,98	0,07	-7,59	3,94	16,5
21	b_1	-4,25	3,17	0,05	-11,22	-3,96	1,58
21	c_1	-2,91	2,27	0,05	-7,85	-2,83	1,63
21	d_1	0,48	3,26	0,06	-6,58	0,55	6,93
21	e_1	-0,94	3,48	0,06	-7,94	-1,03	6,41
22	cte	61,97	7,67	0,27	45,78	62,01	77,57
22	a_1	-15,28	7,51	0,18	-31,83	-14,62	-2,28
22	b_1	3,27	4,09	0,11	-4,12	3,01	11,96
22	c_1	-0,45	4,24	0,15	-8,49	-0,77	8,57
22	d_1	0,19	5,03	0,16	-9,2	-0,04	11,08

Tabela E.18: Estimativas das utilidades sob o Modelo Bayesiano II

Indivíduo	Parâmetro	Média	DP	EP_{MC}	val2,5%	Mediana	val97,5%
22	e_1	-2,4	5,02	0,15	-11,87	-2,67	8,17
23	cte	72,55	4,16	0,09	63,86	72,65	80,6
23	a_1	-2,08	5,91	0,06	-13,64	-2,22	10,08
23	b_1	-5,88	3,05	0,04	-12,22	-5,83	0,28
23	c_1	-2,41	2,22	0,05	-7,14	-2,39	2,06
23	d_1	-4,92	3,28	0,06	-12,32	-4,57	0,85
23	e_1	-1,16	3,48	0,06	-8,87	-0,8	5,11
24	cte	69,59	7,99	0,29	52,6	69,57	85,58
24	a_1	-27,86	7,93	0,21	-44,68	-27,44	-13,31
24	b_1	1,35	4,13	0,11	-5,98	1,1	10,1
24	c_1	0,54	4,34	0,15	-7,52	0,27	10,06
24	d_1	-0,1	5,05	0,16	-8,57	-0,68	11,34
24	e_1	3,96	5,44	0,17	-4,35	3,29	16,3
25	cte	57,36	3,94	0,08	49,85	57,2	66,03
25	a_1	12,59	5,86	0,05	1,4	12,35	24,84
25	b_1	-3,54	2,99	0,04	-10,02	-3,34	2,19
25	c_1	-0,33	2,11	0,04	-4,96	-0,22	3,78
25	d_1	-4,89	3,15	0,05	-11,6	-4,85	1,42
25	e_1	-0,73	3,31	0,04	-7,82	-0,63	5,81
26	cte	65,4	4,07	0,09	57,64	65,23	74,26
26	a_1	14,52	5,9	0,06	3,09	14,3	26,82
26	b_1	0,02	3,09	0,05	-6,65	0,21	5,93
26	c_1	-3,72	2,21	0,05	-8,41	-3,66	0,77
26	d_1	-5,53	3,2	0,05	-12,34	-5,49	0,86
26	e_1	-1,82	3,41	0,05	-8,8	-1,87	5,26
27	cte	63,73	4,19	0,1	54,46	63,94	71,57
27	a_1	2,51	5,92	0,06	-9,55	2,55	14,29
27	b_1	2,88	3,12	0,05	-2,81	2,61	9,8
27	c_1	1,24	2,21	0,05	-3,28	1,18	5,91
27	d_1	3,77	3,26	0,06	-3,37	3,98	10,07
27	e_1	-8,2	3,53	0,06	-15,6	-8,07	-1,15
28	cte	78,03	4,19	0,1	69,73	77,97	86,58
28	a_1	-22,13	5,98	0,07	-34,25	-22,06	-10,29
28	b_1	-8	3,09	0,05	-14,1	-8,12	-1,43
28	c_1	-2,86	2,19	0,05	-7,39	-2,83	1,7
28	d_1	2,23	3,18	0,05	-3,56	1,93	9,32
28	e_1	2,65	3,5	0,06	-3,48	2,25	10,51
29	cte	58,29	4,21	0,1	49,18	58,47	66,26
29	a_1	1,74	6,03	0,07	-9,96	1,6	14,16
29	b_1	1,24	3,16	0,05	-5,26	1,27	7,7
29	c_1	5,46	2,39	0,06	0,42	5,48	10,28
29	d_1	-0,35	3,35	0,06	-7,78	-0,05	5,76
29	e_1	-3,93	3,68	0,07	-12,17	-3,55	2,46

Tabela E.19: Estimativas das 8 avaliações de perfil-completo sob o Modelo Bayesiano II

indivíduo	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>	<i>VII</i>	<i>VIII</i>
1	59	49	56	56	54	49	49	44
2	88	89	76	65	66	59	70	67
3	68	83	60	51	33	39	47	48
4	68	74	71	62	40	55	64	46
5	67	75	62	62	56	60	60	65
6	68	77	52	59	68	45	38	77
7	62	61	59	52	59	52	58	58
8	70	72	68	53	56	52	67	58
9	55	62	41	45	33	30	27	40
10	78	75	67	61	76	63	69	73
11	85	75	86	26	39	27	88	28
12	71	63	71	64	73	56	63	65
13	53	53	41	73	77	72	40	77
14	64	63	59	58	59	55	56	59
15	63	65	45	37	44	28	35	46
16	71	73	85	64	58	56	76	59
17	78	76	58	60	60	58	56	58
18	50	47	60	82	79	85	63	76
19	39	41	22	43	59	42	20	61
20	95	97	80	67	65	56	68	67
21	88	86	85	80	84	77	81	82
22	44	46	50	65	60	62	47	62
23	69	68	60	67	66	63	56	65
24	46	42	43	71	73	75	47	70
25	69	70	62	54	52	53	60	52
26	78	76	74	65	58	60	69	56
27	58	67	73	67	59	60	66	69
28	59	53	50	70	83	70	50	77
29	56	65	61	60	54	61	62	63

Table 1. The number of cases of *Salmonella* infection in the United Kingdom, 1990-1999, by serotype and age group

Age group	Serotype									
	1	2	3	4	5	6	7	8	9	10
0-4	10	10	10	10	10	10	10	10	10	10
5-14	10	10	10	10	10	10	10	10	10	10
15-24	10	10	10	10	10	10	10	10	10	10
25-34	10	10	10	10	10	10	10	10	10	10
35-44	10	10	10	10	10	10	10	10	10	10
45-54	10	10	10	10	10	10	10	10	10	10
55-64	10	10	10	10	10	10	10	10	10	10
65-74	10	10	10	10	10	10	10	10	10	10
75-84	10	10	10	10	10	10	10	10	10	10
85-94	10	10	10	10	10	10	10	10	10	10
95-104	10	10	10	10	10	10	10	10	10	10
105-114	10	10	10	10	10	10	10	10	10	10
115-124	10	10	10	10	10	10	10	10	10	10
125-134	10	10	10	10	10	10	10	10	10	10
135-144	10	10	10	10	10	10	10	10	10	10
145-154	10	10	10	10	10	10	10	10	10	10
155-164	10	10	10	10	10	10	10	10	10	10
165-174	10	10	10	10	10	10	10	10	10	10
175-184	10	10	10	10	10	10	10	10	10	10
185-194	10	10	10	10	10	10	10	10	10	10
195-204	10	10	10	10	10	10	10	10	10	10
205-214	10	10	10	10	10	10	10	10	10	10
215-224	10	10	10	10	10	10	10	10	10	10
225-234	10	10	10	10	10	10	10	10	10	10
235-244	10	10	10	10	10	10	10	10	10	10
245-254	10	10	10	10	10	10	10	10	10	10
255-264	10	10	10	10	10	10	10	10	10	10
265-274	10	10	10	10	10	10	10	10	10	10
275-284	10	10	10	10	10	10	10	10	10	10
285-294	10	10	10	10	10	10	10	10	10	10
295-304	10	10	10	10	10	10	10	10	10	10
305-314	10	10	10	10	10	10	10	10	10	10
315-324	10	10	10	10	10	10	10	10	10	10
325-334	10	10	10	10	10	10	10	10	10	10
335-344	10	10	10	10	10	10	10	10	10	10
345-354	10	10	10	10	10	10	10	10	10	10
355-364	10	10	10	10	10	10	10	10	10	10
365-374	10	10	10	10	10	10	10	10	10	10
375-384	10	10	10	10	10	10	10	10	10	10
385-394	10	10	10	10	10	10	10	10	10	10
395-404	10	10	10	10	10	10	10	10	10	10
405-414	10	10	10	10	10	10	10	10	10	10
415-424	10	10	10	10	10	10	10	10	10	10
425-434	10	10	10	10	10	10	10	10	10	10
435-444	10	10	10	10	10	10	10	10	10	10
445-454	10	10	10	10	10	10	10	10	10	10
455-464	10	10	10	10	10	10	10	10	10	10
465-474	10	10	10	10	10	10	10	10	10	10
475-484	10	10	10	10	10	10	10	10	10	10
485-494	10	10	10	10	10	10	10	10	10	10
495-504	10	10	10	10	10	10	10	10	10	10
505-514	10	10	10	10	10	10	10	10	10	10
515-524	10	10	10	10	10	10	10	10	10	10
525-534	10	10	10	10	10	10	10	10	10	10
535-544	10	10	10	10	10	10	10	10	10	10
545-554	10	10	10	10	10	10	10	10	10	10
555-564	10	10	10	10	10	10	10	10	10	10
565-574	10	10	10	10	10	10	10	10	10	10
575-584	10	10	10	10	10	10	10	10	10	10
585-594	10	10	10	10	10	10	10	10	10	10
595-604	10	10	10	10	10	10	10	10	10	10
605-614	10	10	10	10	10	10	10	10	10	10
615-624	10	10	10	10	10	10	10	10	10	10
625-634	10	10	10	10	10	10	10	10	10	10
635-644	10	10	10	10	10	10	10	10	10	10
645-654	10	10	10	10	10	10	10	10	10	10
655-664	10	10	10	10	10	10	10	10	10	10
665-674	10	10	10	10	10	10	10	10	10	10
675-684	10	10	10	10	10	10	10	10	10	10
685-694	10	10	10	10	10	10	10	10	10	10
695-704	10	10	10	10	10	10	10	10	10	10
705-714	10	10	10	10	10	10	10	10	10	10
715-724	10	10	10	10	10	10	10	10	10	10
725-734	10	10	10	10	10	10	10	10	10	10
735-744	10	10	10	10	10	10	10	10	10	10
745-754	10	10	10	10	10	10	10	10	10	10
755-764	10	10	10	10	10	10	10	10	10	10
765-774	10	10	10	10	10	10	10	10	10	10
775-784	10	10	10	10	10	10	10	10	10	10
785-794	10	10	10	10	10	10	10	10	10	10
795-804	10	10	10	10	10	10	10	10	10	10
805-814	10	10	10	10	10	10	10	10	10	10
815-824	10	10	10	10	10	10	10	10	10	10
825-834	10	10	10	10	10	10	10	10	10	10
835-844	10	10	10	10	10	10	10	10	10	10
845-854	10	10	10	10	10	10	10	10	10	10
855-864	10	10	10	10	10	10	10	10	10	10
865-874	10	10	10	10	10	10	10	10	10	10
875-884	10	10	10	10	10	10	10	10	10	10
885-894	10	10	10	10	10	10	10	10	10	10
895-904	10	10	10	10	10	10	10	10	10	10
905-914	10	10	10	10	10	10	10	10	10	10
915-924	10	10	10	10	10	10	10	10	10	10
925-934	10	10	10	10	10	10	10	10	10	10
935-944	10	10	10	10	10	10	10	10	10	10
945-954	10	10	10	10	10	10	10	10	10	10
955-964	10	10	10	10	10	10	10	10	10	10
965-974	10	10	10	10	10	10	10	10	10	10
975-984	10	10	10	10	10	10	10	10	10	10
985-994	10	10	10	10	10	10	10	10	10	10
995-1004	10	10	10	10	10	10	10	10	10	10
1005-1014	10	10	10	10	10	10	10	10	10	10
1015-1024	10	10	10	10	10	10	10	10	10	10
1025-1034	10	10	10	10	10	10	10	10	10	10
1035-1044	10	10	10	10	10	10	10	10	10	10
1045-1054	10	10	10	10	10	10	10	10	10	10
1055-1064	10	10	10	10	10	10	10	10	10	10
1065-1074	10	10	10	10	10	10	10	10	10	10
1075-1084	10	10	10	10	10	10	10	10	10	10
1085-1094	10	10	10	10	10	10	10	10	10	10
1095-1104	10	10	10	10	10	10	10	10	10	10
1105-1114	10	10	10	10	10	10	10	10	10	10
1115-1124	10	10	10	10	10	10	10	10	10	10
1125-1134	10	10	10	10	10	10	10	10	10	10
1135-1144	10	10	10	10	10	10	10	10	10	10
1145-1154	10	10	10	10	10	10	10	10	10	10
1155-1164	10	10	10	10	10	10	10	10	10	10
1165-1174	10	10	10	10	10	10	10	10	10	10
1175-1184	10	10	10	10	10	10	10	10	10	10
1185-1194	10	10	10	10	10	10	10	10	10	10
1195-1204	10	10	10	10	10	10	10	10	10	10
1205-1214	10	10	10	10	10	10	10	10	10	10
1215-1224	10	10	10	10	10	10	10	10	10	10
1225-1234	10	10	10	10	10	10	10	10	10	10
1235-1244	10	10	10	10	10	10	10	10	10	10
1245-1254	10	10	10	10	10	10	10	10	10	10
1255-1264	10	10	10	10	10	10	10	10	10	10
1265-1274	10	10	10	10	10	10	10	10	10	10
1275-1284	10	10	10	10	10	10	10	10	10	10
1285-1294	10	10	10	10	10	10	10	10	10	10
1295-1304	10	10	10	10	10	10	10	10	10	10
1305-1314	10	10	10	10	10	10	10	10	10	10
1315-1324	10	10	10	10	10	10	10	10	10	10
1325-1334	10	10	10	10	10	10	10	10	10	10
1335-1344	10	10	10	10	10	10	10	10	10	10
1345-1354	10	10	10	10	10	10	10	10	10	10
1355-1364	10	10	10	10						

Apêndice F

Estatísticas para avaliação do ajuste do modelo

Este apêndice contém as fórmulas das estatísticas utilizadas para a comparação do ajuste dos diferentes modelos. São elas: $\sqrt{EQM}$ e MAD (Chapman e Staelin, 1982), Coeficiente de Correlação (Netter et al., 1996).

Seja $\hat{y}_{ijk}$ e y_{ijk} as avaliações estimadas e observadas, respectivamente, referentes ao k -ésimo estímulo do bloco j , e pertencentes ao indivíduo i . E sejam, r o número de estímulos por bloco, b o total de blocos apresentados por indivíduo, e n o total de indivíduos entrevistados. As estatísticas utilizadas medem o quão próximos, ou relacionados, estão esses dois vetores.

$$\sqrt{EQM}$$

A medida $\sqrt{EQM}$ é a raiz do erro quadrático médio, e pode ser calculada por

$$\sqrt{EQM} = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^b \sum_{k=1}^r (\hat{y}_{ijk} - y_{ijk})^2}{(n \times b \times r)}}$$

MAD

MAD (mean absolute deviation) é o desvio médio absoluto, que considera o módulo da diferença entre as medidas $\hat{y}_{ijk}$ e y_{ijk} , ou seja, a diferença absoluta entre elas. Sua fórmula pode ser dada por

$$MAD = \frac{\sum_{i=1}^n \sum_{j=1}^b \sum_{k=1}^r |\hat{y}_{ijk} - y_{ijk}|}{(n \times b \times r)}$$

Coeficientes de Correlação

O coeficiente de Correlação de Pearson é dado por

$$R = \frac{s_{\hat{y}y}}{s_{\hat{y}}s_y}$$

sendo

$$s_{\hat{y}y} = \frac{1}{(nbr)} \sum_{i=1}^n \sum_{j=1}^b \sum_{k=1}^r (\hat{y}_{ijk} - \bar{\hat{y}})(y_{ijk} - \bar{y}),$$

$$s_{\hat{y}} = \sqrt{\frac{1}{(nbr)} \sum_{i=1}^n \sum_{j=1}^b \sum_{k=1}^r (\hat{y}_{ijk} - \bar{\hat{y}})^2} \text{ e } s_y = \sqrt{\frac{1}{(nbr)} \sum_{i=1}^n \sum_{j=1}^b \sum_{k=1}^r (y_{ijk} - \bar{y})^2}$$

no qual, $s_{\hat{y}y}$ é a covariância entre os valores $\hat{y}_{ijk}$ e y_{ijk} ; $s_{\hat{y}}$ e s_y são os desvios padrão dos valores $\hat{y}_{ijk}$ e y_{ijk} , respectivamente; $\bar{\hat{y}}$ é a média dos valores $\hat{y}_{ijk}$ e $\bar{y}$ a média dos valores y_{ijk} .

O coeficiente de Correlação de Spearman é equivalente ao de Pearson, porém calculado sobre os postos dos valores.

Apêndice G

Implementação dos Modelos Bayesianos no BRugs

Neste apêndice descrevemos os passos utilizados na obtenção das estimativas dos parâmetros, sob os Modelos Bayesianos. Como referência para auxiliar a utilização desse programa, citamos o livro de Congdon (2001), que nos foi bastante esclarecedor. Exemplos de estimação no contexto Bayesiano são encontrados, também, em Robert (2001).

Para utilizar o pacote BRugs, que é uma interface desenvolvida para o software R, foi preciso construir, para cada modelo, dois arquivos com extensão txt. O primeiro arquivo (chamado de *model*) deve conter a implementação do modelo desejado, no qual devemos informar para o pacote quais são as distribuições *a priori*, a verossimilhança e todos os cálculos necessários para compor uma ou outra. O pacote faz o cálculo das distribuições *a posteriori* internamente, logo, não há necessidade de informá-lo sobre isso.

A Figura G.1 contém o arquivo *model* desenvolvido para os Modelos Bayesianos que possuem o vetor α (modelos I e II). Note que informamos as distribuições *a priori* de μ_i , de σ^2 , de Ω e de α . Informamos, também, o cálculo que deve ser feito com o vetor α na composição da média que servirá como informação *a priori* de μ_i .

Para os Modelos Bayesianos sem este vetor de correção (modelos III e IV), a implementação fica simplificada, veja a Figura G.2.

A única diferença é a exclusão da distribuição *a priori* de α e, é claro, do cálculo intermediário da média *a priori* de μ_i .

O segundo arquivo txt (*list*) deve conter os dados que irão compor a função de verossimilhança. No nosso caso, devemos introduzir os vetores observados y_i que, por sua vez, são organizados em uma única matriz. Além desses dados, este arquivo contém os valores dos parâmetros que não são estimados

no processo de simulação, como os vetores u_i , os graus de liberdade das distribuições *a priori*, o número de parâmetros, o número de indivíduos etc. Os arquivos *list* referentes aos Modelos Bayesianos I, II, III e IV, são apresentados nas figuras G.3, G.4, G.5 e G.6, respectivamente.

Figura G.1: Implementação utilizada nos Modelos Bayesianos com parâmetro alpha

```
model
{
  for(i in 1:N)
  {
    Y[i,1:m]~dlnorm(meday[i,1:m],IS[i]);
    for(j in 1:m)
    {
      meday[i,j]<-X[j,1]*m[i,1]+X[j,2]*m[i,2]+X[j,3]*m[i,3]+X[j,4]*m[i,4]+X[j,5]*m[i,5]+X[j,6]*m[i,6];
    }
    m[i,1:q]~dlnorm(Za[i,1:q],IC[i]);
    for(j in 1:q)
    {
      Za[i,j]<-alpha[1,1]+(alpha[1,2]*U[i,j]);
    }
  }
  for(i in 1:m)
  {
    IS[i,j]<-(1/sigma); for(r in (i+1):m){IS[r,j]<0; IS[i,r]<-IS[r,i,j]}
  }
  sigma~dlnorm(gzero);
  IC[1:q,1:q]~dwish(IDzero[,],dzero);
  alpha[1,1,2]~dlnorm(Azero[1,1,2],SAzero[,]);
  sigma<-1/sigma;
  omega[1:q,1:q]~inverse(IC[,]);
}
```

Figura G.2: Implementação utilizada nos Modelos Bayesianos sem parâmetro alpha

```
model
{
  for(i in 1:N)
  {
    Y[i,1:m]~dlnorm(meday[i,1:m],IS[i]);
    for(j in 1:m)
    {
      meday[i,j]<-X[j,1]*m[i,1]+X[j,2]*m[i,2]+X[j,3]*m[i,3]+X[j,4]*m[i,4]+X[j,5]*m[i,5]+X[j,6]*m[i,6];
    }
    m[i,1:q]~dlnorm(U[i,1:q],IC[i]);
  }
  for(i in 1:m)
  {
    IS[i,j]<-(1/sigma); for(r in (i+1):m){IS[r,j]<0; IS[i,r]<-IS[r,i,j]}
  }
  sigma~dlnorm(gzero);
  IC[1:q,1:q]~dwish(IDzero[,],dzero);
  sigma<-1/sigma;
  omega[1:q,1:q]~inverse(IC[,]);
}
```

Observe que a escala das avaliações da fase de decomposição dos modelos III e IV foram alteradas previamente.

Se for de interesse do usuário, é possível construir um terceiro arquivo informando os valores iniciais dos parâmetros que serão estimados (*inits*). Optamos em utilizar a função *modelGenInits()*

disponibilizada pelo BRugs, que gera valores iniciais aleatórios automaticamente.

A primeira parte de comandos utilizados na estimação do Modelo Bayesiano I foi esta:

```
library(BRugs)
oldwd <- getwd()
setwd(system.file("OpenBUGS", "Examples", package="BRugs"))
modelCheck("Modelo Bayesiano com alpha.txt")
modelData("Dados do Modelo Bayesiano I.txt")
modelCompile(numChains=5)
modelGenInits()
```

Os comandos devem ser escritos na janela “Console” do pacote R. Pra cada comando dado há uma “resposta” do pacote. O modelo que escrevemos num arquivo de formato txt, presente na Figura G.1, entra como parâmetro da função *modelCheck()* que verifica se o modelo está adequado. Ela não confere se o cálculo escolhido está correto, mas se as rotinas criadas e as distribuições estão escritas de forma coerente. O pacote só aceitará novos comandos quando o modelo for aceito.

A próxima função, *modelData()*, solicita o arquivo com as observações da verossimilhança e valores que não serão estimados. O arquivo referente ao modelo I, também em formato txt, foi apresentado na Figura G.3. No comando *modelCompile()* informamos quantas cadeias independentes de Markov devem ser construídas, no caso construímos 5 cadeias. A função *modelGenInits()* gera automaticamente valores iniciais para os parâmetros que serão simulados no processo; não é preciso entrar com nenhum valor nessa função.

Após compilar o modelo e entrar com os dados, segue a segunda parte de comandos:

```
modelUpdate(5000, thin=10)
samplesSet(c("mi","sigma","omega","alpha"))
modelUpdate(25000)
```

A função *modelUpdate()* realiza quantas iterações forem indicadas. É possível informar quantos dos valores simulados devem ser armazenados para cálculos posteriores. Neste caso, pedimos que guardasse 1 valor a cada 10 iterações e que totalizasse 5.000 valores guardados (internamente o programa realizou 50.000 iterações). O pacote retorna o número de iterações guardadas e o tempo gasto para realizá-las.

Na próxima função, *sampleSet()*, informamos quais parâmetros devem ser monitorados, ou seja, quais parâmetros devem ter os valores armazenados. Só será possível pedir a construção de gráficos, por exemplo, para os parâmetros declarados nesta função.

Novamente a função *modelUpdate()* é acionada. Note que, neste caso, optamos em informar apenas quantas iterações devem ser guardadas, dessa forma ele não descartará nenhuma. A opção de descartar iterações serve, por exemplo, para amenizar possíveis correlações entre os valores simulados ou para atingir um elevado número de simulações sem sobrecarregar o processo de armazenamento dos valores.

Na última parte, utilizamos para cada um dos parâmetros, que declaramos que deveriam ser monitorados, estas funções:

```
samplesBgr("sigma")
plotHistory("sigma")
samplesDensity("sigma", beg=15000 )
samplesStats("sigma", beg=15000)
```

A primeira, *samplesBgr()*, retorna o gráfico com a estatística de Gelman e Rubin (1992), neste caso exemplificada pela Figura D.1 do Apêndice D. Se introduzirmos na função o comando *PLOT=FALSE*, o gráfico não é mostrado, apenas os valores dessa estatística são impressos na tela. Através desse gráfico identificamos quando termina o aquecimento das cadeias, observando o ponto em que a estatística atingiu valor 1.

O segundo comando, *plotHistory()*, exibe o gráfico com todos os valores simulados nas iterações, para todas as cadeias. Esse gráfico nos permite verificar se as cadeias convergiram para o mesmo ponto. Na Figura 4.1, o terceiro gráfico é um exemplo deste.

Definido o período de aquecimento, pedimos ao programa que construísse um gráfico de densidade utilizando apenas as iterações posteriores a esse período. A função *samplesDensity()* constrói esse gráfico e, nesse exemplo, utiliza apenas os valores simulados após a iteração de número 15.000 que informamos em *beg=15000*

Por fim, pedimos ao pacote que imprimisse as estimativas dos parâmetros, também tomando o cuidado de informar o número da iteração a partir da qual os valores devem ser considerados.

O livro de Congdon (2001) contém diversos exemplos de aplicação e utilização do pacote WinBugs, que é o original do BRugs.

Figura G.3: Arquivo *list* utilizado na estimação dos parâmetros do Modelo Bayesiano I

```
list(
  U= structure(.Data=c(
    622.5, 37.5, 5, -100, -35, 10,
    825, 50, -60, -20, -45, -15,
    744, 80, -8, 8, -20, -4,
    585, 225, 50, 0, -180, -150,
    735, -100, 15, 45, -20, -10,
    824, -420, -128, -8, -12, -80,
    650, 30, -40, -10, 30, 10,
    640, 175, -45, 5, 0, -30,
    650, 0, -75, -45, -90, -60,
    770, 0, -115, 0, 0, 0,
    250, 600, 0, -50, 50, 30,
    760, 30, -60, -100, 30, -30,
    960, -300, -100, -10, -30, 15,
    730, 40, -50, -20, 0, 0,
    675, 100, -175, -50, 0, 0,
    665, 250, 25, -35, 0, -100,
    810, -30, -60, -20, -60, 60,
    881.82, -54.55, 9.09, -4.55, -9.09, -22.73,
    780, 0, -300, -20, -30, -50,
    880, 120, -100, -50, -90, -60,
    940, 0, -30, -30, 0, -20,
    775, 0, -25, -20, 0, -40,
    855, -30, -60, -30, -30, 15,
    805, -150, -60, 0, -60, -30,
    635, 100, -25, 0, -60, -10,
    720, 120, 15, -40, -70, -30,
    760, 60, 20, 10, 60, -80,
    870, -250, -100, -30, 0, 0,
    700, 20, 20, 60, 25, -10
  ),..Dim = c(29,6)),
  ...

  Y= structure(.Data=c(
    60, 50, 55, 45, 55, 52, 58, 50,
    95, 90, 80, 75, 60, 60, 65, 75,
    70, 90, 50, 30, 10, 30, 40, 50,
    70, 70, 80, 70, 60, 50, 70, 80,
    65, 80, 70, 60, 55, 50, 40, 45,
    80, 90, 60, 70, 70, 40, 50, 50,
    70, 60, 50, 40, 60, 50, 60, 50,
    70, 75, 65, 55, 57, 56, 63, 57,
    40, 90, 20, 30, 50, 15, 10, 80,
    80, 80, 70, 70, 80, 70, 70, 80,
    90, 85, 75, 40, 45, 40, 75, 40,
    70, 80, 60, 60, 70, 80, 85, 75,
    50, 50, 40, 50, 80, 60, 50, 70,
    70, 60, 50, 40, 60, 50, 50, 60,
    70, 70, 30, 30, 30, 30, 30, 40,
    70, 60, 100, 60, 65, 50, 50, 50,
    70, 80, 70, 50, 50, 50, 50, 90,
    30, 40, 70, 80, 90, 80, 60, 80,
    30, 40, 10, 20, 60, 50, 10, 30,
    100, 100, 80, 60, 70, 50, 90, 50,
    100, 90, 70, 80, 90, 80, 70, 80,
    40, 50, 30, 60, 60, 70, 30, 70,
    70, 65, 60, 65, 60, 60, 50, 55,
    50, 40, 30, 60, 80, 90, 20, 70,
    70, 70, 65, 60, 55, 50, 50, 50,
    85, 80, 70, 60, 55, 70, 70, 55,
    60, 60, 80, 70, 40, 70, 80, 60,
    65, 50, 50, 75, 85, 75, 60, 85,
    50, 60, 70, 40, 50, 50, 60, 80
  ),..Dim = c(29,8)),
  ...

  dzero=6,
  IDzero= structure(.Data=c(
    1,0,0,0,0,0,
    0,1,0,0,0,0,
    0,0,1,0,0,0,
    0,0,0,1,0,0,
    0,0,0,0,1,0,
    0,0,0,0,0,1
  ),..Dim = c(6,6)),
  gzero=1,
  X= structure(.Data=c(
    1,1,0,0,0,1,
    1,1,0,1,0,0,
    1,1,1,0,1,0,
    1,0,1,0,0,0,
    1,0,0,0,1,1,
    1,0,1,1,0,1,
    1,1,1,1,1,1,
    1,0,0,1,1,0
  ),..Dim = c(8,6)),
  Azero= structure(.Data=c(
    0, 0.1
  ),..Dim = c(1,2)),
  SAzero= structure(.Data=c(
    10, 0,
    0, 10
  ),..Dim = c(2,2)),
  N=29, m=8, q=6)
```


Figura G.4: Arquivo *list* utilizado na estimação dos parâmetros do Modelo Bayesiano II

```

list(
  U= structure(.Data=c(
    622.5, 37.5, 5, -100, -35, 10,
    825, 50, -60, -20, -45, -15,
    744, 80, -8, 8, -20, -4,
    585, 225, 50, 0, -180, -150,
    735, -100, 15, 45, -20, -10,
    824, -420, -128, -8, -12, -80,
    650, 30, -40, -10, 30, 10,
    640, 175, -45, 5, 0, -30,
    650, 0, -75, -45, -90, -60,
    770, 0, -115, 0, 0, 0,
    250, 600, 0, -50, 50, 30,
    760, 30, -60, -100, 30, -30,
    960, -300, -100, -10, -30, 15,
    730, 40, -50, -20, 0, 0,
    675, 100, -175, -50, 0, 0,
    665, 250, 25, -35, 0, -100,
    810, -30, -60, -20, -60, 60,
    881.82, -54.55, 9.09, -4.55, -9.09, -22.73,
    780, 0, -300, -20, -30, -50,
    880, 120, -100, -50, -90, -60,
    940, 0, -30, -30, 0, -20,
    775, 0, -25, -20, 0, -40,
    855, -30, -60, -30, -30, 15,
    805, -150, -60, 0, -60, -30,
    635, 100, -25, 0, -60, -10,
    720, 120, 15, -40, -70, -30,
    760, 60, 20, 10, 60, -80,
    870, -250, -100, -30, 0, 0,
    700, 20, 20, 60, 25, -10
  ), Dim = c(29,6)),
  ...

  Y= structure(.Data=c(
    60, 50, 55, 55, 52,
    95, 90, 80, 60, 60,
    70, 90, 50, 10, 30,
    70, 70, 80, 60, 50,
    65, 80, 70, 55, 50,
    80, 90, 60, 70, 40,
    70, 60, 50, 60, 50,
    70, 75, 65, 57, 56,
    40, 90, 20, 50, 15,
    80, 80, 70, 80, 70,
    90, 85, 75, 45, 40,
    70, 80, 60, 70, 80,
    50, 50, 40, 80, 60,
    70, 60, 50, 60, 50,
    70, 70, 30, 30, 30,
    70, 60, 100, 65, 50,
    70, 80, 70, 50, 50,
    30, 40, 70, 90, 80,
    30, 40, 10, 60, 50,
    100, 100, 80, 70, 50,
    100, 90, 70, 90, 80,
    40, 50, 30, 60, 70,
    70, 65, 60, 60, 60,
    50, 40, 30, 80, 90,
    70, 70, 65, 55, 50,
    85, 80, 70, 55, 70,
    60, 60, 80, 40, 70,
    65, 50, 50, 85, 75,
    50, 60, 70, 50, 50
  ), Dim = c(29,5)),
  ...

  dzero=6,
  IDzero= structure(.Data=c(
    1,0,0,0,0,0,
    0,1,0,0,0,0,
    0,0,1,0,0,0,
    0,0,0,1,0,0,
    0,0,0,0,1,0,
    0,0,0,0,0,1
  ), Dim = c(6,6)),
  gzero=1,
  X= structure(.Data=c(
    1,1,0,0,0,1,
    1,1,0,1,0,0,
    1,1,1,0,1,0,
    1,0,0,0,1,1,
    1,0,1,1,0,1
  ), Dim = c(5,6)),
  Azero= structure(.Data=c(
    0, 0.1
  ), Dim = c(1,2)),
  SAzero= structure(.Data=c(
    10, 0,
    0, 10
  ), Dim = c(2,2)),
  N=29, m=5, q=6)

```

Figura G.5: Arquivo *list* utilizado na estimação dos parâmetros do Modelo Bayesiano III

```

list(
  U= structure(.Data=c(
    622.5, 37.5, 5, -100, -35, 10,
    825, 50, -60, -20, -45, -15,
    744, 80, -8, 8, -20, -4,
    585, 225, 50, 0, -180, -150,
    735, -100, 15, 45, -20, -10,
    824, -420, -128, -8, -12, -80,
    650, 30, -40, -10, 30, 10,
    640, 175, -45, 5, 0, -30,
    650, 0, -75, -45, -90, -60,
    770, 0, -115, 0, 0, 0,
    250, 600, 0, -50, 50, 30,
    760, 30, -60, -100, 30, -30,
    960, -300, -100, -10, -30, 15,
    730, 40, -50, -20, 0, 0,
    675, 100, -175, -50, 0, 0,
    665, 250, 25, -35, 0, -100,
    810, -30, -60, -20, -60, 60,
    881.82, -54.55, 9.09, -4.55, -9.09, -22.73,
    780, 0, -300, -20, -30, -50,
    880, 120, -100, -50, -90, -60,
    940, 0, -30, -30, 0, -20,
    775, 0, -25, -20, 0, -40,
    855, -30, -60, -30, -30, 15,
    805, -150, -60, 0, -60, -30,
    635, 100, -25, 0, -60, -10,
    720, 120, 15, -40, -70, -30,
    760, 60, 20, 10, 60, -80,
    870, -250, -100, -30, 0, 0,
    700, 20, 20, 60, 25, -10
  ), Dim = c(29,6)),
  ...

  Y= structure(.Data=c(
    600, 500, 550, 450, 550, 520, 580, 500,
    950, 900, 800, 750, 600, 600, 650, 750,
    700, 900, 500, 300, 100, 300, 400, 500,
    700, 700, 800, 700, 600, 500, 700, 800,
    650, 800, 700, 600, 550, 500, 400, 450,
    800, 900, 600, 700, 700, 400, 500, 500,
    700, 600, 500, 400, 600, 500, 600, 500,
    700, 750, 650, 550, 570, 560, 630, 570,
    400, 900, 200, 300, 500, 150, 100, 800,
    800, 800, 700, 700, 800, 700, 700, 800,
    900, 850, 750, 400, 450, 400, 750, 400,
    700, 800, 600, 600, 700, 800, 850, 750,
    500, 500, 400, 500, 800, 600, 500, 700,
    700, 600, 500, 400, 600, 500, 500, 600,
    700, 700, 300, 300, 300, 300, 300, 400,
    700, 600, 1000, 600, 650, 500, 500, 500,
    700, 800, 700, 500, 500, 500, 500, 900,
    300, 400, 700, 800, 900, 800, 600, 800,
    300, 400, 100, 200, 600, 500, 100, 300,
    1000, 1000, 800, 600, 700, 500, 900, 500,
    1000, 900, 700, 800, 900, 800, 700, 800,
    400, 500, 300, 600, 600, 700, 300, 700,
    700, 650, 600, 650, 600, 600, 500, 550,
    500, 400, 300, 600, 800, 900, 200, 700,
    700, 700, 650, 600, 550, 500, 500, 500,
    850, 800, 700, 600, 550, 700, 700, 550,
    600, 600, 800, 700, 400, 700, 800, 600,
    650, 500, 500, 750, 850, 750, 600, 850,
    500, 600, 700, 400, 500, 500, 600, 800
  ), Dim = c(29,8)),
  ...

  dzero=6,
  IDzero= structure
  (.Data=c(
    1,0,0,0,0,0,
    0,1,0,0,0,0,
    0,0,1,0,0,0,
    0,0,0,1,0,0,
    0,0,0,0,1,0,
    0,0,0,0,0,1
  ), Dim = c(6,6)),
  gzero=1,
  X= structure(.Data=c(
    1,1,0,0,0,1,
    1,1,0,1,0,0,
    1,1,1,0,1,0,
    1,0,1,0,0,0,
    1,0,0,0,1,1,
    1,0,1,1,0,1,
    1,1,1,1,1,1,
    1,0,0,1,1,0
  ), Dim = c(8,6)),
  N=29, m=8, q=6)

```

Figura G.6: Arquivo *list* utilizado na estimação dos parâmetros do Modelo Bayesiano IV

```
list(
  U= structure(.Data=c(
    622.5, 37.5, 5, -100, -35, 10,
    825, 50, -60, -20, -45, -15,
    744, 80, -8, 8, -20, -4,
    585, 225, 50, 0, -180, -150,
    735, -100, 15, 45, -20, -10,
    824, -420, -128, -8, -12, -80,
    650, 30, -40, -10, 30, 10,
    640, 175, -45, 5, 0, -30,
    650, 0, -75, -45, -90, -60,
    770, 0, -115, 0, 0, 0,
    250, 600, 0, -50, 50, 30,
    760, 30, -60, -100, 30, -30,
    960, -300, -100, -10, -30, 15,
    730, 40, -50, -20, 0, 0,
    675, 100, -175, -50, 0, 0,
    665, 250, 25, -35, 0, -100,
    810, -30, -60, -20, -60, 60,
    881.82, -54.55, 9.09, -4.55, -9.09, -22.73,
    780, 0, -300, -20, -30, -50,
    880, 120, -100, -50, -90, -60,
    940, 0, -30, -30, 0, -20,
    775, 0, -25, -20, 0, -40,
    855, -30, -60, -30, -30, 15,
    805, -150, -60, 0, -60, -30,
    635, 100, -25, 0, -60, -10,
    720, 120, 15, -40, -70, -30,
    760, 60, 20, 10, 60, -80,
    870, -250, -100, -30, 0, 0,
    700, 20, 20, 60, 25, -10),
    .Dim = c(29,6)),
  ...

  Y= structure(.Data=c(
    600, 500, 550, 550,520,
    950, 900, 800, 600,600,
    700, 900, 500, 100,300,
    700, 700, 800, 600,500,
    650, 800, 700, 550,500,
    800, 900, 600, 700,400,
    700, 600, 500, 600,500,
    700, 750, 650, 570,560,
    400, 900, 200, 500,150,
    800, 800, 700, 800,700,
    900, 850, 750, 450,400,
    700, 800, 600, 700,800,
    500, 500, 400, 800,600,
    700, 600, 500, 600,500,
    700, 700, 300, 300,300,
    700, 600, 1000, 650,500,
    700, 800, 700, 500,500,
    300, 400, 700, 900,800,
    300, 400, 100, 600,500,
    1000, 1000, 800, 700,500,
    1000, 900, 700, 900,800,
    400, 500, 300, 600,700,
    700, 650, 600, 600,600,
    500, 400, 300, 800,900,
    700, 700, 650, 550,500,
    850, 800, 700, 550,700,
    600, 600, 800, 400,700,
    650, 500, 500, 850,750,
    500, 600, 700, 500,500
    ),.Dim = c(29,5)),
  ...

  dzero=6,
  IDzero= structure(.Data=c(
    1,0,0,0,0,0,
    0,1,0,0,0,0,
    0,0,1,0,0,0,
    0,0,0,1,0,0,
    0,0,0,0,1,0,
    0,0,0,0,0,1
    ),.Dim = c(6,6)),
  gzero=1,
  X= structure(.Data=c(
    1,1,0,0,0,1,
    1,1,0,1,0,0,
    1,1,1,0,1,0,
    1,0,0,0,1,1,
    1,0,1,1,0,1
    ),.Dim = c(5,6)),
  N=29, m=5, q=6)
```

Referências Bibliográficas

- [1] Albert, J. H. e Chib, S. (1993). Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, 88, 669-679.
- [2] Allenby, G. M., Arora, N. e Ginter, J. L. (1995). Incorporating Prior Knowledge into the Analysis of Conjoint Studies. *Journal of Marketing Research*, 32, 152-162.
- [3] Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. 2.ed., John Wiley & Sons.
- [4] Artes, R. (1991). *Análise de Preferência 'Conjoint Analysis'*. Dissertação de Mestrado. IME/USP, São Paulo.
- [5] Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of The Royal Statistical Society*, 36, 192-236.
- [6] Brooks, S. P. e Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7, 434-455.
- [7] Carrol, J. D. (1972). Individual differences and multidimensional scaling em Shepard, R. N.; Romney, A. K. e Nerlone, S. B., eds. *Multidimensional scaling: theory and applications in the behavioral sciences*, New York: Seminar Press, 105-57.
- [8] Chapman, R. G. e Staelin, R. (1982). Exploiting Rank Ordered Choice Set Data Within the Stochastic Utility Model. *Journal of Marketing Research*, 19, 288-301.
- [9] Congdon, P. (2001). *Bayesian statistical modelling*. New York: Wiley & Sons.
- [10] Dawes, R. M. e Corrigan, B. (1974). Linear Models in Decision Making. *Psychology Bulletin*, 81, 95-106.

- [11] Dean, A. e Voss, D. (1999). Design and Analysis of Experiments. *Spring Texts in Statistics*, New York: Springer-Verlag.
- [12] DeGroot, M. H. (1974). *Optimal Statistical Decisions*. New York: McGraw-Hill.
- [13] Fishbein, M. (1967). A Behavior Theory Approach to the Relations Between Beliefs About an Object and the Attitude Towards the Object em Fishbein, M.(Ed.), *Readings in Attitude and Theory Measurement*, New York: Wiley & Sons.
- [14] Friedmann, L. S. (1998). *Análise de Preferência*. Dissertação de Mestrado. EAESP/FGV, São Paulo.
- [15] Gelfand, A. E. e Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85, 398-409.
- [16] Gelman, A. e Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7, 457-511.
- [17] Geman, S. e Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.
- [18] Graybill, F. A. (1983). *Matrices with applications in Statistics*. 2.ed., Wadsworth.
- [19] Green, P. E. (1984). Hybrid Models for Conjoint Analysis: An Expository Review. *Journal of Marketing Research*, 21, 155-169.
- [20] Green, P. E. e Carmone, F. J. (1970). Multidimensional scaling and scaling and related techniques in marketing research. *Marketing Science Institute Series*. Boston: Allyn.
- [21] Green, P. E., Carmone, F. J. e Wind, Y. (1972). Subjective Evaluation Models and Conjoint Measurement. *Behavioral Science*, 17, 288-299.
- [22] Green, P. E., Goldberg, S. M. e Montemayor, M. (1981). A hybrid utility estimation model for conjoint analysis". *Journal of Marketing*, 45(1), 33-41.
- [23] Green, P. E. e Krieger, A. M. (1996). Individualized Hybrid Models for Conjoint Analysis. *Management Science*, 42, 850-867.

- [24] Green, P. E. e Rao, V. R. (1971). Conjoint Measurement for Quantifying Judgmental Data. *Journal of Marketing Research*, 8, 355-363.
- [25] Green, P. E. e Srinivasan, V. (1978). Conjoint Analysis in Consumer Research: Issues and Outlook. *Journal of Consumer Research*, 5, 103-23.
- [26] Green, P. E. e Srinivasan, V. (1990). Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice. *Journal of Marketing*, 3-18.
- [27] Green, P. E. e Wind, J. (1975). New Way to Measure Consumers' Judgments. *Harvard Business Review*, July-August, 107-117.
- [28] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97-109.
- [29] Johnson, R. M. (1974). Trade-off Analysis of Consumer Values. *Journal of Marketing Research*, 22, 121-127.
- [30] Johnson, R. M. (1987). Adaptive Conjoint Analysis. In Proceedings of the Sawtooth software conference on perceptual mapping, conjoint analysis, and computer interviewing. Sun Valley. 253-66.
- [31] Kruskal, J. B. (1965). Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data. *Journal of The Royal Statistical Society*. 27, 251-263.
- [32] Lenk, P. J., DeSarbo, S., Green, P. E. e Young, M. R. (1996). Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs. *Marketing Science*, 15, 173-191.
- [33] Luce, R.D. e Tukey, J. W.,(1964). Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement. *Journal Math. Psychology*, 1, 1-27.
- [34] Malhotra, N. K. (1986). An Approach to the Measurement of Consumer Preferences Using Limited Information. *Journal of Marketin Research*, 23, 33-40.
- [35] Mardia, K. V., Kent, J. T., Bibby, J. M. (1979). *Multivariate Analysis*. London: Academic Press.

- [36] Marshall, P. e Bradlow, E. T. (2002). A Unified Approach to Conjoint Analysis Models. *Journal of the American Statistical Association*, 97, 674- 681.
- [37] McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behaviour. *Frontiers of Econometrics*, ed. P. Zarembka, New York: Academic Press, 105-142.
- [38] Neter, J., Kutner, M. H., Nachtsheim, C. J. e Wasserman, W. (1996). *Applied Linear Statistical Models*. 4.ed., Boston: Mc Graw Hill.
- [39] Paulino, C. D., Turkman, M. A. A. e Murteira, B. (2003). *Estatística Bayesiana*. Lisboa: Fundação Calouste Gulbenkian.
- [40] Rao, V. R. e Hauser, J. R. (2002). Conjoint Analysis, Related Modeling, and Applications. *Advances in Marketing Research: Progress and Prospects*, Setembro, 2002.
- [41] Robert, C. P. (2001). *The Bayesian Choice*. 2.ed., Springer-Verlag.
- [42] Smith, A. F. M. e Roberts, G. O. (1993). Bayesian Computatios via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods. *Journal of The Royal Statistical Society*, 55, 3-23.
- [43] Srinivasan, V. e Shocker, A. D. (1973). Linear Programming Techniques for Multidimensional Analysis of Preferences. *Psychometrika*, 38, 337-369.
- [44] Srinivasan, V. e Park, C. S. (1997). Surprising Robustness of the Self-Explicated Approach to Customer Preference Structure Measurement. *Journal fo Marketing Research*, 34, 286-291.
- [45] Tanner, M. A. e Wong, W. H. (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, 82, 528-550.
- [46] Zellner, A. (1996). *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley & Sons.
- [47] Wind, J., Green, P. E., Shifflet, D. e Scarbrough, M. (1989). Courtyard by Marriott: Designing a Hotel Facility with Consumer-Based Marketing Models. *Interfaces*, pp.25-47.
- [48] Wittink, D.R. e Cattin, P. (1989). Commercial Use ofConjoint Analysis: An Update. *Journal of Marketing*, 53, 91-96.