

UNIVERSIDADE FEDERAL DE JUIZ DE FORA

DEPARTAMENTO DE ESTATÍSTICA

CURSO DE ESTATÍSTICA

Thiago Otoni Matos

**Conjoint Analysis: Uma Aplicação ao Marketing**

JUIZ DE FORA

2011

Thiago Otoni Matos

## CONJOINT ANALYSIS: UMA APLICAÇÃO AO MARKETING

Monografia apresentada ao curso de Estatística da Universidade Federal de Juiz de Fora, como requisito para a obtenção do grau de Bacharel em Estatística.

Orientador: Lupércio França Bessegato

Doutor em Estatística – UFMG

JUIZ DE FORA

2011

# CONJOINT ANALYSIS: UMA APLICAÇÃO AO MARKETING

Monografia apresentada ao curso de Estatística da Universidade Federal de Juiz de Fora, como requisito para a obtenção do grau de Bacharel em Estatística.

Aprovada em 30 de novembro de 2011.

## BANCA EXAMINADORA

---

Lupércio França Bessegato (orientador)

Doutor em Estatística – Universidade Federal de Minas Gerais

---

Alfredo Chaoubah

Doutora em Engenharia Elétrica – PUC-Rio de Janeiro

---

Roberto Quinino

Doutor em Engenharia de Produção - Universidade de São Paulo

Matos, Thiago Otoni – Juiz de Fora, 2011

Conjoint Analysis: Uma Aplicação ao Marketing/ Thiago Otoni Matos

125.p

Monografia – Universidade Federal de Juiz de Fora e Instituto de  
Ciências Exatas

Orientador: Lupércio França Bessegato

À Deus por sempre ter indicado o caminho certo .

E aos meus pais.

“Jamais se desespere em meio às sombrias  
aflições de sua vida , pois das nuvens mais negras  
cai água límpida e fecunda”.

*Provérbio Chinês*

## AGRADECIMENTOS

Primeiramente agradeço a Deus por ter me ajudado a superar todas as dificuldades encontradas pelo caminho, a vencer todos os desafios e sempre me indicar o caminho certo.

Agradeço aos meus pais, pelo incentivo, apoio e principalmente por terem me ensinado a estudar, por terem me colocado em uma rotina de estudos que me fez chegar até aqui.

A galera da faculdade, por todos os momentos cômicos, pelos desesperos e pelas noites viradas nos últimos períodos.

Em especial ao Vitão, pela irmandade e por ter largado do meu pé e ter posto em prática todos os ensinamentos que lhe proporcionei desde o ensino médio até aqui. Ao Samuel, meu grande amigo Samuel, pelas caronas, pelas conversas fiadas, pelas forças, e mentiras contadas. Ao Thales, Marcão, Jarbão por completarem a irmandade musical, e pela grande amizade formada ao longo destes 5 anos.

À Laura, Priscila e Carol, por serem as melhores e mais “palosas” amigas que se pode ter. Muito obrigado pelas risadas, pelos papos e pelos estudos em conjunto.

Ao Lu, Luís Gustavo, Iago, Raquel, Leiliane e Sarah, pelo companheirismo, apoio e pelas listas de exercícios que chegavam bem explicadinhas na nossa mão.

À Karol, pelo companheirismo ao longo de quase toda minha faculdade.

‘

Aos Professores do departamento, pelos ensinamentos passados, pela paciência e pelas aulas tão bem administradas. Em especial ao Clécio que além de ser um ótimo professor, se tornou ao longo deste tempo um grande amigo. Ao professor Alfredo pelo tempo de trabalho em conjunto no Laboratório de Ensinos de Saúde. Ao professor Marcel, por sempre querer o melhor para este curso, sempre visando o bem dos alunos e ao professor Lupércio por ser o orientador deste trabalho, e também pelos grandes conselhos dados que vão muito além de meros estudos.

Ao professor Roberto Quinino, por ter se mostrado disposto a participar da banca avaliadora deste trabalho.

A todos vocês, muito obrigado.



## RESUMO

A atividade de marketing, há tempos, reconhece a importância de se medir a qualidade e o valor percebido pelos consumidores e sua interdependência com a satisfação (Zeithaml, 1988). Uma mensuração da estrutura de preferência do consumidor bem feita pode resultar em benefícios mercadológicos às unidades empresariais, como por exemplo, a fidelização do cliente com a marca, o que, em um período extremamente competitivo, se transforma em uma das mais valiosas estratégias mercadológicas e vantagens competitivas.

A estrutura de preferências do consumidor pode ser entendida como as atividades diretamente relacionadas ao comportamento de escolha, decisão de compra e de consumo de bens e/ou serviços, incluindo os processos de decisão que antecedem e/ou sucedem estas ações.

É nesse intuito que a *conjoint analysis* (CA) ou Análise Conjunta de Fatores (ANCF) se destaca, seu objetivo é ajudar aos profissionais de marketing a entender a complexidade do processo de escolha e decisão de compra do consumidor, denominada estrutura de preferência dos consumidores. Portanto, a *conjoint analysis* é umas das técnicas estatísticas mais indicadas para: lançar um novo produto, melhorar um existente, descobrir, testar e medir mercados de interesse, saber como um produto já existente está posicionado em um mercado alvo e saber suas vantagens e desvantagens em relação aos concorrentes, ou seja, identificar os principais fatores que apresentam uma maior influência na decisão de compra a fim de torna-lo mais competitivo.

Neste trabalho serão apresentadas duas vertentes de *conjoint analysis*, a tradicional, *conjoint analysis* baseada em notas e em postos (*rank*) e a *conjoint analysis* baseada em escolhas, uma metodologia alternativa que vêm sendo cada vez mais utilizada pelos pesquisadores. Serão apresentadas suas vantagens e desvantagens, suas aplicações, suas particularidades e suas inferências. Por fim, apresentaremos um *case* da área de telecomunicações.

**Palavras-chave:** *Conjoint Analysis*, Modelos de regressão, Marketing,

Preferência do consumidor

## Sumário

Lista de Figuras.....	14
Lista de Tabelas .....	15
1- Introdução.....	16
1.1- Motivação e Introdução.....	16
1.2- Objetivos. ....	19
1.3- Estrutura do Trabalho. ....	19
2- Fundamentação Teórica.....	21
2.1- Comportamentos de Compra do Consumidor.....	21
2.2- Produtos e Serviços .....	22
2.3- Customização .....	27
2.4 – Valor percebido .....	29
2.5- Terminologia .....	31
3- Conjoint Analysis: Definição e Classificação .....	36
3.1- Problemas na mensuração da estrutura da preferência do consumidor....	43
3.2- Vantagens da <i>Conjoint Analysis</i> , nos problemas de Administração. ....	45
3.3- Revisão Bibliográfica .....	47
3.4- Condução de um estudo em CA (HAIR JR., 2005).....	52
3.4.1- Objetivos da <i>Conjoint Analysis</i> .....	53
3.4.2- O projeto de uma <i>Conjoint Analysis</i> .....	54
3.4.3- Suposições da Conjoint Analysis .....	59
3.4.4- Estimação do modelo e avaliação da qualidade do ajuste.....	60

3.4.5-	Interpretação dos resultados.....	61
3.4.6-	Validação dos resultados .....	62
4-	Metodologia.....	64
4.1-	<i>Conjoint Analysis</i> baseada em notas .....	64
4.1.1-	Modelo para a Utilidade.....	64
4.1.2-	Importâncias Relativas (IR) dos atributos. ....	67
4.1.3-	Um exemplo de <i>conjoint analysis</i> baseada em notas.....	69
4.2-	<i>Conjoint Analysis</i> baseada em postos(rank) .....	75
4.2.1-	Modelo para utilidade.....	75
4.2.2-	Cálculo das utilidades parciais ou coeficientes de preferência.....	77
4.2.3-	Importâncias relativas. ....	78
4.2.4-	Um exemplo de <i>conjoint analysis</i> baseada em postos. ....	79
4.3-	<i>Conjoint Analysis</i> baseada em escolhas. ....	83
4.3.1-	Modelo para a Utilidade.....	85
4.3.2-	O Modelo Logit Multinomial.....	87
4.3.2.1	– Estimação dos parâmetros do modelo.....	90
4.3.2.2	– Razão de escolhas .....	96
5-	Aplicação .....	97
5.1-	Contextualização do Problema .....	97
5.2-	Descrições dos dados .....	98
5.3-	Aplicação da <i>Conjoint Analysis</i> baseada em escolhas para estimação de participação de mercado ou <i>share</i> de preferência.....	103
6-	Conclusão e recomendações futuras.....	111
	Referências Bibliográficas .....	113

Apêndice 1- *Conjoint Analysis* tradicional, baseada em notas e Postos no SPSS.

..... 118

## Lista de Figuras

Figura 1- Modelo do comportamento do consumidor.....	21
Figura 2 - Características dos serviços.....	24
Figura 3 - Os níveis de um produto.....	25
Figura 4 - Níveis de desenvolvimento de produtos .....	26
Figura 5 - Escolha de uma técnica multivariada .....	42
Figura 6- Utilidades Parciais dos níveis dos fatores para o exemplo do celular .	74
Figura 8- Utilidades parciais dos níveis dos fatores do detergente.....	83
Figura 9 - Classificação dos modelos probabilísticos de escolha discreta .....	85
Figura 10- base de planejamento no spss para o exemplo do celular .....	119
Figura 11- base de respostas no SPSS para o exemplo do celular.....	119
Figura 12- base de planejamento no SPSS para o exemplo do detergente .....	120
Figura 13- base de respostas no SPSS para o exemplo do detergente.....	120
Figura 14- Coeficientes de preferencia e Importância relativa para o exemplo do celular .....	125
Figura 15- Coeficientes de preferência e importâncias relativas.....	126

## Lista de Tabelas

Tabela 1- Tabela de Importância de desempenho.....	46
Tabela 2- Uma comparação de metodologias conjuntas alternativas .....	55
Tabela 3- níveis dos fatores do exemplo celular .....	69
Tabela 4-notas do consumidor 1 para os 8 estímulos de celular .....	70
Tabela 5- Resumo dos resultados da Conjoint Analysis baseada em notas.....	72
Tabela 6- Níveis dos fatores para o detergente .....	79
Tabela 7- Descrição dos estímulos de detergente e ordenação do consumidor 1 80	
Tabela 8- Utilidades parciais estimadas para o respondente 1.....	82
Tabela 9- Importâncias relativas dos fatores .....	83
Tabela 10 - Função indicadora da função de verossimilhança .....	91
Tabela 11- Níveis dos fatores utilizados na aplicação .....	101
Tabela 12- Entrevista 1, Cenário 1.....	102
Tabela 13- Coeficientes estimados pelo método <i>multinomial logistic</i> do SPSS	104
Tabela 14- Participação de Mercado das Operadoras .....	106
Tabela 15- Comparativo de probabilidades das 10 configurações com maior chance de aquisição das operadoras 2 e 3 .....	107
Tabela 16- Probabilidades de aquisição de planos com preço maior que 180 reais .....	109

# 1- Introdução

## 1.1- Motivação e Introdução

Segundo Kotler & Armstrong (2008), o marketing é um processo administrativo e social pelo qual indivíduos e organizações obtêm o que necessitam e desejam por meio da criação e troca de valor uns com os outros. Em um contexto mais específico dos negócios, o marketing envolve construir relacionamentos lucrativos de longo prazo e de valor com os clientes. Assim, marketing é o processo pelo qual as empresas criam valor para os clientes e constroem fortes relacionamentos com eles, oferecendo mais que o esperado, e recebendo a fidelização e seu valor em troca.

No século XIX, as evoluções no setor da produção foram mais notórias. A busca por processos que reduziam os custos, fez com que a chave para o sucesso das grandes empresas fosse a padronização e a produção em massa. Já com as evoluções tecnológicas, a popularização da internet, a maior capacidade das pessoas absorverem as informações e tirarem suas conclusões do que é melhor ou pior, fez com que o cenário de produção mudasse, passando de produção padronizada para produção customizada, onde as empresas buscam atingir de forma mais direta os clientes que ela pode atender melhor, em busca de sua satisfação e de sua fidelização com a marca.

A atividade de marketing, há tempos, reconhece à importância de se medir a qualidade e o valor percebido pelos consumidores e sua interdependência com a satisfação (ZEITHAML, 1988). Uma mensuração da



estrutura de preferencia do consumidor bem feita pode resultar em benefícios mercadológicos às unidades empresariais, como por exemplo, a fidelização do cliente com a marca, o que, em um mercado extremamente competitivo, se transforma em uma das mais valiosas estratégias mercadológicas e vantagens competitivas.

A estrutura de preferências do consumidor pode ser entendida como as atividades diretamente relacionadas ao comportamento de escolha, decisão de compra e de consumo de bens e/ou serviços, incluindo os processos de decisão que antecedem e sucedem estas ações. Definir os fatores que influenciam esta estrutura e mensurá-los corretamente é uma necessidade vital não apenas para as empresas, mas para toda a organização que se assuma como orientada para o mercado (BASTOS, 2010).

Ainda segundo Kotler e Armstrong (2008), faz parte do processo de marketing a segmentação de mercado, e a definição de mercado-alvo, pois empresas boas em marketing sabem que não podem servir a todos os clientes de todas as formas. Em vez disso, elas precisam concentrar seus recursos nos clientes a que podem servir melhor e de maneira mais lucrativa.

Em ambientes instáveis e caracterizados por intensa concorrência, o custo da implementação de decisões mal avaliadas é, geralmente, alto. Assim é desejável que o processo decisório seja amparado por técnicas que reduzam a incerteza a ele associada (BARQUETTE & CHAOUBAH, 2007).

Com o passar dos anos e com desenvolvimentos de metodologias estatísticas de análises de dados, como modelos de regressão múltipla, análise de correlação, modelos de planejamentos de experimentos, além das técnicas multivariadas fizeram com que os estudos em *marketing* fossem mais enriquecidos (PRETTO & ARTES, 2009). Conseqüentemente estes estudos deixaram de ser em sua maioria subjetiva, com métodos observacionais, e passaram a ser mais quantitativos podendo assim fazer uma mensuração mais fidedigna.

A tendência atual da pesquisa em *marketing* aponta para estudos sobre a análise de preferência do consumidor e o processo de decisão presente no momento da compra de um bem ou na contratação de um serviço (FRIEDMANN, 1998).

É nesse intuito que a *conjoint analysis* (*CA*) ou Análise Conjunta de Fatores (ANCF) se destaca. Seu objetivo é ajudar os profissionais de marketing a entender a complexidade do processo de escolha e decisão de compra do consumidor, denominada estrutura de preferencia dos consumidores. Portanto, a *CA* é umas das técnicas estatísticas mais indicadas para lançar um novo produto, melhorar um existente, descobrir e testar novos mercados, saber como um produto já existente está posicionado em um mercado alvo, saber suas vantagens e desvantagens em relação aos concorrentes, ou seja, identificar os principais fatores que apresentam uma maior influencia na decisão de compra a fim de torna-lo mais competitivo.

## 1.2- Objetivos.

Este trabalho tem como objetivos apresentar a metodologia *conjoint analysis* tradicional e a *conjoint analysis* baseada em escolhas , bem como suas aplicações ,suas vantagens e desvantagens, e uma detalhada análise dos modelos estatísticos aplicados. Uma rápida revisão teórica sobre teorias de *marketing* e economia, para situar os leitores sobre termos técnicos, definições e uma visão geral de onde veio as ideias para as formulações dos modelos estatísticos e por fim uma aplicação a um conjunto de dados.

## 1.3- Estrutura do Trabalho.

No capítulo 1, é apresentada a introdução, a motivação e a estrutura do trabalho.

No capítulo 2, é apresentada uma fundamentação teórica sobre assuntos da área de marketing e economia, como as teorias do (i) comportamento do consumidor, (ii) definição de produtos e serviços, (iii) customização, (iv) valor percebido e (v) uma lista de alguns termos comuns da área.

No capítulo 3, é apresentada a definição de *Conjoint Analysis*, os problemas possíveis gerados pelo seu uso, as vantagens administrativas que ela pode gerar uma revisão bibliográfica e por fim, os passos para a sua aplicação segundo Hair Jr.(2005).

O capítulo 4 é reservado para a metodologia, onde apresentamos os detalhes estatísticos das metodologias tradicionais e baseadas em escolhas.

‘

O capítulo 5 é formado pelas informações geradas através da aplicação da metodologia a um banco de dados.

O capítulo 6 tratará da conclusão, onde além dela, são sugeridos alguns trabalhos futuros e considerações finais.

## 2- Fundamentação Teórica

O objetivo deste capítulo é apresentar contribuições teóricas para gerar sustentabilidade e suporte conceitual para a *Conjoint Analysis*.

### 2.1- Comportamentos de Compra do Consumidor

O mercado consumidor consiste em todos os indivíduos e lares que compram ou adquirem bens e serviços para o consumo pessoal. A maioria das grandes empresas pesquisa como, quando, quanto e por que os consumidores compram.

As decisões de compra dos consumidores são altamente influenciadas por características culturais, sociais, pessoais e psicológicas.

O modelo mais simples de comportamento do consumidor é o de estímulo e resposta.

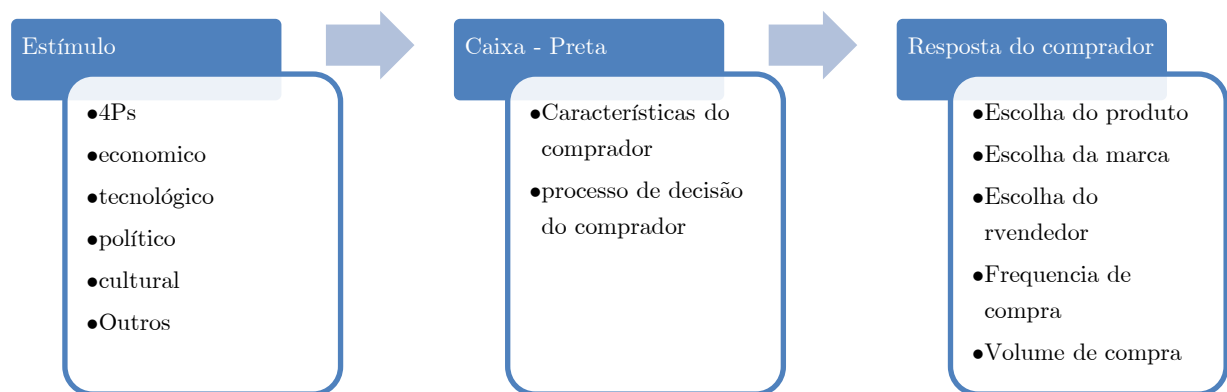


Figura 1- Modelo do comportamento do consumidor

Adaptado de (KOTLER & ARMSTRONG, 2008).

De acordo com esse modelo os estímulos de marketing (4Ps: Praça, Produto, Preço e Promoção) e outras forças importantes penetram na “caixa-preta” do consumidor e produzem determinadas respostas: reações, escolhas, frequências e volumes de compra (KOTLER & ARMSTRONG, 2008).

Para Kotler(1995), o campo do comportamento do consumidor estuda como as pessoas, grupos e organizações selecionam, compram, usam e descartam artigos, serviços, ideias ou experiências para satisfazer as necessidades de seus desejos.

Engel et. ali(2000) definem o comportamento do consumidor como as atividades diretamente envolvidas em obter, consumir e dispor de produtos e serviços, incluindo os processos decisórios que antecedem e sucedem estas ações.

Mowner e Minor(2003) definem comportamento do consumidor como o estudo das unidades compradoras e os processos de troca envolvidos na aquisição, consumo e utilização de mercadorias, serviços, experiências e ideias.

Vemos que o ponto em comum nas definições é o consumidor enquanto pessoa física/instituição relacionado ao ato de adquirir bens ou serviços.

## **2.2- Produtos e Serviços**

Kotler e Armstrong(2008) definem um produto como algo que pode ser oferecido a um mercado para apreciação, aquisição, uso ou consumo e que pode satisfazer um desejo ou uma necessidade. Produtos incluem mais do que apenas

bens tangíveis. Definidos amplamente incluem objetos físicos, serviços, eventos pessoas, lugares, organizações, ideias ou um misto de todas essas entidades.

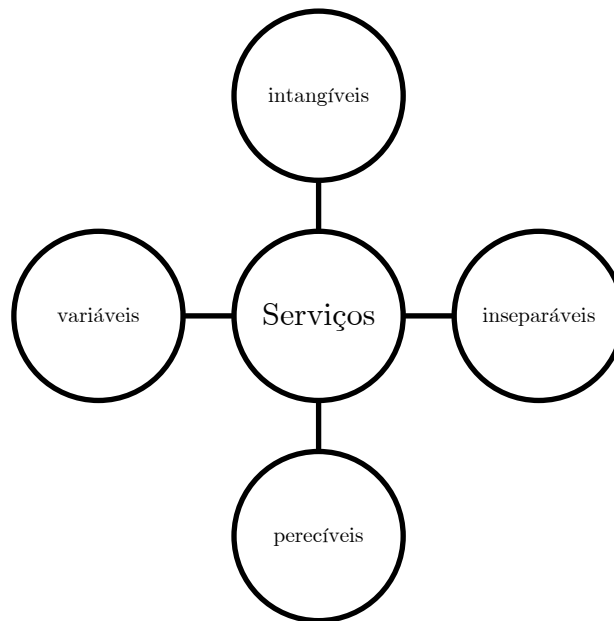
Eles ainda reiteram que, o produto é um elemento chave da oferta ao mercado. O planejamento do *mix* de marketing começa com a formulação de uma oferta que proporcione valor aos clientes-alvo. Essa oferta se torna a base sobre a qual a empresa constrói relacionamentos lucrativos com os clientes.

Nos dias atuais, as ofertas, são mais do que meros bens tangíveis e serviços, são uma combinação de ambos com a finalidade de criar e gerenciar experiências para os clientes.

Devido à sua importância no mundo da economia, daremos uma atenção especial aos serviços. Os serviços são um tipo de produto que consiste em atividades, benefícios, ou satisfações oferecidas para a venda que são essencialmente intangíveis e não resultam na posse de nada como serviços bancários, de viagens aéreas, de saúde entre outros.

De maneira geral os serviços são caracterizados por quatro características principais: são intangíveis, inseparáveis, variáveis e perecíveis.

Figura 2 - Características dos serviços



- **Intangibilidade:** Os serviços não podem ser vistos, provados, ouvidos ou cheirados antes da compra. Por exemplo, quem faz uma cirurgia plástica não pode ver o resultado antes da hora.
- **Variabilidade:** Significa que a qualidade deles depende de quem os fornecem e de quem os executa, bem como quando, onde e como são executados. Por exemplo, duas pizzarias de uma mesma franquia podem ter certa diferença no sabor de seus produtos.
- **Percibilidade:** Os serviços não podem ser armazenados para a venda ou uso posterior.
- **Inseparabilidade:** Os serviços não podem ser separados de seus prestadores. Por exemplo, para realizar um reparo na rede elétrica é necessário um operador especializado, e o desempenho do reparo

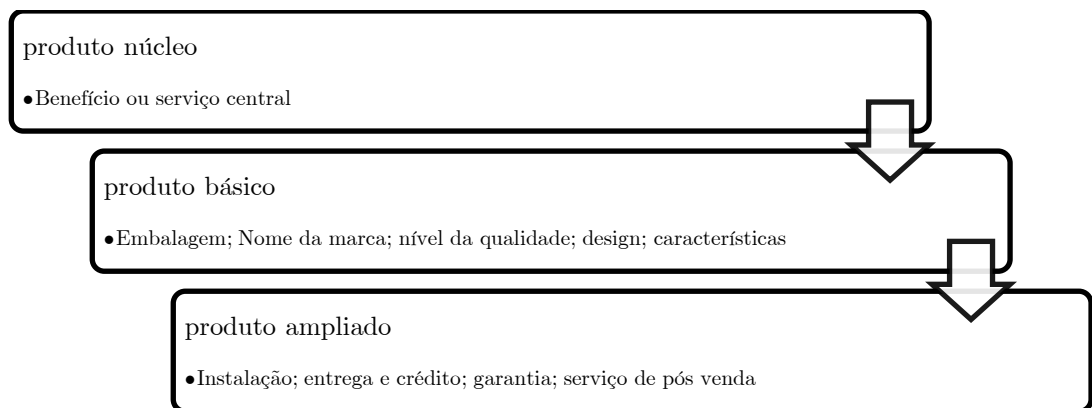


está altamente relacionado com a eficiência do operador (KOTLER & ARMSTRONG, 2008).

As empresas tentam descobrir maneiras de tornar os serviços mais tangíveis para aumentar a produtividade de quem os executa e que é inseparável de seus produtos, padronizar a qualidade diante da variabilidade e aprimorar os movimentos da demanda e as capacidades de ofertas em relação à precibilidade dos serviços.

Os planejadores de produtos precisam pensar em três níveis de produtos e serviços. Cada nível agrega mais valor para o cliente. São eles: O produto núcleo, o produto básico e o produto ampliado.

Figura 3 - Os níveis de um produto



O produto núcleo consiste nos benefícios fundamentais que solucionam o problema e que é o que os clientes buscam quando compram um produto ou serviço. O produto básico existe ao redor do produto núcleo e inclui o nível de qualidade, as características, o design, o nome da marca e a embalagem. O

produto ampliado é o produto básico somado aos diversos serviços e benefícios oferecidos com ele, como garantia, entrega grátis, instalação e manutenção.

Desenvolver um produto ou serviço envolve definir os benefícios que ele oferecerá. Esses benefícios são comunicados e entregues por meio de atributos de produto, o branding e a estratégia de branding que o acompanha, a embalagem, a rotulagem e os serviços de apoio e assistência.



Figura 4 - Níveis de desenvolvimento de produtos

Adaptado de (KOTLER & ARMSTRONG, 2008).

As decisões de atributos de produto envolvem qualidade, características, estilo e design. Entre as decisões de branding estão a escolha de um nome da marca e o desenvolvimento de uma estratégia de marca. A embalagem oferece muitos benefícios importantes, como proteção, economia, praticidade e promoção. As decisões de embalagem quase sempre incluem o desenho dos rótulos, que identificam, descrevem e possivelmente promovem o produto. As empresas também desenvolvem os serviços de apoio e assistência, que aprimoram o serviço ao consumidor e a satisfação dele e são uma defesa contra a concorrência (KOTLER & ARMSTRONG, 2008).

O interessante é que a *conjoint analysis*, pode auxiliar o profissional em todas essas etapas acima, pois todas elas têm contato com o consumidor, e suas ações são representadas em preferências.

### 2.3- Customização

A palavra customizada vem do inglês "*Customers*", significa cliente ou freguês. Então produção customizada é qualquer produto ou serviço que é personalizado para o cliente, em outras palavras, qualquer produto ou serviço que é produzido focado em um tipo específico de consumidor. As empresas procuram ver quais são as necessidades das pessoas e procuram responder com o produto e serviço mais adequado.

Os principais aspectos que caracterizam a produção customizada é a existência de mercados fragmentados, com nichos heterogêneos e demanda instável.

O alvo principal da customização é produzir grandes variedades de produtos ou serviços de forma que quase todos encontram exatamente o que querem a um preço razoável (KUNDER, 2010). Nesse contexto, vemos que muitas empresas utilizam-se de linhas de produtos.

Uma linha de produtos é um grupo de produtos intimamente relacionados com relação às funções que executam, às necessidades de compra dos clientes e/ou aos canais de distribuição (KOTLER & ARMSTRONG, 2008).

Uma linha de produtos bem estruturada é capaz de fornecer as melhores soluções para as necessidades ou desejos de diversos clientes, com gostos diferentes e com poderes aquisitivos diferentes, portanto atingindo segmentos e nichos diferentes. Por exemplo, os tênis Nike, que variam desde calçados para prática esportiva (amadores, casuais e profissionais.) até para cliente casuais onde podem ser encontrados calçados com diferentes tecnologias e preços. Outro possível objetivo pode ser persuadir o cliente a comprar outros produtos da linha. Assim a BMW pode querer que seus clientes saltassem dos carros da serie 3 para a serie 5, ou serie 7. Outro objetivo pode ser em realizar vendas cruzadas, uma empresa pode vender impressoras e ao mesmo tempo vender os cartuchos de tinta. Ainda outro objetivo pode ser proteção quanto às oscilações econômicas.

É fácil ver que a customização está fortemente presente nos dias atuais, basta ver a indecisão na hora da compra de um de aparelho eletrônico frente a grande quantidade de opções de marcas e configurações. Por exemplo, as diferentes composições e utilidades dos cremes dentais ou as inúmeras possibilidades de se satisfazer ao assistir TV por assinatura, ou a TV em geral que está ganhando mais formas de satisfazer os clientes com as inovações da TV digital e até redes de *fast food* já aderiram a esta nova tendência.

Em contrapartida, há uma tendência mundial de globalização, isto é, padrões e tendências se convergindo de forma que objetos e produtos da mesma classe tendem a ser muito parecidos podendo simplificar o planejamento estratégico de algumas empresas (LEVIIT, 1995).

## **2.4 – Valor percebido**

A noção do comportamento do consumidor traz à luz a importância de entender as premissas dos indivíduos no contexto socioeconômico das empresas. À continuação deste constructo conceitual, apresenta-se a teoria do valor percebido, ou seja, dado o plano no qual o indivíduo está inserido (ambiente e sociedade) e suas inter-relações individuais e coletivas, busca-se entender como o indivíduo avalia o bem, produto ou serviço (CASTRO, 2006).

Para Kotler(1995), o valor percebido indica que uma oferta de marketing para o mercado tem um valor agregado entregue ao consumidor, resultado da diferença entre o valor total esperado (conjunto de benefícios) e o custo total para o consumidor, e que pode ser interpretado como o lucro do consumidor na transação de troca. Dessa forma, o valor percebido é o valor atribuído pelos clientes ao produto ou serviço, baseado na relação entre os benefícios que trará, segundo a ótica do consumidor, e os custos percebidos para sua aquisição, comparativamente à concorrência.

Woodruff(1997) adota o conceito de valor do cliente como sendo a percepção do cliente sobre as preferências e as avaliações dos atributos do

produto, do desempenho desses atributos e das consequências originadas pelo uso. Os clientes enxergam o produto como um conjunto de atributos e desempenham esses atributos; quando compram e usam o produto, desenvolvem preferências e desejos por certos atributos, que lhes proporcionam as consequências desejadas nas situações de uso (conceito de valor em uso), atendendo seus objetivos e gerando satisfações com valor recebido. O valor do cliente é a qualidade percebida pelo mercado, ajustada pelo preço relativo de seu produto.

Castro(2006) afirma que, dado que exista uma alternativa competitiva de uma dada oferta de marketing, os autores resumem a essência do conceito do valor percebido por meio da expressão.

$$(\text{valor}_{\text{oferta}} - \text{preço}_{\text{oferta}}) > (\text{valor}_{\text{alternativa}} - \text{preço}_{\text{alternativa}}) \quad (1.1)$$

Esta expressão compara o valor e o preço da oferta de mercado com o valor e o preço da próxima melhor alternativa. Dessa forma, o cliente deve perceber um efeito diferencial de benefícios superior na oferta (em relação a melhor alternativa concorrente) para optar por ela (CASTRO, 2006).

Um ponto convergência mostra-se evidente entre os diversos autores: o conceito de valor percebido para o cliente está associado ao uso do produto ou serviço (conceito de utilidade), à sua percepção (que determina seu comportamento) e à relação de troca custo por benefícios.

Portanto quanto mais fidedigna for a mensuração do valor percebido, mais uma empresa pode focar nos mínimos detalhes que os clientes julgam essenciais para escolher um produto e/ou um serviço dentro as inúmeras possibilidades que lhe são oferecidas todos os dias.

## 2.5- Terminologia

Como muitos textos que tratam de *Conjoint Analysis* são voltados para a área de marketing e administração, é comum o emprego de termos provenientes da área tais como estímulos, utilidade parcial e *tradeoff*.

Em Da Silva, et al. (2010) vemos uma interessante conceituação dos principais termos básicos.

- **Consumidor, respondente ou julgador** - Correspondem às pessoas que participam da avaliação dos tratamentos, respondem questionários, etc. Em geral, trata-se de uma amostra aleatória de consumidores.
- **Fator ou atributo** - São características, que compõem o tratamento (produto ou serviço), objeto de estudo da análise conjunta. Também são denominados fatores controláveis, variáveis independentes ou atributos e, geralmente, são representados por letras maiúsculas. Atributos podem ser qualitativos (categóricos ordinais ou nominais) ou quantitativos.
- **Nível do atributo ou do fator** – São desmembramentos ou alternativas dos fatores que servem para qualificá-los ou quantificá-los.

- **Tratamentos, produtos** – Os tratamentos são os produtos ou serviços ou conceitos etc. Em geral são especificados pela combinação de um dos níveis de cada atributo.
- **Estímulos** – É o conjunto de tratamentos apresentados ao consumidor para avaliar, julgar, escolher, atribuir uma nota, responder perguntas a respeito e etc. Podem ser instruções, questionários, objetivos reais, bens de consumo, protótipos de produtos, fotos. Em geral na *Conjoint Analysis* baseada em notas são apresentados todos os estímulos (*full-profile* ou fatorial completo) ou uma parte deles (*partial-profiles* ou fatorial fracionário). Já na *Conjoint Analysis* baseada em escolhas, cada consumidor é apresentado a diversos conjuntos de estímulos (*choice-sets*), cada um formado por poucos (4 a 6) tratamentos, e ele é pedido para escolher um em cada conjunto.
- **Ortogonalidade** – É uma restrição matemática imposta para que os efeitos dos atributos ou fatores sejam estimados de forma independente uns aos outros. Dessa forma, a estimativa do efeito de um não é alterada por variações dos outros efeitos.
- **Método do perfil completo** (*full-profile*) – Método de formação dos tratamentos, que consiste na completa descrição do tratamento pela



combinação de um nível de cada atributo, sem haver atributos faltantes. Quando há atributos faltantes na definição dos tratamentos apresentados, denomina-se perfil incompleto.

- **Planejamento fatorial fracionário ou fracionado** – Planejamentos empregados para reduzir o número de tratamentos a serem avaliados com a abordagem do perfil completo. Apresenta-se uma parte da totalidade dos tratamentos aos consumidores. Por exemplo, se há cinco atributos com dois níveis cada, então um fatorial completo resulta em  $2^5 = 32$  tratamentos. Um plano fracionado poderia tentar selecionar a fração  $\frac{1}{4}$  destes, ou 8 tratamentos para comporem o estímulo.
- **Utilidade Parcial ou *part-worth*** – É a estimativa para as preferencias ou utilidades associadas a cada nível dos atributos. Alguns autores propõem a tradução como coeficiente de preferencia (CP) na *Conjoint analysis* tradicional.
- **Efeitos principais** – São os efeitos individuais de cada atributo sobre a preferencia dos consumidores.
- **Utilidade total** – Se refere ao valor atribuído pelo consumidor ao tratamento. O valor deve ser informativo de sua preferencia. Pode ser uma nota de intenção de compra em uma escala numérica, 1 a 5 ou 1 a

10, ou em uma escala de *Lichert*, por exemplo, com definitivamente compraria em um extremo e definitivamente não compraria em outro extremo. O valor informado pelo consumidor pode ser um posto ou *rank*, neste caso variando de 1 a N, em que N é o número de tratamentos que formam o estímulo. Na *Conjoint Analysis* baseada em notas adota-se um modelo no qual a utilidade do tratamento é formada pela soma das utilidades parciais de um específico conjunto de níveis de fatores. Na *Conjoint Analysis* baseada em escolhas os valores da utilidade ou preferência atribuídos pelo consumidor são em geral 0 e 1, respectivamente para os tratamentos escolhido e não escolhido.

- **Modelo de composição** – Classe de modelos multivariados que relaciona a variável dependente às independentes. É o modelo estatístico adotado na *Conjoint Analysis*. Pode ou não incluir interações entre fatores. A escolha do modelo depende dos objetivos e restrições do estudo.
- **Trade off** – Uma característica desejável nas avaliações de preferência. Há um *trade off* quando se informa a preferência em comparação com as outras opções, ou seja, o quanto se prefere cada alternativa em relação às demais. Não confundir este termo com o método *trade off* de apresentação de estímulos que consiste em apresentar os atributos aos pares, um par por vez até esgotar todos os pares. Esta técnica tem sido pouco empregada por apresentar muitas desvantagens.



### 3- Conjoint Analysis: Definição e Classificação

*Conjoint Analysis* é uma técnica estatística multivariada pela qual as preferências dos respondentes por diferentes ofertas são decompostas para determinar a função utilidade para cada atributo (KOTLER, 1995). Segundo Hair Jr. (2005) a *conjoint analysis* é uma técnica multivariada usada especificamente para entender como os respondentes desenvolvem preferências por produtos ou serviços.

A *conjoint analysis* se preocupa em estimar modelos que expressem as preferências individuais de um consumidor com a função de um conjunto de características pré-definidas, como por exemplo, em um produto. Sendo possível ainda mensurar o impacto de cada uma dessas características sobre a preferência do consumidor. Este produto pode ser qualquer objeto tal como um aparelho eletrônico, automóvel, alimentos, produtos de beleza, ou, então, pode representar um serviço, como, por exemplo, a composição de um plano de saúde, forma de atendimento em uma agência bancária ou estrutura de um curso de pós-graduação. (PRETTO & ARTES, 2009)

Analisando as definições acima vemos algo em comum: os autores levam em consideração a avaliação da preferência de consumidores sobre um determinado produto/serviço, este composto por atributos específicos que podem ser expressas separadamente.

Ao entender exatamente como as pessoas tomam decisões e o que elas valorizam em produtos e serviços desejados, as empresas podem focar seus

esforços nessas áreas, o que valoriza o cliente e diminui os custos para os seus cofres.

Segundo Hair Jr. (2005), a flexibilidade e a unicidade da Análise Conjunta surgem a partir de:

- (1) Da sua habilidade em acomodar tanto uma variável dependente métrica quanto não métrica;
- (2) Do uso de variáveis preditoras categóricas;
- (3) Das suposições gerais sobre as relações de variáveis independentes com a variável dependente.

O uso da técnica possibilita que qualquer conjunto de objetos (marcas, empresas) ou de conceitos (posicionamento, benefícios, imagens) seja avaliado como uma coleção de atributos. Desta forma, “*a flexibilidade da Análise Conjunta viabiliza sua aplicação em praticamente qualquer área na qual as decisões são estudadas*” (HAIR JR., 2005).

Um estudo com *Conjoint analysis* requer que os tratamentos/produtos avaliados pelos consumidores sejam formados pela combinação de fatores ou atributos, assim se o produto alvo for um *notebook*, o processador, o disco rígido (HD), a memória, o leitor de *Blu-ray*, o preço e até a própria marca podem ser possíveis atributos na caracterização do produto, estes atributos com diversas alternativas ou variações denominadas níveis, o que no nosso exemplo poderiam ser: processador core i3 ou i5, HD de 300mb ou de 600mb, possui leitor de *Blu-ray* ou não, preço alto ou preço baixo.

Esses atributos e os níveis podem ser usados para definir diferentes perfis e (ou) estímulos para os consumidores avaliarem, cada um destes representando um produto.

Ao analisar os itens que foram escolhidos ou que apresentaram maiores preferências a partir dos perfis de produtos oferecidos ao cliente, é possível avaliar o que está impulsionando a preferência dos mesmos, ou seja, é possível saber quais níveis e atributos mostrados são importantes positivamente e negativamente, e o mais importante, assim é possível avaliar numericamente a importância de cada atributo e cada nível.

Deste modo, a *conjoint analysis* avalia a utilidade total de um produto ou serviço, que é a medida de avaliação da preferência do consumidor, como a soma das utilidades dos níveis dos atributos.

$$Utilidade\ total = \sum Utilidades\ dos\ atributos \quad (3.1)$$

Assim vemos que a metodologia é decomposicional, pois, objetiva-se estimar a importância de cada fator (utilidades dos níveis dos fatores) a partir da avaliação do efeito conjunto do tratamento. Desta forma, mediante a avaliação de combinações de níveis, quer-se conhecer a contribuição de cada fator na formação da preferência do consumidor (BASTOS, 2010), já que a única contribuição do respondente é a nota, ordenação ou escolha dos tratamentos que ele realiza.

Para medida de comparação, a seguir a definição das metodologias composicional e híbrida:

Na abordagem composicional, o consumidor avalia os níveis de cada fator em forma sequencial e isolada dos outros fatores. Partindo-se das avaliações em separado dos fatores e seus níveis, deseja-se conhecer a preferência global por um tratamento (produto/ serviço/ conceito) (BASTOS, 2010), como na análise discriminante por exemplo.

A abordagem híbrida, formada pela união dos modelos de preferência composicionais e decomposicionais, envolve duas tarefas: uma que submete o entrevistado na avaliação de fatores independentemente dos outros e outra, posterior, onde este desenvolve um exercício conjunto. (CÓRDOVA, 2002)

Existem vários métodos de aplicação da *Conjoint analysis* (*CA*), no mais tradicional, a utilidade que cada consumidor atribui a cada produto é definida por notas ou postos. Alternativa que vem sendo cada vez mais utilizada é a *CA* baseada em escolhas, onde, como o próprio nome diz a utilidade que cada consumidor atribui a cada produto, é dada por escolhas. Logicamente existem algumas diferenças tanto na estimação quanto na inferência das duas metodologias, mais a frente veremos com mais detalhes.

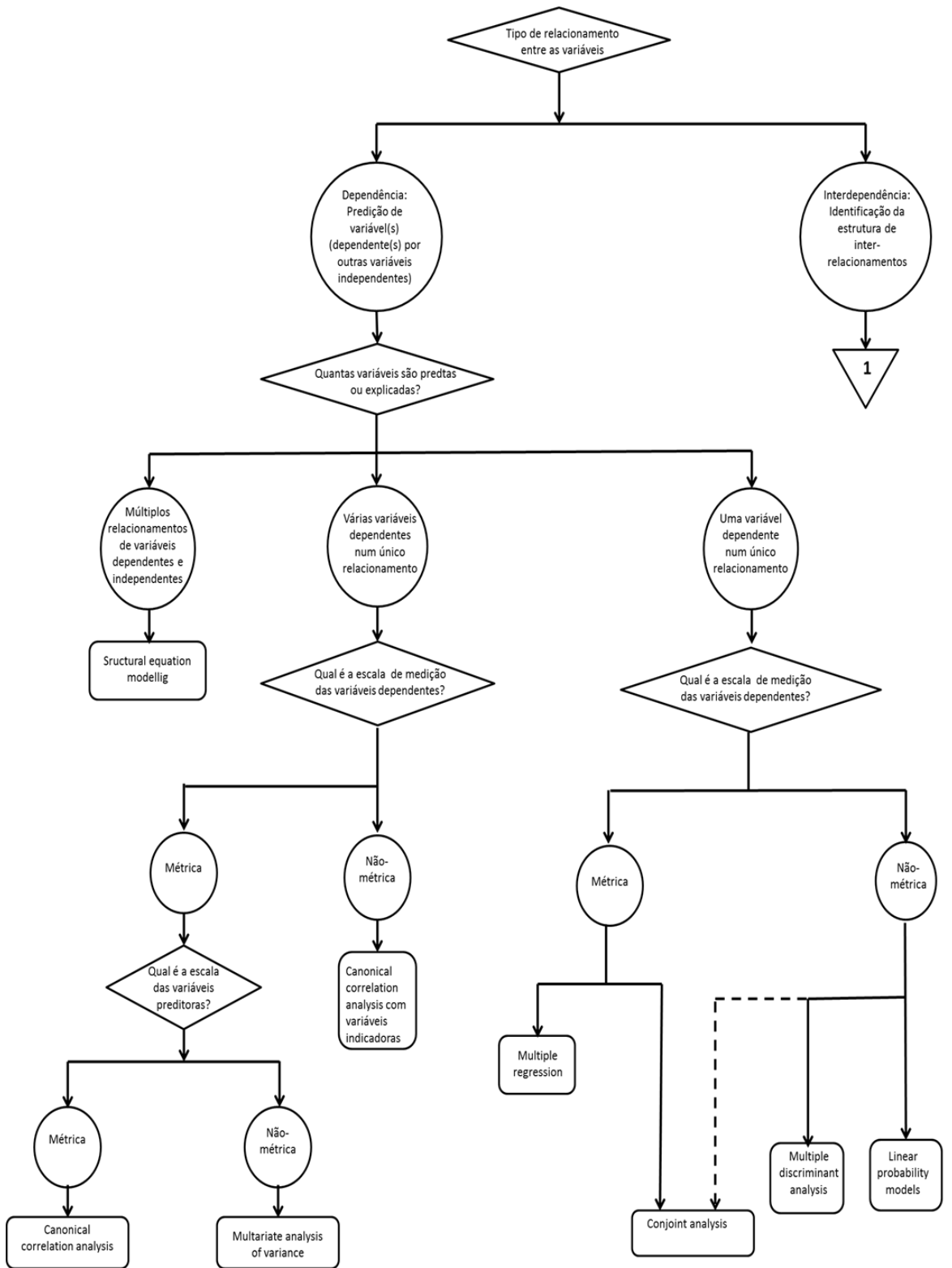
A seguir temos um fluxograma que nos permite fazer certa comparação entre as técnicas multivariadas.

Pela figura 5, vemos que a *Conjoint Analysis* é classificada como uma técnica de dependência, isto é, uma técnica que trata modelos com variável dependente, e concluímos que realmente é uma técnica de dependência, pois se deseja explicar a preferência do indivíduo pelos níveis dos fatores do produto/tratamento.

A variável dependente em CA pode ser métrica (escala intervalar ou de razão) ou não métrica (escala ordinal). Pode assumir o valor de uma nota (discreta ou contínua) ou um posto atribuído a um estímulo (SIQUEIRA, 2000).

Atualmente, o termo *conjoint analysis* traduzido como Análise Conjunta de Fatores (ANCF), é muitas vezes confundido com o de Preferência Declarada (PD), mas na verdade é um dos métodos que fazem parte das técnicas empregadas na PD (BRITO, 2007).





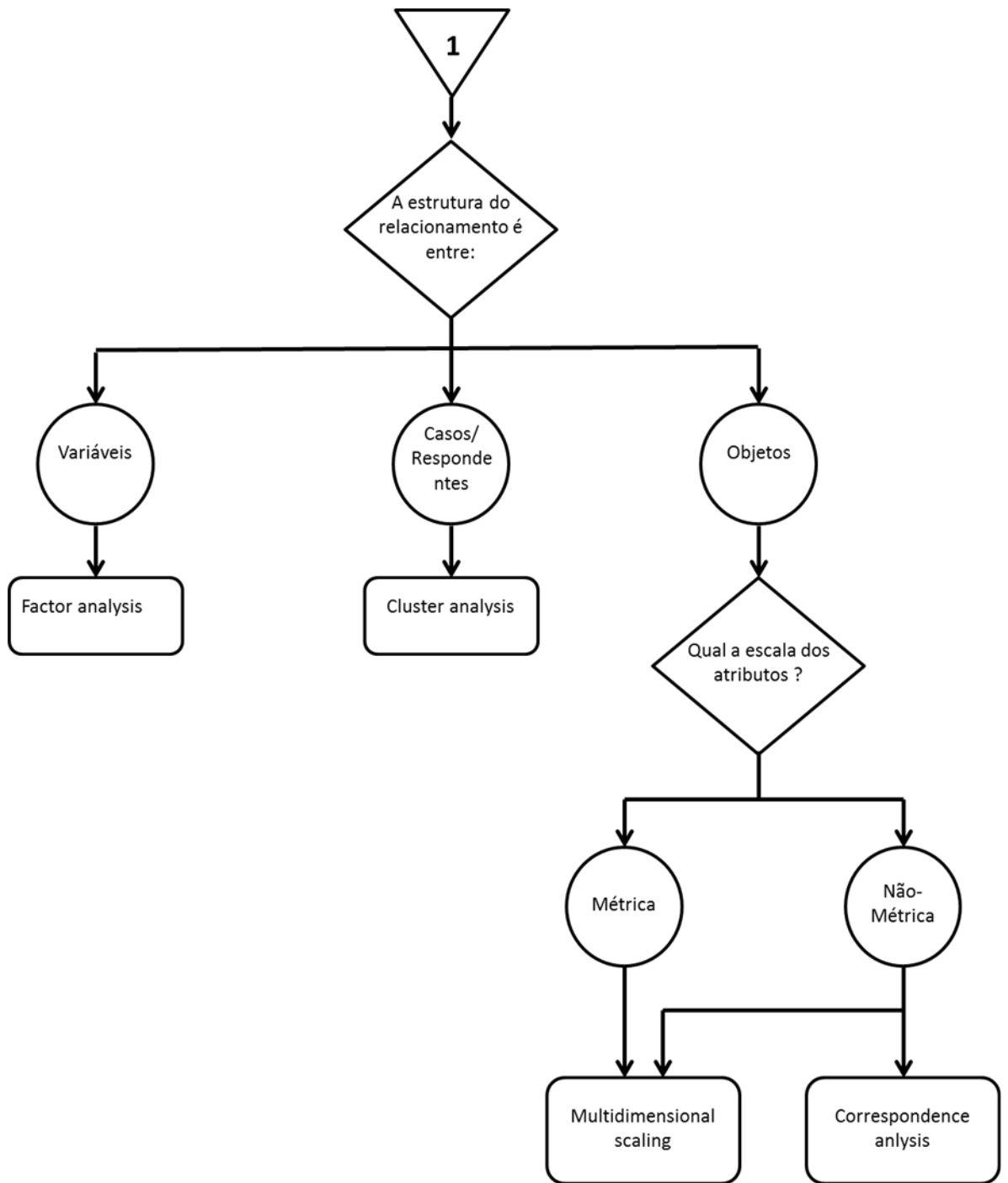


Figura 5 - Escolha de uma técnica multivariada

### **3.1- Problemas na mensuração da estrutura da preferência do consumidor**

Logo, como todo empreendimento tem como objetivo final a aceitação e a satisfação de um consumidor, as empresas para competirem com sucesso, devem entender e estar cada vez mais próximas dos clientes, ou seja, é de fundamental importância que a empresa conheça e compreenda o seu consumidor e o comportamento dele (BASTOS, 2010).

Em geral, ao se decidir sobre a preferência sobre um determinado produto, o consumidor não considera característica a característica deste produto, mas o conjunto de características simultâneas que o produto possui. É uma decisão muitas vezes não consciente e de ser difícil de ser manifestada pelo decisor (SIQUEIRA, 2000), conseqüentemente difícil de ser mensurada.

Porém a quantidade excessiva de características ou atributos é um dos pontos mais críticos neste tipo de estudo, pois quanto mais detalhes de um produto forem avaliados simultaneamente, um maior número de estímulos pode ser gerado, o que aumenta significativamente o tempo de resposta do questionário e a fadiga dos respondentes (PRETTO & ARTES, 2009). Ou seja, se o produto for composto por 10 atributos, cada um com 2 níveis cada, será possível formar um total de 1024 produtos, o que pode gerar resultados menos fidedignos.

Além de que, no entender de Miller (1956), a mente humana tem em média uma capacidade de processar conjuntamente cerca de sete variáveis.

Para contornar esse problema, *Green e Srinivasan*(1978) propõem o uso de delineamentos experimentais fatoriais fracionários que podem ser encontrados em Montgomery(1996).

Hair Jr.(2005) também menciona este artifício. “Quando o número de estímulos é muito grande para que o respondente possa avaliar, são criados planos sistemáticos (delineamento fatorial fracionário), que emprega um subconjunto dos possíveis estímulos”.

Porém há sempre a possibilidade de quando se realiza um planejamento fatorial fracionário, alguma combinação importante ficar de fora. Quando um atributo muito importante é omitido na descrição de um estímulo, a percepção de um indivíduo sobre o produto pode ser afetada, e conseqüentemente seu julgamento sobre a preferência dos demais também pode sofrer alterações.

Há situações, no entanto, em que mesmo esse artifício não consegue reduzir satisfatoriamente o número de produtos hipotéticos, sendo necessária a utilização de planejamentos de experimentos fracionários combinados com a omissão de um ou mais atributos na descrição do estímulo. Nesse contexto Pretto e Artes (2009) apresentam um estudo da influência de atributos faltantes quando estímulos incompletos (*partial profiles*) são apresentados.

Outra solução é apresentar os tratamentos a cada indivíduo e ele escolherá uma, nenhuma ou mais de uma alternativa. Esse método de análise é denominado *Choice Based Conjoint Analysis*(ANCFE), e pode ser visto com detalhes em (BASTOS, 2010).

### 3.2- Vantagens da *Conjoint Analysis*, nos problemas de Administração.

Sem dúvida o grande apelo e atrativo da *conjoint analysis* são suas aplicações e resultados para o meio empresarial. A flexibilidade da *conjoint analysis* viabiliza a sua implementação em praticamente toda a teoria vista até aqui. Entre as várias aplicações, podemos citar:

- Ações de **lançamentos de novos produtos**, com **testes de produtos conceitos**, em busca de uma combinação ótima de níveis de fatores.
- Um complemento para a decisão de projeto é a **análise de lucratividade**, se o custo de cada produto é conhecido, este pode ser combinado com a participação de mercado e o volume de vendas esperados para avaliar sua viabilidade. Esse processo pode gerar conclusões como apontar que um conjunto de atributos com participação menor pode ser a mais lucrativa, ou vice-versa e adjacentes. Um adjunto da análise de lucratividade, é a avaliação de **sensibilidade ao preço**.
- Mostrar as contribuições relativas de cada atributo e cada nível para avaliação geral do objeto e identificar oportunidades de *marketing*

explorando o potencial de mercado para combinações de características indisponíveis no momento, podendo fazer um estudo de concorrência, sabendo os pontos fortes e fracos do produto em relação a um concorrente em específico.

Tabela de importância de desempenho			
Importância do atributo	Nossa <i>performance</i>	Desempenho do concorrente	Resultado
<b>Alta</b>	Fraca	Fraca	Oportunidade negligenciada
		Boa	Desvantagem competitiva
	Boa	Fraca	Vantagem competitiva
		Boa	Competição acirrada
<b>Baixa</b>	Fraca	Fraca	Oportunidade nula
		Boa	Alarme falso
	Boa	Fraca	Vantagem falsa
		Boa	Competição Falsa

Tabela 1- Tabela de Importância de desempenho

Da tabela 1 percebe-se que o estudo de concorrência por atributos é de grande valia, um atributo importante onde a empresa tem uma alta *performance* e um concorrente tem um desempenho fraco, resulta em uma grande **vantagem competitiva**. O estudo de concorrência pode também mostrar que um atributo que está em desvantagem em relação ao desempenho do

concorrente, deve ser melhorado ou não devido a sua importância para o consumidor, ou ainda, pode mostrar que certo esforço em um determinado produto subjugado como importante na realidade para o consumidor não é, e então esses esforços podem ser direcionados para outras áreas.

- **Segmentação de mercado**, essa é sem dúvidas, um dos usos mais comuns dos resultados da *conjoint analysis* a nível individual, agrupando respondentes com utilidades parciais semelhantes é possível criar segmentos e mercados-alvo. É possível segmentar também em relação a outras covariáveis como variáveis socioeconômicas. E assim direcionar melhor as estratégias de propaganda, melhorar a maneira de como abordar esses clientes-alvo, entre outras inúmeras práticas.

### 3.3- Revisão Bibliográfica

Esta seção tem por objetivo fazer uma revisão histórica dos avanços da metodologia desde sua criação em meados de 1970, uma revisão mais aprofundada pode ser vista no ótimo *paper* da empresa de softwares para *Conjoint Analysis*, Sawtooth Software Inc. “A Short History of Conjoint Analysis “. (ORME, 2010)

O marketing, a partir da década de 1970, viu crescer a influência do desenvolvimento das ciências comportamentais, principalmente da psicometria nos métodos de pesquisa (GREEN, KRIEGER, & WIND, 2001). Neste contexto, três técnicas foram introduzidas na área de marketing, agregando mais importância na pesquisa de segmentação de mercado, posicionamento e

desenvolvimento de novos produtos: a Análise de Cluster (*Cluster Analysis*), o Escalonamento Multidimensional (*Multidimensional Scaling*) e a Análise Conjunta (*Conjoint Analysis*) (HENRIQUE & DE SOUZA, 2006).

Luce e Tukey(1964), um psicólogo matemático e um estatístico, lançaram a primeira aplicação na área econômica, denominada de *Conjoint Measurement*. Após a publicação deste trabalho houve uma grande evolução das técnicas, principalmente, com o estudo de pesquisadores da área de marketing, os quais estavam interessados em modelos e técnicas que enfatizassem a transformação de respostas subjetivas aos estímulos objetivos em parâmetros estimados.

As formas funcionais propostas pela teoria de *Conjoint Measurement* forneceram a base teórica para relacionar a utilidade do produto aos valores dos fatores do produto e, adotou-se o nome de *Conjoint Analysis* para ressaltar esta distinção (BASTOS, 2010).

Green e Rao(1971) propuseram a análise conjunta aplicada ao marketing, enfatizando a estimação de um modelo de preferência baseado no planejamento fatorial completo (*full-profile*).

Green e Wind (1975) publicaram um artigo na Harvard Business Review sobre a avaliação dos consumidores sobre materiais de limpeza de tapetes. Difundiu-se então o uso do método de estimação por mínimos quadrados ordinários (MQO).



No fim dos anos 70, Green e Srinivassan(1978), resumiram o uso da análise conjunta na indústria, informando novos desenvolvimentos, e dando conselhos sobre as melhores práticas desenvolvidas até aquele momento.

Os anos 80 foram responsáveis pelo crescimento de popularidade da *conjoint analysis* entre os pesquisadores, conhecedores de estatística e computação, além de uma notória popularização frente ao mercado. Isso graças à criação de alguns softwares para uso em computadores pessoais.

Por sua vez Green(1984) apresentou estudos de modelos com estruturas híbridas, que buscam minimizar o número de estímulos respondidos pelos participantes quando um grande número de atributos é avaliado simultaneamente. Estes modelos incorporam métodos composicionais e decomposicionais.

Alicerçada pelas recentes divulgações de estudos em análise conjuntas, em 1985, a empresa Sawtooth lançou um software, também para computadores pessoais, chamado *Adaptative Conjoint Analysis (ACA)*. Frente à dificuldade em que os respondentes tinham em dar respostas realistas perante o grande número de tabelas para a avaliação dos estímulos, a Sawtooth fez um software para administrar a pesquisa e a coleta dos dados. O software era capaz de se adaptar à pesquisa em tempo real capturando apenas as opiniões sobre os *trade offs* mais relevantes levando a respostas mais fidedignas.

‘

Ao mesmo tempo, o modelo de escolhas discretas já estava em processo de desenvolvimento. A base para sua empregabilidade foi o modelo logístico multinomial desenvolvido por Mcfadden(1970), que era uma adaptação e extensão do modelo logístico para respostas binárias. O conceito de escolha era atraente, os consumidores necessitavam apenas de escolher um entre vários tratamentos e fazer a sua escolha. O modelo parecia mais natural e realista para os respondentes, e ofereceu alguns benefícios e para as pesquisas de *Marketing*.

Green e Srinivasan(1978, 1990) descreveram as etapas fundamentais para a realização de um estudo de *conjoint analysis* enfatizando aspectos, teóricos e práticos, importantes a serem considerados neste tipo de pesquisa. No fim da década já havia softwares para análise conjunta baseada em escolhas(CBC – choise-based conjoint), novas variações de softwares para análise conjunta para perfil-completo e ACA. Outro fato importante que marcou a década de 90 foi a introdução de métodos Bayesianos.

A década de 90 representou uma década de crescimento forte para a *conjoint analysis* e sua aplicação em uma grande variedade de áreas. Até então a *conjoint analysis* tinha tradicionalmente sido aplicada a produtos de grande consumo, produtos de tecnologia e eletrônicos, bens duráveis (especialmente automóveis), e uma variedade de serviços baseados em produtos tais como telefones celulares, cartões de crédito e serviços bancários. Outras áreas interessantes de crescimento para análise conjunta surgiram. Incluindo concepção de *web sites*, avaliação de danos, recursos humanos e de pesquisa

empregada. Até o final da década, os analistas se tornaram tão confiantes da técnica que algumas pesquisas de análise conjunta eram usadas para fins pessoais, como, por exemplo, para ajudá-los a decidir entre opções de carros ou na busca por membros do sexo oposto.

Atualmente, os estudos buscam maneiras de deixar os experimentos em *CA* mais realistas, novas formas de apresentação de informações, pesquisas administradas por computador, entre outros. Há também uma tendência de novos desenvolvimentos em metodologias alternativas de *CA*.

De modo geral a literatura sobre *Conjoint Analysis* é muito extensa e contempla muitos aspectos de um estudo com a aplicação da metodologia. Textos voltados para aplicações procuram ilustrar o problema pesquisado, as alternativas de coleta de dados e a interpretação dos resultados, sendo que, quase sempre a análise estatística é realizada por softwares e os detalhes matemáticos não são mencionados. Rao(2008) apresenta um histórico dessa metodologia, exemplos de aplicação, perspectivas futuras e diversas referências.

Por outro lado, dentre os trabalhos nacionais, Siqueria(2000) apresenta uma importante revisão histórica, além dos detalhes dos métodos implementados, inclusive em diversos softwares, apresentando um exemplo real na área de marketing. Em Bastos(2010) pode se ver ótima revisão da metodologia, principalmente na *Conjoint analysis* baseada em escolhas.

Citamos também Castro(2006) que apresenta uma aplicação de *Conjoint Analysis* na indústria hoteleira, com objetivo de avaliar os pacotes de serviços preferidos pelos clientes de um hotel e Della Lucia(2010) que aplicou conjoint analysis nas duas modalidades, *CA* baseada em notas e *CA* baseada em escolhas, tanto para avaliar os fatores mais importantes quanto para avaliar intenção de compra, ambas para iogurte ligh sabor morango.

### **3.4- Condução de um estudo em CA (HAIR JR., 2005).**

Segundo Hair Jr.(2005) a *conjoint analysis* difere das outras técnicas multivariadas em três áreas distintas: (1) sua natureza decomposicional, (2) o fato de que as estimativas podem ser feitas a nível individual e (3) sua flexibilidade em termos de relações entre variáveis dependentes e independentes.

Na CA o pesquisador define as variáveis independentes (fatores) e seus valores (níveis), restando ao respondente apenas fornecer o valor da variável dependente. Esse aspecto mostra que a CA está muito próxima dos planejamentos de experimentos. Portanto o planejamento do experimento de CA é crítico para o seu sucesso.

Hair jr(2005), define a experimentação em CA em 6 estágios como veremos a seguir.

### 3.4.1- Objetivos da *Conjoint Analysis*

Como ocorre em qualquer análise estatística, o ponto de partida é a questão de pesquisa e o enfoque na análise de decisão do consumidor tem dois objetivos.

- 1) Determinar as contribuições de variáveis preditoras e seus níveis na determinação de preferência do consumidor.
- 2) Estabelecer um modelo válido de julgamentos do consumidor.

O respondente reage apenas ao que o pesquisador fornecer em termos de estímulos. Portanto o ideal é que a pesquisa seja contextualizada em torno de dois aspectos importantes, pois fornecerão uma orientação fundamental para as decisões-chaves dos próximos estágios.

- 1) É possível descrever todos os atributos que conferem utilidade ou valor ao produto ou serviço em estudo?
- 2) Quais são os critérios-chaves de decisão envolvida no processo de escolha para esse tipo de produto ou serviço?

#### ***3.4.1.1- Definição da utilidade total do objeto e especificação dos fatores determinantes***

Para representar o processo de julgamento do respondente com precisão, o pesquisador tem que incluir os principais atributos que potencialmente criam ou diminuem a utilidade geral do produto ou serviço, pois:

- 1) Concentrar-se em apenas fatores positivos irá distorcer seriamente os julgamentos dos respondentes.
- 2) Os respondentes podem empregar subconscientemente os fatores negativos, mesmo quando fornecidos, e assim acarretar viés inviabilizando o experimento.

Além disso, o pesquisador deve incluir todos os fatores determinantes. A meta é incluir os fatores que diferenciam melhor entre os objetos, ou seja, muitos atributos podem ser considerados importantes, mas também podem não diferenciar um produto do outro, pois não variam substancialmente entre as alternativas. Por exemplo, a segurança em automóveis é um atributo importante, mas não seria determinante na maioria dos casos porque todos os carros respeitam padrões mínimos regidos por leis. Já atributos como quilômetros por litro, desempenho ou preço são igualmente importantes e muito mais prováveis de serem utilizados como critério de comparação entre as escolhas.

### 3.4.2- O projeto de uma *Conjoint Analysis*

Nessa etapa o pesquisador volta suas atenções para as questões particulares envolvidas no delineamento e na execução do experimento de análise conjunta, como resolver questões como, qual das metodologias de *conjoint analysis* deve ser usada? Quais combinações de níveis e atributos usar? Qual o tamanho da amostra? Como medir a preferência e coletar os dados e qual o procedimento de estimação a ser empregado? Assim as questões de

delineamento talvez sejam uma das mais importantes na aplicação da *conjoint analysis*.

### 3.4.2.1- Seleção de uma metodologia de Conjoint Analysis.

Depois de determinar os atributos básicos que constituem a utilidade do produto ou serviço, uma questão fundamental deve ser resolvida: Qual das três metodologias conjuntas básicas, tradicional, baseada em escolha, ou adaptativas, deve ser usada. A escolha da metodologia gira em torno do número de atributos, o nível da análise e a forma de modelo permitida.

Características	Metodologia Conjunta		
	Tradicional	Adaptativa	Baseada em escolhas
Número máximo de atributos	9	30	6
Nível de análise	Individual	Individual	Agregada
Forma de modelo	Aditiva	Aditiva	Aditiva + Interação

Tabela 2- Uma comparação de metodologias conjuntas alternativas

Segundo Hair Jr(2005), a *conjoint analysis* tradicional é caracterizada por um modelo simples aditivo, com até nove fatores estimados sem pesar muito a experimentação. O *adaptive conjoint analysis* foi desenvolvido para acomodar um grande número de fatores, cerca de 30, que não seriam praticáveis na *conjoint analysis* tradicional. O método *Conjoint analysis* baseada em escolhas difere no ato da escolha já que o entrevistado escolhe um entre vários cenários

disponíveis, mas também difere no sentido que o modelo permite incluir interações, porém deve ser estimado no nível agregado.

#### ***3.4.2.2- Planejamento de estímulos***

Os fatores e níveis têm algumas características gerais: Devem ser operacionais, sendo comunicáveis, ou seja, conseguir passar ao respondente tudo aquilo que é necessário e o que foi o propósito do nível ou fator, e operacionais, ou seja, os fatores e níveis devem ter condições de ser colocados em prática, o que significa que os atributos devem ser distintos e representar um conceito que possa ser representado precisamente.

O número de fatores incluídos na análise afeta diretamente a eficiência estatística e a confiabilidade dos resultados, como já vimos muitos níveis e fatores geram uma grande combinação de estímulos, o que gera inconfiabilidade das respostas. Ao mesmo tempo os pesquisadores devem tentar equilibrar ou equalizar o número de níveis nos fatores, pois uma diferença significativa pode resultar em resultados distorcidos.

Uma das questões mais importantes relacionadas à definição dos fatores é a chamada **multicolinearidade fatorial**. A multicolinearidade entre fatores, ou a correlação entre eles, é um problema conceitual que geralmente envolve o conhecimento das características do que se está estudando, a colinearidade entre atributos resulta em combinações inacreditáveis de dois ou mais fatores. Por exemplo, imagina-se que potência e quilômetros por litro sejam atributos automotivos correlacionados negativamente, então um carro com os mais altos



níveis de potência e os mais baixos níveis de consumo não tem grandes chances de ser real.

Um atributo em específico merece maior atenção, o preço. O preço é um fator incluído em muitos experimentos de CA por representar uma característica muito influenciadora na tarefa de decisão de compra. Porém o preço, em sua maioria, tem um grau elevado de correlação inter-atributos com outros fatores. Para muitos produtos, um aumento na quantia do atributo é associado com um aumento de preço, e, portanto uma queda de preço nesse caso pode não ser realista. Além do que o preço pode interagir com fatores intangíveis, como nome da marca.

#### ***3.4.2.3- Coleta de dados***

O objetivo é transmitir ao respondente as combinações de estímulos da maneira mais realística e eficiente possível.

A escolha do método de apresentação pode variar pelo método de troca, perfil completo e comparação aos pares.

- **O método de apresentação de trocas**, ou método *trade-off*, compara atributos aos pares, classificando todas as combinações de níveis. Ele tem as vantagens de ser simples para o respondente e fácil de administrar. Porém o seu uso vem reduzindo drasticamente, pois ele é considerado pouco realístico ao usar só dois fatores por vez, necessita um grande número de julgamentos para um baixo numero de níveis, tendência de padrão de resposta por fadiga e só é viável com respostas não métricas.

- **O método de apresentação de perfil completo** é o método mais popular devido ao seu evidente realismo e de sua possibilidade de utilização de planejamentos fatoriais fracionários que reduzem o número de estímulos. Este método representa o estímulo/tratamento pela combinação de um nível de cada fator, este método permite um maior número de julgamentos como intenção de compra, chances de troca de produtos, probabilidade de experimentação entre outros. Porém apresenta algumas limitações. A alta quantidade de fatores tende a sobrecarregar as informações, e atrapalhar os julgamentos dos respondentes, e a ordem que os fatores são expostos pode causar algum tipo de viés.
- **O método de comparação aos pares** é uma combinação dos métodos anteriores, se assemelha ao método de apresentação de trocas, mas ao invés de apresentar fatores para comparação par a par ele apresenta tratamentos incompletos e a medida de resposta é o nível de intensidade. Este método é usado no software ACA.

#### ***3.4.2.4- Delineamento fatorial fracionário***

Quando todas as combinações de níveis e fatores são empregadas, temos um delineamento fatorial completo, mas na medida em que o número de fatores e níveis aumenta o experimento fica impraticável, e as respostas pouco fidedignas como visto anteriormente. Nesses casos pode ser empregado um método que desenvolva um subconjunto dos estímulos totais que podem ser avaliados e que possam ter informações necessárias para gerar estimativas de utilidades parciais fidedignas. Este método é o delineamento fatorial fracionário

e a seleção dos tratamentos é feita através de uma amostra. Delineamentos ótimos são aqueles que são ortogonais e balanceados.

A criação de delineamentos ótimos, não significa que todos os estímulos naquele delineamento serão aceitáveis para a avaliação, como a criação de estímulos inacreditáveis, como visto anteriormente no caso da multicolinearidade inter-atributos (em termos práticos as correlações entre atributos devem ser minimizadas, mas não necessariamente nulas. Correlações na ordem de 0,20 ou menos são perfeitamente aceitáveis), ou a excessividade dos estímulos composto por extremos de níveis. Assim restrições podem ser colocadas sobre as combinações de atributos. Em qualquer desses casos, os estímulos inacreditáveis devem ser eliminados do delineamento para garantir um processo de escolha mais realista. Neste caso serão realizados delineamentos quase ortogonais, e estes não violam nenhum dos pressupostos da *conjoint analysis*, e as estimativas serão paralelamente tão precisas quanto desde que as eliminações não ocorram em grande escala, deixando o experimento muito desbalanceado.

### 3.4.3- Suposições da Conjoint Analysis

A *conjoint analysis* tem o menor conjunto restritivo de suposições que envolvem a estimação do modelo conjunto. O delineamento experimental estruturado e a natureza generalizada do modelo tornam a maioria dos testes realizados em outros métodos de dependência desnecessários. Portanto testes de normalidade, homocedasticidade e independência não são necessários.

Ainda que existam menos suposições estatísticas, as suposições conceituais talvez sejam em maior número que em qualquer outra técnica multivariada.

#### **3.4.4- Estimação do modelo e avaliação da qualidade do ajuste**

As opções disponíveis ao pesquisador em termos de técnicas de estimação aumentaram nos últimos anos. A escolha da melhor técnica de estimação depende diretamente do método de aplicação da *conjoint analysis*.

##### ***3.4.4.1- Técnicas de estimação***

Na *conjoint analysis* tradicional, avaliada ou por ordenação (postos) ou por notas, exigem diferentes tipos de técnicas estatísticas para estimação. Quando a avaliação dos consumidores é dada por ordenação a estimação é dada por uma forma modificada de ANOVA, planejada para dados ordinais. Quando a avaliação dos consumidores é dada por notas, são utilizados os modelos de regressão múltipla para a estimação das utilidades parciais.

Já em relação aos métodos mais recentes, podemos citar a *conjoint analysis* baseada em escolhas, que é um dos enfoques deste trabalho. Seu método de estimação são os modelos de regressão logística multinomial.

De forma geral a maioria dos programas disponíveis no mercado hoje em dia pode acomodar qualquer forma de estimação.

#### ***3.4.4.2- Avaliação da Qualidade do ajuste do modelo***

Os resultados da *conjoint analysis* são avaliados quanto à precisão tanto em nível individual quanto agregado. O objetivo é determinar o quão consistentemente o modelo prevê as avaliações de preferência dadas pelos consumidores. Para os dados ordenados, sugerem-se testes de correlação nas classificações reais e previstas, como os de Spearman ou de Kendall. Para os dados métricos uma correlação de Pearson, juntamente como uma comparação entre os dados reais e os previstos é adequado. Em nível agregado, pode ser usada uma amostra de validação para avaliar a precisão preditiva.

Malhotra(2005) cita como procedimentos para avaliar a confiabilidade de uma *conjoint analysis* a aderência do modelo, como  $R^2$ , ou um teste-reteste, que nada mais é do que pedir ao entrevistado que avalie novamente alguns tratamentos, e assim correlacionar as suas notas e avaliar os coeficientes de correlação.

#### **3.4.5- Interpretação dos resultados**

##### ***3.4.5.1- Análise agregada X Análise individual***

A abordagem mais tradicional na *conjoint analysis* é a nível individual. Cada respondente é modelado separadamente e os resultados deste modelo são analisados separadamente.

O corriqueiro é estimar as utilidades parciais de cada fator, avaliando seu padrão e sua magnitude. Quanto maior a utilidade parcial maior o impacto daquele fator sobre a utilidade total. Posteriormente pode ser feita ou não uma

agregação dos indivíduos pelas utilidades parciais utilizando, por exemplo, análise de cluster.

A interpretação também pode ser a nível agregado, obviamente os resultados serão mais pobres no sentido de tentar prever o que um respondente faria quando ou na impotência de analisar as utilidades parciais de um respondente em especial, porém, muitas vezes a análise agregada prevê mais precisamente o comportamento agregado, como participação de mercado. Desse modo o importante é identificar o propósito primário do estudo e empregar o nível de análise apropriado.

#### ***3.4.5.2- Avaliação da importância relativa***

Como apresentado anteriormente, as utilidades parciais nos dão o impacto de cada nível na utilidade total, como evolução desta análise surge a importância relativa. Que nada mais é do que a conversão das utilidades parciais para uma escala em comum.

Na maioria das vezes as utilidades parciais são convertidas em percentuais que somam 100%. Isso viabiliza a comparação entre respondentes em uma escala comum bem como dá sentido ao escore gerado pela utilidade parcial.

#### **3.4.6- Validação dos resultados**

A validação externa em geral envolve a habilidade da *conjoint analysis* para prever escolhas reais e, em termos específicos, a questão da representatividade da amostra. Apesar de não haver avaliação de erro amostral

nos modelos de nível individual, o pesquisador sempre deve garantir que a amostra seja representativa da população de estudo. Isso se torna extremamente importante quando os resultados da *conjoint analysis* são usados para fins de segmentação de mercado e simulação de escolhas.

## 4- Metodologia

Neste capítulo apresentaremos duas alternativas de modelagem da preferência do consumidor, a *conjoint analysis* tradicional baseada em notas (*ratings-based conjoint analysis*) e postos (*ranking-based conjoint analysis*) e a *conjoint analysis* baseada em escolhas (*choise-based conjoint analysis(CBC)*).

### 4.1- *Conjoint Analysis* baseada em notas

#### 4.1.1- Modelo para a Utilidade

Como vimos, na *conjoint analysis* os produtos/tratamentos são formados pela combinação de níveis dos fatores. Então suponhamos um produto que possa ser definido pela combinação de  $r$  fatores, cada um com  $m_s$  níveis.

Assim é possível definir uma quantidade  $J$  de tratamentos possíveis, onde  $J = \prod_{s=1}^r m_s$ . Quando o experimento for balanceado isso se resume a  $J = (m_s)^r$ .

Para uniformizar o desenvolvimento das expressões utilizaremos a seguinte notação de índices:

- $j$ : índice que representa os tratamentos/produtos, com  $j = 1, 2, 3, \dots, J$ .
- $s$ : índice que representa os fatores/atributos, com  $s = 1, 2, 3, \dots, r$ .
- $i$ : índice que representa os níveis, com  $i = 1, 2, 3, \dots, m_s$ .
- $n$ : índice que representa os respondentes/consumidores, com  $n = 1, 2, 3, \dots, N$ .



Seja  $U_{nj}$ , a utilidade atribuída pelo n-ésimo consumidor ao j-ésimo tratamento. Assim o modelo para a utilidade pode ser descrito como:

$$U_{nj} = \beta_0 + \sum_{s=1}^r \sum_{i=1}^{ms} X_{si}^j \beta_{si} + \varepsilon_{nj} \quad (4.1)$$

Onde:

- $U_{nj}$ , representa a utilidade total que o n-ésimo respondente determina para o j-ésimo tratamento, em termos mais práticos  $U_{nj}$  é a nota dada ao tratamento j pelo indivíduo n.

Podem ser usadas notas centradas na média, neste caso para uma análise em nível individual  $U_{nj} = U_{nj} - \bar{U}_n$ , e para uma análise agregada  $U_{nj} = U_{nj} - \bar{U}_{..}$ . A vantagem de se usar as notas na média é que  $\beta_0=0$ , sem perda de generalidade.

- $X_{si}^j$ , é uma variável binária, que funciona como indicadora definindo quais níveis dos fatores estarão presentes nos tratamento j, assim:

$$X_{si}^j = \begin{cases} 1, & \text{se o } i - \text{ésimo nível do } s - \text{ésimo fator estiver presente no tratamento } j \\ 0, & \text{caso contrário.} \end{cases}$$

- $\beta_0$ , é o intercepto ou constante do modelo, e corresponde à nota média de todos os tratamentos.
- $\beta_{si}$ , é a utilidade parcial ou coeficiente de preferência (CP), associado ao i-ésimo nível do s-ésimo fator.

- $\varepsilon_{nj}$ , é o erro aleatório não observável do modelo. Usualmente se assume independência, entre e dentro de consumidores, e normalidade.

$$\varepsilon_{nj} \sim N(0, \sigma^2)$$

Define-se inicialmente  $\mathbf{X}_n$  como uma matriz formada pelos vetores de incidência  $[\mathbf{x}^j]^t$ , onde  $\mathbf{x}^j$  possui dimensão  $1 \times (1 + \sum_{s=1}^r ms)$  com  $j = 1, 2, 3, \dots, J$ , e é formado por valores 0 e 1 referentes a  $X_{si}^j$ , que define quais níveis dos fatores estarão presentes no  $j$ -ésimo tratamento. Assim  $\mathbf{X}_n$  é uma matriz com dimensões  $J \times (1 + \sum_{s=1}^r ms)$ .

$$\mathbf{X}_n = \begin{bmatrix} 1 & X_{11}^1 & X_{12}^1 & \cdots & X_{1m_1}^1 & \cdots & X_{s1}^1 & X_{s1}^1 & \cdots & X_{sm_s}^1 \\ 1 & X_{11}^2 & X_{11}^2 & \cdots & X_{1m_1}^2 & \cdots & X_{s1}^2 & X_{s2}^2 & \cdots & X_{sm_s}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{11}^J & X_{11}^J & \cdots & X_{m_1}^J & \cdots & X_{s1}^J & X_{s2}^J & \cdots & X_{sm_s}^J \end{bmatrix}$$

Constrói-se essa matriz para todos os indivíduos, ou seja, para  $n = 1, 2, 3, \dots, N$ .

Prosseguindo, é necessário definir o vetor  $\boldsymbol{\beta}$  de dimensão  $(1 + \sum_{s=1}^r ms) \times 1$ , que é composto pelos  $\beta_{si}$ , que representam as utilidades parciais associadas ao  $i$ -ésimo nível do  $s$ -ésimo fator:

$$\boldsymbol{\beta} = (\beta_0, \beta_{11}, \beta_{12}, \dots, \beta_{1m_1}, \dots, \beta_{2m_2}, \dots, \beta_{rms})^t$$

Considerando um estudo com  $n = 1, 2, 3, N$  consumidores, onde cada um avalia todos os  $j = 1, 2, 3, \dots, J$  tratamentos, ou,  $J^* < J$  tratamentos convenientemente selecionados (fatorial fracionado). Define-se o modelo de utilidade total matricial dado por:

$$\mathbf{U} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (4.2)$$

Onde o vetor de utilidades é dado por:

$$\mathbf{U} = [U_{11}, \dots, U_{1j}, \dots, U_{nj}, \dots, U_{NJ}]^t$$

Cujas componentes são as utilidades atribuídas pelos  $n$  respondentes,  $n = 1, 2, \dots, N$ , aos  $j$  tratamentos apresentados,  $j = 1, 2, \dots, J$ .

E a matriz  $\mathbf{X}$ , formada por várias matrizes  $\mathbf{X}_n$ , que se repetem para todos  $N$  indivíduos. Assim  $\mathbf{X}$  tem dimensões  $NJ \times (1 + \sum_{s=1}^r m_s)$ .

$$\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n, \dots, \mathbf{X}_N]^t$$

#### 4.1.2- Importâncias Relativas (IR) dos atributos.

O objetivo principal da *conjoint analysis* baseadas em notas de preferência é estimar o vetor  $\boldsymbol{\beta}$ . Uma das alternativas para alcançar estas estimativas é através do método de mínimos quadrados ordinários (MQO) com as restrições

$$\sum_{i=1}^{m_s} \beta_{si} = 0 \quad , \text{ para } s = 1, 2, 3, \dots, r$$

De maneira que o sistema de equações normais  $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{U}$  possa ter solução única e permita interpretações importantes para as estimativas dos  $\beta_{si}$ .

Quando trabalhamos com softwares, essas restrições são automaticamente realizadas pela maneira da estruturação do experimento.

Na forma matricial estas restrições irão completar posto da matriz  $\mathbf{X}$ . Sendo expressas como:

$$\mathbf{B}'\boldsymbol{\beta} = \mathbf{0}$$

Onde  $\mathbf{B}$  tem dimensão  $(1 + \sum_{s=1}^r m_s) \times r$ , e  $\mathbf{0}$  uma matriz com os componentes todos iguais a 0.

De forma que o sistema de equações normais ficará:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{B} \\ \mathbf{B}' & \mathbf{0} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\mathbf{i}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{U} \\ \mathbf{0} \end{bmatrix} \quad (4.3)$$

O vetor  $\widehat{\mathbf{i}}$  é adicionado ao vetor  $\widehat{\boldsymbol{\beta}}$  para completar posto.

Assim o estimador dos coeficientes do modelo de regressão e consequentemente o estimador das utilidades parciais segue a expressão:

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\mathbf{i}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{B} \\ \mathbf{B}' & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}'\mathbf{U} \\ \mathbf{0} \end{bmatrix} \quad (4.4)$$

Estimados as utilidades parciais ou coeficientes de preferência (cp) podemos estimar as importâncias relativas (IR) dos fatores. Portanto se:

$\beta_{si} < 0$  indica efeito desfavorável, ou seja, que diminuem a nota de preferência pelo produto, enquanto;

$\beta_{si} > 0$  indica efeito favorável na preferência do consumidor, ou seja, aumentam a nota de preferência do consumidor pelo produto.

Os valores  $\beta_{si}$ , para  $i = 1, 2, \dots, m_s$ , fornecem a estimativa da importância do s-ésimo fator dada por:

$$\widehat{I}_s = \max(\widehat{\beta}_{si}) - \min(\widehat{\beta}_{si}) \quad (4.5)$$

Logo, a importância relativa (IR) de cada fator é estimada como:

$$\widehat{IR}_s(\%) = \frac{\widehat{I}_s}{\sum_{s=1}^r \widehat{I}_s} \times 100 \quad (4.6)$$

Que pode ser interpretado como o impacto atribuído na preferência do consumidor a um j-ésimo produto.

#### 4.1.3- Um exemplo de *conjoint analysis* baseada em notas

Um exemplo hipotético, imaginemos que uma empresa que fabrica telefones celulares esteja estudando em lançar um novo aparelho de *smartphone* no mercado, e frente aos modelos concorrentes, a empresa quer se decidir em quais fatores apostar em seu novo aparelho. Três fatores são considerados chave para o sucesso deste novo aparelho, o teclado, o sistema operacional e a cor. Os gerentes de marketing têm dúvidas se aplicam teclado *touch screen* ou *qwerty*, se utilizam como sistema operacional Android ou Windows, e se utilizam das cores branca ou preta.

Teclado		Sistema Operacional		Cor	
<i>Touch</i>	<i>Qwerty</i>	Android	Windows	Branco	Preto
1	0	1	0	1	0

Tabela 3- níveis dos fatores do exemplo celular

Foi decidido fazer uma análise individual, pois assim poderia ser feita uma possível segmentação, avaliar se caberia a fabricação de mais de um tipo de produto, em qual focar nas ações publicitárias entre outras funções.

Também foi decidido fazer um planejamento fatorial completo, onde temos 8 tratamentos/produtos diferentes e possíveis sendo formados por esses atributos.

Como se trata de um exemplo hipotético, vamos analisar apenas o resultado de um consumidor.

Teclado	Sistema operacional	Cor	Nota
1	1	1	9,0
1	1	0	8,0
1	0	1	7,5
1	0	0	6,5
0	1	1	7,0
0	1	0	6,0
0	0	1	5,0
0	0	0	4,0

Tabela 4-notas do consumidor 1 para os 8 estímulos de celular

Usando o SPSS, podemos estimar os coeficientes, as utilidades parciais e as importâncias dos fatores.

Foi feito um planejamento ortogonal a partir do comando:

*Data > orthogonal design > generate*

Este comando permite a criação de um experimento, onde de maneira simples o software inclui os fatores e seus níveis, permite realizar experimentos fatoriais completos e fatoriais fracionários de maneira a gerar experimentos ortogonais, ou quase ortogonais. Permite também gerar casos para validação e simulação.

O ônus de se trabalhar com o SPSS para *conjoint analysis* é que o modulo CONJOINT SPSS, só está disponível por meio de sintaxe, e este só realiza análises na abordagem tradicional.

Além do que, a difusão destes passos é pouco comentada, e, portanto gera algumas dificuldades para leigos e iniciantes na prática desta metodologia, forçando o pesquisador que não conhece os comandos a adaptar os comandos existentes para realizar as análises, sendo que os resultados provenientes do módulo CONJOINT são extremamente simples e completos.

O procedimento Conjoint do SPSS requer dois arquivos (dois bancos de dados), o arquivo de dados, onde é especificada a nota ou posto que cada consumidor atribui a cada tratamento/produto, e o arquivo de planejamento, consiste no conjunto de perfis de produto a ser avaliado pelos sujeitos e devem ser gerados usando o procedimento de projeto ortogonal que vimos mais acima.

Tendo os dois bancos de dados preenchidos corretamente, podemos “rodar” o comando na sintaxe do SPSS.

A função Utilidade para a análise desse exemplo é:

$$U_{1j} = \beta_0 + (\beta_{11}X_{11}^j + \beta_{12}X_{12}^j) + (\beta_{21}X_{21}^j + \beta_{22}X_{22}^j) + (\beta_{31}X_{31}^j + \beta_{32}X_{32}^j) + \varepsilon_{1j}$$

A sintaxe do exemplo segue abaixo.

```

CONJOINT PLAN='C:\Users\Thiago\Documents\monografia\banco de
dados\celular\celular.plano.sav'
/ATA='C:\Users\Thiago\Documents\monografia\banco de
dados\celular\celular.data.sav'
/SCORES=pref01 to pref08
/SUBJECT=ID
/FACTORS=
teclado (DISCRETE)
sistema (DISCRETE)
cor (DISCRETE)
/PRINT=SUMMARYONLY
/PLOT=SUMMARY.

```

#### 4.1.3.1 – Interpretar os resultados

Os resultados da saída do SPSS seguem a tabela abaixo.

Fatores	Níveis	Coefficientes de preferencia	Importâncias relativas (%)
Teclado	<i>Touch</i>	1,125	45
	<i>Qwerty</i>	-1,125	
Sistema Operacional	Android	0,875	35
	Windows	-0,875	
Cor	Branco	0,5	20
	Preto	-0,5	

Tabela 5- Resumo dos resultados da Conjoint Analysis baseada em notas

O modelo ajustado para a função utilidade deste consumidor é:



$$\widehat{U}_{1j} = 6,625 + (1,125X_{11}^j + (-1,125X_{12}^j)) + (0,875X_{21}^j + (-0,875X_{22}^j)) \\ + (0,5X_{31}^j + (-0,5X_{32}^j))$$

Podemos conferir que as estimativas das importâncias relativas realmente estão corretas. Para calcularmos a Importância relativa temos que calcular as importâncias.

$$\widehat{I}_s = \text{máx}(\widehat{\beta}_{si}) - \text{mín}(\widehat{\beta}_{si})$$

Neste caso como só temos dois níveis para cada fator as importâncias do teclado, sistema operacional e cor são respectivamente:

$$\widehat{I}_1 = 1,125 - (-1,125) = 2,250$$

$$\widehat{I}_2 = 0,875 - (-0,875) = 1,750$$

$$\widehat{I}_3 = 0,5 - (-0,5) = 1$$

Assim a importância relativa de cada fator é dada por  $\widehat{IR}_s(\%) =$

$$\frac{\widehat{I}_s}{\sum_{s=1}^r \widehat{I}_s} \times 100$$

$$\widehat{IR}_1(\%) = \frac{2,250}{5} \times 100 = 45\%$$

$$\widehat{IR}_2(\%) = \frac{1,750}{5} \times 100 = 35\%$$

$$\widehat{IR}_3(\%) = \frac{1}{5} \times 100 = 20\%$$

Para a análise dos resultados pode ser conveniente fazer o gráfico dos coeficientes de preferência.

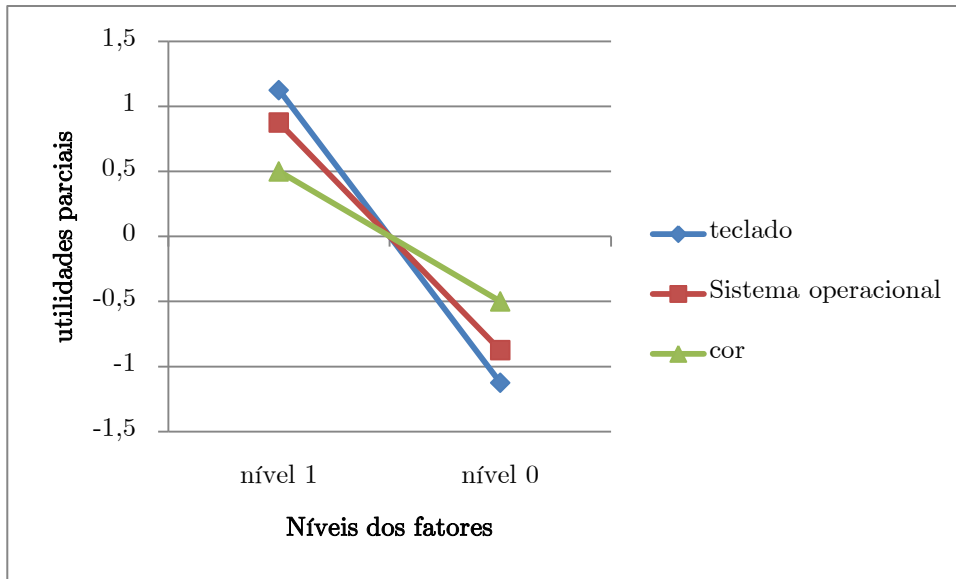


Figura 6- Utilidades Parciais dos níveis dos fatores para o exemplo do celular

Como se pode ver na tabela 3 e nos gráficos da figura 6, o fator que apresentou maior impacto na intenção de compra deste entrevistado foi o teclado com 45% de importância relativa, seguido pelo sistema operacional com 35% de importância relativa e a cor foi o fator que gerou menor importância relativa de 20%.

Este entrevistado acusa maior preferência por teclado *touch*, sistema operacional android e cor branca cuja configuração é estimada somando os coeficientes de preferências de seus níveis.

$$\widehat{U}_{11} = 6,625 + (1,125 \times 1 + (-1,125 \times 0)) + (0,875 \times 1 + (-0,875 \times 0)) + (0,5 \times 1 + (-0,5 \times 0))$$

$$\widehat{U}_{11} = 6,625 + 1,125 + 0,875 + 0,5 = 9,125$$

E a menor preferência foi pelo oitavo tratamento, cuja configuração é teclado *qwerty*, sistema operacional Windows e cor preta, e sua estimativa é dada por:

$$\widehat{U}_{18} = 6,625 - 1,125 - 0,875 - 0,5 = 4,125$$

Em relação à qualidade de ajuste do modelo, o teste de correlação de Pearson entre as notas do consumidor e as notas estimadas apontou uma correlação altíssima de 99,7%, e o coeficiente de aderência  $R^2$  ajustado foi de 98,8%.

Vemos que este consumidor prefere as últimas evoluções do mercado, então poderíamos classificá-lo como um consumidor sensível a novos produtos, novas tecnologias, um consumidor mais inovador.

Em estudos reais as estimativas dos coeficientes de preferências podem ser utilizadas para estabelecer segmentos de consumidores com padrões semelhantes de preferência e não somente por variáveis socioeconômicas.

## 4.2- *Conjoint Analysis* baseada em postos(rank)

### 4.2.1- Modelo para utilidade

Como no modelo baseado em notas podemos descrever  $J = \prod_{s=1}^r m_s$  tratamentos/produtos distintos pela combinação de  $s$  fatores, cada um com  $m_s$  níveis. Onde:

- $j$ : índice que representa os tratamentos/produtos, com  $j = 1, 2, 3, \dots, J$ .
- $s$ : índice que representa os fatores/atributos, com  $s = 1, 2, 3, \dots, r$ .
- $i$ : índice que representa os níveis, com  $i = 1, 2, 3, \dots, m_s$ .
- $n$ : índice que representa os respondentes/consumidores, com  $n = 1, 2, 3, \dots, N$ .

O modelo para utilidade segue o mesmo da *conjoint analysis* baseada em notas, a diferença se dará na estimação dos coeficientes de preferência ou utilidades parciais.

$$U_{nj} = \beta_0 + \sum_{s=1}^r \sum_{i=1}^{ms} X_{si}^j \beta_{si} + \varepsilon_{nj} \quad (4.1)$$

Onde:

- $U_{nj}$ , representa a utilidade total que o n-ésimo respondente determina para o j-ésimo tratamento, em termos mais práticos  $U_{nj}$  é o posto dado ao tratamento j pelo indivíduo n.
- $X_{si}^j$ , é uma variável binária, que funciona como indicadora definindo quais níveis dos fatores estarão presentes nos tratamento j, assim:

$$X_{si}^j = \begin{cases} 1, & \text{se o } i - \text{ésimo nível do } s - \text{ésimo fator estiver presente no tratamento } j \\ 0, & \text{caso contrário.} \end{cases}$$

- $\beta_0$ , é o intercepto ou constante do modelo, e corresponde a ordenação média de todos os tratamentos.
- $\beta_{si}$ , é a utilidade parcial ou coeficiente de preferência (CP), associado ao i-ésimo nível do s-ésimo fator.
- $\varepsilon_{nj}$ , é o erro aleatório não observável do modelo. Usualmente se assume independência, entre e dentro de consumidores, e normalidade.

$$\varepsilon_{nj} \sim N(0, \sigma^2)$$

#### 4.2.2- Cálculo das utilidades parciais ou coeficientes de preferência.

Hair jr(2005) cita alguns passos para obtenção das estimativas das utilidades parciais ou coeficientes de preferência para modelo baseado em postos ou *ranks*.

- Calcular a ordenação média do experimento

$$\bar{U}_{..} = \sum_{j=1}^J \frac{U_{.j}}{J}$$

São as somas das ordenações divididas pela quantidade de estímulos presentes.

- Calcular a ordenação média dos níveis

$$\bar{u}_{nsi} = \sum_{j=1}^J \frac{u_{nj} X_{si}^j}{J}$$

É calculado da mesma maneira que a ordenação média do experimento a diferença é que só serão somadas as ordenações que o i-ésimo nível estiver presente no j-ésimo tratamento. Onde:

$$X_{si}^j = \begin{cases} 1, & \text{se o } i - \text{ésimo nível do } s - \text{ésimo fator estiver presente no tratamento } j \\ 0, & \text{caso contrário.} \end{cases}$$

- Calcular o desvio de cada ordenação média de nível em relação a ordenação média do experimento

$$d_{nsi} = \bar{u}_{nsi} - \bar{U}_{..}$$

Quando a medida de preferência é inversamente relacionada à preferência, ou seja, usar números menores para indicar postos mais elevados e um estímulo mais preferido, invertemos os sinais dos desvios de modo que desvios positivos serão associados às utilidades parciais que indicam maior preferência.

- Calcular um valor de padronização
  - Elevar os desvios ao quadrado e somá-los por todos os níveis

- o Valor de padronização é dado pelo número total de níveis dividido pela soma dos desvios ao quadrado

$$\text{valor de padronização}(vp) = \frac{m_s}{\sum_{s=1}^r \sum_{i=1}^{m_s} (d_{nsi})^2}$$

- Calcular os desvios padronizados multiplicando cada desvio quadrado pelo valor de padronização

$$D_{nsi} = \frac{(d_{nsi})^2 m_s}{\sum_{s=1}^r \sum_{i=1}^{m_s} (d_{nsi})^2} = vp \times (d_{nsi})^2$$

- As estimativas das utilidades parciais ou coeficientes de preferência são obtidos através da raiz quadrada dos desvios padronizados.

$$\widehat{\beta}_{si} = \sqrt{D_{nsi}}$$

#### 4.2.3- Importâncias relativas.

Estimados as utilidades parciais ou coeficientes de preferência (cp) podemos estimar as importâncias relativas (IR) dos fatores. A estimação ocorre da mesma maneira que na *conjoint analysis* baseada em notas. Portanto se:

$\beta_{si} < 0$  indica efeito desfavorável, ou seja, que diminuem a nota de preferência pelo produto, enquanto;

$\beta_{si} > 0$  indica efeito favorável na preferência do consumidor, ou seja, aumentam a nota de preferência do consumidor pelo produto.

Os valores  $\beta_{si}$ , para  $i = 1, 2, \dots, m_s$ , fornecem a estimativa da importância do s-ésimo fator dada por:

$$\widehat{I}_s = \max(\widehat{\beta}_{si}) - \min(\widehat{\beta}_{si}) \quad (4.5)$$

Logo, a importância relativa (IR) de cada fator é estimada como:

$$\widehat{IR}_s(\%) = \frac{\hat{I}_s}{\sum_{s=1}^r \hat{I}_s} \times 100 \quad (4.6)$$

Que pode ser interpretado como o impacto atribuído na preferência do consumidor a um j-ésimo produto.

#### 4.2.4- Um exemplo de *conjoint analysis* baseada em postos.

Para ilustrar uma *conjoint analysis* baseada em postos ou ordenação, utilizaremos um exemplo que pode ser encontrado em HAIR. JR(2005).

Suponhamos que uma empresa chamada HATCO esteja tentando desenvolver um novo detergente. Após conversas com grupos de interesse decide-se que três atributos são importantes: Ingredientes, Forma e Nome da marca com dois níveis cada.

Forma		Ingredientes		Marca	
Líquido	Pó	Sem fosfato	Com fosfato	HATCO	Genérico
1	0	1	0	1	0

Tabela 6- Níveis dos fatores para o detergente

Será utilizado um planejamento fatorial completo. Com um total de 8 estímulos para a avaliação, onde o entrevistado vai ordená-los de acordo com sua preferência de compra ( 1=preferido e 8= última opção).

A descrição dos estímulos e a ordenação do consumidor 1 para os mesmo segue a tabela 6.

Descrição dos estímulos				
Estímulo	Forma	Ingredientes	Marca	Ordenação
1	Líquido	Sem Fosfato	HATCO	1
2	Líquido	Sem Fosfato	Genérico	2
3	Líquido	Com Fosfato	HATCO	5
4	Líquido	Com Fosfato	Genérico	6
5	Pó	Sem Fosfato	HATCO	3
6	Pó	Sem Fosfato	Genérico	4
7	Pó	Com Fosfato	HATCO	7
8	Pó	Com Fosfato	Genérico	8

Tabela 7- Descrição dos estímulos de detergente e ordenação do consumidor 1

Examinado as respostas do consumidor 1, percebemos uma ordenação mais alta para os detergentes compostos por ingredientes sem fosfato (1,2,3,4) enquanto detergentes compostos por ingredientes com fosfato tem postos mais altos (5,6,7,8).

Assumindo que um modelo básico se aplique (modelo aditivo) podemos calcular o impacto de cada nível como desvios da ordenação média geral (HAIR JR., 2005).

Vamos visualizar como encontrar a utilidade parcial para apenas um nível por questão de exemplificação.

- Calcular a ordenação média do experimento

$$\bar{U}_{..} = \sum_{j=1}^J \frac{U_j}{J}$$

A ordenação média é dada pela soma de todos os ordenamentos divididos pela quantidade de estímulos. A ordenação média é de 4,5.

$$\bar{U}_{..} = [(1 + 2 + 3 + 4 + 5 + 6 + 7 + 8)/8] = 4,5$$

- Calcular a ordenação média dos níveis

$$\bar{u}_{nsi} = \sum_{j=1}^J \frac{u_{nj} X_{si}^j}{J}$$



As ordenações médias para os dois níveis de ingredientes dos detergentes para o respondente 1 são.

Sem fosfato:  $\bar{u}_{121} = (1 + 2 + 3 + 4)/4 = 2,5$

Com fosfato:  $\bar{u}_{122} = (5 + 6 + 7 + 8)/4 = 4,5$

- Calcular o desvio de cada ordenação média de nível em relação a ordenação média do experimento. Para o nível sem fosfato temos que:

$$d_{nsi} = u_{nsi} - \bar{U}_{..}$$

$$d_{121} = 2,5 - 4,5 = -2$$

Como utilizamos postos inferiores para representar preferencias elevadas, temos que inverter o sinal do desvio, portanto  $d_{121} = 2$ .

- Calcular um valor de padronização. Para o exemplo temos um total de 6 níveis onde a soma dos desvios ao quadrado nos dá um valor de 10,5.

$$\text{valor de padronização} = \frac{m_s}{\sum_{s=1}^r \sum_{i=1}^{m_s} (d_{nsi}^2)} = \frac{6}{10,5} = 0,571$$

- Calcular os desvios padronizados multiplicando cada desvio quadrado pelo valor de padronização

$$D_{nsi} = \frac{(d_{nsi})^2 m_s}{\sum_{s=1}^r \sum_{i=1}^{m_s} (d_{nsi}^2)} = vp \times (d_{nsi})^2$$

$$D_{121} = 0,571 \times 2^2 = 2,284$$

- As estimativas das utilidades parciais ou coeficientes de preferência são obtidos através da raiz quadrada dos desvios padronizados.

$$\hat{\beta}_{si} = \sqrt{D_{nsi}}$$

$$\hat{\beta}_{21} = \sqrt{2,284} = 1,511$$

A estimação das importâncias relativas ocorre da mesma maneira que na *conjoint analysis* baseada em notas.

Abaixo, nas tabelas 7, 8 e 9, podemos ver todos os passos para a obtenção das utilidades parciais através da *conjoint analysis* baseada em postos

e o cálculo das importâncias relativas para o exemplo do detergente. Observamos que como previsto, o atributo que detém a maior importância são os ingredientes com 57,1% de importância, onde o ingrediente sem fosfato representa a preferência positiva, seguido pela forma com 28,6% de importância onde a forma líquida apresenta preferência positiva. Com a menor importância o atributo marca com 14,3%, porém o entrevistado ainda prefere a marca HATCO a uma marca genérica.

<b>Utilidades parciais estimadas</b>							
<b>Nível do fator</b>	<b>Ordenações de estímulos</b>	<b>Ordenação média de nível</b>	<b>Desvio</b>	<b>Desvio invertido</b>	<b>Desvio quadrado</b>	<b>Desvio padronizado</b>	<b>Utilidade parcial estimada</b>
<b>Forma</b>							
<b>Líquido</b>	1,2,5,6	3,5	-1	1	1	0,576	0,756
<b>Pó</b>	3,4,7,8	5,5	1	-1	-1	-0,576	-0,756
<b>Ingredientes</b>							
<b>Sem fosfato</b>	1,2,3,4,	2,5	-2	2	4	2,284	1,511
<b>Com fosfato</b>	5,6,7,8	6,5	2	-2	-4	-2,284	-1,511
<b>Marca</b>							
<b>HATCO</b>	1,3,5,7	4	-0,5	0,5	0,25	0,143	0,378
<b>Genérico</b>	2,4,6,8	5	0,5	-0,5	-0,25	-0,143	-0,378
<b>Ordenação média total</b>		4,5					
<b>Soma de desvios quadrado</b>					10,5		
<b>Valor de padronização</b>					0,576		

Tabela 8- Utilidades parciais estimadas para o respondente 1.

Como o objetivo do estudo é obter um produto ideal, se este entrevistado representasse a população alvo o produto ideal seria: detergente com forma líquida, com ingredientes sem fosfato e marca HATCO.

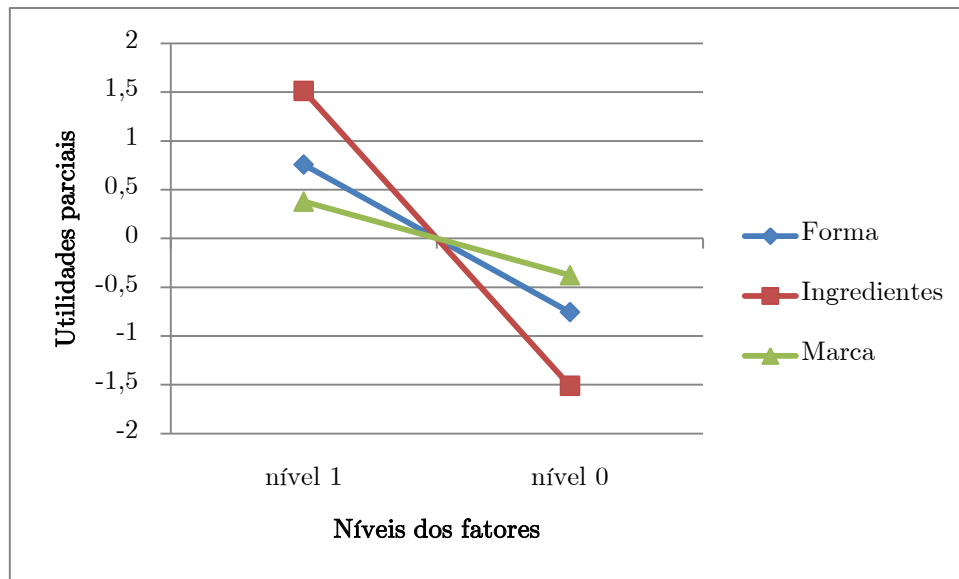


Figura 7- Utilidades parciais dos níveis dos fatores do detergente

Importâncias relativas dos fatores		
Fatores	Amplitude de utilidades parciais	Importância do fator
<b>Forma</b>	1,512	28,60%
<b>Ingredientes</b>	3,022	57,10%
<b>Marca</b>	0,756	14,30%

Tabela 9- Importâncias relativas dos fatores

Usando o SPSS, podemos estimar os coeficientes, as utilidades parciais e as importâncias dos fatores de maneira igual ao modelo baseado em notas. Melhores informações podem ser vistas no Apêndice 1.

#### 4.3- *Conjoint Analysis* baseada em escolhas.

A natureza básica da *conjoint analysis* baseada em escolhas tem conduzido outra perspectiva de avaliação. Talvez a maior vantagem da *conjoint analysis* baseada em escolhas seja o processo realista de escolha, ao fazer o indivíduo confrontar uma escolha a outra.

Neste processo de escolha, decisões de compra são tomadas não só com base em fatores racionais, mas também são influenciadas por fatores subjetivos que o consumidor não consegue verbalizar diretamente. Esta subjetividade estaria relacionada aos valores associados nas características dos fatores que compõem um produto, como, por exemplo, sua Marca. Um dos principais objetivos dos modelos de escolha discreta é justamente medir a preferência do consumidor sobre produtos e serviços competitivos, expressando matematicamente a importância dos diferentes fatores de escolha (BASTOS, 2010).

O resultado sempre nos trará respostas como, “qual a probabilidade da marca X ser escolhida?”, “Qual cor será escolhida?”, “Qual a probabilidade do produto com preço mais alto ser o escolhido?”, etc.

Nos modelos probabilísticos as utilidades das alternativas são variáveis aleatórias, logo o modelo é utilizado para estimar a probabilidade de uma alternativa ter a máxima utilidade e, por conseguinte, a probabilidade com a qual será escolhida (LOUVIERE & GAETH, 1988).

Dentre os modelos probabilísticos de escolha discreta podemos citar os modelos baseados na maximização da utilidade (MEYER & KAHN, 1991), estes são usados *conjoint analysis* baseada em escolhas.

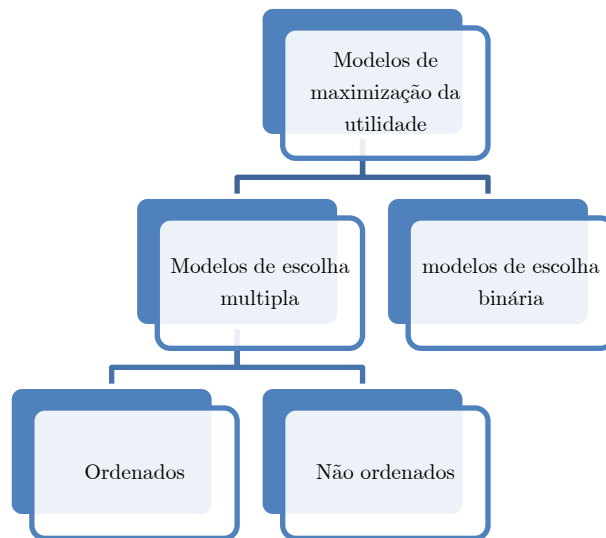


Figura 8 - Classificação dos modelos probabilísticos de escolha discreta

Os modelos probabilísticos de escolha discreta baseados na maximização da utilidade são classificados em duas categorias: Modelos de Escolha binária (duas alternativas para escolha) e Modelos de Escolha múltipla (mais de duas alternativas). Adicionalmente, os modelos de escolha múltipla abrangem modelos não ordenados no qual o entrevistado efetua uma única escolha do conjunto de alternativas e ordenados no qual o entrevistado efetua uma ordenação total ou parcial do conjunto de escolha.

Um aspecto particular que ainda não está resolvido é a propriedade de IIA (Independência de alternativas irrelevantes), uma suposição que torna problemática a previsão de alternativas muito semelhantes (HAIR JR., 2005).

#### 4.3.1- Modelo para a Utilidade

Consideremos novamente a possibilidade de descrever um tratamento/produto pela combinação de  $s$  fatores, cada um com  $m_s$  níveis, de modo que tenhamos  $J = \prod_{s=1}^r m_s$  tratamentos/produtos distintos.

Definimos a função Utilidade como:

$$U_{nj} = \mathbf{X}_j' \boldsymbol{\beta} + \varepsilon_{nj} \quad (4.7)$$

Onde:

- $U_{nj}$ , é a utilidade atribuída pelo  $n$ -ésimo consumidor ao  $j$ -ésimo tratamento. Neste caso se o respondente escolher o tratamento  $j$  a resposta será 1, caso contrário 0.
- $\mathbf{X}_j' = [x_1^j, x_2^j, \dots, x_r^j]$ , é o vetor cujas entradas são os códigos dos níveis dos  $r$  fatores presentes no tratamento  $j$ , ou seja,  $x_s^j$  representa o nível do fator  $s$  presente no tratamento  $j$ .
- $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_s]^t$ , é o vetor de parâmetros a serem estimados, sendo que na *conjoint analysis* baseada em escolha apenas um parâmetro é estimado por fator.
- $\varepsilon_{nj}$ , é o erro aleatório associado à mensuração da escolha do tratamento  $j$  pelo indivíduo  $n$ .

No modelo Logit, o erro se comporta segundo uma distribuição Gumbel (BEN-AKIVA & LERMAN, 1985).

$$F(\varepsilon_{ij}) = e^{-e^{-\varepsilon_{ij}}}$$

Considerando  $n=1, 2, \dots, N$  indivíduos podemos expandir o vetor  $\mathbf{X}_j'$ , para a matriz  $\mathbf{X}$  formada pelos vetores de incidência, assim  $\mathbf{X}$  é dado por:

$$\mathbf{X} = [\mathbf{X}_1', \mathbf{X}_2', \dots, \mathbf{X}_J']^t$$

Na *conjoint analysis* baseada em notas a estimação dos parâmetros é feita através do Modelo Logit Multinomial.

#### 4.3.2- O Modelo Logit Multinomial

Dado o modelo de regressão logística tradicional, ou seja, quando a variável resposta é dicotômica ou binária, o modelo de regressão logística multinomial é uma generalização do modelo tradicional para a situação onde a variável resposta é nominal com mais de duas alternativas.

Essa adaptação do modelo logístico foi proposta por McFadden(1974) na qual ele denominou como modelo de escolha discreta.

Existem diversas maneiras de se obter o modelo logístico multinomial Hosmer & Lemeshow(1989) utilizam da premissa de utilizar a maneira mais simples, ou seja, 3 categorias, em função de facilitar a notação.

Suponhamos 3 categorias possíveis para variável resposta 0, 1 e 2. Importante frisar que neste desenvolvimento o zero representa uma categoria, e não uma não resposta ou algo do tipo.

No modelo logístico binário a variável resposta é parametrizada em termos de sucesso e fracasso ou,  $Y = 1$  *versus*  $Y = 0$ . Para desenvolver o modelo com três categorias nós precisaremos de duas funções logísticas, tornando uma das alternativas como base na confrontação. Neste caso usaremos  $Y = 0$  como categoria referência comparando-a com  $Y = 1$  e  $Y = 2$ .

,

Para desenvolver o modelo assumimos  $x_k$  covariáveis, com  $k=1, 2, \dots, p$ , e um termo constante. Define-se o vetor  $\mathbf{x}$  com dimensão  $p+1$ . Denotemos as funções a seguir.

$$\begin{aligned}g1(\mathbf{x}) &= \ln \left[ \frac{P(Y = 1|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] \\g1(\mathbf{x}) &= \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \dots + \beta_{1p}x_p \\g1(\mathbf{x}) &= \mathbf{x}'\boldsymbol{\beta}_1\end{aligned}$$

$$\begin{aligned}g2(\mathbf{x}) &= \ln \left[ \frac{P(Y = 2|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] \\g2(\mathbf{x}) &= \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \dots + \beta_{2p}x_p \\g2(\mathbf{x}) &= \mathbf{x}'\boldsymbol{\beta}_2\end{aligned}$$

Somando  $g1(\mathbf{x})$  e  $g2(\mathbf{x})$ , podemos obter um sistema de equações e assim isolar o termo  $P(Y = 0|\mathbf{x})$ .

$$\begin{cases} g1(\mathbf{x}) = \ln \left[ \frac{P(Y = 1|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] \\ g2(\mathbf{x}) = \ln \left[ \frac{P(Y = 2|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] \end{cases}$$

Aplicando o exponencial em ambos os lados temos:

$$\begin{cases} e^{g1(\mathbf{x})} = \frac{P(Y=1|\mathbf{x})}{P(Y=0|\mathbf{x})} \\ e^{g2(\mathbf{x})} = \frac{P(Y=2|\mathbf{x})}{P(Y=0|\mathbf{x})} \end{cases} \quad (4.8)$$

Somando:

$$e^{g1(\mathbf{x})} + e^{g2(\mathbf{x})} = \frac{P(Y=1|\mathbf{x}) + P(Y=2|\mathbf{x})}{P(Y=0|\mathbf{x})} \quad (4.9)$$

Sabendo-se que:



$$P(Y = 0|\mathbf{x}) + P(Y = 1|\mathbf{x}) + P(Y = 2|\mathbf{x}) = 1 \quad (4.10)$$

Temos que:

$$P(Y = 1|\mathbf{x}) + P(Y = 2|\mathbf{x}) = 1 - P(Y = 0|\mathbf{x}) \quad (4.11)$$

Substituindo (4.11) em (4.9) temos:

$$\begin{aligned} e^{g^1(\mathbf{x})} + e^{g^2(\mathbf{x})} &= \frac{1 - P(Y = 0|\mathbf{x})}{P(Y = 0|\mathbf{x})} \\ (e^{g^1(\mathbf{x})} + e^{g^2(\mathbf{x})}) \cdot P(Y = 0|\mathbf{x}) &= 1 - P(Y = 0|\mathbf{x}) \\ P(Y = 0|\mathbf{x}) + P(Y = 0|\mathbf{x}) \cdot (e^{g^1(\mathbf{x})} + e^{g^2(\mathbf{x})}) &= 1 \\ P(Y = 0|\mathbf{x}) \cdot (e^{g^1(\mathbf{x})} + e^{g^2(\mathbf{x})} + 1) &= 1 \end{aligned}$$

Logo:

$$P(Y = 0|\mathbf{x}) = \frac{1}{(e^{g^1(\mathbf{x})} + e^{g^2(\mathbf{x})} + 1)} \quad (4.12)$$

Para acharmos  $P(Y = 1|\mathbf{x})$  e  $P(Y = 2|\mathbf{x})$  basta substituirmos (4.12) em (4.8)

$$e^{g^1(\mathbf{x})} = \left[ \frac{\frac{P(Y = 1|\mathbf{x})}{1}}{(e^{g^1(\mathbf{x})} + e^{g^2(\mathbf{x})} + 1)} \right]$$

Logo:

$$P(Y = 1|\mathbf{x}) = \frac{e^{g^1(\mathbf{x})}}{(e^{g^1(\mathbf{x})} + e^{g^2(\mathbf{x})} + 1)} \quad (4.13)$$

E:

$$e^{g^2(\mathbf{x})} = \left[ \frac{\frac{P(Y = 2|\mathbf{x})}{1}}{(e^{g^1(\mathbf{x})} + e^{g^2(\mathbf{x})} + 1)} \right]$$

Logo:

$$P(Y = 2|\mathbf{x}) = \frac{e^{g_2(\mathbf{x})}}{(e^{g_1(\mathbf{x})} + e^{g_2(\mathbf{x})} + 1)} \quad (4.14)$$

Seguindo a convenção do modelo logístico binário temos que:

$$\pi_j(\mathbf{x}) = P(Y = j|\mathbf{x}), \text{ para } j = 1, 2, \dots, J.$$

Assim a expressão geral para o modelo com 3 categorias é:

$$P(Y = j|\mathbf{x}) = \frac{e^{g_j(\mathbf{x})}}{\sum_{k=0}^2 e^{g_k(\mathbf{x})}}$$

Sem perda de generalidade, podemos estender o modelo acima para  $J$  tratamentos distintos.

$$P(Y = j|\mathbf{x}) = \frac{e^{g_j(\mathbf{x})}}{\sum_{k=0}^J e^{g_k(\mathbf{x})}} \text{ ou } P(Y = j|\mathbf{x}) = \frac{e^{\mathbf{x}'_j \boldsymbol{\beta}}}{\sum_{k=0}^J e^{\mathbf{x}'_k \boldsymbol{\beta}}} \quad (4.15)$$

Onde  $\beta_0 = 0$  e  $g_0(\mathbf{x}) = 0$ .

O modelo dado em (4.15) é denominado de modelo logit multinomial e representa a probabilidade de escolha da  $j$ -ésima alternativa para o  $n$ -ésimo consumidor.

#### 4.3.2.1 – Estimação dos parâmetros do modelo

Dada uma amostra aleatória das escolhas de  $N$  consumidores,  $(Y_1 = i, Y_2 = j, Y_3 = k, \dots, Y_N = l)$ ,  $i, j, k, l \in \{1, 2, \dots, J\}$ . Define-se como função de verossimilhança a probabilidade conjunta desta amostra. Os estimadores de

máxima verossimilhança fornecem as estimativas dos parâmetros  $(\beta_1, \beta_2, \dots, \beta_s)$  que maximizam a probabilidade da amostra (após conhecidas as escolhas).

$$l(\boldsymbol{\beta}) = \prod_{j=1}^J P(Y = j | \mathbf{X})^{I_{nj}} \quad (4.16)$$

Onde  $I_{nj}$  é uma função indicadora da escolha do  $j$ -ésimo tratamento pelo  $n$ -ésimo indivíduo.

	$I_{n0}$	$I_{n1}$	$I_{n2}$	...	$I_{nJ}$
$Y = 0$	1	0	0	...	0
$Y = 1$	0	1	0	...	0
$Y = 2$	0	0	1	...	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$
$Y = J$	0	0	0	...	1

Tabela 10 - Função indicadora da função de verossimilhança

A função de verossimilhança pode ser expressa da seguinte forma:

$$l(\boldsymbol{\beta}) = [\pi_0(\mathbf{x}_n)^{I_{n0}} \times \pi_1(\mathbf{x}_n)^{I_{n1}} \times \dots \times \pi_J(\mathbf{x}_n)^{I_{nJ}}]$$

Sob a suposição de independência de escolha entre os consumidores temos que:

$$l(\boldsymbol{\beta}) = \prod_{n=1}^N [\pi_0(\mathbf{x}_n)^{I_{n0}} \times \pi_1(\mathbf{x}_n)^{I_{n1}} \times \dots \times \pi_J(\mathbf{x}_n)^{I_{nJ}}]$$

Tomando o log da função e sabendo que  $\sum_{j=1}^J Y_{jn} = 1$ , temos que a função de verossimilhança é dada por:

$$L(\boldsymbol{\beta}) = \ln l(\boldsymbol{\beta}) = \sum_{n=1}^N [I_{n0} \cdot \ln[\pi_0(\mathbf{x}_n)] + I_{n1} \cdot \ln[\pi_1(\mathbf{x}_n)] + \dots + I_{nJ} \cdot \ln[\pi_J(\mathbf{x}_n)]]$$

Sabendo que:

$$\pi_j(\mathbf{x}_n) = P(Y = j|\mathbf{x}_n) = \frac{e^{g_j(\mathbf{x}_n)}}{\sum_{k=0}^J e^{g_k(\mathbf{x}_n)}} = \frac{e^{g_j(\mathbf{x}_n)}}{1 + e^{g^1(\mathbf{x}_n)} + e^{g^2(\mathbf{x}_n)} + \dots + e^{g^J(\mathbf{x}_n)}}$$

Então:

$$\ln[\pi_j(\mathbf{x}_n)] = \ln(e^{g_j(\mathbf{x}_n)}) - \ln(1 + e^{g^1(\mathbf{x}_n)} + e^{g^2(\mathbf{x}_n)} + \dots + e^{g^J(\mathbf{x}_n)})$$

$$\ln[\pi_j(\mathbf{x}_n)] = g_j(\mathbf{x}_n) - \ln(1 + e^{g^1(\mathbf{x}_n)} + e^{g^2(\mathbf{x}_n)} + \dots + e^{g^J(\mathbf{x}_n)})$$

Generalizando temos:

$$\begin{aligned} L(\boldsymbol{\beta}) = \sum_{n=1}^N & \left[ I_{n0} \cdot \ln[\pi_0(\mathbf{x}_n)] + [I_{n1} \cdot (g_1(\mathbf{x}_n) - \ln(1 + e^{g^1(\mathbf{x}_n)} + e^{g^2(\mathbf{x}_n)} + \dots + e^{g^J(\mathbf{x}_n))})] \right. \\ & + [I_{n2} \cdot (g_2(\mathbf{x}_n) - \ln(1 + e^{g^1(\mathbf{x}_n)} + e^{g^2(\mathbf{x}_n)} + \dots + e^{g^J(\mathbf{x}_n))})] + \dots \\ & \left. + [I_{nJ} \cdot (g_J(\mathbf{x}_n) - \ln(1 + e^{g^1(\mathbf{x}_n)} + e^{g^2(\mathbf{x}_n)} + \dots + e^{g^J(\mathbf{x}_n))})] \right] \end{aligned}$$

Sabendo que  $\ln[\pi_0(\mathbf{x}_n)] = 0$ .

$$\begin{aligned} L(\boldsymbol{\beta}) = \sum_{n=1}^N & [I_{n1} \cdot g_1(\mathbf{x}_n) + I_{n2} \cdot g_2(\mathbf{x}_n) + \dots + I_{nJ} \cdot g_J(\mathbf{x}_n) - \ln(1 + e^{g^1(\mathbf{x}_n)} + e^{g^2(\mathbf{x}_n)} + \dots \\ & + e^{g^J(\mathbf{x}_n)})] \end{aligned}$$

Para ajustar o modelo de regressão devemos estimar os parâmetros do modelo. Os estimadores de máxima verossimilhança para os parâmetros  $\beta_{jk}$  são os valores  $\widehat{\beta}_{jk}$  que maximizam o logaritmo da função de verossimilhança. Onde  $j = 1, 2, \dots, J$  e  $k = 1, 2, \dots, p$ .

A função de verossimilhança tem máximo, pois  $0 < P(Y = j|\mathbf{x}_n) < 1$  mesmo a função logaritmo sendo função estritamente crescente (BASTOS, 2010).

Para maximizar a função de máxima verossimilhança basta tomarmos as primeiras derivadas parciais de  $L(\boldsymbol{\beta})$  em relação a cada parâmetro desconhecido.

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_{jk}} = \frac{\partial}{\partial \beta_{jk}} \left\{ \sum_{n=1}^N [I_{n1} \cdot g1(\mathbf{x}_n) + I_{n2} \cdot g2(\mathbf{x}_n) + \dots + I_{nj} \cdot g_j(\mathbf{x}_n) - \ln(1 + e^{g1(\mathbf{x}_n)} + e^{g2(\mathbf{x}_n)} + \dots + e^{g_j(\mathbf{x}_n)})] \right\}$$

Dado que:

$$g_j(\mathbf{x}) = \beta_{j0} + \beta_{j1}x_1 + \dots + \beta_{jk}x_k + \dots + \beta_{jp}x_p$$

$$\pi_j(\mathbf{x}_n) = \frac{e^{g_j(\mathbf{x}_n)}}{1 + e^{g1(\mathbf{x}_n)} + e^{g2(\mathbf{x}_n)} + \dots + e^{g_j(\mathbf{x}_n)}}$$

temos que:

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_{jk}} = \sum_{n=1}^N I_{nj} \cdot x_{kn} - \frac{e^{g_j(\mathbf{x}_n)} \cdot x_{kn}}{1 + e^{g1(\mathbf{x}_n)} + e^{g2(\mathbf{x}_n)} + \dots + e^{g_j(\mathbf{x}_n)}}$$

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_{jk}} = \sum_{n=1}^N I_{nj} \cdot x_{kn} - \pi_j(\mathbf{x}_n) \cdot x_{kn}$$

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_{jk}} = \sum_{n=1}^N x_{kn} [I_{nj} - \pi_j(\mathbf{x}_n)] \quad (4.17)$$

Os estimadores de máxima verossimilhança,  $\hat{\boldsymbol{\beta}}$ , são obtidos igualando as equações das derivadas acima à zero. Porém estas equações são não lineares nos parâmetros e para sua resolução é necessário utilizar de métodos computacionais iterativos. O método Newton-Raphson (GALLANT, 1984), é uma alternativa e é bastante utilizada nos softwares atuais. Como é um método de implementação computacional, e dificilmente iremos reproduzi-lo na “mão”, optamos por não demonstrá-lo, porém em Bastos(2010), vemos todos os passos de seu desenvolvimento.

O método de estimar as variâncias e covariâncias dos coeficientes estimados segue a teoria da estimação da máxima verossimilhança, que assegura que os estimadores são obtidos da matriz de segundas derivadas parciais.

$$\begin{aligned}
\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_{jk} \partial \beta_{j'k'}} &= \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_{jk} \partial \beta_{j'k'}} \left( \sum_{n=1}^N I_{nj} \cdot x_{kn} - \pi_j(\mathbf{x}_n) \cdot x_{kn} \right) \\
&= - \sum_{n=1}^N x_{kn} \cdot \left( \frac{(e^{g_j(x_n)} \cdot x_{k'n}) \cdot (1 + e^{g^1(x_n)} + e^{g^2(x_n)} + \dots + e^{g^J(x_n)}) - (e^{g_j(x_n)}) \cdot (e^{g_j(x_n)}) \cdot (x_{k'n})}{(1 + e^{g^1(x_n)} + e^{g^2(x_n)} + \dots + e^{g^J(x_n)})^2} \right) \\
&= - \sum_{n=1}^N x_{kn} \cdot \left[ \frac{(e^{g_j(x_n)} \cdot x_{k'n}) \cdot (1 + e^{g^1(x_n)} + \dots + e^{g^J(x_n)})}{(1 + e^{g^1(x_n)} + \dots + e^{g^J(x_n)})^2} - \frac{(e^{g_j(x_n)}) \cdot (e^{g_j(x_n)}) \cdot (x_{k'n})}{(1 + e^{g^1(x_n)} + \dots + e^{g^J(x_n)})^2} \right] \\
&= - \sum_{n=1}^N x_{kn} \cdot (x_{k'n} \pi_j(\mathbf{x}_n) - \pi_j(\mathbf{x}_n) \cdot \pi_j(\mathbf{x}_n) \cdot x_{k'n}) \\
&= - \sum_{n=1}^N x_{kn} \cdot x_{k'n} \cdot \pi_j(\mathbf{x}_n) - x_{kn} \cdot x_{k'n} \cdot \pi_j(\mathbf{x}_n) \cdot \pi_j(\mathbf{x}_n) \\
&= - \sum_{n=1}^N x_{kn} \cdot x_{k'n} \cdot \pi_j(\mathbf{x}_n) \cdot (1 - \pi_j(\mathbf{x}_n)) \tag{4.18}
\end{aligned}$$

E

$$\begin{aligned}
\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_{jk} \partial \beta_{j'k'}} &= \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_{jk} \partial \beta_{j'k'}} \left( \sum_{n=1}^N I_{nj} \cdot x_{kn} - \pi_j(\mathbf{x}_n) \cdot x_{kn} \right), \quad \forall j \neq j' \ e \ j' = 1, 2, \dots, J \\
&= - \sum_{n=1}^N x_{kn} \cdot \left[ \frac{0 \cdot (1 + e^{g^1(x_n)} + e^{g^2(x_n)} + \dots + e^{g^J(x_n)}) - (e^{g_j(x_n)}) \cdot (e^{g_{j'}(x_n)}) \cdot (x_{k'n})}{(1 + e^{g^1(x_n)} + e^{g^2(x_n)} + \dots + e^{g^J(x_n)})^2} \right] \\
&= - \sum_{n=1}^N x_{kn} \cdot \left[ \frac{-(e^{g_j(x_n)}) \cdot (e^{g_{j'}(x_n)}) \cdot (x_{k'n})}{(1 + e^{g^1(x_n)} + e^{g^2(x_n)} + \dots + e^{g^J(x_n)})^2} \right] \\
&= \sum_{n=1}^N x_{kn} \cdot x_{k'n} \cdot \pi_j(\mathbf{x}_n) \cdot \pi_{j'}(\mathbf{x}_n) \tag{4.19}
\end{aligned}$$

A matriz de informação de fisher observada,  $\mathbf{I}(\hat{\boldsymbol{\beta}})$ , é uma matriz simétrica cujos elementos são os negativos das derivadas segundas parciais. As variâncias e covariâncias dos coeficientes estimados são obtidos da inversa da matriz de informação de fisher.

$$\widehat{Var}(\widehat{\boldsymbol{\beta}}) = \mathbf{I}(\widehat{\boldsymbol{\beta}})^{-1}$$

A matriz de informação de fisher estimada,  $\widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})$ , pode ser encontrada de maneira similar ao modelo binário onde:

$$\widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}}) = \begin{bmatrix} \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{11} & \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{12} & \dots & \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{1J} \\ \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{21} & \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{22} & \dots & \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{J1} & \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{J2} & \dots & \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{JJ} \end{bmatrix} \quad (4.20)$$

Onde:

$$\begin{aligned} \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{11} &= (\mathbf{X}'\mathbf{V}_{11}\mathbf{X}) \\ \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{12} &= -(\mathbf{X}'\mathbf{V}_{12}\mathbf{X}) = \widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{21} \end{aligned}$$

Generalizando:

$$\widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{jj} = (\mathbf{X}'\mathbf{V}_{jj}\mathbf{X}) \quad (4.21)$$

$$\widehat{\mathbf{I}}(\widehat{\boldsymbol{\beta}})_{jj} = -(\mathbf{X}'\mathbf{V}_{jj}\mathbf{X}) \quad (4.22)$$

Onde,  $\forall j$ , com  $j = 1, 2, \dots, J$ :

$$\mathbf{V}_{jj} = \begin{bmatrix} \pi_j(\mathbf{x}_n) \cdot (1 - \pi_j(\mathbf{x}_n)) & 0 & \dots & 0 \\ 0 & \pi_j(\mathbf{x}_n) \cdot (1 - \pi_j(\mathbf{x}_n)) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \pi_j(\mathbf{x}_n) \cdot (1 - \pi_j(\mathbf{x}_n)) \end{bmatrix}$$

E

$$\mathbf{V}_{jj} = \begin{bmatrix} \pi_j(\mathbf{x}_n) \cdot \pi_j(\mathbf{x}_n) & 0 & \dots & 0 \\ 0 & \pi_j(\mathbf{x}_n) \cdot \pi_j(\mathbf{x}_n) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \pi_j(\mathbf{x}_n) \cdot \pi_j(\mathbf{x}_n) \end{bmatrix}$$

#### 4.3.2.2 – Razão de escolhas

A *conjoint analysis* baseada em escolhas, oferece um resultado interessante que é a razão de escolhas, nome adaptado à área de *marketing*, porém se trata da mesma razão de chances do modelo logístico binário.

Portanto suponhamos que  $\mathbf{X}'_p = [x_1^p, x_2^p, \dots, x_k^p, \dots, x_s^p]$ , é o vetor dos níveis dos fatores presentes no tratamento  $p$  e,  $\mathbf{X}'_q = [x_1^q, x_2^q, \dots, x_k^q, \dots, x_s^q]$ , é o vetor dos níveis dos fatores presentes no tratamento  $q$ , então  $x_i^p \neq x_i^q$  para um único  $i \in \{1, 2, \dots, k-1, k, k+1, \dots, s\}$ . A razão de escolha pode ser denotada como:

$$RE_k(p, q) = \frac{P(Y_n=p|x)}{P(Y_n=q|x)} = e^{(x_k^p - x_k^q)\beta_k} \quad (4.23)$$



## 5- Aplicação

### 5.1- Contextualização do Problema

O mercado de telefonia há tempos abrange outros mercados. As empresas de telefonia ou operadoras de telefonia, em sua maioria, não têm a telefonia como única e exclusiva fonte de renda. Nos últimos anos, estas empresas vêm abrangendo outros serviços a fim de aumentar sua lucratividade e expandir seu mercado em uma área que estava defasada no início dos anos 2000.

Com a popularização da internet e a viabilização da banda larga, as empresas de telefonia, que antes eram responsáveis pela transmissão de internet discada, passaram a disponibilizar serviços de internet banda larga com maior frequência e com preços cada vez mais acessíveis.

A partir do momento em que os serviços de banda larga ficaram muito parecidos entre os concorrentes, novos meios de diferenciação foram realizados para agregar valor ao cliente. A partir da viabilização em termos de custo, foram agregados serviços de TV por assinatura, formas alternativas de disponibilização da internet, como internet à rádio ou internet 3G e também, a criação de planos econômicos e com características pré-definidas para telefones fixos e móveis.

Nos dias atuais, isso se traduz em um mercado extremamente customizado e fragmentado, onde existem milhares de configurações para atender melhor e mais especificadamente os consumidores.

Este mercado é caracterizado por ser altamente restrito a grandes companhias, em alguns casos quase monopolistas, como nas regiões mais interiores do país. E ao mesmo tempo é visto como um mercado muito lucrativo e de extrema concorrência, já que, cada vez mais as empresas vêm adotando estratégias agressivas para conquistar mercado.

A participação de mercado que uma empresa tem é uma informação extremamente importante e que pode auxiliar nas estratégias a serem seguidas de acordo com os objetivos.

## 5.2- Descrições dos dados

Para ilustrar o desenvolvimento de uma aplicação possível da *conjoint analysis*, utilizaremos um banco de dados de uma pesquisa no setor de telefonia, que obteve as escolhas de 67 consumidores da cidade de Araxá por diferentes configurações de combinações de serviços do tipo: telefone + internet + TV por assinatura.

A variável de interesse é “Escolha” onde o tratamento escolhido recebe o valor 1 e os tratamentos não escolhidos recebem valor 0.

Os 67 entrevistados foram divididos entre 33 homens e 34 mulheres, com idades entre 18 e 65 anos categorizadas em 6 níveis.

Através de alguns cruzamentos de dados obtivemos algumas informações sobre o perfil socioeconômico e demográfico dos entrevistados

- 34,3% dos entrevistados tinham entre 18 a 24 anos onde destes, 43,5% são homens e 56,5% são mulheres.
- 40,3% dos entrevistados tinham entre 25 a 34 anos onde destes, 51,9% são homens e 48,1% são mulheres.
- 20,9% dos entrevistados tinham entre 35 a 49 anos onde destes, 66,7% são homens e 33,3% são mulheres.
- 4,5% dos entrevistados tinham entre 50 a 64 anos onde destes, 49,3% são homens e 50,7% são mulheres.

Obtivemos também que:

- 88,1% dos entrevistados estão dentro da faixa de renda 3, e 11,9% dentro da faixa de renda 4. Dos que estão na faixa de renda 3, 45,8% são homens e 54,2% são mulheres, e dos que estão na faixa de renda 4, 75% são homens e 25% são mulheres.

São considerados Fatores considerados importantes para a escolha de um pacote de serviços são: Operadora, Plano de TV, Pacotes de TV Premium, Serviços de Internet, Velocidade da Internet, Pacotes de Minutos do Fixo, Pacotes de Minutos do Fixo para Outras Operadoras, Preço Adaptado.

<b>Código</b>	<b>Operadora</b>	<b>Plano de Tv</b>	<b>Pacotes Premium (Tv)</b>	<b>Tipo do serviço de internet</b>	<b>Velocidade internet</b>	<b>Pacote de minutos no fixo****</b>	<b>Pacote de minutos no fixo para outra operadora</b>
<b>1</b>	1	Nenhum plano de Tv	Nenhum canal premium (1)	Nenhum	Nenhuma	Nenhum pacote de minutos incluído	Nenhum pacote de minutos incluído
<b>2</b>	2	20 canais	2	Banda larga fixa (ADSL ou Cabo)	200 Kbps sem limite mensal de download	Paga apenas pelo uso	100 minutos de ligações locais para outra operadora
<b>3</b>	3	40 canais	3	3G	500 Kbps sem limite mensal de download	150 minutos incluídos para ligações fixo-fixo local	150 minutos de ligações locais para outra operadora
<b>4</b>	4	80 canais	4	Banda larga wireless (Rádio)	500 Kbps com pacote mensal de 3GB para download	250 minutos incluídos para ligações fixo-fixo local	200 minutos de ligações locais para outra operadora
<b>5</b>	5		5		1 Mbps sem limite mensal de download	1000 minutos incluídos para ligações fixo-fixo local	
<b>6</b>	6				1 Mbps com pacote mensal de 5GB para	2000 minutos incluídos para	

		download	ligações fixo-fixo local
<b>7</b>	7	2Mbps sem limite mensal de download	ilimitado para chamadas para telefone fixo
<b>8</b>	8	2Mbps com pacote mensal de 7GB para download	
<b>9</b>		10 Mbps sem limite mensal de download	

Tabela 11- Níveis dos fatores utilizados na aplicação

Como podemos ver um planejamento fatorial completo seria impossível de administrar, pois assim teríamos  $(4)(5)(4)(9)(7)(4) = 20160$  tratamentos possíveis. Então foi adotado um planejamento fatorial fracionário onde cada um dos 67 respondentes avalia 20 cenários com 4 tratamentos cada escolhendo um, nenhum ou mais de um tratamento.

<b>OPERADORA</b> 4	<b>OPERADORA</b> 2	<b>OPERADORA</b> 7	<b>OPERADORA</b> 2
<b>PLANO DE TV</b> Nenhum plano de Tv	<b>PLANO DE TV</b> Nenhum plano de Tv	<b>PLANO DE TV</b> 20 canais	<b>PLANO DE TV</b> Nenhum plano de Tv
<b>PACOTES PREMIUM (TV)</b> 1	<b>PACOTES PREMIUM (TV)</b> 1	<b>PACOTES PREMIUM (TV)</b> 2	<b>PACOTES PREMIUM (TV)</b> 1
<b>TIPO DE SERVIÇO DE INTERNET</b> Banda larga fixa (ADSL ou Cabo)	<b>TIPO DE SERVIÇO DE INTERNET</b> Banda larga fixa (ADSL ou Cabo)	<b>TIPO DE SERVIÇO DE INTERNET</b> Nenhum	<b>TIPO DE SERVIÇO DE INTERNET</b> Banda larga fixa (ADSL ou Cabo)
<b>VELOCIDADE INTERNET</b> 2Mbps sem limite mensal de download	<b>VELOCIDADE INTERNET</b> 1 Mbps sem limite mensal de download	<b>VELOCIDADE INTERNET</b> Nenhuma	<b>VELOCIDADE INTERNET</b> 1 Mbps sem limite mensal de download
<b>MINUTOS NO FIXO</b> Nenhum pacote de minutos incluído	<b>MINUTOS NO FIXO</b> Paga apenas pelo uso	<b>MINUTOS NO FIXO</b> Nenhum pacote de minutos incluído	<b>MINUTOS NO FIXO</b> Nenhum pacote de minutos incluído
<b>MINUTOS NO FIXO OUTRA OPERADORA</b> Nenhum pacote de minutos incluído	<b>MINUTOS NO FIXO OUTRA OPERADORA</b> 200 minutos de ligações locais para outra operadora	<b>MINUTOS NO FIXO OUTRA OPERADORA</b> Nenhum pacote de minutos incluído	<b>MINUTOS NO FIXO OUTRA OPERADORA</b> Nenhum pacote de minutos incluído
<b>MENSALIDADE</b> R\$ 140,00	<b>MENSALIDADE</b> R\$ 170,00	<b>MENSALIDADE</b> R\$ 60,00	<b>MENSALIDADE</b> R\$ 120,00
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>

Tabela 12- Entrevista 1, Cenário 1

### 5.3- Aplicação da *Conjoint Analysis* baseada em escolhas para estimação de participação de mercado ou *share* de preferência.

Como dito anteriormente a *conjoint analysis* baseada em escolhas nos trará respostas como, “qual a probabilidade da marca X ser escolhida?”, “Qual cor será escolhida?”, “Qual a probabilidade do produto com preço mais alto ser o escolhido?”, etc.

Bastos(2010) afirma que estas probabilidades podem ser interpretadas como estimativa de participação de mercado em um estudo envolvendo produtos de distintos fabricantes envolvendo os mesmos fatores.

Neste intuito podemos obter uma boa *proxy* da participação de mercado das operadoras ao estimar as probabilidades de escolha dos tratamentos disponíveis e agrupá-los para cada operadora. Este resultado é usualmente chamado de *share* de preferência no meio empresarial.

No nosso caso, temos 3159 tratamentos distintos, cujas probabilidades de escolha são muito pequenas analisadas individualmente, porém somadas, essas probabilidades nos dão a participação de mercado estimada para as operadoras que as ofertam.

Esta análise foi feita com a utilização dos softwares SPSS e EXCELL.

Os coeficientes estimados foram obtidos através do procedimento *multinomial logistic* no software SPSS. A variável resposta é “Escolha”, e as variáveis explicativas foram declaradas como covariáveis.

Coeficientes estimados	
Variáveis	Coeficientes
Operadora	-0,05
Plano de TV	-0,091
Pacotes de TV <i>premium</i>	-0,146
Tipo de serviço de internet	0,069
Velocidade da internet	-0,059
Pacote de minutos para o fixo	-0,129
Pacote de minutos do fixo para outras operadoras	0,065
Preço adaptado	-0,004

Tabela 13- Coeficientes estimados pelo método *multinomial logistic* do SPSS

O cálculo das probabilidades de cada tratamento foi obtido através da fórmula

$$P(Y = j|\mathbf{x}) = \frac{e^{\mathbf{x}'_j\boldsymbol{\beta}}}{\sum_{k=0}^J e^{\mathbf{x}'_k\boldsymbol{\beta}}}$$

Através do comando *compute* foi calculado o valor de  $e^{\mathbf{x}'_j\boldsymbol{\beta}}$  para todas as configurações de plano possíveis. Após este passo, criamos uma chave para identificar cada configuração pelas suas características a fim de identificar os casos duplicados, ou seja, configurações ou tratamentos que aparecem mais de uma vez no experimento.



A chave foi feita através da seguinte fórmula:

```
Operadora*1 + PlanodeTv * 10 + PacotesPremiumTv * 100 +  
Tipodoserviçodeinternet * 1000 + Velocidadeinternet * 10000 +  
Pacotedeminutosnofixo * 100000 +  
Pacotedeminutosnofixoparaoutraoperadora * 1000000 + PreçoAdaptado *  
10000000
```

A leitura da chave pode ser facilitada se feita de trás pra frente, de forma que a configuração 301111111, representa operadora 1, sem plano de TV(nível 1), sem pacotes de TV *premium* (nível 1), sem internet (nível 1), sem velocidade de internet (nível 1), pacote de minutos do fixo onde se paga só pelo que usou (nível 1), 100 minutos para ligações do fixo para outras operadoras (nível 1) e preço adaptado de 30 reais.

De cerca de 5000 configurações de planos na base de dados, filtramos os casos duplicados e chegamos a um numero de 3159 configurações distintas. A finalidade desta filtragem é calcular o valor  $\sum_{k=0}^J e^{X'_k \beta}$ .

Dividindo o valor  $e^{X'_j \beta}$  de cada configuração por  $\sum_{k=0}^J e^{X'_k \beta}$  podemos encontrar as probabilidades de escolha para cada configuração.

Pela quantidade enorme de configurações, as probabilidades de escolha relativamente pequenas, da ordem que a maior probabilidade de escolha é 0,2882%, que se refere ao plano 302211111, ou seja, plano da operadora 1, sem plano de Tv por assinatura e conseqüentemente sem pacote de TV premium, sem internet e conseqüentemente sem velocidade internet, com plano de minutos de fixo aonde se paga apenas pelo usado, um plano de 100 minutos para ligações locais para outras operadoras e um preço de 30 reais.

Estas probabilidades tendem a ser cada vez menores a maneira que as empresas tendem a diferenciar cada vez mais seus planos para atingir cada vez mais nichos menores e mais específicos.

Portanto a empresa pode saber se suas ações, em um modo geral, estão sendo absorvidas com sucesso, estimando a participação de mercado que ela está conquistando ou o *share* de preferência.

Para tal tarefa basta agruparmos as probabilidades de escolha das configurações de plano para cada operadora.

<b>Estimativa da participação de mercado das operadoras</b>	
<b>Operadora 1</b>	17,81%
<b>Operadora 2</b>	16,06%
<b>Operadora 3</b>	15,73%
<b>Operadora 4</b>	14,63%
<b>Operadora 5</b>	13,42%
<b>Operadora 6</b>	12,39%
<b>Operadora 7</b>	0,39%
<b>Operadora 8</b>	9,57%
<b>Total</b>	100%

Tabela 14- Participação de Mercado das Operadoras

Vemos que a operadora 1 tem a maior fatia de mercado com 17,8% seguida de perto pelas operadoras 2 e 3 com 16,06% e 15,73% respectivamente. A operadora 7 foi a que apresentou a menor participação de mercado com apenas 0,39%.

Podemos também avaliar quais são os carros-chefes das operadoras, ou seja, os planos com maior probabilidade de aquisição. Esta análise pode ser bem proveitosa, se a empresa não sabe qual é seu plano com maior possibilidade de venda. A partir desta informação é possível gerar campanhas de publicidade em cima destas melhores configurações de plano. É possível também, analisar as melhores ofertas dos concorrentes, a fim de, descobrir vantagens competitivas e se igualar ou até superar no mercado.

Como ilustração, vamos avaliar as 10 configurações de plano com maior probabilidade de aquisição de 2 concorrentes, operadora 2 e operadora 3.

<b>Operadora 3</b>		<b>Operadora 2</b>	
<b>Configuração do Plano</b>	<b>Probabilidade</b>	<b>Configuração do Plano</b>	<b>Probabilidade</b>
<b>(chave)</b>		<b>(chave)</b>	
<b>304211113</b>	,002611	<b>302211112</b>	,002410
<b>303211113</b>	,002447	<b>804222112</b>	,002270
<b>302211113</b>	,002293	<b>301211112</b>	,002259
<b>804222113</b>	,002159	<b>601143112</b>	,002192
<b>301211113</b>	,002149	<b>601132112</b>	,002170
<b>601143113</b>	,002085	<b>301311112</b>	,001985
<b>601132113</b>	,002064	<b>1004232112</b>	,001975
<b>802222113</b>	,001896	<b>1003243112</b>	,001870
<b>301311113</b>	,001889	<b>404511112</b>	,001791
<b>802243113</b>	,001805	<b>801243112</b>	,001779

Tabela 15- Comparativo de probabilidades das 10 configurações com maior chance de aquisição das operadoras 2 e 3

Podemos observar que as três maiores probabilidades de aquisição na operadora 3 referem-se à planos econômicos variando apenas nos minutos do fixo para outras operadoras. Porém a operadora 2 tem como segundo plano com

maior probabilidade de aquisição um plano que contém internet, onde este mesmo plano na operadora 3 é apenas o de quarta maior probabilidade de aquisição. Está pode ser uma informação, que caso a empresa não a conheça, que pode nortear campanhas de publicidade de vendas e se tornar uma vantagem competitiva em relação às demais.

Podemos calcular a razão de escolhas do plano 804222112 e 804222113, ou seja, a razão de escolhas dos planos sem pacotes de TV, com internet mais simples, transmissão via ADSL e 200 kb/s de velocidade, com plano econômico de telefone com 150 minutos de telefone no fixo para mesma operadora e com também 150 minutos de telefone fixo para outras operados por um preço de 80 reais entre as operadoras 2 e 3.

Portanto variando os níveis do fator operadora (fator 1) e mantendo-se fixo todos os outros temos que:

$$RE_1(804222112,804222113) = \frac{P(Y_n = 804222112|\mathbf{x})}{P(Y_n = 804222113|\mathbf{x})} = 1,0514$$

Isto quer dizer que um indivíduo  $n$  tem 5,14% de chance a mais de escolher o plano da operadora 2 em relação ao mesmo plano na operadora 3.

A partir daí a operadora 3 pode investigar o porque do plano com a mesma característica está sendo preterido ao da operadora 2.

Da tabela 11 podemos avaliar também que dos 10 planos com maior probabilidade de aquisição nenhum deles possui planos de TV por assinatura. O que prova que este fator não é tão importante na escolha de um plano. Por

outro lado, se as operadoras não estão explorando este fator como se deve, podemos ter uma vantagem a ser explorada.

Podemos avaliar também as probabilidades de aquisição dos planos com preço acima de 180 reais, ou seja, planos com maior diferenciação, e provavelmente com maior margem de lucro.

<b>Probabilidades de aquisição de planos com preço maior que 180 reais</b>		
<b>Plano</b>	<b>Chave</b>	<b>Probabilidade</b>
<b>1</b>	2301192111	,000811
<b>2</b>	2303292112	,000772
<b>3</b>	2301192112	,000772
<b>4</b>	2301192113	,000734
<b>5</b>	2004732111	,000730
<b>6</b>	2301292111	,000713
<b>7</b>	2301192114	,000698
<b>8</b>	2301192115	,000664
<b>9</b>	2004732113	,000661
<b>10</b>	2303294118	,000657

Tabela 16- Probabilidades de aquisição de planos com preço maior que 180 reais

Podemos ver que as Operadoras 1 e 2 dominam maior parte deste mercado também, e vemos que a operadora 5 tem apenas o oitavo plano com maior probabilidade de aquisição, que é justamente formado pelas mesmas características do plano 1, cujo tem maior probabilidade de aquisição e é da operadora 1.

Portanto a razão de escolhas entre os planos 1 e 8 da tabela 12, cujas características são sem plano de TV por assinatura, internet distribuída via

ADSL, com velocidade de 10 MB, e nenhum pacote de minutos incluído no telefone fixo, com preço de 230 reais, entre as operadoras 1 e 5, pode ser expressa como:

$$RE_1(2301192111,2301192115) = \frac{P(Y_n = 2301192115|\mathbf{x})}{P(Y_n = 2301192115|\mathbf{x})} = 1,2213$$

Podemos concluir que a chance de um indivíduo  $n$  escolher o plano da operadora 1 é de 22,13% maior do que de escolher o plano da operadora 5.

O banco de dados é muito rico e extenso, permitindo várias interpretações e resultados muito valiosos no meio empresarial. Ainda poderiam ser feitas várias análises, como por exemplo, de qual operadora tem maior probabilidade de escolha considerando plano de internet mais simples ou com preços mais acessíveis, ou considerando algum tipo de pacote de TV por assinatura, entre outros. Ainda, os resultados gerados nas análises acima poderiam ser expandidos para uma comparação entre todas as operadoras e envolver variáveis socioeconômicas.

## 6- Conclusão e recomendações futuras

É evidente a grande utilidade dos modelos de *conjoint analysis* para o mercado empresarial frente ao cenário competitivo em que vivemos. Captar a subjetividade do processo de decisão de compra de um consumidor e transformá-lo em informações quantitativas que podem nortear o processo de decisão empresarial e as estratégias de *marketing* de fato é uma grande vantagem competitiva.

Pode-se observar que as vertentes da técnica, como os modelos de *conjoint analysis* tradicional e os modelos de *conjoint analysis* baseado em escolhas e outros modelos alternativos são complementares e não excludentes. Enquanto o modelo tradicional estima a importância dos atributos de um produto no processo de decisão de compra, o modelo discreto estima probabilidades de escolha de cada produto condicionado a sua configuração, portanto os modelos permitem conclusões e aplicações diferentes, e como ditas acima, conclusões estas que são complementares e que são obtidas somente com cada um dos modelos.

A metodologia é teoricamente recente, e ainda apresenta algumas desvantagens, como a impossibilidade de gerar experimentos com muitos fatores. Porém os avanços tecnológicos vieram a facilitar a aplicação das pesquisas, com a utilização de pesquisas administradas por computador e novos meios de apresentar os estímulos aos respondentes.

‘

Sugerimos trabalhos que possam relacionar a *conjoint analysis* com outras técnicas multivariadas. Por exemplo, a obtenção dos fatores mais importantes pode ser obtida através da análise fatorial, podemos utilizar análise de *cluster* para agrupar os respondentes não por características socioeconômicas apenas, mas sim por características de preferência como os coeficientes de preferência.



## Referências Bibliográficas

- BARQUETTE, S., & CHAOUBAH, A. (2007). *Pesquisa de Marketing*. São Paulo: Saraiva.
- BASTOS, F. S. (2010). *Análise conjunta de Fatores Baseada em Escolhas: Estimativa e Inferências*. Viçosa - MG: Dissertação ( Mestrado em Estatística Aplicada e Biometria ) Departamento de Estatística, Universidade Federal de Viçosa.
- BEN-AKIVA, M., & LERMAN, S. R. (1985). *Discrete Choice Analysis*. Cambridge, MA,,: MIT Press.
- BRITO, A. N. (2007). *Aplicação de um procedimento com preferência declarada para estimativa do valor do tempo de viagem de motoristas em uma escolha entre rotas rodoviárias pedagiadas e não pedagiadas*. São Paulo, 185 p. Dissertação ( Mestrado ): Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Trasnportes.
- CASTRO, L. R. (2006). *Valor percebido como ferramenta para tomada de decisão: Uma aplicação na indústria hoteleira utilizando Analise Conjunta*. São Paulo - SP: Tese ( Mestrado em Engenharia de Produção) - Escola de Engenharia de São Carlos, Universidade de são Paulo.
- CÓRDOVA, L. A. (2002). *Um modelo desagregado de escolha discreta para a análise das preferências do mercado: Uma abordagem Bayesiana, 130 p*. Florianópolis - SC: Dissertação (Doutorado\_ - Universidade Federal de Santa Catarina.
- Da SILVA, C. H., & BASTOS, F. S. (2010). *Introdução à Conjoint Analysis, Minicurso ministrado no IX encontro mineiro de estatística*. Viçosa - MG.

- DELLA LUCIA, S. M. (2010). Análise conjunta de fatores baseada em escolhas no estudo da embalagem de iogurte light sabor morango. *Brazilian Journal of Food technology*, 11 - 18.
- ENGEL, J., BLACKWELL, R. D., & MINIARD, P. W. (2000). *Comportamento do Consumidor, 8. ed.* Rio de Janeiro: LTC.
- FRIEDMANN, L. S. (1998). *Análise de Preferência. Dissertação de mestrado.* São Paulo: Escola de Administração de Empresas da Fundação Getúlio Vargas (FGV - EAE SP).
- GALLANT, A. R. (1984). *Nonlinear Statistical Models.* New York: John Wiley and Sons, Inc, forthcoming.
- GREEN, P. E. (1984). Hybrid Model for Conjoint Analysis: an Expository Review. *Journal of Marketing research*, 155 - 169.
- GREEN, P. E., & SRINIVASA, N. V. (1978). Conjoint Analysis in consumer research: issues and outlook. *Journal of Consumer Research*, pp. v. 5; p. 103 - 123.
- GREEN, P. E., & SRINIVASAN, V. (1990). Conjoint Analysis in Marketing: New Developments with implications for research and practice. *Journal of Marketing Research*, v. 54; p. 3 - 19.
- GREEN, P. E., & WIND, Y. (1975). New Ways to Measure Consumer Judgements. *Harvard Business Review*, v. 53; p. 107 - 117.
- GREEN, P. E., CARROL, J. D., & CARMONE, F. J. (1978). Some New Types of Fractional Factorial Designs for Marketing Experiments. (1 - (99-122)).
- GREEN, P. E., KRIEGER, A. M., & WIND, Y. (2001). Thirty Years of Conjoint Analysis: reflections and prospects. *Interfaces*, v. 31, p. 355 - 363.

- GREEN, P., & RAO, V. (1971). Conjoint Measurement for quantifying judgemental data. *Journal of Marketing Research*, v.8, 355-363.
- HAIR JR, J. F., ANDERSON, R. E., & TATHAN, R. L. (1995). *Multivariate data analysis with readings, 4 ed.* USA: Prentice Hall.
- HAIR JR., J. F. (2005). *Análise multivariada de dados 5ed.* Porto Alegre: Bookman.
- HENRIQUE, J. L., & DE SOUZA, V. R. (2006). *O uso da Técnica de Análise Conjunta na Pesquisa em Marketing: Uma Avaliação das Publicações Brasileiras.* Salvador - BA: 30º EnAPAD.
- HOSMER, D. W., & LEMESHOW, S. (1989). *Applied Logistic Regression.* New York: John Wiley & Sons, INC.
- KOTLER, P. (1988). *Marketing Management: analysis, planning, implamentation, and control. 6 ed.* USA: Prentice Hall.
- KOTLER, P. (1995). *Administração de marketing: análise, planejamento, implementação e controle 4ed.* São Paulo: Atlas.
- KOTLER, P., & ARMSTRONG, G. (2008). *Princípios de marketing, 12ª edição.* Rio de Janeiro: Prentice Hall do Brasil LTDA.
- KUNDER, W. G. (15 de 04 de 2010). *Blog Gestão da Produção e da Qualidade; Post: Produção customizada uma alternativa para pequenas empresas.* Acesso em 14 de 09 de 2011, disponível em Portal SEBRAE: <http://portal.pr.sebrae.com.br/blogs/posts/gestaoproducao?c=1100>
- LEVIIT, T. (1995). *"La globalización de los mercados" La esencia del marketing: estratégia. Vol I.* Seleção de Robert J. Dalan. Colômbia: Havard Business School Press: Grupo editorial Norma.
- LOUVIERE, J., & GAETH, G. (1988). A comparison of rating and choise responses in conjoint analysis. *In R. M. Johnson (ED.), Proceedings of*

- the Sawtooth Software Conference of Perceptual Mapping, Conjoint Analysis, and Computer Interviewing*, No. 2, pp.59-73.
- LUCE, R. D., & TUKEY, J. W. (1964). Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement. *v. 1; p. 1 - 27*.
- MALHOTRA, N. K. (2005). *Pesquisa de Marketing: Uma orientação aplicada*. Porto Alegre: Prentice Hall.
- MCFADDEN, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, 105 - 142.
- MEYER, R. J., & KAHN, B. E. (1991). Probabilistic models of consumer choice behavior. In: T. S. Roberston, & H. Kassarjian, *Handbook of Consumer Behavior*. Englewood Cliffs: Prentice-Hall.
- MILLER, G. A. (1956). The magical number seven, plus or minus two: some limits on own capacity for processing information. (63, p. 81-97).
- MONTGOMERY, D. C. (1996). *Design and analysis of Experiments. 4 ed.* Jonh & Wiley Sons.
- MOWEN, I. C., & MINOR, M. S. (2003). *Comportamento do Consumidor*. São Paulo: Prentice Hall.
- ORME, B. (2010). *Getting started with Conjoint Analysis: Strategies for Product Desing and Pricing research*. Madison, Wis.: Research Publishers LLC.
- PRETTO, K., & ARTES, R. (2009). Análise de Prederência Conjunta: Um estudo sobre Omissão de Atributos. *Revista Brasileira de Estatística*, 7-31, v.70.
- RAO, V. R. (2008). Developments in Conjoint Analysis. In: S. V. Berend Wierenga, *Handbook of Marketing Decision Models, Iternational Series in Operation Research & Management* (pp. 23 - 53). New York.

- SIQUEIRA, J. d. (2000). *Mensuração da estrutura de preferência do consumidor: Uma aplicação de Conjoint Analysis*. São Paulo - SP: Dissertação (Mestrado em Administração, concentração em métodos quantitativos) Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo - SP.
- WOODRUFF, R., & B. (1997). Customer value: the next source for the competitive advantage. *Journal of the Academy of Marketing Science*, v. 25, n. 2, p. 139 - 153.
- Zeithaml, V. A. (1988). Consumers perceptions of price, value, and quality: a means-end model and synthesis of evidence. *v. 52; p. 2-22*.
- ZEITHAML, V. A. (1988). Consumers perceptions of price, value, and quality: a means-end model and synthesis of evidence. *v. 52; p. 2-22*.

## Apêndice 1- *Conjoint Analysis* tradicional, baseada em notas e Postos no SPSS.

Como vimos no capítulo 4, o SPSS é capaz de gerar as respostas de *conjoint analysis* tradicional tanto por meio de notas quanto por postos.

O procedimento Conjoint do SPSS requer dois arquivos (dois bancos de dados), o arquivo de dados, onde é especificada a nota ou posto que cada consumidor atribui a cada tratamento/produto, e o arquivo de planejamento, que consiste no conjunto de perfis de produto a ser avaliado pelos sujeitos e devem ser gerados usando o procedimento de projeto ortogonal.

O planejamento ortogonal pode ser obtido partir do comando:

*Data > orthogonal design > generate*

Este comando permite a criação de um experimento, onde de maneira simples o software inclui os fatores e seus níveis, permite realizar experimentos fatoriais completos e fatoriais fracionários de maneira a gerar experimentos ortogonais, ou quase ortogonais. Permite também gerar casos para validação e simulação.

A seguir podemos ver, como ficam as bases do planejamento(ceular.plan) e dos dados(ceular.data) para o exemplo administrado em 4.1.3.

celular.plano.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

18 : sistema

	teclado	sistema	cor	STATUS	CARD	var	var
1	touch	android	branco	Design	1		
2	touch	android	preto	Design	2		
3	touch	windows	branco	Design	3		
4	touch	windows	preto	Design	4		
5	qwerty	android	branco	Design	5		
6	qwerty	android	preto	Design	6		
7	qwerty	windows	branco	Design	7		
8	qwerty	windows	preto	Design	8		
9							
10							
11							
12							
13							
14							

Figura 9- base de planejamento no spss para o exemplo do celular

celular.data.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

2 : ID

	ID	pref01	pref02	pref03	pref04	pref05	pref06	pref07	pref08	var
1	1,0	9,0	8,0	7,5	6,5	7,0	6,0	5,0	4,0	
2										
3										
4										
5										

Figura 10- base de respostas no SPSS para o exemplo do celular

	Forma	Ingredientes	Marca	STATUS_	CARD_	var
1	liquido	Sem Fosfato	HATCO	Design	1	
2	liquido	Sem Fosfato	Genérico	Design	2	
3	liquido	Com Fosfato	HATCO	Design	3	
4	liquido	Com Fosfato	Genérico	Design	4	
5	pó	Sem Fosfato	HATCO	Design	5	
6	pó	Sem Fosfato	Genérico	Design	6	
7	pó	Com Fosfato	HATCO	Design	7	
8	pó	Com Fosfato	Genérico	Design	8	
9						
10						

Figura 11- base de planejamento no SPSS para o exemplo do detergente

	id	rank01	rank02	rank03	rank04	rank05	rank06	rank07	rank08	var
1	1	1	2	5	6	3	4	7	8	
2										
3										

Figura 12- base de respostas no SPSS para o exemplo do detergente

Para rodar a *Conjoint analysis* no SPSS, é necessária programação da sintaxe para acessar o modulo CONJOINT.

O modelo geral da sintaxe para *conjoint analysis* no SPSS é;

```

CONJOINT [PLAN={*
                }]]
                {'savfile'|'dataset'}

[/DATA={*
                }]]
                {'savfile'|'dataset'}

/{SEQUENCE}=varlist
{RANK    }
{SCORE  }

```



```

[/SUBJECT=variable]

[/FACTORS=varlist['labels'] ( [{DISCRETE [{MORE}]}] )
                                {          {LESS} }
                                {LINEAR [{MORE}] }
                                {          {LESS} }
                                {IDEAL          }
                                {ANTIIDEAL      }

                                [values['labels']] )

varlist...

[/PRINT={ALL**          } [SUMMARYONLY]]
        {ANALYSIS       }
        {SIMULATION     }
        {NONE           }

[/PLOT={ [SUMMARY] [SUBJECT] [ALL]}]
        { [NONE**]          }

```

O comando começa com a introdução da base de planejamento e de dados, através dos comando PLAN e DATA. Podem ser considerados datasets abertos ou arquivos salvos em formato SPSS.

Depois é necessária a identificação de como os respondentes vão classificar os tratamentos, SEQUENCE(sequencias), RANK(postos) ou SCORES(notas). No exemplo do celular, utilizamos SCORES(notas)(figura9). Então neste caso as células irão receber as notas dos consumidores referentes a cada tratamento(variáveis pref01 à pref08), variando de 1 à 8. Se fosse o caso de postos, como no exemplo do detergente, podemos usar ou o comando SEQUENCE, ou o comando RANK, se utilizarmos o comando SEQUENCE

(figura 11) as variáveis pref01 à pref08 não representariam um tratamento em específico, mas sim o posto que um tratamento qualquer receba. Por exemplo, se o tratamento 3 foi o mais preferido, em pref01 o valor atribuído a célula referente a um consumidor em específico será 3. Se Utilizarmos o comando RANK, atribuímos as variáveis pref01 à pref08 se comportam da mesma maneira que utilizando notas, ou seja pref01 representa o estímulo1 e pref08 representa o estímulo 8, preenchendo assim o posto relacionado ao estímulo.

Em SUBJECT, iremos identificar o nome da variável na base de dados (celular.data) que identifica os respondentes, no caso em específico, ID.

O próximo passo é definir quais são os fatores. Devemos lista-los e classifica-los de acordo com a relação esperada entre os fatores e *rankings* ou pontuações. O padrão é identificar os níveis dos fatores como discretos, mas também existem as classificação linear, ideal e anti-ideal.

O subcomando PRINT permite-lhe controlar o conteúdo do output. Por exemplo, se temos um grande número de indivíduos, podemos optar por limitar a saída para os resultados resumidos apenas, omitindo saída detalhada para cada assunto. É possível também escolher se a saída incluirá a análise dos dados experimentais e os resultados para os casos de simulação incluídos no arquivo de planejamento.

O subcomando PLOT define os gráficos que irão sair no output. Assim como no PRINT, também há a opção de um resumo dos resultados ou agregando simulações e dados experimentais ou todos juntos.

‘

Para o exemplo administrado em 4.1.3, exemplo do celular, a sintaxe foi elaborada da seguinte forma:

```
CONJOINT          PLAN='C:\Users\Thiago\Documents\monografia\banco de
dados\celular\celular.plano.sav'
/ DATA='C:\Users\Thiago\Documents\monografia\banco de
dados\celular\celular.data.sav'
/ SCORES=pref01 to pref08
/ SUBJECT=ID
/ FACTORS=
teclado (DISCRETE)
sistema (DISCRETE)
cor (DISCRETE)
/ PRINT=SUMMARYONLY
/ PLOT=SUMMARY.
```

E os resultados podem ser vistos pelo output do SPSS

```
Factor   Model Levels  Label
teclado  d         2
sistema  d         2
cor      d         2
(Model: d=discrete, l=linear, i=ideal, ai=antiideal, <=less, >=more)
```

All the factors are orthogonal.

SUBFILE SUMMARY

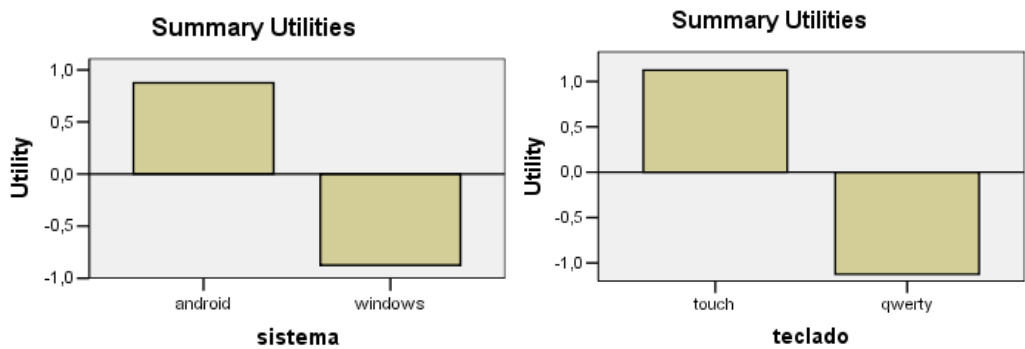
Averaged Importance	Utility	Factor
		teclado
45,00	1,1250	---- touch
	-1,1250	----  qwerty
		sistema
35,00	,8750	--- android
	-,8750	---  windows
		cor
20,00	,5000	-- branco
	-,5000	--  preto
	6,6250	CONSTANT

Pearson's R = ,997 Significance = ,0000

Kendall's tau = 1,000 Significance = ,0003

SUBFILE SUMMARY

No reversals occurred in this split file group.



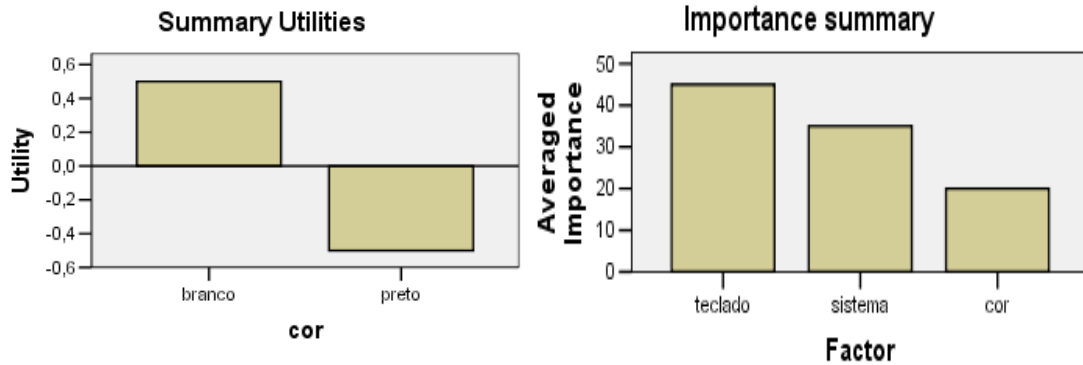


Figura 13- Coeficientes de preferência e Importância relativa para o exemplo do celular

Para o exemplo administrado em 4.2.4, exemplo do detergente, a sintaxe foi elaborada da seguinte forma:

```

CONJOINT      PLAN='C:\Users\Thiago\Documents\monografia\banco de
dados\detergnte\detergente_plan.sav'
/ DATA='C:\Users\Thiago\Documents\monografia\banco de dados\
detergnte\detergente_data.sav'
/ SEQUENCE=pref01 to pref08
/ SUBJECT=id
/ FACTORS=
Forma (DISCRETE)
Ingredientes (DISCRETE)
Marca (DISCRETE)
/ PRINT=SUMMARYONLY
/ PLOT=SUMMARY.

```

Os resultados do Output são:

	Value	Sig.
Pearson's R	1,000	.
Kendall's tau	1,000	,000

a. Correlations between observed and estimated preferences

Forma	28,571
Ingredientes	57,143
Marca	14,286

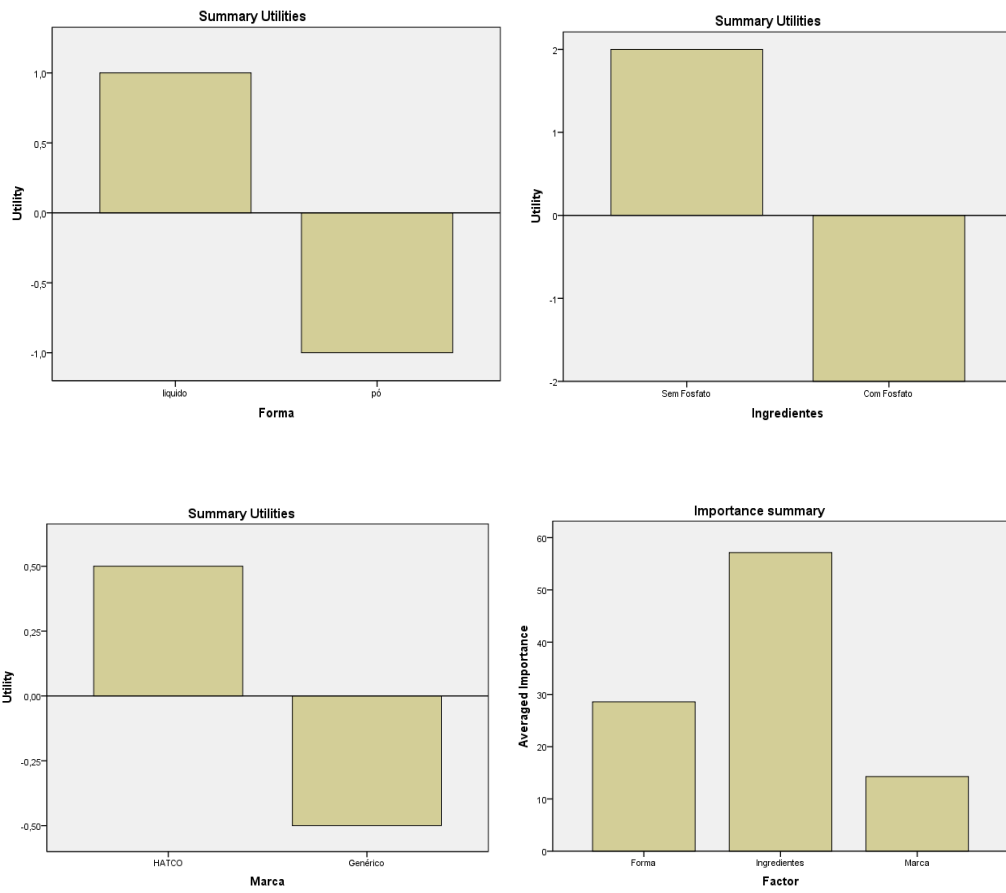


Figura 14- Coeficientes de preferência e importâncias relativas