



FUNDAÇÃO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA



ESCOLA NACIONAL DE CIÊNCIAS ESTATÍSTICAS

**TIPOLOGIA DAS FAMÍLIAS RESIDENTES NO ESTADO DO RIO DE
JANEIRO SEGUNDO A ESTRUTURA DAS DESPESAS DE CONSUMO**

por

RACHEL GLÓRIA MELLO RAGAGLIA GOMES

*Dissertação apresentada para obtenção do título de Mestre em Estudos
Populacionais e Pesquisas Sociais, Área de Concentração em
Estatística Social.*

Orientador: Prof. Dr. Mauricio Teixeira Leite de Vasconcellos

Rio de Janeiro, dezembro de 2005.

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



FUNDAÇÃO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA



ESCOLA NACIONAL DE CIÊNCIAS ESTATÍSTICAS

**TIPOLOGIA DAS FAMÍLIAS RESIDENTES NO ESTADO DO RIO DE
JANEIRO SEGUNDO A ESTRUTURA DAS DESPESAS DE CONSUMO**

por

RACHEL GLÓRIA MELLO RAGAGLIA GOMES

foi avaliada pela Banca Examinadora composta pelos seguintes membros.

Prof. Dr. Victor Hugo de Carvalho Gouvêa

Prof. Dr. Pedro Luis do Nascimento Silva

Prof. Dr. Mauricio Teixeira Leite de Vasconcellos

Dissertação defendida em 28 de dezembro de 2005.

Gomes, Rachel Glória Mello Ragaglia

Tipologia das famílias residentes no estado do Rio de Janeiro segundo a estrutura das despesas de consumo, Rio de Janeiro: Escola Nacional de Ciências Estatísticas, Fundação Instituto Brasileiro de Geografia e Estatística, 2005.

75p.,il.

Inclui bibliografia.

Dissertação (Mestrado) – Fundação Instituto Brasileiro de Geografia e Estatística, Escola Nacional de Ciências Estatísticas, 2005.

1. Consumo. 2. Orçamentos familiares. 3. Componentes principais. 4. Análise de agrupamento. 4. Amostragem complexa. I. Dissertação. II. ENCE. III. IBGE. IV. Título.

*Dedico esta dissertação aos meus pais,
Fernando e Silvia, que sempre incentivaram
meus estudos e me ofereceram condições
para que eu chegasse até aqui.*

AGRADECIMENTOS

Agradeço a Deus por todas as coisas boas que me tem acontecido e pelas pessoas amigas que colocou em meu caminho.

Agradeço ao Professor Maurício Teixeira Leite de Vasconcellos, que, além de um profissional competente, é uma pessoa amiga e dedicada.

Agradeço ao Professor Pedro Luis do Nascimento Silva pelas valiosas ajudas prestadas no decorrer dessa dissertação, dentre as quais destaco as técnicas para a análise multivariada.

Ao Professor Djalma Galvão Carneiro Pessoa envio meu muito obrigado pela ajuda e apoio no desenvolvimento desta dissertação e pelo presente, que permitiu vencer os problemas de memória enfrentados no uso da linguagem R.

Ao Professor André Wallace Nery da Costa, pela atenção e presteza com que me auxiliou em cálculos fundamentais dessa dissertação.

Ao coordenador de pesquisas do Instituto Fecomércio-RJ, o economista Paulo Brück, pela confiança que sempre depositou em meu trabalho por todos esses anos que soube compreender minhas ausências, necessárias para a realização deste mestrado.

Ao meu namorado Yuri Alexander, por todo amor, carinho, dedicação e amizade nos 10 anos que, juntos, completamos em 2005.

Aos meus irmãos Sarah Helena, André Luiz, Victor Hugo e Priscilla Esther, pelos apoios e alguns “poucos momentos” que precisei de tranquilidade.

À vovó Estelita por me acolher em sua casa nos momentos finais dessa dissertação.

Aos meus amigos do mestrado: Ana Luiza Azevedo, Flávia Faria, José Rodrigo Morais, Rita Maria Passos, e do Instituto Fecomércio-RJ: Flávia Fialho, Gabriel Fernandes e Juliana Campos, pela amizade nos momentos de ânimo e desânimo e por toda a torcida que fizeram por mim.

Enfim, agradeço a todos os amigos, colegas e profissionais cujos nomes não foram citados, mas que fazem parte da minha vida e contribuíram com sua amizade sincera e com favores prestados ao longo desse curso.

SUMÁRIO

1. Introdução	1
2. Mercado consumidor e estratégias para projeção de demanda	5
3. Objetivos	17
4. Material e métodos	
4.1 Pesquisa de Orçamentos Familiares 2002-2003 - POF	18
4.2 Análise Multivariada.....	25
4.2.1 Análise de componentes principais – ACP	25
4.2.2 ACP para amostras complexas.....	29
4.2.3 Análise de agrupamentos não hierárquica - o método das k-médias	30
4.3 Descrição dos procedimentos usados	32
4.3.1 Leitura dos dados e criação de arquivo com todos os tipos de despesas .	32
4.3.2 Definição dos grupos e criação das variáveis de valor da despesa	33
4.3.3 Aplicação dos métodos de análise multivariada	35
5. Resultados e discussão	
5.1. Resultados sobre os grupos de produtos	38
5.2. Resultados do método ingênuo	41
5.3. Resultados do método com desenho	47
5.4. Resultados comparativos	58
6. Conclusões	61
Referências Bibliográficas	63
ANEXOS	
I Relação das variáveis da POF	67
II Criação dos códigos de despesa em registros com mais de uma variável de valor	72
III Porcentagem média de cada grupo de despesa em cada tipo de família – Método ingênuo.....	74
IV Porcentagem média de cada grupo de despesa em cada tipo de família – Método com desenho.....	75

ÍNDICE DE FIGURAS

Figura 2.A: Representação gráfica da distribuição do PIB por região no estado do Rio de Janeiro – 2003	7
Figura 2.B: Percentual de estabelecimentos de comércio de bens e serviços do estado do Rio de Janeiro segundo o porte - 1992 a 2002	10
Figura 2.C: Curva de Engel log-log-inversa.....	12
Figura 4.A: Esquema de amostragem da POF: Conglomerado em dois estágios.....	20
Figura 4.B: Representação da despesa monetária total.....	22
Figura 5.A: Scree-plot – Método ingênuo	41
Figura 5.B: Scree-plot – Método com desenho	47

ÍNDICE DE QUADROS

Quadro 4.A: Tipos de registros coletados pela POF.....	23
---	----

ÍNDICE DE TABELAS

Tabela 2.A: Participação percentual do produto interno bruto (PIB), segundo as regiões de governo do estado do Rio de Janeiro – 2003.....	7
Tabela 2.B: Participação da população e arrecadação de ICMS segundo as regiões de governo do estado do Rio de Janeiro – 2002	7
Tabela 2.C: Participação percentual do PIB segundo setores – Estado do Rio de Janeiro – 2003.....	8
Tabela 2.D: Número de empregados segundo setores – Estado do Rio de Janeiro –2003.....	9
Tabela 2.E: Número de estabelecimentos de comércio de bens e serviços segundo as regiões de governo do estado do Rio de Janeiro - 2000 a 2002.....	10
Tabela 2.F: Número de estabelecimentos de comércio de bens e serviços do estado segundo o porte - 1992 a 2002	10
Tabela 5.A: Grupos de despesas de consumo e setores de comércio ou serviço associados	38
Tabela 5.B: Estimativa dos valores das despesas (monetária e total) e respectivos percentuais, segundo os grupos de despesas, no estado do Rio de Janeiro ..	40
Tabela 5.C: Autovalores e porcentagem de explicação da variação total, segundo os componentes – Método ingênuo	41
Tabela 5.D: Informações sobre os grupos considerados estáveis e não-estáveis após três execuções consecutivas do Procedimento Fastclus do SAS – Método ingênuo.....	42
Tabela 5.E: Descrição dos tipos de família – Método ingênuo	43
Tabela 5.F: Informações adicionais segundo os tipos de famílias – Método ingênuo	44

Tabela 5.G: Autovalores e porcentagem de explicação da variação total, segundo os componentes – Método com desenho.....	47
Tabela 5.H: Informações sobre os grupos considerados estáveis e não-estáveis após três execuções consecutivas do procedimento Fastclus do SAS – Método com desenho.....	48
Tabela 5.I: Descrição dos tipos de família – Método com desenho.....	49
Tabela 5.J: Informações adicionais segundo os tipos de famílias – Método com desenho.....	50
Tabela 5.K: Comparação dos autovalores e da porcentagem de explicação da variabilidade total entre os dois métodos.....	58
Tabela 5.L: Distribuição do número de domicílios segundo as tipologias elaboradas pelos dois métodos – dados amostrais.....	59
Tabela 5.M: Estimativa do número de domicílios segundo as tipologias elaboradas pelos dois métodos.....	60

RESUMO

Esta dissertação nasceu no contexto da discussão sobre a interiorização do comércio no estado do Rio de Janeiro, da qual a autora participa como funcionária do Instituto Fecomércio-RJ (Federação do Comércio do Estado do Rio de Janeiro), motivada pelas profundas desigualdades observadas entre a capital e o interior do estado. Essa discussão requer informações sobre a demanda das famílias por produtos e serviços nos diversos municípios do estado para apoiar programas e ações. No entanto, a projeção de demanda, requerida por essa discussão, implica o uso de sistemas de equações de demanda que devem ter a propriedade aditiva, para que a soma dos valores projetados por produto ou serviço seja igual à projeção para o total.

No entanto, as curvas de Engel com propriedade aditiva não são as que melhor se ajustam aos dados, gerando um conflito estatístico: manter a propriedade aditiva errando na projeção por produto ou estimar bem a demanda por produto e errar no total de renda disponível para consumo? Dentre as soluções encontradas na literatura científica, a construção de uma tipologia de famílias e a projeção da demanda para cada tipo de família da tipologia construída apresenta-se como uma das mais promissoras, visto que esta aproximação conduziria a grupos mais homogêneos e com menor variação do nível de renda, permitindo um melhor ajuste aos dados das funções de Engel com propriedade aditiva.

A construção de tipologias de famílias requer análise de agrupamentos, geralmente precedida pela análise de componentes principais para construção de uma base de fatores ortogonais e para a redução do número de dimensões estudadas. No entanto, com dados provenientes de uma amostra complexa, que inclui, em seu desenho, estratificação e conglomeração, a análise de componentes principais não pode ser feita pelas rotinas disponíveis nos sistemas estatísticos usuais, visto que a estimativa da matriz de variância seria enviesada pelo pressuposto de independência e igual distribuição contido nas rotinas programadas. Uma solução proposta na literatura consiste em estimar corretamente a matriz de variância e usá-la como entrada para as rotinas usuais de componentes principais.

Seguindo essas recomendações, esta dissertação buscou a construção de tipos de famílias, definidos em função da estrutura de gastos de consumo, por meio das técnicas referidas, a partir dos dados da Pesquisa de Orçamentos Familiares 2002-2003 (POF), realizada pelo IBGE.

Para mostrar a importância de utilizar as informações estruturais de um desenho de amostra complexo, duas tipologias foram elaboradas e seus resultados comparados. Uma das tipologias foi construída pressupondo, ingenuamente, que os dados da POF eram independentes e igualmente distribuídos, enquanto a outra considerou o plano de amostragem utilizado na POF.

Os resultados indicaram a existência de alterações importantes nas tipologias construídas, indicando que a estimação correta da métrica utilizada para medir as distâncias entre os pontos (ou famílias) é um aspecto fundamental do método utilizado.

Palavras-chave: Consumo, Orçamentos familiares, Componentes principais, Análise de agrupamento, Amostragem complexa.

Abstract

This dissertation is inserted in the context of the discussion about plans and actions to reduce the inequalities of the commerce volume between the Capital and the interior of the state of Rio de Janeiro of which the author participates as an employee of the Institute of the Commerce Federation of Rio de Janeiro. This discussion requires information on family demand for goods and services at the different municipalities of the state to support plans and actions. Nevertheless, the demand projection needed implies the use of a system of equations with additive property, since the sum of projected values by product or service must equal the projection of the total demand.

Since additive Engel curves are not the ones that better fit the data, a statistical conflict appears: keep the additive property with errors in the projected value for each product and service or have a good demand estimate for each product and service making errors in the projected value for total demand. The more interesting solution found in scientific literature is to create a family typology based on expenditure structure and to fit Engel curves for each family type, since within each type the families are more homogeneous and have smaller income range than in the whole population, allowing a better fit of additive Engel curves.

The creation of family typologies requires a cluster analysis, generally preceded by a principal component analysis to establish an orthogonal base and to reduce the number of dimensions analyzed. But principal component analysis based on data from a complex survey, designed with clustering and stratification, cannot be performed by procedures available in the usual statistical systems, since the estimated variance matrix will be biased by the assumptions of independence and equal distribution behind those procedures. A proposed solution is to obtain an unbiased estimate of the variance matrix and use it as input to the principal component analysis.

Following these recommendations, this dissertation aimed at the definition of a family typology according the structure of family expenditure using the referred technique and data from the Family Budget Survey 2002-2003 (POF), conducted by the Brazilian Institute of Geography and Statistics.

To show the importance of using structural information from a complex sample design two family typologies were built and their results compared. The first was built assuming the data from POF were independent and equally distributed, while the other considered the real sample design used in the survey.

Results showed the existence of important differences between the two family typologies, indicating that the use of the correct metric to measure the distance between the points (or families) is a relevant issue in the method used.

Key words: Consumption, Family budgets, Principal Components, Cluster analysis, Complex Sampling.

Capítulo 1

Introdução

Esta dissertação nasceu no contexto da discussão sobre a interiorização do comércio no estado do Rio de Janeiro, da qual a autora participa por trabalhar no Instituto Fecomércio-RJ (Federação do Comércio do Estado do Rio de Janeiro). Esta discussão é motivada pelas profundas desigualdades observadas entre a capital e o interior do estado e, sem dúvida, requer informações sobre a demanda das famílias nos diversos municípios do estado, visto que o comércio não pode ser interiorizado se não houver demanda e esta, se existe, acaba sendo atendida pela capital ou por municípios maiores.

Encontra-se na literatura científica uma discussão que vem de muito tempo sobre a estimação de sistemas de curvas de Engel (Prais&Houthaker, 1955), visto que a demanda não pode ser projetada isoladamente por produto ou serviço, uma vez que a renda familiar é limitada e a soma dos valores projetados não pode superá-la. Assim, o valor total da demanda corresponde à renda familiar, que é gasta em função de decisões familiares motivadas por necessidades, escolhas e julgamentos de valores desconhecidos. Portanto, a projeção de demanda, requerida pela discussão referida, implica o uso de sistemas de equações de demanda que devem ter a propriedade aditiva, para que a soma dos valores projetados por produto ou serviço seja igual ao total projetado. No entanto, as curvas com propriedade aditiva não são as que melhor se ajustam aos dados, gerando um conflito estatístico: manter a propriedade aditiva errando na projeção por produto ou estimar bem a demanda por produto e errar no total de renda disponível para consumo?

Prais & Houthaker (1955) sugerem estratégias de agregação dos dados para reduzir o problema. François (1977), em curso de projeção da demanda alimentar, ministrado na Universidade Montpellier I para alunos de língua portuguesa, mostrou que a projeção da demanda alimentar pressupunha que as equações do sistema fossem aditivas porque um dos

objetivos da projeção era atender aos requerimentos nutricionais, considerando que erros para mais nas estimativas por produto poderiam causar sobrepeso e obesidade, e que erros para menos poderiam causar subnutrição e todas as doenças derivadas da carência de nutrientes.

Na realidade, o autor reflete o pensamento de Marcel Autret, farmacêutico francês e ex-diretor do Departamento de Nutrição da *Food and Agriculture Organization of the United Nations* (FAO), que havia realizado cursos de formação de nutricionistas na Universidade de Marselha, em 1952 e 1955 (Autret & Ganzin, 1952). Este pensamento foi relato por Julian Chacel, quando este era Diretor de Pesquisas do Instituto Brasileiro de Economia da Fundação Getulio Vargas:

“Como técnica, o planejamento envolve, agora, conhecimentos muito além do âmbito do economista e do estatístico. Assim, por exemplo, ao determinar objetivos da produção agrícola para fins alimentares, é indispensável a colaboração de médicos e bioquímicos, especialistas da nutrição. Estes é que podem dizer da coerência das projeções e do acerto da política alimentar, implícita nos planos de expansão da produção agrícola. E ainda, no planejamento, estimadas as necessidades e avaliados os recursos, passa-se à etapa das projeções para futuro, próximo ou remoto, de acordo com o horizonte do tempo adotado. Só então é que se busca, numa etapa final, definir uma estratégia que tornem compatíveis necessidades e recursos, em função dos rumos de política adotados. Na essência, a idéia da abordagem nutricional no planejamento agrícola é a de fornecer um teste de coerência para diminuir os riscos inerentes ao método das projeções. Sem essa precaução de verificar a coerência das metas de produção, que levem à melhoria dos padrões nutricionais, os objetivos fixados podem resultar em monumental erro fisiológico e econômico, na expressão de Autret. É sob esse aspecto que a nutrição torna-se uma técnica operacional para o planejamento econômico.”(FGV, 1975, v.I:2).

Em suas sugestões para a busca de um sistema de equações de demanda aditivas, François (1977) propõe a construção de uma tipologia de famílias e a projeção da demanda para cada tipo de família da tipologia construída. Segundo ele, esta aproximação conduziria a grupos mais homogêneos, com menor amplitude dos valores de renda que os de toda a população, permitindo um melhor ajuste aos dados das funções de Engel com propriedade

aditiva. Para tanto, indicou que a construção de tipologias de famílias deveria ser feita por meio de uma análise de agrupamentos, precedida pela análise de componentes principais para a construção de uma base de fatores ortogonais e a redução das dimensões estudadas, trabalhando-se, apenas, com os fatores de maior proporção de explicação da variância total.

Como os dados usados nesta dissertação provêm da Pesquisa de Orçamentos Familiares 2002-2003 (POF) realizada pelo IBGE, cuja amostra inclui em seu desenho estratificação e conglomeração, sendo, portanto, considerada uma amostra complexa, a análise de componentes principais não pode ser feita pelas rotinas disponíveis nos sistemas estatísticos usuais, visto que a estimativa da matriz de variância e covariância seria enviesada pelo pressuposto de independência e igual distribuição contido nas rotinas usuais de componentes principais (Pessoa & Silva, 1998). Em Smith & Holmes (1989) há uma proposta de se estimar corretamente a matriz de variância e covariância e usá-la como entrada para as rotinas usuais de análise de componentes principais.

Dentro deste contexto, esta dissertação visa à discussão de estratégias para a construção de tipos de famílias definidos em função da estrutura de despesas de consumo em produtos e serviços típicos dos segmentos de mercado de interesse da Fecomércio-RJ e de outros setores do mercado.

Duas ferramentas de análise multivariada foram utilizadas: análise de componentes principais (ACP) e análise de agrupamentos. A partir do detalhamento dos grupos de despesas de consumo, previamente definidos na POF pelo IBGE, outros grupos foram considerados levando em conta os tipos de atividades específicas desenvolvidas pelos estabelecimentos do ramo de comércio de bens e serviços considerados, atualmente, de razoável representatividade. Com os novos grupos, a ACP foi aplicada como técnica de análise intermediária, visando apenas resumir os dados e fornecer uma base para a criação de um novo conjunto de variáveis

de maior impacto analítico que foi, posteriormente, utilizado na análise de agrupamentos, diferenciando, de maneira mais efetiva, os padrões de consumo existentes no estado.

Os procedimentos multivariados foram aplicados em duas situações distintas. A primeira supôs que os dados da POF foram selecionados por amostragem aleatória simples e que as observações eram realizações de variáveis aleatórias independentes e identicamente distribuídas. Na segunda, o caráter complexo da amostra foi reconhecido e as técnicas foram aplicadas utilizando as informações estruturais do desenho da amostra da POF e os pesos das observações para os cálculos das estimativas.

Esta dissertação foi estruturada em seis capítulos, sendo o primeiro esta introdução.

O capítulo 2 apresenta a dados sobre a concentração econômica do estado fluminense, revelando suas raízes históricas e evidenciando a importância do conhecimento do mercado consumidor como forma de nortear a tomada de decisões por parte dos empresários. Aborda, também, as estratégias encontradas na literatura científica para a projeção da demanda.

No capítulo 3 são descritos os objetivos desta dissertação.

O capítulo 4 trata do material e dos métodos utilizados para atender aos objetivos definidos. Descreve aspectos da Pesquisa de Orçamentos Familiares 2002-2003, a fonte de dados deste trabalho, apresenta os métodos de análise multivariada usados e os procedimentos estabelecidos para a obtenção dos resultados.

No capítulo 5, os principais resultados obtidos são apresentados e discutidos.

Por fim, o capítulo 6 aborda as principais conclusões e sugestões de trabalhos futuros que não puderam ser incluídos nesta dissertação. As referências bibliográficas e alguns anexos complementam esta dissertação.

CAPÍTULO 2

Mercado consumidor e estratégias para projeção de demanda

A forte concentração das atividades econômicas na Região Metropolitana do Rio de Janeiro (RMRJ), identificada na estrutura estadual fluminense, é uma situação bastante antiga, com origem na condição de capital imperial e republicana exercida pela cidade do Rio de Janeiro por quase dois séculos, visto que a infraestrutura instalada e a qualificada mão-de-obra existente, assim como a presença de um grande número de trabalhadores com altos e estáveis salários formavam um excelente mercado consumidor (IPEA/IBGE/NESUR/SEADE, 2001). Este mercado conferiu à cidade grandes vantagens locais e, portanto, questões sobre crescimento econômico não eram fontes de preocupação das elites, mais envolvidas com assuntos de caráter nacional, enquanto o interior do estado participava de duas grandes culturas agrícolas: a cafeeira e a açucareira.

Basicamente pela falta de implantação de novas técnicas de cultivo do café, após o século XIX, e pela não-diversificação da base produtiva e falta de modernização do parque industrial do cultivo da cana-de-açúcar, a economia do interior entrou em decadência. O mesmo estudo indica que a escolha da localização da Companhia Siderúrgica Nacional (CSN), em Volta Redonda, região do Médio Paraíba; da Refinaria de Petróleo Duque de Caxias, na região da Baixada Fluminense; e da Companhia Nacional de Álcalis, na região das Baixadas Litorâneas, foram ações diretas do governo federal com vistas ao desenvolvimento do interior do estado.

Santos (2003), em livro dedicado à compreensão da atual dinâmica econômica do estado do Rio de Janeiro, aponta que a situação atual é um reflexo tanto dos efeitos do fenômeno de globalização que atingiu, de maneira geral, a todas as economias, quanto das questões históricas que são particulares do estado fluminense.

Perdendo em 1960, a condição de capital federal para Brasília, a cidade do Rio de Janeiro, ao invés de ser unificada com o resto do estado do Rio, passou a constituir uma nova unidade da Federação, o estado da Guanabara, mantendo a separação entre núcleo e periferia e trazendo conseqüências negativas para a distribuição dos investimentos públicos, concentrando-os no núcleo da região metropolitana, enquanto que sua periferia sofria com a falta de infraestrutura em um contexto de grande crescimento demográfico. Nem mesmo a ampliação da refinaria de Duque de Caxias, a modernização do porto do Rio de Janeiro, a construção das usinas nucleares em Angra dos Reis e os incentivos à indústria naval, foram suficientes para manter a participação fluminense na produção industrial nacional (Santos, 2003).

Na tentativa de corrigir o erro histórico, o novo estado do Rio de Janeiro é criado em 1975, no governo do Almirante Faria Lima, pelo projeto de fusão dos estados da Guanabara e do Rio de Janeiro. Santos (2003) definiu o projeto como a junção de duas unidades que, apesar de próximas geograficamente, guardavam – e ainda guardam – profundas diferenças entre si, concluindo, também, que a ação, buscando reverter o quadro de crise da economia, não surtiu os efeitos esperados e hoje, o estado se consolida como unidade da federação cuja Região Metropolitana é responsável por quase toda a economia estadual.

De fato, estima-se que quase 65% do PIB do estado seja gerado na capital estadual e na sua periferia (o restante da RMRJ), que concentram, aproximadamente, 70% da população e são responsáveis por aproximadamente 91% da arrecadação do ICMS, dados que confirmam as disparidades existentes entre a RMRJ e o interior do estado (Tabelas 2.A e 2.B).

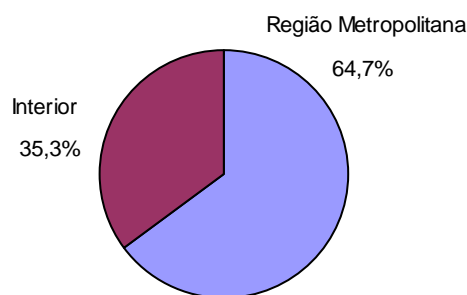
“Superar a excessiva polarização presente na economia do estado do Rio de Janeiro é uma premissa importante para dinamizar sua economia” no ver de Santos (2003:72). Partindo dessa conclusão, derivada dos estudos que realizou, o autor ressaltou o importante papel que as cidades médias fluminenses (população entre 100 mil e 500 mil habitantes, que entre os censos de 1980 e o de 2000, passaram de 10 para 17, com 10 localizadas no interior) poderiam

desempenhar nesse processo, assumindo posições de centros regionais de desenvolvimento rumo ao interior do estado, dando suporte à estrutura social em rede.

Tabela 2.A: Participação percentual do produto interno bruto (PIB), segundo as regiões de governo do estado do Rio de Janeiro - 2003.

Região	PIB (1.000 R\$)	Participação percentual
ESTADO	209.213.788	100,0%
Metropolitana	135.271.048	64,7%
Interior Fluminense	73.942.740	35,3%
Norte Fluminense*	44.023.159	21,0%
Médio Paraíba	14.793.181	7,1%
Serrana	5.271.526	2,5%
Baixadas Litorâneas	3.880.695	1,9%
Costa Verde	3.044.805	1,5%
Sul-Fluminense	1.535.395	0,7%
Noroeste Fluminense	1.393.979	0,7%

Figura 2.A: Representação gráfica da distribuição do PIB por região no estado do Rio de Janeiro – 2003.



Fonte: Fundação CIDE
*Inclui a Baía de Campos

Tabela 2.B: Participação da população e arrecadação de ICMS segundo as regiões de governo do estado do Rio de Janeiro – 2002.

Regiões do Estado	População estimada ⁽¹⁾	Participação percentual	Arrecadação ICMS (R\$)	Participação percentual
Estado	14.768.969	100,0%	7.716.601.223	100,0%
Metropolitana	10.954.050	74,2%	7.008.026.229	90,8%
Noroeste	803.571	5,4%	14.640.208	0,2%
Norte	719.582	4,9%	83.774.972	1,1%
Serrana	767.560	5,2%	162.071.226	2,1%
Baixada Litorânea	685.290	4,6%	27.350.063	0,4%
Médio Paraíba	306.881	2,1%	382.335.771	5,0%
Centro-Sul	260.222	1,8%	19.427.472	0,3%
Costa Verde	271.813	1,8%	18.975.282	0,2%

Fonte: Elaboração da Fecomércio-RJ

⁽¹⁾ População residente estimada para 2002

O aumento do número de municípios de porte médio torna mais viável a possibilidade de interiorização das atividades econômicas, uma vez que estas localidades podem atender a um mercado consumidor de tamanho considerável, atraindo, dessa forma, investimentos industriais que buscam por custos menores do que os que são despendidos nos grandes centros urbanos (Santos, 2003).

O resgate, revitalização ou motivação de atividades econômicas, particularmente o comércio de bens e serviços, tem sido tema discutido entre entidades do ramo, dentre as quais se destaca a Federação do Comércio do Estado do Rio de Janeiro (Fecomércio-RJ), que através de diversos encontros, entre outras ações, vem revelando os anseios daqueles a quem representa e contribuindo, de certa forma, com propostas para o processo de redução dos desequilíbrios regionais e atendimento a interesses empresariais.

A idéia de que comércio traz desenvolvimento é inspirada em fatos que remontam às grandes navegações, quando continentes foram descobertos e a demanda de mercado de novos produtos urgia em ser atendida (Santos, 2001). Os navegadores-comerciantes abriram novos caminhos no País e a figura do caixeiro-viajante, cuja profissão foi praticada através do comércio de mercadorias, atravessou os séculos e hoje, lojas de armário e pequenos armazéns convivem com grandes lojas de departamento e luxuosos centros de compras, responsáveis por milhares de empregos formais. É possível observar o quanto o ramo de comércio de bens e serviços representa no PIB do estado do Rio de Janeiro: uma fatia significativa de quase 30% (Tabela 2.C).

Tabela 2.C: Participação percentual do PIB segundo setores – Estado do Rio de Janeiro - 2003.

Setor	PIB (1.000 R\$)	Participação percentual
ESTADO	209.213.788	100,00%
Comércio de bens e serviços	58.362.320	27,9%
Extração de petróleo	37.399.432	17,9%
Indústria de transformação	34.339.776	16,4%
Aluguéis	19.289.428	9,2%
Administração pública	16.291.933	7,8%
Construção civil	11.200.654	5,4%
Transportes	9.669.046	4,6%
Serviços industriais de utilidade pública	8.444.755	4,0%
Instituições financeiras	7.912.261	3,8%
Comunicações	5.229.668	2,5%
Agropecuária	938.325	0,4%
Extração de outros minerais	136.190	0,1%

Fonte: Fundação CIDE

O ramo de comércio de bens e serviços é, também, a maior fonte geradora de empregos, empregando, aproximadamente, 66,20% dos trabalhadores com emprego formal, segundo dados da Relação Anual de Informações Sociais - RAIS 2003, acentuando a importância do setor frente à economia do estado (Tabela 2.D).

Tabela 2.D: Número de empregados segundo setores – Estado do Rio de Janeiro - 2003.

Setor	Número de trabalhadores	Participação percentual
E S T A D O	4.029.962	100,0%
Comércio de Bens e Serviços	2.667.996	66,2%
Administração pública	632.281	15,7%
Indústria de Transformação	416.689	10,3%
Construção Civil	202.283	5,0%
Serviços industriais de utilidade pública	48.190	1,2%
Agropecuária	37.319	0,9%
Extrativa mineral	25.204	0,6%

Fonte: Fundação CIDE

A Tabela 2.E indica, em uma análise da evolução do número de estabelecimentos comerciais existentes, segundo as regiões do estado, o movimento crescente em regiões como a Norte (crescimento de 6,94%, de 2001 para 2002) e a das Baixadas Litorâneas (7,42%). Tais variações podem ser consideradas, de certa forma, como conseqüências de importantes fenômenos ocorridos em ambas as regiões. A região Norte sendo favorecida pela presença da Petrobras e dos efeitos decorrentes da exploração de petróleo, consolidando Macaé, e também Campos, “*como importantes elos na cadeia de fluxos no âmbito da economia fluminense*” (Santos, 2003 p.107). E a exploração de atividades turísticas nas Baixadas Litorâneas, especialmente em Cabo Frio e Araruama, tem desempenhado papel importante na dinamização de sua economia. Já a região do Médio Paraíba, que responde pelo terceiro maior percentual de crescimento de estabelecimentos (2,76%), foi beneficiada pelo complexo metal-mecânico fluminense e foi fortalecida pela instalação de duas montadoras de veículos: a Volkswagen, em Resende, e a Peugeot-Citroën, em Porto Real.

Tabela 2.E: Número de estabelecimentos de comércio de bens e serviços segundo as regiões de governo do estado do Rio de Janeiro - 2000 a 2002.

Regiões de Governo	Estabelecimentos de Comércio de Bens e Serviços ⁽¹⁾				
	2000	2001	Varição 2001 / 2000	2002	Varição 2002 / 2001
Estado	175 704	179 141	2,0%	183 816	2,6%
Região das Baixadas Litorâneas	7 750	8 237	6,3%	8 848	7,4%
Região Norte Fluminense	6 921	7 180	3,7%	7 678	6,9%
Região do Médio Paraíba	8 992	9 282	3,2%	9 538	2,8%
Região Metropolitana	131 923	133 692	1,3%	136 648	2,2%
Região Serrana	11 421	11 680	2,3%	11 917	2,0%
Região Centro-Sul Fluminense	2 888	2 923	1,2%	2 978	1,9%
Região Noroeste Fluminense	2 948	3 084	4,6%	3 122	1,2%
Região da Costa Verde	2 861	3 063	7,1%	3 087	0,8%

Fonte: Fundação CIDE

(1) Exclusive administração pública (Nota citada pela Fonte).

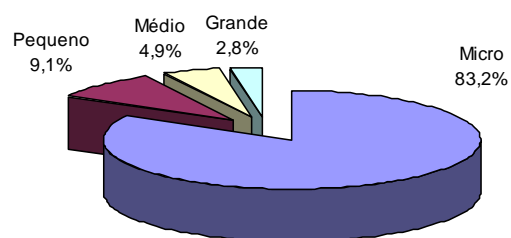
De acordo com dados da Fecomércio-RJ, mais da metade dos empreendimentos de porte micro e pequeno não chega a completar três anos de existência e, aproximadamente, 30% encerram suas atividades antes de um ano de instalação. É evidente que existem diversos fatores responsáveis por estes elevados índices como, por exemplo, a dificuldade de acesso ao crédito, exigências burocráticas e elevada carga tributária, porém, quando conhecimentos sobre as condições de negócio e demanda esperada são ignorados, os riscos de insucesso aumentam. Os micro e pequenos estabelecimentos têm uma representatividade bastante elevada no total de estabelecimentos do estado (83,19%), como pode ser observado, através da Tabela 2.F e da Figura 2.B.

Tabela 2.F: Número de estabelecimentos de comércio de bens e serviços do estado segundo o porte - 2003.

Porte do estabelecimento do Comércio de Bens e Serviços		Número de estabelecimentos	Participação Percentual
Estado		176 981	100,0%
Micro	Até 9 empregados	147 222	83,2%
Pequeno	De 10 a 19 empregados	16 171	9,1%
Médio	De 20 a 49 empregados	8 672	4,9%
Grande	Acima de 50 empregados	4 916	2,8%

Fonte: Dados da Fundação CIDE e classificação de porte do Sebrae-RJ

Figura 2.B: Percentual de estabelecimentos de comércio de bens e serviços do estado do Rio de Janeiro segundo o porte -2003.



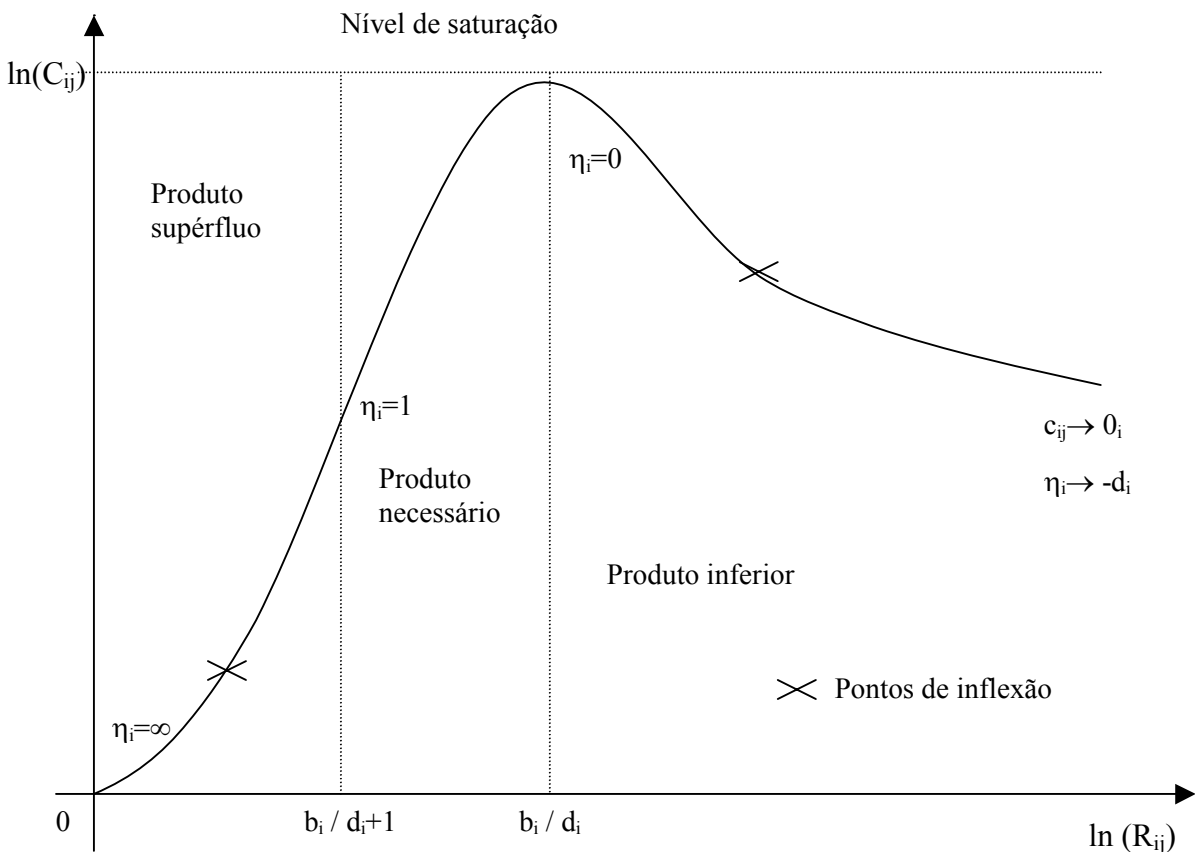
O conhecimento do comportamento do mercado consumidor, assim como das atividades comerciais já existentes no local cria vantagens competitivas para as empresas, possibilitando a comparação de seus desempenhos com o do setor a que pertencem (IPDC, 2004). As informações contidas na Pesquisa de Orçamentos Familiares 2002-2003 (POF) tornam possível o desenvolvimento de estudos úteis na definição de estratégias de investimentos em que o conhecimento do perfil do consumidor e da demanda por bens e serviços seja determinante, constituindo uma de suas múltiplas aplicações. A possibilidade de cruzamento da composição dos gastos com características socioeconômicas das famílias é uma contribuição do IBGE para os que buscam uma visão mais realista das possibilidades de investimentos, já que, por exemplo, o fato de uma região possuir uma vasta população não indica, diretamente, que seus integrantes pertençam à sociedade mercantil. É importante lembrar, por exemplo, a forte relação existente entre consumo e renda, idéia amplamente aceita pela teoria econômica (Neto,1990).

Um mesmo bem ou serviço disponível em lugares ou circunstâncias diversas pode ser valorizado de maneira muito diferente pelo consumidor, conforme a situação. Qualquer projeto ou pesquisa, portanto, visando ao investimento em atividades produtivas deve respeitar e privilegiar as vocações e características locais e de sua demanda por consumo. Esta idéia foi bem definida por Aranha (1998), quando declarou que o volume de vendas de um determinado produto ou serviço em uma região é alterado em consonância com as modificações da economia e da sociedade. Ou seja, depende da renda corrente da população, do padrão de distribuição de renda entre as classes sociais, da sua riqueza, do estágio da economia no ciclo econômico e do nível de desenvolvimento da região. O volume de vendas é, também, influenciado por aspectos culturais, como costumes e valores, e reflete a maneira como o mercado estruturou-se em termos de número e tamanho relativo das empresas participantes, variedade e abundância da oferta, tipo de distribuição, gastos em propaganda, nível de preços,

existência de produtos substitutos, nível de competição, etc. (Aranha, 1998). Neste contexto, o conhecimento da demanda real e potencial por produtos e serviços surgem como alternativa técnica mais viável para a descoberta de novos mercados e para a redução dos riscos do negócio.

Para este fim, o ajuste de curvas de Engel tem sido apontado na literatura como uma solução técnica adequada para conhecimento da demanda em função da renda dos consumidores. Estas curvas são devidas ao estatístico alemão Ernest Engel (1821-1896), cuja lei mais conhecida estabelece que “quanto menor for a renda familiar maior será a proporção gasta com alimentos”, e representam uma denominação genérica para as curvas (ou modelos) que retratam a relação entre as quantidades que um consumidor deseja comprar de um produto em diferentes níveis de renda. Seus principais aspectos são descritos a partir da Figura 2.C.

Figura 2.C: Curva de Engel log-log-inversa



A expressão algébrica da curva na Figura 2.C pode ser escrita da seguinte forma:

$$\ln(C_{ij}) = a_i + \frac{b_i}{R_j} - d_i \ln(R_j) + \varepsilon_{ij},$$
 onde C refere-se ao nível de consumo; R é o valor da

renda; a, b e d são e os parâmetros do modelo; e i e j são os índices do produto e da família, respectivamente.

Os pontos em que o coeficiente de elasticidade de renda, representado por $\eta_i = \frac{b_i}{R_j} - d_i$, atinge os valores de corte 0 e 1 estão apresentados na Figura 2.C

A curva de Engel log-log inversa descreve bem, por exemplo, o consumo de pão francês, um bem supérfluo (elasticidade de renda maior que 1, ou seja, $\eta_i > 1$) para rendas muito baixas. Com pequeno acréscimo de renda, o pão francês passa a ser um bem necessário ($0 \leq \eta_i \leq 1$) e, para os valores maiores de da renda, transforma-se em um bem inferior ($\eta_i \leq 0$), em função de sua substituição por outros tipos de pão.

No entanto, o uso de sistemas de curvas de Engel e sua aplicação para mercados locais tem sido grandes desafios para a econometria e para a estatística (François, 1977). O desafio para a econometria reside na coerência interna de um sistema de funções de demanda que deve garantir que a soma dos valores projetados por serviço e produto coincida com o valor projetado para o total da demanda. Em suas aplicações para funções de consumo de produtos alimentares, o autor justifica esta necessidade a partir da constatação de que a quantidade de cada produto alimentar ingerido deve coincidir com metas globais de nutrientes, definidas para satisfazer aos requerimentos nutricionais, sem subestimá-las para não causar desnutrição, nem superestimá-las para não provocar obesidade (apenas no caso da energia, visto que os demais nutrientes, quando ingeridos em quantidades superiores aos requerimentos, são expelidos pelo organismo através da urina e das fezes).

No entanto, a economia apresenta um outro argumento para o mesmo objetivo de coerência do sistema de equações: o valor de demanda de todos os produtos e serviços em uma região depende do volume de renda de que as famílias (da região ou consumidoras eventuais) dispõem (Aranha, 1998).

A solução técnica para este problema está na utilização de sistemas de equações de demanda que garantam a coerência das parcelas com o total e, a forma mais simples de garantir esta coerência é usando equações que tenham a propriedade aditiva (François, 1977).

O desafio para a estatística, apontado por François (1977), refere-se à necessidade de dados locais sobre os padrões de consumo das famílias e sobre sua renda, que geralmente são obtidos em pesquisas sobre orçamentos familiares, cujo custo obriga os órgãos de estatística a serem bem conservadores na definição dos estratos geográficos. A solução técnica neste caso seria a aplicação de técnicas de estimação em pequenas áreas, apoiadas por informações obtidas em censos de população, combinadas com dados locais obtidos pelas entidades de classes empresariais, por agregação de informações fornecidas pelos estabelecimentos locais de comércio e de serviços.

Na proposta de François (1977), a busca de um sistema aditivo de equações de demanda seria facilitada pela determinação de tipos de famílias, estabelecidos em função da estrutura de consumo observada em estratos geográficos mais abrangentes, por meio de dados coletados em pesquisas sobre orçamentos familiares, aplicando-se as funções de demanda de forma independente em cada grupo ou tipo de família. O autor ressalta que, em cada tipo de família, haveria uma redução do número de produtos para projeção de demanda, associado à redução esperada do intervalo de variação da renda, exceto nos tipos compostos pelas famílias mais ricas, em que, por um lado, não haveria grandes simplificações, e, por outro, o erro cometido no total seria menos relevante dado o reduzido número de famílias ricas que realmente afetam o consumo dos principais produtos e serviços.

François (1977) indica, ainda, que outra vantagem desta estratégia seria a identificação de características socioeconômicas das famílias de cada tipo, que seriam importantes para ajudar na estimação para as pequenas áreas.

Observa-se, pelo exposto, que o autor concebeu um sistema de informações abrangente, envolvendo fontes diversas (algumas, possivelmente, não-disponíveis) e uma gama considerável de desafios para a estatística.

A construção de um sistema de informações, tal como o concebido por François, transcende o objeto de uma dissertação de mestrado, devendo ser enfrentado dentro de uma linha de pesquisa mais ampla. De fato, o problema até aqui descrito é um desafio da Fecomércio-RJ, que não pode ser resolvido no prazo de um programa de Mestrado. No entanto, parte deste problema pode ser alvo de uma dissertação de mestrado.

Assim, é dentro deste quadro mais geral, que foi feita a opção pelo recorte conceitual desta dissertação, limitando-o à definição de tipos de famílias com estrutura homogênea de despesas.

O uso de análise multivariada, em particular a análise de componentes principais e análise de agrupamento, têm sido aplicados em diferentes campos do saber. Na área de Marketing, por exemplo, a análise multivariada foi utilizada para definir perfis de consumidores de um *shopping centre* (Andrade, Oliveira & Antonialli, 2004); e para a identificação de perfis de empresas de sementes de trigo e soja (Acosta, Barros & Peske, 2002).

Nas áreas de avaliação e monitoramento de programas e políticas públicas, a análise multivariada tem sido utilizada para definir grupos homogêneos ou fatores determinantes da população estudada para facilitar os processos de compreensão, avaliação e monitoramento. Trabalhos, neste grupo, incluem avaliação tecnológica de piscicultores (Araújo, 2006); monitoramento da qualidade da água baseada na tipologia de fontes hídricas (Brito et al, 2006); uma tese de doutorado que avaliou a política educacional a partir de uma tipologia de

estabelecimentos de ensino (Cerqueira, 2004); além de uma tipologia de eleitores para avaliar a obrigatoriedade do voto no Brasil. (Elkins, 2000).

Na área de análise do espaço geográfico, diversos trabalhos buscam a definição de regiões homogêneas a partir de distintas óticas, tais como: agricultura (Llanillo et al, 2006); dinâmica urbana (Lemos et al, 2001); economia (Pereira, 2002); e oferta de serviços de saúde (Simões et al, 2004).

Alguns dos estudos baseados em análise multivariada, referidos acima, utilizaram dados do IBGE que eram provenientes de amostras complexas, mas não consideraram a fonte de aleatorização devida ao desenho da amostra.

Assim, a construção da tipologia de famílias fluminenses, na ótica do consumo de bens e serviços, além de constituir-se em um subsídio relevante ao projeto referido da Fecomércio-RJ, justifica-se pela incorporação da fonte de aleatorização devida ao desenho da amostra na aplicação dos métodos de análise multivariada.

De fato, os dados utilizados são provenientes de uma amostra complexa, desenhada para a Pesquisa de Orçamentos Familiares 2002-2003 (IBGE, 2004a). Um desenho de amostra que combina estratificação e conglomeração, e que, portanto, pode ser informativo em relação às estimativas de variâncias e covariâncias, na qual se baseiam os métodos a serem utilizados. O que era um problema que, à época de François (1977), ainda não estava na ordem do dia da estatística, mas que precisa ser enfrentado nos dias de hoje.

CAPÍTULO 3

Objetivos

A presente dissertação tem por objetivo principal discutir estratégias para a definição de tipos de famílias baseados na estrutura de gastos de consumo e, tomando por base os dados da Pesquisa de Orçamentos Familiares 2002-2003, construir tipologias e analisar as características dos diferentes tipos de famílias.

Este objetivo principal pode ser desmembrado nos objetivos específicos enumerados a seguir.

1. Identificar os grupos de produtos ou serviços demandados pelas famílias e que representam os diferentes setores da economia.
2. Descrever os métodos de análise multivariada a serem usados na construção da tipologia de famílias, ou seja, os métodos de componentes principais e de agrupamento.
3. Descrever uma solução proposta na literatura para realização da análise multivariada referida a partir de dados provenientes de amostras complexas.
4. Criar duas tipologias de famílias segundo a estrutura de gastos de consumo, uma ignorando o desenho da amostra da POF e outra considerando as informações estruturais do seu plano de amostragem.
5. Descrever os tipos de famílias criados em cada tipologia e comparar os resultados obtidos para avaliar a importância de considerar as informações estruturais do desenho da amostra da POF.

CAPÍTULO 4

Material e métodos

4.1 Pesquisa de Orçamentos Familiares 2002-2003 - POF

A Pesquisa de Orçamentos Familiares 2002-2003 (POF), quarta pesquisa sobre orçamentos familiares realizada pelo IBGE, “visa mensurar as estruturas de consumo, dos gastos e dos rendimentos das famílias e possibilita traçar um perfil das condições de vida da população brasileira a partir da análise de seus orçamentos domésticos.” (IBGE, 2004a:17). A POF é importante fonte para conhecimento do mercado consumidor tendo em vista que revela e identifica os bens adquiridos pelas famílias, assim como os serviços utilizados nas regiões de interesse. De posse de tais informações, podem ser realizados estudos sobre hábitos de consumo, gastos por classes de rendimento familiar, locais de compra, alocação de gastos, dentre outros de relevante utilidade para o ramo de comércio.

A POF de 2002-2003 foi realizada no período compreendido entre julho de 2002 e junho de 2003, tendo 15 de janeiro de 2003 como a data de referência para a expansão da amostra e para a referência dos preços, conforme usada para a apresentação dos resultados. Levando em consideração que a frequência de aquisição e o nível do valor gasto diferem por tipo de despesa, quatro períodos de referência foram definidos (sete dias, 30 dias, 90 dias e 12 meses), em função da frequência provável de consumo e do valor da despesa. As despesas de menor valor, como, por exemplo, os alimentos, costumam ser realizadas com maior frequência e as de maior valor, como é o caso de automóveis, pelo contrário, são realizadas com uma frequência muito menor.

Em função do tipo do informante e do nível da informação, a POF utilizou seis instrumentos de coleta ou questionários: domicílio; despesa coletiva; despesa individual; rendimento individual; condições de vida; e caderneta de despesa coletiva.

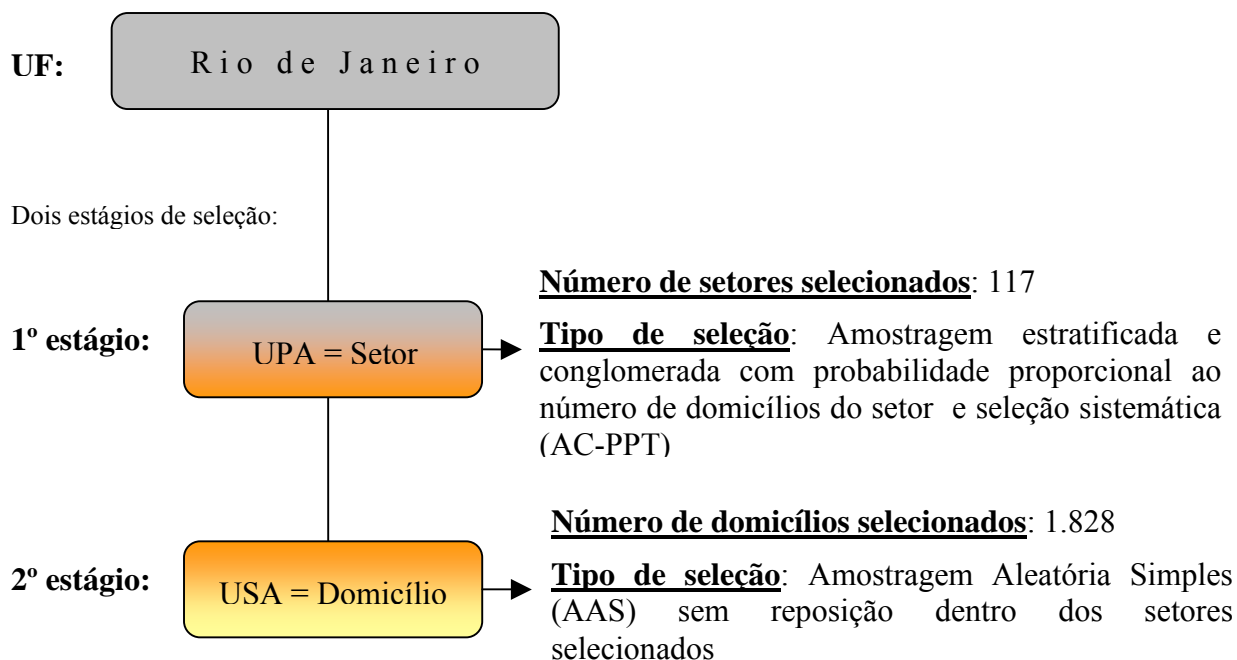
A POF é uma pesquisa baseada em uma amostra selecionada em dois estágios, que empregou estratificação geográfica e estatística (socioeconômica). Trata-se, portanto, de uma pesquisa amostral complexa e, portanto, deve ser tratada de maneira adequada de modo a incorporar as informações sobre pesos e plano amostral utilizado. De acordo com Pessoa & Silva (1998), existem aspectos importantes a serem considerados nesses tipos de amostragem, tais como: probabilidades distintas de seleção, conglomeração das unidades, estratificação e não-resposta. Ignorar tais aspectos traria impacto nas estimativas pontuais dos parâmetros da população, que são influenciadas pelos pesos das observações, assim como produziria estimativas incorretas de variâncias, já que as mesmas são influenciadas pela conglomeração, estratificação e pesos.

Na POF, os setores da base Geográfica do Censo Demográfico 2000 foram tratados como unidades primárias de amostragem (UPA) e selecionados de forma sistemática e com probabilidade proporcional ao número de domicílios de cada setor. Os domicílios particulares permanentes correspondem às unidades secundárias de amostragem (USA) e foram selecionados por amostragem aleatória simples sem reposição, dentro dos setores já selecionados no primeiro estágio. No dimensionamento da amostra, foram prefixados os números de domicílios a serem selecionados por situação do setor (urbano e rural) e macrorregião do País, que foram aumentados em cerca de 25% para considerar a não-resposta esperada. Em seguida, foi feita uma listagem para atualização do cadastro de domicílios dos setores selecionados e o número de domicílios a selecionar por setor foi alterado de forma proporcional ao aumento do número de domicílios do setor entre o momento do censo e o da listagem.

Com este desenho, foram selecionados, para o Rio de Janeiro, Unidade da Federação de interesse desta dissertação, 117 setores no primeiro estágio e 1.828 domicílios no segundo

estágio. Dos domicílios selecionados, foram efetivamente entrevistados 1.285 domicílios no estado. A Figura 4.A, ilustra o esquema de amostragem aplicado para o Rio de Janeiro.

Figura 4.A: Esquema de amostragem da POF: Conglomerado em dois estágios.



Entre julho de 2002 e junho de 2003, período da pesquisa, os setores e respectivos domicílios foram distribuídos de forma que, em todos os trimestres, fosse feita coleta de dados em todos os estratos geográficos e estatísticos.

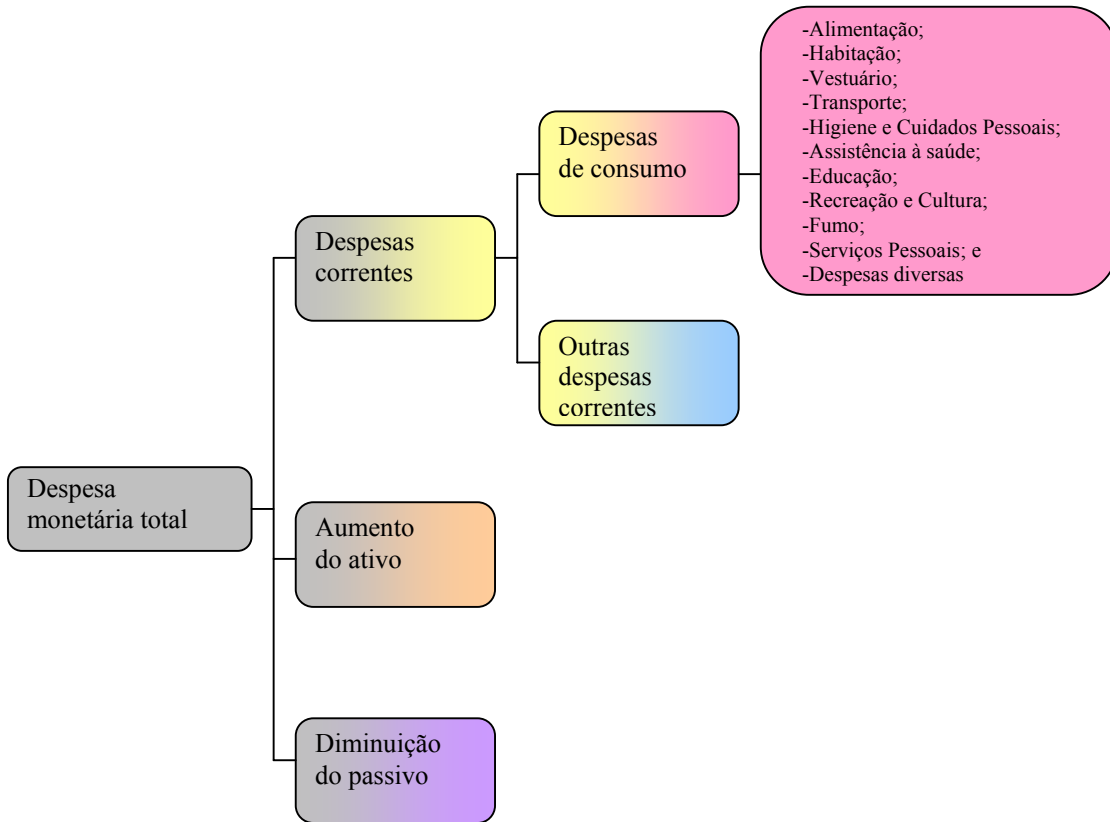
A estratificação geográfica garantiu a participação da amostra em todas as partes do território nacional. Na área urbana de cada unidade da federação (UF) foram definidos três estratos (município da capital, Região Metropolitana sem o município da capital e restante da área urbana). A área rural não-metropolitana, não foi definida por UF, sendo dividida em cinco estratos, um para cada grande região do país (Norte, Nordeste, Sudeste, Sul e Centro-Oeste). Já a estratificação estatística, construída em cada estrato geográfico, foi feita usando a média de anos de estudo dos responsáveis pelos domicílios do setor com base nos dados da Contagem de

População de 1996. A utilização desta variável é uma das grandes diferenças desta POF para as anteriores, que usaram informações sobre rendimentos, fornecidas pelo Censo Demográfico logo anterior.

Identificar todas as despesas das unidades de consumo é, de maneira simplificada, o objetivo principal da POF. A despesa total de uma família inclui todas as despesas correntes (despesas de consumo e outras despesas correntes), o aumento do ativo e a diminuição do passivo. Cada um destes conceitos de despesa pode ser subdividido em monetário e não-monetário, através da informação sobre a forma de obtenção dos produtos e serviços.

Apenas as despesas de consumo obtidas monetariamente, ou seja, bens e serviços utilizados para atender necessidades diretas e desejos pessoais dos consumidores obtidos com pagamento à vista ou a prazo, em dinheiro, cheque ou cartão de crédito, foram consideradas como de interesse nesta dissertação, visto que seu objetivo maior é estabelecer tipos de famílias para serem futuramente usados em estudos de projeção da demanda. As despesas monetárias de consumo foram classificadas de acordo com 11 grupos: alimentação; habitação; vestuário; transporte; higiene e cuidados pessoais; assistência à saúde; educação; recreação e cultura; fumo; serviços pessoais; e despesas diversas (despesas de consumo não classificadas anteriormente). As despesas monetárias totais podem ser representadas conforme o esquema apresentado na Figura 4.B.

Figura 4.B: Representação da despesa monetária total.



No produto de microdados da POF, os dados foram estruturados em 13 tipos de registros correspondentes aos 13 arquivos, cada um com formato fixo por variável e gravado como arquivo texto, conforme indicação a seguir: (1) domicílio; (2) pessoas, (3) condições de vida, (4) inventário de bens duráveis; (5) despesas de 90 dias; (6) despesas de 12 meses; (7) outras despesas; (8) despesas com serviços domésticos; (9) despesa coletiva; (10) despesas individuais; (11) despesas com veículos; (12) rendimentos e deduções; e (13) outros rendimentos (Quadro 4.A).

Quadro 4.A: Tipos de registros coletados pela POF.

Número do registro	Tipo do registro	Número de variáveis*	Descrição
1	Domicílio	25	Registros gerais sobre o domicílio
2	Pessoas	34	Registros das características de todos os moradores do domicílio
3	Condições de vida	33	Registros das condições de vida da Unidade de Consumo
4	Inventário de bens duráveis	17	Registros sobre as quantidade e aquisição de bens duráveis
5	Despesas de 90 dias	19	Registros das despesas em um período de referência de 90 dias
6	Despesas de 12 meses	20	Registros das despesas em um período de referência de 12 meses
7	Outras despesas	20	Registros das despesas não incluídas nos registros do tipo 5 e 6
8	Despesas com serviços domésticos	23	Registros das despesas com serviços domésticos
9	Caderneta de despesa coletiva	24	Registro das aquisições de uso comum no domicílio com alimentação, higiene e limpeza
10	Despesas individuais	20	Registros das despesas caracterizadas de uso ou finalidade individual
11	Despesas com veículos	21	Registros das despesas com aquisição de veículos
12	Rendimentos e deduções	27	Registros sobre rendimentos recebidos e deduções do trabalho principal
13	Outros rendimentos	21	Registros sobre rendimentos e recebimentos não derivados do trabalho principal

Fonte: Documentação da POF

*Existem variáveis comuns nos 13 tipos de registros, possibilitando o cruzamento de dados.

Como em toda amostra, cada unidade da amostra representa um determinado número de unidades da população de pesquisa, que no caso é constituída pelos domicílios particulares permanentes da população brasileira. O número de unidades da população que a unidade da amostra representa é chamado de peso amostral. O peso amostral natural do desenho é igual ao inverso da probabilidade de inclusão da unidade na amostra. No desenho da POF, as probabilidades de inclusão são diferentes por setor e, portanto, o peso natural do desenho difere de setor para setor, sendo constante para domicílios num mesmo setor.

Se não houvesse atualização do número de domicílios nos setores, antes da seleção dos domicílios; se o número de domicílios a serem selecionados por setor fosse constante por estrato; se não ocorresse não-resposta diferenciada por setor; e se as medidas de tamanho usadas na seleção dos setores fossem estritamente iguais ao número de domicílios observados no Censo Demográfico 2000, os pesos amostrais naturais do desenho de amostra da POF

seriam iguais em cada estrato e o plano de amostragem seria considerado como autoponderado no estrato. No entanto, este não é o caso da POF, como indicado anteriormente.

Atribuindo os pesos amostrais às variáveis pesquisadas, é possível obter estimativas das quantidades de interesse para a população, o que pode ser simplesmente demonstrado através do emprego do estimador de total de Horvitz-Thompson (Cochran, 1977). No entanto, os pesos naturais do desenho nem sempre conduzem a totais da população conhecidos por outras fontes de dados e é comum, nas pesquisas do IBGE, o emprego de técnicas de calibração dos pesos amostrais.

Segundo Silva (2004), a justificativa mais comum para a calibração dos pesos é a manutenção da coerência com dados populacionais já divulgados. De fato, a calibração permite, até certo ponto, corrigir vieses de pesquisa e tornar as estimativas coerentes com totais populacionais obtidos por outras fontes. Nessa técnica, os pesos do estimador usual de total de Horvitz-Thompson são modificados, obtendo-se um novo estimador com pesos que sejam os mais próximos possíveis dos pesos naturais do desenho ao mesmo tempo em que, quando associados às variáveis auxiliares, retornam os totais populacionais conhecidos dessas variáveis. Os valores das variáveis auxiliares, também conhecidas como variáveis de calibração, ou seus respectivos valores totais populacionais são conhecidos através de censos ou registros administrativos (Pessoa, 2005).

No caso da POF, o processo de calibração consistiu em obter estimativas para os totais de pessoas em determinados pós-estratos iguais às respectivas projeções populacionais obtidas para 15 de janeiro de 2003, por meio da técnica conhecida como *integrated household weighting*. Os pesos foram calibrados por meio do *Generalized Estimation System*, versão 4.0, desenvolvido pelo *Statistics Canada* (Pessoa, 2005).

O arquivo de microdados da POF, segunda divulgação, fornece tanto o peso natural do desenho, denominado por “Fator de expansão1 – peso do desenho”, quanto o peso calibrado,

chamado de “Fator de expansão² – peso final”. Nesta dissertação, foram usados os pesos naturais do desenho para as análises que consideram as informações estruturais do plano de amostragem da POF, como indicado na seção 4.3.

4.2 Análise multivariada

O termo análise multivariada, de um modo geral, refere-se a todos os métodos estatísticos que, simultaneamente, analisam múltiplas medidas sobre cada indivíduo ou unidade sob investigação (Hair et al, 2005). Com o conjunto de técnicas multivariadas é possível trabalhar e relacionar inúmeras observações correspondentes a medidas de diferentes variáveis; criar tipologias das variáveis e das observações; construir grupos de observações similares; dentre outras opções. A teoria básica sobre os dois tipos de análise multivariada aplicadas nesta dissertação, análise de componentes principais e análise de agrupamentos, será desenvolvida ao longo desta seção.

4.2.1 Análise de Componentes Principais - ACP¹

A Análise de Componentes Principais é uma técnica estatística de análise multivariada que transforma linearmente um conjunto original de variáveis num subconjunto substancialmente menor de variáveis não correlacionadas que contém a maior parte da informação do conjunto original (Fleck & Bourdel, 1998). Esta técnica está relacionada com a explicação da estrutura de covariância ou correlação por meio de poucas combinações lineares das variáveis em estudo. Em geral, a explicação de toda a variabilidade do sistema determinado por p variáveis só pode ser efetuada por p componentes principais. No entanto, algumas vezes, uma grande parte dessa variabilidade pode ser explicada por um número r menor de componentes, onde $r < p$. Os componentes principais são uma técnica de análise intermediária

¹ Descrição baseada, principalmente, em Johnson & Wichern (1988).

e, portanto, não se constituem em um método final e conclusivo. Esta técnica desempenha um importante papel complementar para outras técnicas multivariadas por representar uma forma de resumo e redução de dados, visto ser capaz de fornecer a base para a criação de um novo conjunto de variáveis que incorpora o caráter e a natureza das variáveis originais em um número muito menor de variáveis que, realmente, terão impacto na análise. Matematicamente, a análise de componentes principais representa uma projeção ortogonal para uma variedade linear de menor dimensão.

Algebricamente, os componentes principais representam combinações lineares de p variáveis aleatórias X_1, X_2, \dots, X_p . Geometricamente, essas combinações lineares representam a seleção de novos eixos ortogonais, que são obtidos por rotações do sistema de eixos original, representados por X_1, X_2, \dots, X_p . Os novos eixos representam as direções de máxima variabilidade.

Os componentes principais dependem somente da matriz de covariância² Σ (ou da matriz de correlação ρ) de X_1, X_2, \dots, X_p . Seu desenvolvimento não requer pressuposições de normalidade multivariada. Os conceitos, a seguir definidos para os componentes principais populacionais, podem ser estendidos para a situação amostral.

Seja o vetor aleatório $\tilde{X}^t = [X_1 \ X_2 \ \dots \ X_p]$ de uma população com matriz de covariância Σ , cujos autovalores são $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Então os componentes principais (X_1, X_2, \dots, X_p) são as combinações lineares dadas por (1.1)

$$\begin{aligned}
 Y_1 &= \mathbf{e}_1^t \tilde{X} = e_{11}X_1 + e_{12}X_2 + \dots + e_{1p}X_p \\
 Y_2 &= \mathbf{e}_2^t \tilde{X} = e_{21}X_1 + e_{22}X_2 + \dots + e_{2p}X_p \\
 &\vdots \\
 &\vdots \\
 &\vdots
 \end{aligned} \tag{1.1}$$

² Johnson & Wichern (1988) refere-se à matriz de variância e covariância simplesmente como matriz de covariância.

$$Y_p = \mathbf{e}_p^t \mathbf{X} = e_{p1}X_1 + e_{p2}X_2 + \dots + e_{pp}X_p$$

sendo,

$$\text{Var}(Y_i) = \text{Var}(\mathbf{e}_i^t \mathbf{X}) = \mathbf{e}_i^t \text{Var}(\mathbf{X}) \mathbf{e}_i = \mathbf{e}_i^t \Sigma \mathbf{e}_i \quad (1.2)$$

$$\text{Cov}(Y_i, Y_k) = \text{Cov}(\mathbf{e}_i^t \mathbf{X}, \mathbf{e}_k^t \mathbf{X}) = \mathbf{e}_i^t \Sigma \mathbf{e}_k \quad (1.3)$$

Pode-se, então, definir o i -ésimo componente principal (Y_i) por (1.4), pressupondo que o vetor \mathbf{X} possui covariância Σ , com pares de autovalores e autovetores $(\lambda_i, \mathbf{e}_i)$, $i=1,2,\dots,p$, em que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$.

$$Y_i = \mathbf{e}_i^t \mathbf{X} = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p \quad i=1,2,\dots,p \quad (1.4)$$

O máximo da forma quadrática dada por $\lambda = \frac{\mathbf{e}^t \Sigma \mathbf{e}}{\mathbf{e}^t \mathbf{e}}$, é obtido pela solução da equação

(1.5). A maximização de uma forma quadrática resulta na solução dada pelo conjunto de todos os pares de autovalores e autovetores da matriz núcleo, sendo os autovetores restritos ao comprimento unitário. Dessa forma, chega-se à relação dada por: $\Sigma \mathbf{e}_i = \lambda_i \mathbf{e}_i$ e, portanto, à variância e covariância de Y_i , por (1.6) e (1.7):

$$(\Sigma - \lambda_i \mathbf{I}) \mathbf{e}_i = \mathbf{0} \quad (1.5)$$

$$\text{Var}(Y_i) = \mathbf{e}_i^t \Sigma \mathbf{e}_i = \mathbf{e}_i^t \lambda_i \mathbf{e}_i = \lambda_i \mathbf{e}_i^t \mathbf{e}_i = \lambda_i \quad (1.6)$$

$$\text{Cov}(Y_i, Y_k) = \mathbf{e}_i^t \Sigma \mathbf{e}_k = \mathbf{e}_i^t \lambda_k \mathbf{e}_k = \lambda_k \mathbf{e}_i^t \mathbf{e}_k = 0 \quad i \neq k \quad (1.7)$$

Utilizando propriedades matriciais é possível demonstrar que $\sum_{i=1}^p \text{Var}(X_i) = \sum_{i=1}^p \text{Var}(Y_i)$,

ou seja, $\sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \lambda_1 + \lambda_2 + \dots + \lambda_p$. A variação total existente nas variáveis X_i , $i=1,2,\dots,p$,

portanto, é igual à variação existente nos p componentes principais e a porcentagem da variação total explicada pelo k -ésimo componente principal é dada por (1.8).

$$\%Var(Y_k) = \frac{\lambda_k}{\sum_{i=1}^p \lambda_i} \times 100 \quad (1.8)$$

A determinação da porcentagem de variação explicada por um número r de componentes que seja suficiente para que estes substituam ou resumam as p variável original, sem grande perda de informação, é um dos problemas que dificultam o uso desta técnica. Não existe um critério único para a definição do número de fatores a ser analisado. Heck e Bourdel (1998) citam alguns métodos que auxiliam nesse processo de escolha: o critério de Kaiser, que consistiria na seleção apenas dos autovalores com valores superiores à unidade; o método dos fatores interpretáveis, que sugere o uso somente dos fatores com significado facilmente identificável; e, a análise de uma representação gráfica conhecida por “scree plot”, na qual seriam considerados os autovalores até o ponto em que sua diferença para o próximo fosse significativa (ponto de inflexão da curva). Apesar desses critérios, Hair (2005) recomenda que os fatores escolhidos não representem menos de 60% da variabilidade total.

Padronização:

Os componentes principais podem ser obtidos pela padronização das variáveis originais de acordo com (1.9), representados, em notação matricial, por (1.10).

$$Z_i = \frac{X_i - \mu_i}{\sqrt{\sigma_{ii}}} \quad (1.9)$$

$$\tilde{Z} = V^{-1/2} \left(\tilde{X} - \tilde{\mu} \right) \quad (1.10)$$

em que $V^{-1/2}$ é uma matriz diagonal com os elementos da diagonal dados por $1/\sqrt{\sigma_{ii}}$ e,

$$(1.11)$$

$$E(\underline{Z}) = \underline{0} \quad e \quad Cov(\underline{Z}) = V^{-1/2} \Sigma V^{-1/2} = \rho$$

Ou seja, os componentes principais de \underline{Z} são dados pelos autovalores e autovetores de ρ , matriz de correlação de \underline{X} . Os autovalores e autovetores de Σ são, em geral, diferentes daqueles derivados de ρ .

Sejam as variáveis padronizadas Z_1, Z_2, \dots, Z_p dispostas no vetor \underline{Z} com $Cov(\underline{Z}) = \rho$.

Então os componentes principais de Z são dados por (1.11):

$$Y_i = \mathbf{e}_i^t \underline{Z} = \mathbf{e}_i^t V^{-1/2} (\underline{X} - \underline{\mu}), \quad i = 1, 2, \dots, p \quad (1.11)$$

Da mesma forma, verifica-se que:

$$\sum_{i=1}^p Var(Y_i) = \sum_{i=1}^p Var(Z_i) = \sum_{i=1}^p \lambda_i = p \quad (1.12)$$

$$\rho_{Y_i Z_k} = e_{ik} \sqrt{\lambda_i} \quad (1.13)$$

4.2.2 ACP para amostras complexas³

Como já indicado, o método de Análise de Componentes Principais (ACP) é baseado, fundamentalmente, no exame da estrutura da matriz de variância e covariância dos dados. Smith & Holmes (1989) consideram os problemas de estimar as funções conhecidas dessas matrizes, como autovalores e autovetores, para o caso de amostras complexas.

Os parâmetros de interesse considerados neste caso são as funções de autovalores e autovetores da matriz de variância e covariância de uma superpopulação, definida por Σ_{xx} .

³ Baseado em SKINNER et al, 1989.

Pressupondo que uma população finita seja gerada pelo modelo de superpopulação por amostragem aleatória simples e que sua matriz de variância e covariância, definida por $V_{\sim xxU}$, difira de $\Sigma_{\sim xx}$ apenas por termos de $O_p(N_0^{-1/2})$, onde N_0 é o tamanho da população finita, somente quando N_0 for suficientemente grande, as medidas de inferência sobre $V_{\sim xxU}$ e $\Sigma_{\sim xx}$ serão similares.

Smith & Holmes (1989) compararam o estimador da matriz de variância covariância, segundo alguns diferentes desenhos de amostra, facilitando a visualização dos vieses que esses estimadores produzem quando a amostragem complexa, pela qual os dados foram obtidos, é ignorada.

Por meio de simulações com amostras de dados normais multivariados, os autores verificaram que o estimador não-ponderado, que é a opção padrão da maioria dos pacotes estatísticos clássicos, são extremamente enviesados para desenhos de amostra não-autoponderados e não devem ser usados nestas situações. A melhor opção que se apresenta, portanto, para considerar os efeitos do desenho complexo de amostra seriam as seguintes:

- (a) Estimar a matriz de variância e covariância (ou a matriz de correlações) utilizando os pesos amostrais e as informações estruturais do desenho da amostra; e
- (b) Utilizar esta matriz como arquivo de entrada para uma rotina clássica de componentes principais.

4.2.3 Análise de agrupamentos não hierárquica e o método das k-médias⁴

Os agrupamentos não-hierárquicos buscam a partição de n objetos em k grupos. Os métodos exigem a pré-determinação de certas condições para produzir medidas sobre a

⁴ Estudo baseado em Anderberg, 1973 e informações contidas no manual SAS (Proc Fastclus, 2001).

qualidade do agrupamento efetuado. Pelo fato de a matriz de distâncias não precisar ser determinada e dos dados não serem armazenados no processo de execução do procedimento, os métodos não-hierárquicos aplicam-se melhor a uma quantidade maior de informações do que os métodos hierárquicos. Nos métodos não-hierárquicos, o ponto de partida é a divisão dos objetos em grupos ou a especificação de um conjunto de sementes que representariam o núcleo dos agrupamentos. Para que tais escolhas estejam livres de tendência, pode-se optar pela seleção aleatória da partição inicial ou das sementes que deverão ser utilizadas em um primeiro momento (SAS, 2001:PROC FASTCLUS).

O método das k-médias é um dos métodos não-hierárquicos mais populares desenvolvido através de um algoritmo que determina que cada objeto deve possuir um centróide mais próximo. O centróide, no caso das k-médias, seria a média de todos os objetos contidos em um mesmo agrupamento. De maneira bem simplificada, o algoritmo pode ser executado de acordo com os seguintes passos:

1. Dividir, arbitrariamente (ou aleatoriamente), os objetos em k grupos iniciais ou especificar k centróides iniciais (sementes);
2. Percorrer a lista de objetos e calcular as distâncias de cada um deles para o centróide dos grupos formados no primeiro passo, alocando cada item ao grupo cujo centróide esteja mais perto (menor distância) e recalculando os centróides dos grupos que ganharam ou perderam os objetos; e
3. Repetir o passo 2 até que nenhum objeto troque de grupo.

No passo 2, as medidas de distância utilizadas podem ser a distância euclidiana ou a distância de Mahalanobis, dentre outras. A formação final dos objetos agrupados depende da partição inicialmente realizada ou das sementes utilizadas, ou seja, cada alteração no passo 1 poderá gerar agrupamentos distintos dos que foram calculados em um momento anterior.

A distância euclidiana é uma medida de comprimento de um segmento de reta ligando dois objetos e a de Mahalanobis é uma forma padronizada da distância euclidiana, obtida por escalonamento em termos de desvios-padrão que padroniza os dados, com ajustes feitos para intercorrelação entre as variáveis. Dessa forma, sejam $\tilde{X}_1^t = [X_{11} X_{12} \dots X_{1p}]$ e $\tilde{X}_2^t = [X_{21} X_{22} \dots X_{2p}]$ observações de dois objetos. Então, a distância euclidiana entre eles é dada por (1.14) e a distância de Mahalanobis dada por (1.15). Neste caso, S^{-1} seria a inversa da matriz de variância e covariância amostral.

$$d_e(\tilde{X}_1, \tilde{X}_2) = \sqrt{(X_{11} - X_{21})^2 + (X_{12} - X_{22})^2 + \dots + (X_{1p} - X_{2p})^2} = \sqrt{(\tilde{X}_1 - \tilde{X}_2)^t (\tilde{X}_1 - \tilde{X}_2)} \quad (1.14)$$

$$d_M(\tilde{X}_1, \tilde{X}_2) = \sqrt{(\tilde{X}_1 - \tilde{X}_2)^t S^{-1} (\tilde{X}_1 - \tilde{X}_2)} \quad (1.15)$$

Para garantir grupos estáveis, ou seja, grupos de unidades que permaneçam sempre juntas, independentemente das sementes iniciais ou partição inicial, Vasconcellos (1987) propõe aplicar o procedimento de agrupamento, repetidas vezes e identificar as unidades que ficaram sempre juntas nas diferentes execuções do procedimento. As unidades que ficaram em distintos grupos nas diferentes aplicações são alocadas a um grupo especial, particularmente não-homogêneo.

4.3 Descrição dos procedimentos usados

Os procedimentos utilizados nesta dissertação são descritos de forma mais pormenorizada nesta seção.

4.3.1 Leitura dos dados e criação de arquivo com todos os tipos de despesas

Os 13 tipos de registros em que os dados da POF foram estruturados (ver Quadro 4.A) foram processados pelo SAS[®], para gerar um só arquivo com os registros dos domicílios e de

todos os seus itens de despesa. No caso dos registros com mais de uma variável de valor, foram gerados tantos registros quantas fossem as variáveis com valor positivo, sendo feita uma alteração no código do item de despesa, de forma a identificar de forma unívoca todos os tipos de despesas de um mesmo domicílio (Anexo II). Neste arquivo, foram gravadas duas variáveis de valor anual para cada item de despesa, ambas deflacionadas para preços de 15 de janeiro de 2003, independente do período de referência originalmente utilizado na coleta dos dados. Uma variável referia-se ao valor monetário e a outra ao valor não-monetário, apesar de a maioria dos resultados ser baseada na variável de valor monetário.

4.3.2 Definição dos grupos e criação das variáveis de valor da despesa

A criação de grupos de despesas de consumo é um ponto delicado do procedimento utilizado, visto ser necessário cobrir uma gama importante de setores do comércio e da prestação de serviços, sobretudo por que alguns englobam outros. A criação destes grupos tinha que permitir que o empresário de certo setor pudesse observar a proporção e o valor gasto pelas famílias com produtos que comercializa ou com serviços que oferece.

Para exemplificar, um grupo que incluísse todos os tipos de carnes (natural, industrializada, defumada ou até mesmo em conserva) não atenderia as necessidades de um empresário de um micro ou pequeno estabelecimento, como é o caso da maioria dos açougues, apesar de não fazer grande diferença para o empresário de um supermercado. Assim, a solução foi criar dois grupos distintos: um com as carnes vermelhas, e outro com enlatados, alimentos em conserva, preparados ou industrializados. Dessa forma, para o empresário do setor de açougue a análise seria limitada ao primeiro grupo, enquanto que o de supermercados e hipermercados poderia somar as despesas dos dois grupos.

É certo que a determinação destes grupos de despesa afeta toda a análise realizada e que grupos diferentes conduziram a outras tipologias de famílias. Além disto, a estatística

recomenda uma análise de correlação para reduzir o número de dimensões. No entanto, a opção pela definição de 60 grupos de produtos baseados nos setores de mercado se justifica pela necessidade de conduzir os resultados no sentido do projeto mais amplo de projeção da demanda. A princípio, pensou-se que, apesar do número expressivo de grupos de despesa criados, a amostra da POF no estado do Rio de Janeiro seria suficientemente grande para comportar esta análise, visto serem 117 unidades primárias de amostragem, garantindo, assim, os graus de liberdade necessários à estimação realizada. No entanto, ao longo do desenvolvimento deste trabalho, foi percebido que, na verdade, $((60*61)/2)=1.830$ parâmetros seriam estimados, ou seja, um número bastante superior aos graus de liberdade, indicando, portanto certa limitação do método, dada a quantidade de grupos necessária à análise.

Para a construção das variáveis que foram usadas na análise multivariada, uma primeira aproximação poderia ser feita com os valores monetários das despesas efetuadas em cada grupo definido, visto serem todos expressos na mesma unidade de medida. No entanto, esses valores guardam relação com a composição das famílias, que determina quantidades compradas e estabelece restrições a despesas que não são de subsistência ou de primeira necessidade. A idéia de expressar as despesas em termos *per capita* surge, então, de forma natural. No entanto, mesmo esta correção apresenta problemas, pois o consumo e as necessidades dos membros da família mudam em função de seu sexo e sua idade.

Desta forma, a solução adotada foi trabalhar com a estrutura dos gastos de despesa de cada família, expressa pela porcentagem de despesa monetária que a família gasta com cada um dos 60 grupos de despesa definidos em relação ao total da despesa monetária familiar. Definindo-se que x_k seja o valor, em reais, que determinada família gasta com o k-ésimo produto ou serviço, então, o percentual desta despesa em relação a todos os gastos de consumo familiar seria dado por (1.16). O denominador indica o valor total dos gastos com despesas de consumo familiar.

$$\%X_k = \frac{X_k}{\sum_{i=1}^{60} X_i} \times 100, \quad (1.16)$$

4.3.3 Aplicação dos métodos de análise multivariada

A análise multivariada foi aplicada com duas hipóteses distintas. A primeira supõe que os dados da POF foram selecionados por amostragem aleatória simples e que suas observações eram, portanto, realizações de variáveis aleatórias independentes e identicamente distribuídas, referida nesta dissertação como “método ingênuo”. Na segunda hipótese, as técnicas multivariadas foram aplicadas a partir de uma matriz de variância e covariância estimada considerando as informações estruturais do desenho da amostra da POF e o peso natural do desenho, referida simplesmente como “método com desenho”.

O método com desenho é um método que poderia ser executado por um usuário do arquivo de microdados da POF, apesar de ser mais provável que fosse usado o peso calibrado, visto que é chamado na documentação dos arquivos, de “Fator de expansão2 – peso final”. No entanto, o uso do peso calibrado implicaria considerar, também, a fonte de aleatorização derivada das regressões de calibração.

É fato que a estimação da matriz de variância e covariância populacional usando o peso calibrado, as informações estruturais, e os resíduos das regressões de calibração configura-se como a forma mais precisa realizar a estimação a partir dos dados da POF. No entanto, por não estar disponível uma rotina para este tipo de estimação da matriz de variância covariância populacional e diante dos prazos para defesa desta dissertação, não foi possível considerar esta hipótese.

Com a técnica de componentes principais foi possível representar a nuvem de 1.285 pontos (domicílios) de um espaço 60-dimensional (grupos de despesas) em um subespaço de

dimensão reduzida, de modo a manter no subespaço a relação entre as distâncias dos pontos no espaço original, com a maior precisão possível.

Para o método ingênuo, o arquivo de dados com as 60 variáveis de porcentagem dos gastos por domicílio foi tratado com o procedimento PRINCOMP do SAS, que foi executado sem opção de ponderação. O arquivo de saída do procedimento continha as variáveis na nova base e os autovalores. Considerando apenas as componentes que explicavam aproximadamente 60% da variabilidade total, ou seja, 10 componentes com explicação de 60,54%, o arquivo de saída da PRINCOMP serviu de entrada para o procedimento FASTCLUS do SAS.

No caso do método com desenho, o arquivo de dados com as 60 porcentagens por domicílio e o peso natural do desenho, foi usado para criação de um objeto no R, que além dos dados, continha informações sobre a estrutura do plano de amostra. Por meio de uma rotina desenvolvida pelo Professor Djalma Galvão Carneiro Pessoa, este objeto foi usado para estimar corretamente a matriz de covariância populacional. A matriz, estimada por este procedimento, serviu de entrada para o procedimento PRINCOMP do SAS que produziu os escores para transformações dos valores das variáveis para a nova base. Por meio do procedimento IML do SAS, a matriz de escores foi usada para converter os valores das variáveis do arquivo de dados para a nova base. Usando, apenas, as componentes com variabilidade acumulada de 60%, aproximadamente, conforme aplicado no método anterior, o arquivo de saída do procedimento IML foi usado como entrada para o procedimento FASTCLUS do SAS.

Com a técnica de agrupamento foi possível reunir observações em grupos não definidos anteriormente, de tal modo que as observações de um mesmo grupo fossem similares entre si, mas diferentes das observações dos outros grupos formados. O objetivo foi fazer com que a distância entre os agrupamentos fosse maximizada garantindo maior similaridade (ou menor distância) entre as observações (ou pontos). O procedimento FASTCLUS do SAS foi utilizado com os seguintes parâmetros: (1) critério de convergência (CONV) definido como 0,01, ou

seja, as interações somente foram interrompidas quando a mudança relativa máxima das sementes dos grupos foi menor ou igual a 0,01; (2) número máximo de agrupamentos (MAXCLUSTER) definido como 10; (3) número máximo de interações para o recálculo das sementes (MAXITER) definido como 100; e (4) REPLACE=random, que gera uma amostra pseudo-aleatória das observações completas, definindo o grupo de sementes iniciais, que são substituídas de forma aleatória.

Para os dois métodos, o procedimento FASTCLUS, com os parâmetros indicados acima, foi executado três vezes conforme sugerem as simulações feitas por Vasconcellos (1987). A cada execução, as famílias foram divididas em grupos e, ao final da terceira execução, foram determinados os grupos estáveis, isto é, os grupos de famílias que permaneceram juntas nas três execuções. O conjunto de grupos estáveis de famílias, para cada método, compõe a tipologia de famílias elaborada pelo método.

Em vez de explicar o significado de cada componente principal, optou-se por rotular os tipos de famílias em função dos grupos de despesa mais importantes em sua estrutura de despesas, como proposto por Vasconcellos (1987). Para compor o rótulo de cada tipo de família, foram consideradas as despesas com média no tipo maior ou igual a 5% da despesa média total do grupo. Além disto, consistentemente com as propostas de François (1977) foram estimadas diversas características socioeconômicas das famílias de cada tipo.

O maior problema da comparação entre as tipologias criadas pelos dois métodos reside na definição de medidas e indicadores. Smith & Holmes (1989) usaram os autovalores e a proporção explicada da variabilidade total como medidas para avaliação das diferenças entre os distintos estimadores que usaram.

CAPÍTULO 5

Resultados e discussão

5.1 Resultados sobre os grupos de produtos

Os grupos de despesa definidos com base nos critérios estabelecidos nesta dissertação constam da Tabela 5.A, com a indicação do setor de comércio ou serviço associado.

Tabela 5.A: Grupos de despesas de consumo e setores de comércio ou serviço associados.

continua

Tipo de bem obtido ou serviço utilizado	Setor de comércio ou serviço
Açúcar, bala, doce, biscoito e sobremesa	Super e Hipermercado, Mini Mercado, Loja de conveniência, Mercadinho, Padaria, Restaurante, Lanchonete, Bomboniere, Cafeteria, Sorveteria, etc
Água**	-
Alimentação fora de casa - almoço e jantar	Restaurante, Lanchonete, Pensão, etc.
Alimentação fora de casa (exceto almoço e jantar)	Lanchonete, Sorveteria, Cafeteria, Bar, Casa de chá, Adega, Loja de petisco, etc.
Aluguel de habitação	Administradoras de imóveis, etc.
Animal doméstico (aquisição, produto, serviço e tratamento)	Super e hipermercado, Loja de ração, Loja especializada em animal, clínica veterinária, etc.
Artigo de limpeza	Super e Hipermercado, Mini Mercado, Mercadinho, etc
Artigo doméstico	Hiper e Supermercado, Loja de departamento, loja de decoração, bazar, loja de cama mesa e banho, etc.
Artigo e equipamento para cuidado e higiene pessoal	Perfumaria, Hiper e supermercado, Loja de departamento, Loja de cosmético, Farmácia, etc.
Auto Peça, acessório e outro gasto com veículo próprio (exceto combustível)	Loja de auto peça e acessório, Loja de pneu, instalação de som, bateria, amortecedor, etc.
Aves e ovos	Super e Hipermercado, Mini Mercado, Mercadinho, Açougue, Abatedouro, etc.
Bebida não alcoólica (refrigerante, suco, chá, mate, café, etc.)	Super e Hipermercado, Mini Mercado, Loja de Conveniência, Mercadinho, Padaria, etc.
Brinquedo e jogo	Super e Hipermercado, Loja de departamento, loja de brinquedo, etc.
Calçado e acessório adulto	Sapataria, loja de departamento, Loja de artigos de couro e viagem, etc.
Carne vermelha	Super e Hipermercado, Açougue, etc.
Cereal (arroz, feijão, grão, etc.)	Super e Hipermercado, Mini Mercado, Mercadinho, Quitanda, Sacolão, etc
Cerveja e chopp	Restaurante, Lanchonete, Bar, Loja de petisco, etc.
Combustível	Posto de gasolina, álcool, diesel e posto de gás natural.
Computador e acessório, serviço de acesso à internet	Loja de eletrodoméstico, Informática, Loja de departamento, Provedor de acesso à internet, etc.
Conserto e manutenção de artigo doméstico e pessoal	Marcenaria, Assistência técnica, Loja técnica de informática, etc.

Tipo de bem obtido ou serviço utilizado	Setor de comércio ou serviço
Curso de formação regular	Faculdade e universidade particular, colégio, creche, pré-vestibular, supletivo, etc
Curso não regular	Academia de ginástica, dança, balé, natação, curso de informática, curso de idioma, etc.
Diversão, esporte e cultura	Cinema, teatro, casa de espetáculo, locadora de dvd, loteria, motel, etc.
Educação, leitura e gasto escolar	Colégio, Jornaleiro, Livraria, Papelaria, Empresa Jornalística, etc.
Eletrodoméstico de grande porte	Super e Hipermercado, Loja de eletrodoméstico, etc.
Eletrodoméstico de pequeno e médio porte, equipamentos domésticos	Super e Hipermercado, Loja de departamento, Loja de eletrodoméstico, etc.
Eletroeletrônico - Equipamento de som, fotografia e tv	Hiper e Supermercado, Loja de departamento, Loja de Cine Foto e Som, Loja de Fotografia, etc.
Empregado doméstico (serviço)*	-
Energia elétrica**	-
Enlatado, conserva, alimento preparado e industrializado	Super e Hipermercado, Mini Mercado, Mercadinho, etc.
Fumo	Mini Mercado, Loja de Conveniência, Mercadinho, Padaria, Charutaria, Lanchonete, Bar, etc.
Gás, aquisição de bujão de gás**	-
Hospedagem	Hotel, Pousada, Hospedaria em geral, etc.
Jóia e bijouteria	Joalheria, Relojoaria, Loja de bijouteria, etc.
Leguminosa, hortaliça e fruta	Super e Hipermercado, Mini Mercado, Mercadinho, Quitanda, Sacolão, etc
Leite, queijo e derivado	Super e Hipermercado, Mini Mercado, Mercadinho, Queijaria, Loja de Laticínios, Padaria, etc.
Massa e farinha	Super e Hipermercado, Mini Mercado, Mercadinho, Padaria, etc
Material de construção	Loja de material de construção, Loja de material elétrico, Loja de tinta, Loja de ferragem, Marmoraria, etc.
Material e artigo escolar e serviço de papelaria	Papelaria, Copiadora, Loja de departamento, Livraria, etc
Mobiliário	Loja de móvel, Loja de colchão, etc.
Óculos, lentes e material oftalmológico	Ótica, Farmácia, etc.
Outra bebida alcoólica (exceto cerveja e chopp)	Adega, Bar, Restaurante, etc.
Outro produto de alimentação (óleo, gordura, margarina, sal, etc.)	Super e Hipermercado, Mini Mercado, Mercadinho, etc.
Pão e bolo de padaria	Super e Hipermercado, Padaria, Confeitaria, etc.
Peixe e fruto do mar	Super e Hipermercado, Mini Mercado, Mercadinho, Peixaria, etc
Plano de saúde e odontológico	Empresa de plano de saúde, Empresa de plano odontológico, etc.
Refeição pronta, alimento preparado para viagem	Restaurante, Lanchonete, etc.
Remédio e material de tratamento	Farmácia, Farmácia de manipulação, Homeopatia, Loja de artigo médico, ortopédico, etc.
Roupa feminina e masculina	Loja de roupa, Loja de departamento, etc.
Roupa, calçado e acessório para criança	Loja de roupa infantil, Sapataria infantil, Loja de departamento, etc.
Seguro em geral (exceto de saúde e odontológico)	Banco, Seguradora de vida, Seguradora de automóvel, etc.

conclusão

Tipo de bem obtido ou serviço utilizado	Setor de comércio ou serviço
Serviço de reforma e manutenção do imóvel	Firma de detetização e desratização, Empresa de engenharia e arquitetura, Marcenaria, Serralheria, etc.
Serviço e aparelho para comunicação (exceto internet)	Empresa de telefonia fixa, Empresa de telefonia móvel, TV por assinatura, Empresa de entrega, etc.
Serviço e tratamento médico e odontológico	Hospital particular, Clínica médica em geral e odontológica, Laboratórios de análise clínica, Clínica de raio-X, etc.
Serviço para cuidado e higiene pessoal	Cabeleireiro, Clínica de estética, Clínica de depilação, etc.
Serviço, taxa e outro gasto com habitação*	-
Transporte com viagem e outro gasto relacionado	Empresa aérea, Empresa de transporte intermunicipal, interestadual, etc.
Transporte urbano	Empresa de ônibus, metrô, trem e barca.
Veículo (aquisição)	Concessionária e Agência de veículo
Outra despesa de consumo	Diversos

*Tipo de despesa cuja associação com estabelecimento formal de comércio varejista ou de serviço não se aplica.

**Serviços industriais de utilidade pública

Outro ponto a destacar é que a decisão de uso da despesa monetária de consumo reduziu a proporção de despesa considerada nesta dissertação para 57,7% da despesa total (Tabela 5.B). Apesar de esta porcentagem poder parecer natural, visto que só as despesas monetárias importam para o comércio, há que se considerar que este ponto representa uma limitação dos dados usados, uma vez que as despesas não-monetárias das famílias pesquisadas podem advir de compras feitas no mercado por outras famílias, empresas ou pelo governo. Neste caso, a demanda por estes produtos e serviços poderia não figurar nas projeções.

Tabela 5.B: Estimativa dos valores das despesas (monetária e total) e respectivos percentuais, segundo os grupos de despesas, no estado do Rio de Janeiro

Grandes grupos e grupos de despesa	Estimativa do total da despesa (R\$)*		Percentual das despesas** das famílias em relação ao total	Percentual das despesas monetárias das famílias em relação ao total das despesas**
	Monetária	Despesa total**		
Despesa total	130.877.936.372	150.225.181.671	100,0%	87,1%
1. Despesas correntes	113.029.967.991	132.266.316.601	88,0%	75,2%
1.1. Despesas de consumo	86.673.212.258	105.886.283.204	70,5%	57,7%
1.2. Outras despesas correntes	26.356.755.733	26.380.033.397	17,6%	17,5%
2. Aumento do ativo	15.340.660.021	15.374.018.962	10,2%	10,2%
3. Diminuição do passivo	2.507.308.360	2.584.846.108	1,7%	1,7%

Fonte: POF 2002-2003 (IBGE).

** Incluindo despesas não monetárias.

5.2 Resultados do método ingênuo

O gráfico que representa as 60 componentes com seus respectivos autovalores (“scree-plot”) gerados pelo método ingênuo é apresentado na Figura 5.A. A porcentagem de explicação da variabilidade total consta na Tabela 5.C. Utilizando o critério sugerido por Hair de considerar componentes que expliquem, pelo menos, 60% da variabilidade total, os 10 primeiros componentes foram extraídos para o procedimento de agrupamento. Estes, juntos, explicam 60,6% da variância.

Figura 5.A: Scree-plot – Método ingênuo.

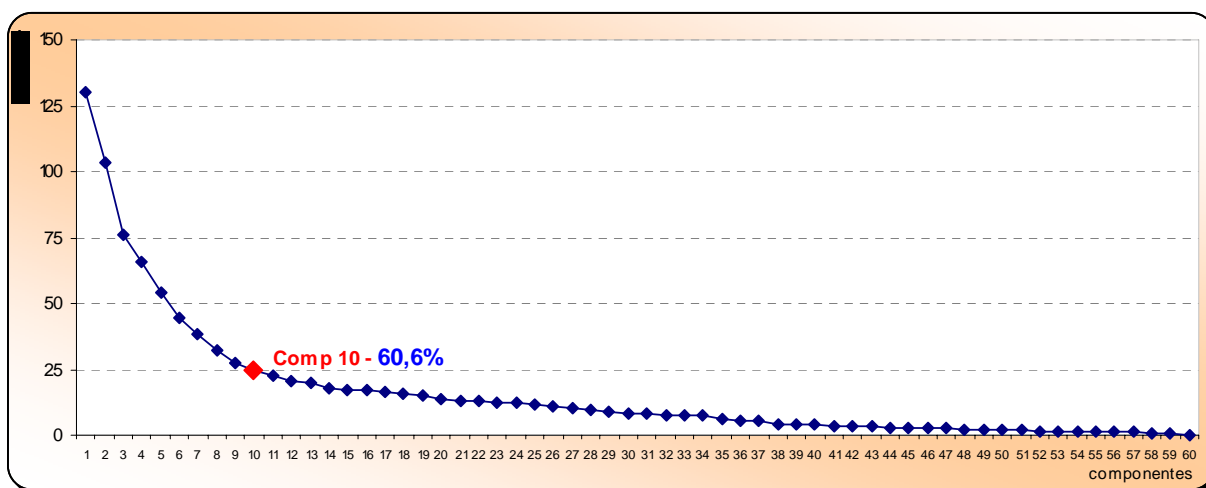


Tabela 5.C: Autovalores e porcentagem de explicação da variação total, segundo os componentes – Método ingênuo.

Componente	autovalor	%	% acum	Componente	autovalor	%	% acum	Componente	autovalor	%	% acum
1	130,2	13,2%	13,2%	21	13,3	1,3%	79,9%	41	3,7	0,4%	96,2%
2	103,4	10,5%	23,7%	22	12,9	1,3%	81,2%	42	3,6	0,4%	96,6%
3	76,2	7,7%	31,5%	23	12,4	1,3%	82,5%	43	3,3	0,3%	96,9%
4	65,7	6,7%	38,1%	24	12,2	1,2%	83,7%	44	3,0	0,3%	97,2%
5	54,2	5,5%	43,6%	25	11,5	1,2%	84,9%	45	2,9	0,3%	97,5%
6	44,2	4,5%	48,1%	26	10,7	1,1%	86,0%	46	2,7	0,3%	97,8%
7	38,6	3,9%	52,0%	27	10,2	1,0%	87,0%	47	2,6	0,3%	98,1%
8	31,9	3,2%	55,3%	28	9,5	1,0%	88,0%	48	2,4	0,2%	98,3%
9	27,5	2,8%	58,1%	29	9,1	0,9%	88,9%	49	2,3	0,2%	98,5%
10	25,0	2,5%	60,6%	30	8,5	0,9%	89,7%	50	2,1	0,2%	98,8%
11	22,3	2,3%	62,9%	31	8,3	0,8%	90,6%	51	1,8	0,2%	98,9%
12	20,8	2,1%	65,0%	32	7,8	0,8%	91,4%	52	1,7	0,2%	99,1%
13	19,7	2,0%	67,0%	33	7,7	0,8%	92,2%	53	1,6	0,2%	99,3%
14	18,1	1,8%	68,8%	34	7,2	0,7%	92,9%	54	1,5	0,2%	99,4%
15	17,4	1,8%	70,6%	35	6,1	0,6%	93,5%	55	1,5	0,2%	99,6%
16	16,9	1,7%	72,3%	36	5,6	0,6%	94,1%	56	1,3	0,1%	99,7%
17	16,7	1,7%	74,0%	37	5,2	0,5%	94,6%	57	1,3	0,1%	99,8%
18	15,9	1,6%	75,6%	38	4,3	0,4%	95,0%	58	0,9	0,1%	99,9%
19	15,4	1,6%	77,2%	39	4,1	0,4%	95,5%	59	0,4	0,0%	100,0%
20	13,6	1,4%	78,5%	40	3,9	0,4%	95,9%	60	0,2	0,0%	100,0%

O procedimento de agrupamento aplicado conduziu à formação de 68 combinações dos grupos nas três execuções, com uma média de 19 domicílios e tamanho variando de 1 a 308 domicílios na amostra (Tabela 5.D). Foram criados 11 grupos estáveis, ou tipos de famílias, com tamanho médio de 102 domicílios na amostra e intervalo de variação entre 22 e 308, envolvendo 88% dos domicílios da amostra (1.127 dos 1 285 domicílios da amostra da POF no Rio de Janeiro).

Tabela 5.D: Informações sobre os grupos considerados estáveis e não-estáveis após três execuções consecutivas do Procedimento Fastclus do SAS – Método ingênuo.

Variáveis	Geral	Tipo de grupo	
		Estável	Não estável
Número de combinações formadas	68	11	57
Freqüência média de domicílios por grupo	19	102	3
Intervalo de freqüência dos domicílios	1 a 308	22 a 308	1 a 11
Quantidade de domicílios	1285	1127	158

Fonte: elaboração própria através dos resultados obtidos na execução do procedimento Fastclus do SAS.

O próximo passo foi identificar os gastos de consumo mais representativos em cada um dos 11 tipos de família definidos, adotando o limite de 5% da despesa média total do grupo, como descrito ao final da seção 4.3.3. A descrição dos tipos de família consta da tabela 5.E e a estimativa da porcentagem média de cada grupo de despesa em cada tipo de família consta do Anexo III.

As características socioeconômicas dos tipos de família são apresentadas na Tabela 5.F. Os dados indicam que as condições de vida tendem a ser consideradas piores nos tipos de menor renda *per capita* e, de um modo geral, os domicílios dos tipos mais ricos tendem a ter mais responsáveis do sexo masculino do que do sexo feminino, com uma exceção para o tipo de maior renda *per capita*.

Tabela 5.E: Descrição dos tipos de família – Método ingênuo

Tipo	Rótulo do grupo de famílias	Número de famílias		Grupos de despesas de consumo mais representativos
		amostra	população	
1	Veículo	85	299.086	Veículo – aquisição (32,6%)
2	Educação regular, comunicação e plano médico	167	518.947	Curso de formação regular (9,5%) Serviço e aparelho para comunicação – exceto internet (5,7%) Plano de saúde e odontológico (5,3%)
3	Alimentação fora de casa e transporte	68	266.540	Alimentação fora de casa – almoço e jantar (17,1%) Transporte urbano (9,0%) Conta de energia elétrica (5,8%) Serviço e aparelho para comunicação – exceto internet (5,2%)
4	Gasto veicular, comunicação e luz	106	374.866	Combustível (20,3%) Serviço e aparelho para comunicação – exceto internet (6,1%) Auto peça, acessório e gasto com veículo – exc. combustível (5,6%) Conta de energia elétrica (5,1%)
5	Comunicação e transporte urbano	74	269.664	Serviço e aparelho para comunicação – exceto internet (20,8%) Conta de energia elétrica (10,9%) Transporte urbano (8,1%)
6	Medicamentos e luz	70	256.173	Remédio e material de tratamento (24,5%) Conta de energia elétrica (10,1%) Serviço e aparelho para comunicação – exceto internet (5,6%)
7	Aluguel, transporte, luz e comunicação	73	305.917	Aluguel de habitação (30,6%) Transporte urbano (9,6%) Conta de energia elétrica (5,7%) Serviço e aparelho para comunicação – exceto internet (5,5%)
8	Contas domésticas	58	176.282	Conta de energia elétrica (39,7%) Conta de gás, aquisição de bujão de gás (10,0%)
9	Transporte urbano	96	373.297	Transporte urbano (34,7%) Conta de energia elétrica (8,1%)
10	Alimentação dentro de casa	308	1.087.496	Alimentação dentro de casa (31,0%) Conta de energia elétrica (5,5%) Transporte urbano (5,3%)
11	Comunicação e alimentação dentro de casa*	22	79.299	Serviço e aparelho para comunicação – exceto internet (13,3%) Conta de energia elétrica (7,7%) Transporte urbano (7,6%) Leite, queijo e derivado (4,3%)

* A utilização do quarto maior percentual foi útil na diferenciação do grupo 5.

Tabela 5.F: Informações adicionais segundo os tipos de famílias – Método ingênuo (continua)

Características	Tipo 2: Educação regular, comunicação e plano médico	Tipo 1: Veículo	Tipo 3: Alimentação fora de casa e transporte	Tipo 4: Gasto veicular, comunicação e luz	Tipo 5: Comunicação, luz e transporte urbano	Tipo 11: Comunicação, luz e transporte urbano	Tipo 7: Aluguel, transporte, luz e comunicação	Tipo 6: Medicamentos e luz	Tipo 10: Alimentação dentro de casa	Tipo 9: Transporte urbano	Tipo 8: Contas domésticas
Renda <i>per capita</i> (emR\$)	2.249	1.953	1.206	1.135	657	623	568	544	448	344	309
INFORMAÇÕES - DOMICÍLIO:											
Número médio de moradores	3	3	3	3	3	4	3	3	4	3	3
Domício com condição de próprio (já quitado)	65,3%	52,9%	69,1%	79,3%	77,0%	90,9%	0,0%	82,9%	64,0%	76,0%	65,5%
Número médio de cômodos	8	7	6	6	6	6	5	6	5	5	5

Tabela 5.F: Informações adicionais segundo os tipos de famílias – Método ingênuo (continua)

Características	Tipo 2: Educação regular, comunicação e plano médico	Tipo 1: Veículo	Tipo 3: Alimentação fora de casa e transporte	Tipo 4: Gasto veicular, comunicação e luz	Tipo 5: Comunicação, luz e transporte urbano	Tipo 11: Comunicação, luz e transporte urbano	Tipo 7: Aluguel, transporte, luz e comunicação	Tipo 6: Medicamentos e luz	Tipo 10: Alimentação dentro de casa	Tipo 9: Transporte urbano	Tipo 8: Contas domésticas
Renda <i>per capita</i> (emR\$)	2.249	1.953	1.206	1.135	657	623	568	544	448	344	309
INFORMAÇÕES - CONDIÇÕES DE VIDA:											
Aproveitamento da renda no mês com algum grau de dificuldade	74,2%	79,3%	73,1%	83,5%	98,6%	81,8%	91,8%	91,2%	91,7%	91,6%	92,9%
Quando a quantidade de alimentos consumidos pela família nunca é totalmente suficiente	31,4%	28,1%	38,8%	32,0%	56,3%	54,6%	54,8%	57,4%	65,2%	60,0%	57,2%
Não possuem a rua pavimentada	12,6%	14,1%	25,0%	32,1%	32,4%	22,7%	12,3%	38,6%	36,4%	46,9%	62,1%
Não tem serviço de água	1,9%	8,5%	13,4%	11,6%	4,2%	4,6%	6,9%	14,7%	20,9%	21,1%	21,4%
Não tem coleta de lixo	1,3%	3,7%	4,5%	12,5%	4,2%	0,0%	1,4%	14,7%	13,9%	7,4%	32,1%
Não tem iluminação na rua	3,1%	7,3%	6,0%	11,7%	9,9%	9,1%	4,1%	19,1%	21,2%	15,8%	35,7%
Não tem serviço de drenagem e escoamento da água da chuva	0,0%	14,6%	19,4%	35,0%	15,5%	18,2%	15,1%	36,8%	40,1%	32,6%	53,6%
Percepção das condições de moradia da família (avaliação modal)	Boas (69,18%)	Boas (64,63%)	Boas (55,22%)	Boas (56,31%)	Boas (49,30%)	Satisfatórias (40,91%)	Satisfatórias (50,68%)	Boas (48,53%)	Satisfatórias (41,06%)	Boas (40,00%)	Boas (42,86%)
Atrasou nos últimos 12 meses o aluguel ou prestação da residência	5,0%	6,1%	1,5%	1,0%	4,2%	0,0%	28,8%	2,9%	3,3%	3,2%	5,4%
Atrasou nos últimos 12 meses despesas com água, eletrecidade, gás.	25,2%	30,5%	19,4%	20,4%	33,8%	18,2%	34,3%	27,9%	38,1%	26,3%	28,6%
Atrasou no pagamento de prestações de bens ou serviços adquiridos	20,1%	22,0%	20,9%	22,3%	28,2%	36,4%	26,0%	17,7%	26,5%	15,8%	14,3%

Tabela 5.F: Informações adicionais segundo os tipos de famílias – Método ingênuo (conclusão)

Características	Tipo 2: Educação regular, comunicação e plano médico	Tipo 1: Veículo	Tipo 3: Alimentação fora de casa e transporte	Tipo 4: Gasto veicular, comunicação e luz	Tipo 5: Comunicação, luz e transporte urbano	Tipo 11: Comunicação, luz e transporte urbano	Tipo 7: Aluguel, transporte, luz e comunicação	Tipo 6: Medicamentos e luz	Tipo 10: Alimentação dentro de casa	Tipo 9: Transporte urbano	Tipo 8: Contas domésticas
Renda <i>per capita</i> (emR\$)	2.249	1.953	1.206	1.135	657	623	568	544	448	344	309
INFORMAÇÕES - CHEFE DA FAMÍLIA:											
Razão de sexo	0,8	1,1	1,2	1,1	0,8	0,9	0,7	1,1	0,9	0,9	0,8
Idade: Média	35	30	32	32	34	30	32	42	28	28	33
Mediana	34	30	31	30	30	27	33	42	25	25	29
Nível de instrução predominante	Ensino superior ou acima (37,32%)	Ensino fundamental (34,01%)	Ensino fundamental (41,34%)	Ensino fundamental (42,74%)	Ensino fundamental (55,08%)	Ensino fundamental (59,26%)	Ensino fundamental (53,76%)	Ensino fundamental (56,91%)	Ensino fundamental (59,98%)	Ensino fundamental (56,47%)	Ensino fundamental (60,51%)
Cor ou Raça: Branca	79,7%	73,5%	55,3%	70,1%	64,0%	44,4%	52,7%	71,8%	40,6%	48,0%	49,7%
Preta ou parda	19,7%	26,5%	44,1%	28,8%	35,6%	53,1%	45,2%	28,2%	59,0%	51,4%	50,3%
Tem cartão de crédito	41,2%	36,4%	28,5%	28,5%	19,1%	22,2%	17,7%	9,0%	10,2%	7,6%	5,1%
Tem cheque especial	33,5%	27,9%	29,6%	21,9%	6,8%	9,9%	7,0%	9,6%	3,9%	1,9%	1,9%
É titular de plano de saúde	35,6%	28,9%	26,8%	20,0%	13,6%	11,1%	9,7%	13,8%	6,5%	4,1%	3,8%
Renda total mensal:											
Faixa: Até 8 SM	12,2%	23,5%	53,6%	39,2%	65,3%	66,7%	79,0%	76,6%	79,9%	87,1%	89,2%
Acima de 8 SM	87,9%	76,5%	46,4%	60,8%	34,8%	33,3%	21,0%	23,4%	20,1%	12,9%	10,8%
Valor: Média (em R\$)	6.594	4.706	2.341	3.245	1.886	2.132	1.243	1.315	1.213	970	853
Mediana (em R\$)	4.962	2.705	1.382	2.396	1.278	1.267	850	869	868	825	529

5.3 Resultados do método com desenho

O “scree-plot”, com os autovalores derivados da análise de componentes principais feita pelo método com desenho e a porcentagem de explicação da variabilidade total são apresentados na Figura 5.B. e Tabela 5.G, respectivamente. Utilizando o mesmo critério, aplicado no método anterior, bastaria extrair as primeiras nove componentes para obter uma explicação de 60%, aproximadamente. Porém, para facilitar a comparação entre métodos, optou-se por seleccionar o mesmo número de componentes, ou seja, as 10 primeiras, que, neste caso, explicam 62,7%.

Figura 5.B: Scree-plot – Método com desenho.

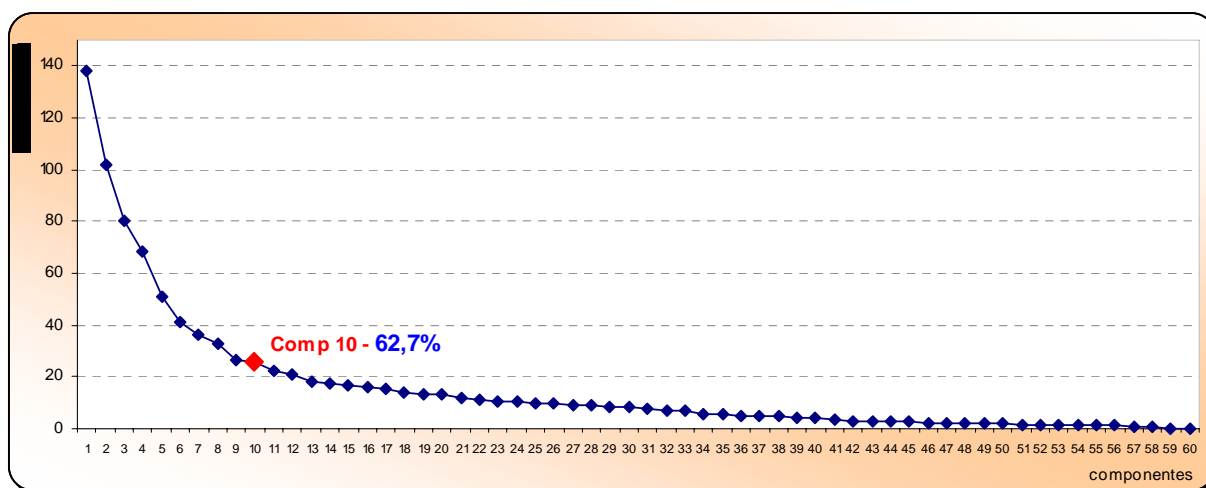


Tabela 5.G: Autovalores e porcentagem de explicação da variação total, segundo os componentes – Método com desenho.

Componente	autovalor	%	% acum	Componente	autovalor	%	% acum	Componente	autovalor	%	% acum
1	137,8	14,4%	14,4%	21	11,7	1,2%	81,3%	41	3,3	0,3%	96,6%
2	101,6	10,6%	24,9%	22	10,9	1,1%	82,5%	42	3,0	0,3%	96,9%
3	80,6	8,4%	33,3%	23	10,8	1,1%	83,6%	43	2,9	0,3%	97,2%
4	68,2	7,1%	40,4%	24	10,6	1,1%	84,7%	44	2,8	0,3%	97,5%
5	50,9	5,3%	45,7%	25	10,0	1,0%	85,8%	45	2,6	0,3%	97,8%
6	41,5	4,3%	50,1%	26	9,6	1,0%	86,7%	46	2,4	0,2%	98,0%
7	36,3	3,8%	53,8%	27	9,3	1,0%	87,7%	47	2,2	0,2%	98,3%
8	32,9	3,4%	57,3%	28	9,1	1,0%	88,7%	48	2,1	0,2%	98,5%
9	26,6	2,8%	60,0%	29	8,6	0,9%	89,6%	49	2,0	0,2%	98,7%
10	25,7	2,7%	62,7%	30	8,2	0,9%	90,4%	50	1,8	0,2%	98,9%
11	22,0	2,3%	65,0%	31	7,8	0,8%	91,2%	51	1,7	0,2%	99,1%
12	21,1	2,2%	67,2%	32	7,3	0,8%	92,0%	52	1,6	0,2%	99,2%
13	18,4	1,9%	69,1%	33	6,9	0,7%	92,7%	53	1,5	0,2%	99,4%
14	17,6	1,8%	70,9%	34	5,9	0,6%	93,3%	54	1,4	0,1%	99,5%
15	16,6	1,7%	72,7%	35	5,6	0,6%	93,9%	55	1,3	0,1%	99,7%
16	16,0	1,7%	74,3%	36	5,1	0,5%	94,4%	56	1,3	0,1%	99,8%
17	15,0	1,6%	75,9%	37	4,8	0,5%	94,9%	57	0,8	0,1%	99,9%
18	14,2	1,5%	77,4%	38	4,6	0,5%	95,4%	58	0,6	0,1%	100,0%
19	13,4	1,4%	78,8%	39	4,2	0,4%	95,9%	59	0,3	0,0%	100,0%
20	13,0	1,4%	80,1%	40	3,9	0,4%	96,3%	60	0,1	0,0%	100,0%

O procedimento de agrupamento aplicado conduziu à formação de 51 combinações dos grupos nas três execuções, com uma média de 25 domicílios e tamanho variando de 1 a 109 domicílios na amostra (Tabela 5.H). Foram criados 10 tipos de famílias, com tamanho médio de 111 domicílios na amostra e intervalo de variação entre 84 e 109, envolvendo 86% dos domicílios da amostra (1.105 dos 1 285 domicílios da amostra da POF no Rio de Janeiro).

Tabela 5.H: Informações sobre os grupos considerados estáveis e não-estáveis após três execuções consecutivas do Procedimento Fastclus do SAS – Método com desenho.

Variáveis	Geral	Tipo de grupo	
		Estável	Não estável
Número de combinações formadas	51	10	41
Frequência média de domicílios por grupo	25	111	4
Intervalo de frequência dos domicílios	1 a 109	1 a 24	84 a 109
Quantidade de domicílios	1285	1105	180

O passo seguinte foi a identificação dos gastos de consumo mais representativos em cada um dos 11 tipos de família definidos, adotando o limite de 5% da despesa média total do grupo, como descrito ao final da seção 4.3.3. A descrição dos tipos de família consta da tabela 5.I e a estimativa da porcentagem média de cada grupo de despesa em cada tipo de família consta do Anexo IV.

As características socioeconômicas dos tipos de família são apresentadas na Tabela 5.J.

Tabela 5.I: Descrição dos tipos de família – Método com desenho.

Tipo	Rótulo do grupo de famílias	Número de famílias		Grupos de despesas de consumo mais representativos
		amostra	população	
1	Gasto veicular, comunicação e luz	109	383.052	Combustível (20,2%) Serviço e aparelho para comunicação – exceto internet (6,1%) Auto peça, acessório e gasto com veículo – exc.combustível (5,7%) Conta de energia elétrica (5,1%)
2	Educação regular, plano médico e comunicação	100	295.181	Curso de formação regular (9,4%) Plano de saúde e odontológico (5,4%) Serviço e aparelho para comunicação – exceto internet (5,2%)
3	Educação regular, gasto veicular e comunicação	31	72.605	Curso de formação regular (14,3%) Combustível (10,9%) Auto peça, acessório e gasto com veículo – exc.combustível (6,1%) Serviço e aparelho para comunicação – exceto internet (5,0%)
4	Transporte urbano e luz	126	505.073	Transporte urbano (32,9%) Conta de energia elétrica (7,6%)
5	Aluguel, transporte, luz e comunicação	52	204.242	Aluguel de habitação (34,3%) Transporte urbano (8,8%) Conta de energia elétrica (6,5%) Serviço e aparelho para comunicação – exceto internet (6,0%)
6	Veículo	84	288.069	Veículo – aquisição (32,7%)
7	Alimentação dentro de casa	337	1.179.839	Alimentação dentro de casa (29,1%) Transporte urbano (5,5%) Conta de energia elétrica (5,4%)
8	Comunicação, luz e transporte	125	471.019	Serviço e aparelho para comunicação – exceto internet (18,3%) Conta de energia elétrica (9,9%) Transporte urbano (8,5%)
9	Contas domésticas	57	172.111	Conta de energia elétrica (39,9%) Conta de gás, aquisição de bujão de gás (10,2%)
10	Medicamentos e luz	84	294.335	Remédio e material de tratamento (23,1%) Conta de energia elétrica (9,4%)

Tabela 5.J: Informações adicionais segundo os tipos de famílias – Método com desenho (continua)

Características	Tipo 2: Gasto veicular, comunicação e luz	Tipo 6: Veículo	Tipo 3: Educação regular, gasto veicular e comunicação	Tipo 1: Gasto veicular, comunicação e luz	Tipo 8: Comunicação, luz e transporte	Tipo 10: Medicamentos e luz	Tipo 7: Alimentação dentro de casa	Tipo 5: Aluguel, transporte, luz e comunicação	Tipo 4: Transporte urbano e luz	Tipo 9: Contas domésticas
Renda <i>per capita</i> (emR\$)	2.717	1.927	1.596	1.160	898	596	517	419	353	337
INFORMAÇÕES - DOMICÍLIO:										
Número médio de moradores	3	4	4	4	3	2	3	3	3	3
Domicílio com condição de próprio (já quitado)	60,0%	53,4%	58,1%	78,9%	80,0%	82,1%	65,0%	0,0%	77,0%	64,9%
Número médio de cômodos	8	7	8	6	6	6	5	4	5	5

Tabela 5.J: Informações adicionais segundo os tipos de famílias – Método com desenho (continua)

Características	Tipo 2: Gasto veicular, comunicação e luz	Tipo 6: Veículo	Tipo 3: Educação regular, gasto veicular e comunicação	Tipo 1: Gasto veicular, comunicação e luz	Tipo 8: Comunicação, luz e transporte	Tipo 10: Medicamentos e luz	Tipo 7: Alimentação dentro de casa	Tipo 5: Aluguel, transporte, luz e comunicação	Tipo 4: Transporte urbano e luz	Tipo 9: Contas domésticas
Renda <i>per capita</i> (emR\$)	2.717	1.927	1.596	1.160	898	596	517	419	353	337
INFORMAÇÕES - CONDIÇÕES DE VIDA:										
Aproveitamento da renda no mês com algum grau de dificuldade	72,9%	79,0%	75,0%	83,0%	92,4%	92,5%	90,9%	96,2%	92,7%	92,7%
Quando a quantidade de alimentos consumidos pela família nunca é totalmente suficiente	33,3%	24,7%	32,1%	31,1%	54,6%	58,8%	64,9%	57,7%	58,9%	56,4%
Não possuem a rua pavimentada	11,0%	13,1%	16,1%	31,2%	25,6%	44,1%	36,5%	9,6%	45,2%	61,4%
Não tem serviço de água	1,0%	8,6%	7,1%	11,3%	3,4%	15,0%	19,4%	7,7%	23,4%	21,8%
Não tem coleta de lixo	2,1%	3,7%	0,0%	12,3%	3,4%	18,8%	13,3%	0,0%	7,3%	32,7%
Não tem iluminação na rua	3,1%	7,4%	3,6%	11,3%	10,1%	21,3%	20,3%	0,0%	14,5%	36,4%
Não tem serviço de drenagem e escoamento da água da chuva	7,3%	13,6%	10,7%	33,0%	13,5%	37,5%	39,4%	13,5%	31,5%	54,6%
Percepção das condições de moradia da família (avaliação modal)	72,9%	65,4%	75,0%	54,7%	50,4%	47,5%	41,5%	48,1%	38,7%	41,8%
Atrasou nos últimos 12 meses o aluguel ou prestação da residência	4,2%	6,2%	3,6%	0,9%	4,2%	2,5%	3,3%	30,8%	2,4%	5,5%
Atrasou nos últimos 12 meses despesas com água, eletricidade, gás.	22,9%	29,6%	35,7%	19,8%	29,4%	26,3%	38,8%	32,7%	25,8%	29,1%
Atrasou no pagamento de prestações de bens ou serviços adquiridos	16,7%	19,8%	28,6%	22,6%	28,6%	17,5%	27,6%	25,0%	18,6%	12,7%

Tabela 5.J: Informações adicionais segundo os tipos de famílias – Método com desenho (conclusão)

Características	Tipo 2: Gasto veicular, comunicação e luz	Tipo 6: Veículo	Tipo 3: Educação regular, gasto veicular e comunicação	Tipo 1: Gasto veicular, comunicação e luz	Tipo 8: Comunicação, luz e transporte	Tipo 10: Medicamentos e luz	Tipo 7: Alimentação dentro de casa	Tipo 5: Aluguel, transporte, luz e comunicação	Tipo 4: Transporte urbano e luz	Tipo 9: Contas domésticas
Renda <i>per capita</i> (emR\$)	2.717	1.927	1.596	1.160	898	596	517	419	353	337
INFORMAÇÕES - CHEFE DA FAMÍLIA:										
Razão de sexo	0,8	1,0	1,0	1,1	0,9	1,0	0,9	0,6	0,8	0,7
Idade: Média	37	29	31	32	33	46	29	32	26	33
Mediana	36	28	27	31	30	48	26	32	25	26
Nível de instrução predominante	Ensino superior (36,31%)	Ensino fundamental (33,33%)	Ensino superior (46,15%)	Ensino fundamental (42,20%)	Ensino fundamental (52,23%)	Ensino fundamental (55,45%)	Ensino fundamental (59,14%)	Ensino fundamental (55,91%)	Ensino fundamental (57,08%)	Ensino fundamental (60,53%)
Cor ou Raça: Branca	81,0%	75,6%	79,5%	70,7%	56,4%	70,9%	41,2%	56,7%	46,2%	48,0%
Preta ou parda	18,2%	24,4%	20,5%	28,2%	42,6%	29,1%	58,4%	40,2%	53,3%	52,0%
Tem cartão de crédito	43,2%	37,5%	42,7%	28,2%	23,3%	7,7%	13,0%	16,5%	9,2%	5,3%
Tem cheque especial	36,6%	30,2%	33,3%	22,0%	9,7%	9,1%	4,7%	2,4%	2,8%	2,0%
É titular de plano de saúde	38,0%	29,2%	35,9%	20,2%	14,9%	11,8%	7,6%	7,9%	5,4%	4,0%
Renda total mensal:										
Faixa: Até 8 SM	13,0%	21,7%	11,1%	38,7%	57,9%	78,2%	77,5%	82,7%	86,8%	92,1%
Acima de 8 SM	87,0%	78,4%	88,9%	61,3%	42,1%	21,8%	22,5%	17,3%	13,2%	7,9%
Valor: Média (em R\$)	8.061	4.208	5.006	3.252	2.136	1.231	1.274	949	1.036	817
Mediana (em R\$)	4.373	2.588	4.941	2.396	1.399	820	904	849	827	522

No **tipo 02**, a renda *per capita* mensal estimada foi de R\$2.717, sendo três o número médio de moradores. Na comparação com os outros tipos, este foi o que concentrou os maiores percentuais de gasto com plano de saúde e odontológico (5,42%), serviços de empregado doméstico (4,88%), serviços e taxa de habitação (4,47%) e serviço e tratamento médico e odontológico (3,66%). Despesas com curso regular, plano médico, comunicação, empregado doméstico, serviço e taxa de habitação, luz, medicamento, roupa, combustível e leite, queijo e derivados representaram mais da metade dos gastos familiares com despesas de consumo. Este tipo também se destacou por ter sido o que registrou os maiores percentuais de gastos em jogos e brinquedos (0,65%), computador e serviços de acesso à internet (0,64%), jóias e bijuterias (0,50%). As famílias residem, com maior frequência, em construção com, em média, oito cômodos. Em mais da metade das famílias (66,67%) declarou-se que a quantidade de alimentos consumidos seja sempre suficiente no mês. O tipo se destacou por ter se revelado como o que mais acessa a serviços públicos na residência, ou próximos a ela, como rede de água, serviços de coleta de lixo, iluminação de rua e serviços de drenagem e escoamento da água da chuva. Os chefes são, com maior frequência, do sexo feminino, de cor predominantemente branca (80,98%) e com nível superior completo (36,31%). Os índices de posse de cartão de crédito, cheque especial e titularidade de plano de saúde foram os mais elevados.

O **tipo 06** foi o que concentrou o maior percentual de gasto com aquisição de veículo. Este chegou a passar de 30% do total de gastos das famílias, enquanto que, nos demais tipos, essa despesa não chegou a atingir 2,5%. Gastos com veículos, combustível, comunicação (exceto internet), conta de energia elétrica, curso de formação regular e transporte urbano representaram aproximadamente metade dos gastos. Este tipo também se destacou por ter sido o que menos gastou, proporcionalmente, com transporte urbano (2,45%). A renda *per capita* estimada das famílias foi de R\$1.927, a segunda mais elevada, sendo quatro o número médio de moradores. O tipo reuniu um elevado percentual de famílias que consideram que a quantidade

de alimentos consumidos no mês seja sempre suficiente (75,31%). Os chefes de domicílio são jovens, com 29 anos, em média, de cor branca (75,60%) e com elevados índices de posse de cartão de crédito, cheque especial e titularidade de plano de saúde.

O **tipo 03** foi o que concentrou não somente o maior percentual de gasto com curso de educação regular (14,27%), mas também com educação como um todo, incluindo cursos não regulares, livros e gastos como materiais escolares e serviços de papelaria. O dispêndio com educação chegou a, aproximadamente, 18% do total de gastos das famílias. Despesas com cursos regulares, combustível, autopeças e gastos com veículo próprio, serviços de comunicação, conta de energia elétrica, plano de saúde e odontológico, serviços e taxas de habitação e transporte urbano representaram, aproximadamente, metade dos gastos. O tipo também foi o que mais gastou com viagem, incluindo transporte e hospedagem (2,71%). As famílias residem, com maior frequência, em construção de, em média, oito cômodos. A renda *per capita* mensal estimada foi de R\$1.596, sendo quatro o número médio de moradores. Em mais da metade das famílias (67,86%) declarou-se que a quantidade de alimentos consumidos seja sempre suficiente no mês. As condições de moradia foram consideradas como boas por 75,00% e o índice de inadimplência em contas domiciliares foi o mais elevado (35,71%). Os chefes são de cor predominantemente branca (78,49%) e com nível superior completo (46,15%). Os índices de posse de cartão de crédito, cheque especial e titularidade de plano de saúde foram uns dos mais elevados.

O **tipo 01** foi o que concentrou o maior percentual de gasto com combustível, atingindo pouco mais de 20% do total de gastos de consumo. Tal despesa, juntamente com comunicação, outros gastos com veículos, conta de energia elétrica, transporte urbano, roupa, remédio e material de tratamento, alimentação fora de casa (almoço e jantar) e aquisição de veículo foram responsáveis por mais da metade dos gastos de consumo. Suas famílias residem, com maior frequência, em construção tipo casa (85,32%), com, em média, seis cômodos. A renda *per*

capita mensal estimada foi de R\$1.160, sendo quatro o número médio de moradores. Em mais da metade das famílias (67,92%) declarou-se que a quantidade de alimentos consumidos seja sempre suficiente no mês. O grupo se destacou por ter se revelado o menos inadimplente no pagamento de aluguel de habitação e contas domiciliares. O grupo uniu famílias cujos chefes são, com maior frequência, do sexo masculino, de cor branca (70,70%) com idade média de 32 anos.

O **tipo 08** foi o que concentrou o maior percentual de gasto com serviços e aparelhos para comunicação, exceto internet. Este tipo de despesa inclui serviços de telefone residencial, móvel e público, TV por assinatura, entre outros e chegou a representar, aproximadamente, 18% do total de gastos familiares. Nos demais nove tipos, serviços de comunicação não ultrapassaram 6,1% dos gastos. Despesas com serviços e aparelhos para comunicação, conta de energia elétrica, transporte urbano, remédio e material de tratamento, conta de água, roupa, pão e bolo de padaria e gás representaram aproximadamente metade dos gastos dessas famílias. A renda *per capita* mensal estimada foi de R\$898 sendo três o número médio de moradores. O tipo reuniu um elevado percentual de famílias que atrasaram em prestações de bens ou serviços adquiridos nos últimos 12 meses (28,57%). Os chefes de domicílio são, com maior frequência, do sexo feminino com 33 anos, em média.

O **tipo 10** foi o que concentrou o maior percentual de dispêndio com remédio e material de tratamento. Este chegou a, aproximadamente, 23% do total de gastos de consumo. Nos demais tipos analisados, este gasto não chegou a representar 3,5% das despesas de consumo. Remédio e material de tratamento, juntamente com conta de energia elétrica, gás, comunicação (exceto internet), pão e bolo de padaria e transporte urbano representaram aproximadamente metade dos gastos domiciliares. É o tipo também que mais gastou, proporcionalmente, com animais domésticos (1,01%). As famílias residem, com maior frequência, em construção de apenas cinco cômodos. A renda *per capita* mensal estimada foi de R\$596. É com muita

dificuldade que grande parte dessas famílias aproveitam a renda ao longo do mês (caso de 40,00%). Declararam, também, que a quantidade de alimentos consumidos no mês às vezes não chega a ser suficiente (42,50%). O percentual dos domicílios que se localizam em ruas sem pavimentação foi um dos mais elevados, assim com os que não tem acesso à serviços públicos de abastecimento de água, coleta de lixo, entre outros. Os chefes de domicílio são, em média, mais velhos, com 50% deles com idade superior a 48 anos. São, predominantemente, do sexo feminino e de cor branca.

O **tipo 07** foi o que apresentou maior diversificação dos gastos, ou seja, não concentrou em nenhum item específico, grande percentual da despesa. Quando a análise é feita de forma mais ampla, observou-se que este foi o que mais, proporcionalmente, gastou com itens de alimentação dentro de casa. Este gasto, no caso, chegou a representar, aproximadamente, 29% do total. Despesas com transporte urbano, luz, leite e derivados, açúcar e sobremesa, carne vermelha, cereal, conta de gás, artigos e equipamentos para higiene pessoal, pão e bolo de padaria, comunicação (exceto internet), bebidas não alcoólicas, aves, ovos, leguminosas, hortaliças e frutas representaram aproximadamente metade dos gastos familiares. A renda *per capita* mensal estimada foi de R\$517. O tipo reuniu um elevado percentual de famílias com falta de acesso a serviços públicos e com o mais alto índice de inadimplência em contas domiciliares como água, energia elétrica e gás. Os chefes de domicílio são jovens, com 29 anos, em média, de cor, predominantemente, preta ou parda (58,39%).

O **tipo 05** foi o que concentrou o maior percentual de gasto com aluguel de habitação, que chegou a quase 34%, aproximadamente, do total de gastos das famílias, enquanto que, nos demais tipos obtidos, este tipo de despesa de consumo não chegou a ultrapassar 2% dos gastos. Despesas com aluguel, transporte urbano e conta de energia elétrica representaram aproximadamente metade dos dispêndios. A renda *per capita* mensal estimada das famílias foi de R\$419, uma das mais baixas registradas, sendo três o número médio de moradores. As

condições de moradia foram consideradas apenas como satisfatórias por 48,08%. O índice de inadimplência em aluguel de habitação, nos últimos 12 meses, foi o mais elevado (30,77%). O tipo uniu famílias cujos chefes são, maior frequência, do sexo feminino, com idade média de 32 anos.

O **tipo 04** concentrou o maior percentual de despesa com transporte urbano. Este chegou a, aproximadamente, 32% do total de gastos de consumo. Nos demais tipos obtidos, esse gasto não chegou a ultrapassar 9%. Despesas com transporte urbano, energia elétrica, comunicação (exceto internet), gás, pães e bolo de padaria, representaram mais da metade dos gastos familiares. O tipo 04, também, se destacou por ter registrado o maior percentual de gasto com roupas masculinas e femininas (3,05%). As famílias reunidas residem, com maior frequência, em construção com cinco cômodos, em média. A renda *per capita* mensal estimada foi de apenas R\$353, uma das mais baixas registradas, sendo três o número médio de moradores. Em somente 39,52% declarou-se que a quantidade de alimentos consumidos seja sempre suficiente no mês. Os chefes são, com maior frequência, do sexo feminino, de cor predominantemente preta ou parda e com apenas 26 anos, em média. Os índices de posse de cartão de crédito, cheque especial e plano de saúde foram uns dos mais baixos.

O **tipo 09** foi o tipo que concentrou o maior percentual de gasto com contas domiciliares como energia elétrica, gás e também água, representando juntos, aproximadamente, 53% dos gastos. A questão da habitação é bem marcante no tipo, já que este reuniu famílias que não possuem gasto com aluguel especificamente, mas gastam bastante com habitação (gastos que chegaram a representar 55% das despesas, quando, juntamente, foram consideradas contas domésticas, despesas com móveis e eletrodomésticos, manutenção, taxas e outros serviços de habitação). Este tipo é também o que mais gasta, proporcionalmente, com fumo (3,07%). A renda *per capita* mensal estimada foi de R\$337, a mais baixa registrada. É com muita dificuldade que grande parte dessas famílias aproveitam a renda ao longo do mês (caso de

45,45%). O percentual dos domicílios que se localizam em ruas com pavimentação é o menor entre todos os tipos analisados (38,60%), assim como o percentual dos que atrasaram no pagamento de prestações de bens ou serviços adquiridos (12,73%). Os chefes de domicílio são, predominantemente, do sexo feminino, de cor preta ou parda, recebendo até oito salários mínimos (92,11% dos casos). Os níveis de aquisição de cartão de crédito, cheque especial e titularidade de plano de saúde configuraram como os mais baixos.

5.4 Resultados comparativos

A Tabela 5.K compara os resultados da análise de componentes principais para os dois métodos. Pode-se observar que todos os autovalores mudaram entre os métodos e que a porcentagem de variância explicada pelo método com desenho é maior para os primeiros quatro fatores.

Tabela 5.K: Comparação dos autovalores e da porcentagem de explicação da variabilidade total entre os dois métodos

Componentes	Análise da comparativa dos métodos							
	Método ingênuo			Método com desenho			Diferença entre os autovalores	
	Autovalor	Proporção da variabilidade	Acumulado	Autovalor	Proporção da variabilidade	Acumulado	Pontos percentuais	Relativa
1	130,2	13,2%	13,2%	137,8	14,4%	14,4%	7,6	5,9%
2	103,4	10,5%	23,7%	101,6	10,6%	24,9%	-1,8	-1,8%
3	76,2	7,7%	31,5%	80,6	8,4%	33,3%	4,4	5,7%
4	65,7	6,7%	38,1%	68,2	7,1%	40,4%	2,5	3,9%
5	54,2	5,5%	43,6%	50,9	5,3%	45,7%	-3,2	-6,0%
6	44,2	4,5%	48,1%	41,5	4,3%	50,1%	-2,7	-6,1%
7	38,6	3,9%	52,0%	36,3	3,8%	53,8%	-2,3	-5,9%
8	31,9	3,2%	55,3%	32,9	3,4%	57,3%	0,9	3,0%
9	27,5	2,8%	58,1%	26,6	2,8%	60,0%	-0,9	-3,3%
10	25,0	2,5%	60,6%	25,7	2,7%	62,7%	0,7	2,8%

Fonte: Resultados dos procedimentos usados nos dois métodos.

Outra forma de avaliar a diferença entre os dois métodos consiste em verificar o número de famílias que ficaram em um mesmo tipo nas tipologias elaboradas pelos dois métodos (Tabela 5.L). Observa-se que o tipo 2, da tipologia com método ingênuo, foi considerado como o menos homogêneo pelo método com peso, sendo desmembrado em dois tipos e com 11 de suas famílias transferidas para o tipo 7, além de quatro famílias alocadas a outros tipos. Além disso, o tipo 3 (Alimentação fora de casa e transporte) foi formado apenas no método ingênuo. Ao considerar o método com desenho, este tipo não foi definido.

Tabela 5.L: Distribuição do número de domicílios segundo as tipologias elaboradas pelos dois métodos – dados amostrais

Método Ingênuo		Tipos de famílias											Tipo não-definido	TOTAL	
		Tipo4: Gasto veicular, comunicação e luz	Tipo2: Educação regular, comunicação e plano médico	Tipo9: Transporte urbano	Tipo7: Aluguel, transporte, luz e comunicação	Tipo1: Veículo	Tipo10: Alimentação dentro de casa	Tipo8: Comunicação, luz e transporte	Tipo11: Comunicação, luz e transporte	Tipo8: Contas domésticas	Tipo6: Medicamentos e luz	Tipo3: Alimentação fora de casa e transporte			
Método com desenho															
Tipos de famílias	Tipo1: Gasto veicular, comunicação e luz	105	1	0	0	0	0	0	0	0	0	0	0	3	109
	Tipo2: Educação regular, plano médico e comunicação	0	98	0	0	0	0	0	0	0	0	0	0	2	100
	Tipo3: Educação regular, gasto veicular e comunicação	0	28	0	0	0	0	0	0	0	0	0	0	3	31
	Tipo4: Transporte urbano	0	0	96	0	0	0	0	0	0	0	0	0	30	126
	Tipo5: Aluguel, transporte, luz e comunicação	0	0	0	51	0	0	0	0	0	0	0	0	1	52
	Tipo6: Veículo	0	0	0	0	82	0	0	0	0	0	0	0	2	84
	Tipo7: Alimentação dentro de casa	0	11	0	0	0	301	0	1	0	2	0	0	22	337
	Tipo8: Comunicação, luz e transporte	0	1	0	0	0	0	74	19	0	0	0	1	30	125
	Tipo9: Contas domésticas	0	0	0	0	0	0	0	0	57	0	0	0	0	57
	Tipo10: Medicamento e luz	0	2	0	0	0	0	0	0	0	66	0	0	16	84
Tipo não-definido		1	26	0	22	3	7	0	2	1	2	67	49	180	
TOTAL		106	167	96	73	85	308	74	22	58	70	68	158	1285	

A Tabela 5.M apresenta dados semelhantes aos da Tabela anterior, indicando que o uso de pesos desiguais não implicou alteração relevante nas estimativas populacionais.

Tabela 5.M: Estimativa do número de domicílios segundo as tipologias elaboradas pelos dois métodos

Método Ingênuo		Tipos de famílias										TOTAL	
		Tipo4: Gasto veicular, comunicação e luz	Tipo2: Educação regular, comunicação e plano médico	Tipo9: Transporte urbano	Tipo7: Aluguel, transporte, luz e comunicação	Tipo1: Veículo	Tipo10: Alimentação dentro de casa	Tipo5: Comunicação, luz e transporte	Tipo11: Comunicação, luz e transporte	Tipo8: Contas domésticas	Tipo6: Medicamentos e luz		Tipo3: Alimentação fora de casa e transporte
Método com desenho													
Tipos de famílias	Grupo1: Gasto veicular, comunicação e luz	370.240	1.490	0	0	0	0	0	0	0	0	0	371.730
	Grupo2: Educação regular, plano médico e comunicação	0	291.882	0	0	0	0	0	0	0	0	0	291.882
	Grupo3: Educação regular, gasto veicular e comunicação	0	64.455	0	0	0	0	0	0	0	0	0	64.455
	Grupo4: Transporte urbano	0	0	373.297	0	0	0	0	0	0	0	0	373.297
	Grupo5: Aluguel, transporte, luz e comunicação	0	0	0	199.749	0	0	0	0	0	0	0	199.749
	Grupo6: Veículo	0	0	0	0	282.764	0	0	0	0	0	0	282.764
	Grupo7: Alimentação dentro de casa	0	40.399	0	0	0	1.057.197	0	4.688	0	6.766	0	1.109.050
	Grupo8: Comunicação, luz e transporte	0	3.099	0	0	0	0	269.664	68.645	0	0	2.968	344.375
	Grupo9: Contas domésticas	0	0	0	0	0	0	0	0	172.111	0	0	172.111
	Grupo10: Medicamento e luz	0	7.964	0	0	0	0	0	0	0	242.138	0	250.102
TOTAL		370.240	409.289	373.297	199.749	282.764	1.057.197	269.664	73.333	172.111	248.904	2.968	3.459.516

CAPÍTULO 6

Conclusões

Esta dissertação procurou, no quadro referencial de uma estratégia para projetar a demanda de bens e serviços comercializados por um conjunto de setores, que interessa direta ou indiretamente ao Instituto Fecomércio-RJ, encontrar uma alternativa para estabelecer uma tipologia de famílias, construída com base na estrutura das despesas monetárias de consumo observadas pela Pesquisa de Orçamentos Familiares 2002-2003 (POF), realizada pelo IBGE.

Foram elaboradas duas tipologias, uma empregando um método ingênuo que desconsidera as probabilidades de seleção desiguais e o caráter complexo do desenho de amostra da POF, e outra que considera os pesos naturais do desenho da amostra e as informações estruturais do plano de amostragem da POF.

Os resultados obtidos indicam que houve alteração importante em pelo menos dois tipos de famílias definidos pelo método ingênuo, quando a estrutura complexa do desenho da amostra é levada em consideração (método com desenho). Excetuando-se estas situações, observou-se uma boa concordância entre os dois métodos utilizados no que diz respeito às tipologias dos domicílios formadas.

Deve-se indicar também, que o método que seria o mais indicado para a construção de uma tipologia com base nos dados da POF, que utilize os pesos amostrais calibrados e considere as fontes de aleatorização derivadas do desenho complexo da amostra e das regressões de calibração, não pôde ser realizado por problemas de ordem operacional referidos ao longo desta dissertação.

Desta constatação, surge como sugestão para trabalhos futuros a construção de uma tipologia de famílias que considere todas as fontes de aleatorização da amostra e da expansão da POF. Além disso, a análise de dados diferenciada para as famílias residentes na Região

Metropolitana e nas demais regiões do estado (o interior), seria relevante para a diferenciação de perfis de consumo e demanda que podem estar relacionados diretamente ao local de residência.

Algumas limitações do método foram observadas ao longo da realização deste trabalho. A primeira deriva da forma como as informações são obtidas nas pesquisas sobre orçamentos familiares e refere-se à utilização exclusiva das despesas monetárias, ignorando que gastos não-monetários podem ser resultantes de obtenção monetária fora do âmbito da família pesquisada.

Outra limitação, decorrente do método empregado em pesquisas domiciliares, é relativa ao consumo feito fora da região de moradia. Este tipo de consumo tem impacto importante na projeção da demanda de regiões de veraneio, como no caso da baixada litorânea fluminense, cuja população flutuante representa importante parcela da demanda.

Referências bibliográficas

- ACOSTA, A.S., BARROS, A.C.S.A. & PESKE, S.T. Tipologias e composto mercadológico de empresas de sementes de trigo e soja do Rio Grande do Sul. *Revista Brasileira de Sementes*, vol. 24, nº. 1, p.81-90, 2002 (LOCAL)
- ANDERBERG, M.R. *Cluster analysis for application*. New York: Academic Press, 1973.
- ANDRADE, D.M., OLIVEIRA, J.L.R. & ANTONIALI, L.M. O perfil de clientes de um shopping center: um estudo exploratório com consumidores do interior. *Organizações Rurais e Agroindustriais* – v.6 – n.2 – julho/dezembro 2004 (LOCAL)
- ARANHA, F. Modelagem do potencial de mercado dos municípios paulistas. Relatório de pesquisa, 12. EAESP/FGV/NPP – Núcleo de Pesquisas e Publicações, 1998.
- ARAÚJO, M. G. Caracterização sócio-econômica e tecnológica dos piscicultores da região central do estado do Tocantins, utilizando técnicas de análise multivariada. Tese de Doutorado em Zootecnia – área de concentração em Produção de Monogástrico. Universidade Federal de Lavras. Minas Gerais, 2006.
- AUTRET, M. & GANZIN, M., 1952. *Cours de Nutrition et Alimentation Tropicale*. Marseille: Organisation des Nations Unies pour l’Agriculture et l’Alimentation / Organisation Mondiale de la Santé.
- BRITO, L.T.L., SILVA, A.S., SRINIVASAN, V.S., GALVÃO, C.O. & GHEYI. Uso de análise multivariada na classificação das fontes hídricas subterrâneas da Bacia Hidrográfica do Salitre. Associação Brasileira de Engenharia Agrícola. Departamento de Engenharia Rural FCAV/UNESP. Jaboticabal, v.26, n.1, p.36-44, jan./abr. 2006.
- CERQUEIRA, A. Tipologia e características dos estabelecimentos escolares brasileiros. Tese de Doutorado em Demografia. Centro de Desenvolvimento e Planejamento Regional da Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais. Minas Gerais, 2004. UFMG/C EDEPLAR
- Cochran W.G. *Sampling techniques*. 3rd Ed. New York: John Wiley & Sons, 1977.
- DIEESE (Departamento Intersindical de Estatísticas e Estudos Sócio-Econômicos). Estrutura dos orçamentos domésticos - Comparação das estruturas orçamentárias nas últimas décadas. São Paulo. <http://www.dieese.org.br/pof/pof3.xml> (acessado em 05.04.2005).

- ELKINS, Z. Opinião Pública, Apr. 2000, vol.6, no.1, p.109-136 – Quem iria votar? Conhecendo as conseqüências do voto obrigatório no Brasil. Campinas: Universidade Estadual de Campinas, 2000.
- FECOMÉRCIO-RJ (Federação do Comércio do Estado do Rio de Janeiro). Pelo desenvolvimento do interior. Reportagem de Consuelo Sánchez. REVISTA SISTEMA Fecomércio Rio de Janeiro,65(9). Rio de Janeiro: FECOMÉRCIO-RJ, 2004.
- FGV (Fundação Getúlio Vargas), 1975. *Pesquisa sobre Consumo Alimentar*. v.I (Conclusões e Análise da Pesquisa), v.II (Metodologia), v.III (Tabelas de Composição dos Alimentos). Rio de Janeiro: FGV.
- FLECK, Marcelo P.A. & BOURDEL, Marie C. Revista de Saúde Pública, 32 (3): 267-72 – Método de simulação e escolha de fatores na análise dos principais componentes. São Paulo: USP - Faculdade de Saúde Pública, 1998.
- FRANÇOIS, P.J. Projeção da demanda alimentar. Notas de aula, Montpellier: Université de Montpellier I, 1977.
- HAIR, J.R.J.F., ANDERSON, R.E., TATHAM, R.L. & BLACK, W.C. Análise Multivariada de Dados; tradução. Adonai Schlup Sant'Anna e Anselmo Chaves Neto. Porto Alegre: Bookman, 2005.
- IBGE (Fundação Instituto Brasileiro de Geografia e Estatística). Pesquisa de Orçamentos Familiares 2002-2003: Primeiros Resultados, Brasil e Grandes Regiões. Rio de Janeiro: IBGE, 2004a.
- IBGE (Fundação Instituto Brasileiro de Geografia e Estatística). Pesquisa de Orçamentos Familiares 2002-2003: Análise de Resultados. Texto disponibilizado na segunda divulgação dos microdados, Rio de Janeiro: IBGE, 2004b.
- IBGE (Fundação Instituto Brasileiro de Geografia e Estatística). Pesquisa de Orçamentos Familiares 2002-2003: Aspectos de Amostragem. Texto disponibilizado na segunda divulgação dos microdados, Rio de Janeiro: IBGE, 2004c.
- IPARDES (Instituto Paranaense de Desenvolvimento Econômico e Social). Tipologia dos municípios paranaenses segundo indicadores socioeconômicos e demográficos. Curitiba: IPARDES, 2003.

- IPDC (Instituto de Pesquisa e Desenvolvimento do Comércio). Pesquisa conjuntural do comércio de Fortaleza - abril de 2004. Fortaleza: Federação do Comércio do Estado do Ceará - Fecomércio-CE, 2004.
- IPEA/IBGE/NESUR/SEADE (Instituto de Pesquisas Econômicas Aplicadas / Fundação Instituto Brasileiro de Geografia e Estatística / Núcleo Economia Social, Urbana e Regional do Instituto de Economia da Universidade Estadual de Campinas / Fundação Sistema Estadual de Análise de Dados). Caracterização e tendências da rede urbana do Brasil: Redes Urbanas Regionais: Sudeste. Brasília: IPEA, 2001.
- JOHNSON, R.A. & WICHERN, D.W. *Applied Multivariate Statistical Analysis. 4th edition.* New Jersey: Prentice Hall, 1988.
- LEMOS, M., MORO, S., BIAZI, E. e CROCCO, M. A dinâmica urbana das regiões metropolitanas brasileiras. In: Anais do XXIX Encontro Nacional de Economia, Salvador, 2001.
- LLANILLO, R.F., GROSSI, M.E.D., SANTOS, F.O., MUNHOS, P.D. & GUIMARÃES, M.F. Regionalização da agricultura no estado do Paraná, Brasil. *Ciência Rural*, janeiro-fevereiro, ano/vol. 36, numero 001. Universidade Federal de Santa Maria. Santa Maria, RS, Brasil. Pp. 120-127
- NETO, F.C. Consumo e Renda: Implicações estocásticas. Dissertação de Mestrado em Economia e Sociologia. Departamento de Economia, Universidade Federal de Pernambuco. Recife: UFPE, 1990.
- PEREIRA, Fabiano M. Cidades Médias Brasileiras: uma tipologia a partir de suas (des)economias de aglomeração. Belo Horizonte: UFMG/CEDEPLAR, tese de mestrado, 2002, pp117.
- PESSOA, D.G.C. & SILVA, P.L.N. *Análise de Dados Amostrais Complexos.* São Paulo: Associação Brasileira de Estatística, 1998.
- PESSOA, D.G.C. Estimadores de Calibração em Pesquisas Amostrais. Relatório Técnico, Coordenadoria de Métodos e Qualidade, Diretoria de Pesquisas, Fundação Instituto Brasileiro de Geografia e Estatística. Rio de Janeiro: IBGE, 2005.

- PRAIS, S. J. & HOUTHAKKER, H. S., 1955. *The Analysis of Family Budgets with an Application to two British Surveys Conducted in 1937-9 and their Detailed Results*. Cambridge: Cambridge University Press.
- Registro Anual de Informações Sociais - RAIS 2003. Ministério do Trabalho e Emprego. Classificação das atividades econômicas.
- SANTOS, A.M.S.P. *Economia, espaço e sociedade no Rio de Janeiro*. Rio de Janeiro: Fundação Getúlio Vargas, 2003.
- SANTOS, A.O. *O comércio e o novo milênio*. Rio de Janeiro: Confederação Nacional do Comércio, 2001.
- SAS (Statistical Analysis System). The SAS On line documentation, Release 8.02., Cary, NC: SAS Institute Inc., 2001.
- SILVA, P.L.N. Calibration estimation: When and why, how much and how. Textos para discussão da Diretoria de Pesquisas, 14. Rio de Janeiro: IBGE; 2004.
- SIMÕES, R., GUIMARÃES, C., GODOY, N., VELLOSO, T., ARAÚJO, T., GALINARI, R., CHEIN, F. Rede urbana da oferta de serviços de saúde: uma análise espacial multivariada para Minas Gerais. In: SEMINÁRIO SOBRE A ECONOMIA MINEIRA, 11, 2004, Diamantina, MG. Anais. Belo Horizonte: UFMG/Cedeplar, 2004. 27p. (Disponível em CDROM)
- SMITH, T.M.F. & HOLMES, D.J. Multivariate analysis. In. Skinner, C.J., Holt D. & Smith T.M.F., *Analysis of Complex Surveys*. Chichester: John Wiley, 1989: 165-190..
- STATISTIC CANADA. *Generalized Estimation System Version 4.0 - Help Guide*. Ottawa: Statistic Canada, 1998.
- VARIAN, H.R. *Microeconomia: princípios básicos*. 5th ed. Tradução de Ricardo Inojosa e Maria José Cyhlar Monteiro. Rio de Janeiro: Campus, 2000.
- VASCONCELLOS, M.T.L. Estudo dos principais tipos alimentares do Brasil. Série de Relatórios de Consultores sobre Nutrição, 81. Roma: FAO, ESN, 1987.

ANEXO I

Relação das variáveis da POF

REGISTRO 01: Registro do domicílio

UF	Unidade da Federação
ID_DOM	Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
FAT_EXP	Fator de expansão 1 (setor) – peso do desenho
V110	Quantidade de moradores
V120	Tipo de domicílio
V130	Quantidade de cômodos
V140	Cômodos servindo de dormitórios
V150	Quantidade de banheiros
V160	Tipo de abastecimento de água
V170	Tipo de escoamento sanitário
V180	Condição ocupação do domicílio
V190	Origem de energia elétrica
V200	Material que predomina no piso
V210	Existe pavimentação na rua
V220	Tempo de moradia
V230	Quantidade de UC, ou seja, de unidade de consumo
V240	Contrato é documentado
V250	Renda total mensal da Unidade de Consumo principal

REGISTRO 02: Registro de pessoas

UF	Unidade da Federação
ID_DOM	Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC	Identificação da unidade de consumo
FAT_EXP	Fator de expansão 1 (setor) – peso do desenho
N_PÉSS	Número do informante
V400	Código de relação com a pessoa de referência
V410	Código da condição de presença
V420	Sexo da pessoa
V430	Dia de nascimento
V440	Mês de nascimento
V450	Ano de nascimento
V460	Idade calculada
V470	Idade calculada em dias
V480	Idade calculada em meses
V490	Frequência à escola
V500	Nível de instrução
V505	Série que frequenta / Última concluída
V510	Anos de estudo
V520	Cor ou raça
V530	Código de religião
V540	Peso

V550	Altura
V560	Orçamento despesa
V570	Orçamento rendimento
V580	Tem cartão de crédito
V590	Tem cheque especial
V600	Titular de plano de saúde
V610	Número de beneficiários do plano de saúde
V620	Renda total mensal da unidade de consumo

REGISTRO 03: Registro condições de vida

UF	Unidade da Federação
ID_DOM	Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC	Identificação da unidade de consumo
FAT_EXP	Fator de expansão 1 (setor) – peso do desenho
N_PESS	Número do informante
V2020	Código da renda familiar
V2030	Código quantidade de alimentos
V2040	Código tipo de alimento consumido
V2050	Código razão não se alimentar
V2060	Código serviço de água
V2070	Código coleta de lixo
V2080	Código iluminação de rua
V2090	Código drenagem e escoamento
V2100	Código energia elétrica
V2110	Código pouco espaço
V2120	Código rua vizinhos barulhentos
V2130	Código casa escura
V2140	Código telhado com goteiras
V2150	Código problemas com umidade
V2160	Código madeiras deterioradas
V2170	Código problemas ambientais
V2180	Código violência
V2190	Código condição de moradia
V2200	Atraso aluguel ou prestação
V2210	Atraso água, eletricidade, gás
V2220	Atraso prestação bens serviços
V2230	Valor mínimo mensal alimentação
V2240	Renda total mensal da unidade de consumo

REGISTRO 04: Registro de inventário

UF	Unidade da Federação
ID_DOM	Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC	Identificação da unidade de consumo
FAT_EXP	Fator de expansão 1 (setor) – peso do desenho
N_QUADRO	Número do quadro
CODIGO	Código do item
QUANT	Quantidade do item

V12020 Última aquisição
ESTADO_AQ Código do estado de aquisição
FORMA Código da forma de obtenção

REGISTRO 05: Registro de despesas de 90 dias - Coletiva

UF Unidade da Federação
ID_DOM Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC Identificação da unidade de consumo
FAT_EXP Fator de expansão 1 (setor) – peso do desenho
N_QUADRO Número do quadro
CODIGO Código do item
LOCAL Código do local de compra
FORMA Código da forma de obtenção
VALOR Valor deflacionado anualizado

REGISTRO 06: Registro de despesas de 12 meses - Coletiva

UF Unidade da Federação
ID_DOM Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC Identificação da unidade de consumo
FAT_EXP Fator de expansão 1 (setor) – peso do desenho
N_QUADRO Número do quadro
CODIGO Código do item
LOCAL Código do local de compra
FORMA Código da forma de obtenção
VALOR Valor deflacionado anualizado

REGISTRO 07: Registro outras despesas

UF Unidade da Federação
ID_UC Identificação da unidade de consumo
FAT_EXP Fator de expansão 1 (setor) – peso do desenho
COD_ITEM Código do item
VALOR_MON Valor monetário do item
VALOR_NAO Valor não-monetário do item

REGISTRO 08: Registro despesas serviços domésticos - Coletiva

UF Unidade da Federação
ID_DOM Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC Identificação da unidade de consumo
FAT_EXP Fator de expansão 1 (setor) – peso do desenho
N_QUADRO Número do quadro
CODIGO Código do item
N_MESES Número de meses
VALOR Valor deflacionado anualizado
V6091 Valor deflacionado anualizado INSS

REGISTRO 09: Registro caderneta de despesa - Coletiva

UF	Unidade da Federação
ID_DOM	Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC	Identificação da unidade de consumo
FAT_EXP	Fator de expansão 1 (setor) – peso do desenho
N_QUADRO	Número do quadro
GRUPO	Número do grupo
CODIGO	Código do item
LOCAL	Código do local de compra
QUANT	Quantidade do item
V7050	Código da unidade de medida
V7060	Código do peso ou volume
FORMA	Código da forma de obtenção
QUANT_KG	Quantidade em kg
VALOR	Valor deflacionado anualizado

REGISTRO 10: Registro de despesas - Individual

UF	Unidade da Federação
ID_DOM	Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC	Identificação da unidade de consumo
FAT_EXP	Fator de expansão 1 (setor) – peso do desenho
N_PESS	Número do informante
N_QUADRO	Número do quadro
CÓDIGO	Código do item
LOCAL	Código do local de compra
FORMA	Código da forma de obtenção
VALOR	Valor deflacionado anualizado

REGISTRO 11: Registro despesas com veículos - Individual

UF	Unidade da Federação
ID_DOM	Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC	Identificação da unidade de consumo
FAT_EXP	Fator de expansão 1 (setor) – peso do desenho
N_PESS	Número do informante
N_QUADRO	Número do quadro
CODIGO	Código do item
LOCAL	Código do local de compra
ESTADO_AQ	Código do estado de aquisição
FORMA	Código da forma de obtenção
VALOR	Valor deflacionado anualizado

REGISTRO 12: Registro de rendimentos e deduções - Individual

UF	Unidade da Federação
ID_DOM	Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
ID_UC	Identificação da unidade de consumo

FAT_EXP Fator de expansão 1 (setor) – peso do desenho
 N_PESS Número do informante
 N_QUADRO Número do quadro
 POS_OCUP Código da posição na ocupação
 CODIGO_REND Código do rendimento bruto
 N_MESES Número de meses recebidos
 VALOR Valor do rendimento deflacionado e anualizado
 VAL_PREV Valor da previdência deflacionado e anualizado
 VAL_IRRF Valor do imposto de renda deflacionado e anualizado
 VAL_OUT_DED Valor de outras deduções deflacionado e anualizado

REGISTRO 13: Registro de outros rendimentos - Individual

UF Unidade da Federação
 ID_DOM Identificação do domicílio, formado pela composição do número seqüencial, divisão do seqüencial e número do domicílio
 ID_UC Identificação da unidade de consumo
 FAT_EXP Fator de expansão 1 (setor) – peso do desenho
 N_PESS Número do informante
 N_QUADRO Número do quadro
 CÓDIGO Código do item
 VALOR Valor do rendimento deflacionado e anualizado
 VALOR_DED Valor da dedução deflacionada e anualizada

ANEXO II

Criação dos códigos de despesa em registros com mais de uma variável de valor

Neste anexo, são descritos os procedimentos adotados para criação dos códigos de despesas em registros com mais de uma variável de valor.

De um modo geral, os códigos de despesa correspondem à concatenação dos valores de duas variáveis: número do quadro (N_QUADRO) e código do item (CODIGO). No entanto, o código do item inclui duas posições finais para identificar diferentes denominações de um mesmo produto (sinonímia regional, ou distintos produtos que são agregados em um mesmo código de despesa). Assim, na criação do código da despesa (COD_DESP), aplicou-se o seguinte algoritmo:

$$\text{COD_DESP} = \text{Parte inteira} [(\text{N_QUADRO} * 100000 + \text{CODIGO}) / 100].$$

No caso do registro de despesas com serviços domésticos, coexistem duas variáveis de valor. A primeira, com o gasto com empregados domésticos (VALOR) e, a segunda, com o valor do INSS do empregador (V6091). Neste caso, aplicou-se o seguinte algoritmo para criar o registro com código e valor do gasto com INSS do empregado, sempre que o valor da V9061 fosse maior que zero:

$$\begin{aligned} \text{COD_DESP} &= \text{Parte inteira} [(\text{N_QUADRO} * 100000 + \text{CODIGO} + 50000) / 100]; \text{ e} \\ \text{VALOR} &= \text{V9061}. \end{aligned}$$

No caso do registro de rendimentos e deduções, a primeira variável de valor não era relevante (era o valor de uma receita e não o de uma despesa). As outras três variáveis indicavam as deduções com impostos de renda (DED_IRRF); contribuições à previdência (DED_PREV), e outras deduções (OUT_DED). Como o código do item de rendimento, neste caso, era criado por $\text{COD_REND} = \text{N_QUADRO} * 1000 + \text{POS_OCUP}$, adotou-se o seguinte algoritmo:

$$\begin{aligned} \text{COD_DESP} &= (\text{N_QUADRO} * 1000 + \text{POS_OCUP} + 3000); \\ \text{VALOR} &= \text{DED_IRRF}; \\ \text{COD_DESP} &= (\text{N_QUADRO} * 1000 + \text{POS_OCUP} + 4000); \\ \text{VALOR} &= \text{DED_PREV}; \\ \text{COD_DESP} &= (\text{N_QUADRO} * 1000 + \text{POS_OCUP} + 5000); \text{ e} \\ \text{VALOR} &= \text{OUT_DED}. \end{aligned}$$

Por fim, no caso do registro de outros rendimentos, a primeira variável de valor não era relevante (era o valor de uma receita e não o de uma despesa). A outra variável (VAL_DED) indicava o valor das deduções associadas com a receita auferida. Neste caso, adotou-se o seguinte algoritmo sempre que VAL_DED era maior que zero:

COD_DESP = Parte inteira [(N_QUADRO * 100000 + CODIGO + 600000) / 100]; e
VALOR = VAL_DED.

ANEXO III

Porcentagem média de cada grupo de despesa em cada tipo de família – Método ingênuo

Descrição	Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8	Grupo9	Grupo10	Grupo11
Açúcar, bala, doce, biscoito e sobremesa	1,20%	1,32%	1,04%	1,43%	1,48%	1,41%	1,57%	1,60%	1,97%	4,15%	2,83%
Alimentação fora de casa - almoço e jantar	1,78%	2,39%	17,06%	2,57%	0,83%	0,70%	1,65%	0,08%	0,57%	0,64%	0,97%
Alimentação fora de casa (exceto almoço e jantar)	1,26%	1,16%	3,07%	1,73%	1,38%	1,51%	1,22%	0,56%	2,87%	2,14%	1,59%
Aluguel de habitação	1,52%	1,42%	1,31%	0,99%	0,43%	0,31%	30,55%	0,00%	0,21%	0,69%	0,23%
Animal doméstico (aquisição, produto, serviço e tratamento)	0,18%	0,23%	0,25%	0,37%	0,35%	0,83%	0,07%	0,50%	0,29%	0,58%	0,31%
Artigo de limpeza	0,53%	0,72%	0,49%	0,45%	0,83%	0,45%	0,38%	0,15%	0,56%	1,80%	0,61%
Artigo doméstico	0,53%	0,75%	1,07%	0,62%	0,60%	0,67%	0,65%	0,40%	0,41%	1,02%	0,25%
Artigo e equipamento para cuidado e higiene pessoal	1,69%	1,93%	1,63%	1,97%	2,31%	1,69%	2,70%	1,16%	1,96%	3,48%	3,49%
Auto Peça, acessório e gasto com veículo(exceto combustível)	2,20%	2,31%	1,20%	5,60%	0,38%	0,24%	0,15%	0,18%	0,09%	0,76%	1,14%
Aves e ovos	0,74%	0,84%	1,30%	1,02%	1,71%	1,42%	1,12%	1,20%	0,88%	3,34%	2,31%
Bebida não alcoólica (refrigerante, suco, chá, mate, café, etc.)	1,13%	1,37%	2,24%	1,57%	1,47%	1,02%	1,14%	2,30%	2,20%	3,23%	2,51%
Brinquedo e jogo	0,54%	0,62%	0,19%	0,50%	0,57%	0,66%	0,30%	0,35%	0,64%	0,56%	0,35%
Calçado e acessório adulto	0,89%	1,07%	1,19%	1,22%	1,05%	1,15%	0,93%	0,92%	1,21%	1,29%	1,32%
Came vermelha	1,09%	1,16%	1,20%	1,88%	1,81%	2,63%	1,01%	0,90%	1,39%	4,11%	3,47%
Cereal (arroz, feijão, grão, etc.)	0,65%	0,44%	0,72%	0,77%	0,94%	1,01%	0,73%	1,30%	1,21%	4,29%	1,23%
Cerveja e chopp	0,92%	0,48%	1,70%	1,22%	0,39%	0,58%	0,78%	0,50%	0,86%	1,93%	2,37%
Combustível	4,54%	3,64%	2,64%	20,33%	1,01%	0,66%	0,67%	0,11%	0,07%	0,81%	0,46%
Computador e acessórios, serviço de acesso à internet	0,64%	0,62%	0,31%	0,36%	0,15%	0,67%	0,02%	0,00%	0,02%	0,09%	0,35%
Conserto e manutenção de artigo doméstico e pessoal	0,53%	0,53%	0,45%	0,14%	0,43%	0,35%	0,32%	0,29%	0,43%	0,76%	0,66%
Conta de água	0,77%	0,71%	1,80%	1,31%	3,63%	3,17%	1,79%	3,29%	2,00%	2,00%	3,29%
Conta de energia elétrica	3,18%	3,86%	5,84%	5,13%	10,94%	10,07%	5,66%	39,72%	8,10%	5,51%	7,70%
Conta de gás, aquisição de bujão de gás	0,98%	0,99%	1,98%	1,73%	2,71%	4,28%	2,40%	10,04%	4,07%	4,12%	2,50%
Curso de formação regular	3,13%	9,52%	0,49%	1,68%	0,32%	0,05%	0,59%	0,18%	0,25%	0,19%	0,23%
Curso não regular	0,58%	1,43%	0,55%	0,44%	0,29%	0,04%	0,14%	0,06%	0,08%	0,23%	0,46%
Diversão, esporte e cultura	1,22%	1,67%	2,12%	1,40%	1,18%	1,50%	0,83%	0,59%	0,92%	1,19%	1,82%
Educação, leitura e gasto escolar	1,07%	1,82%	1,23%	0,96%	1,42%	0,38%	0,94%	0,12%	0,66%	0,68%	0,78%
Eletrodoméstico de grande porte	0,56%	0,56%	0,84%	0,71%	1,52%	1,34%	1,46%	0,34%	1,05%	1,92%	1,96%
Eletrodoméstico de pequeno e médio porte, equipamentos doméstico	0,22%	0,29%	0,39%	0,18%	0,34%	0,20%	0,15%	0,22%	0,46%	0,45%	0,52%
Eletroeletrônico - Equipamento de som, foto e tv	1,18%	1,16%	1,79%	0,84%	1,37%	0,57%	1,02%	1,47%	1,36%	1,88%	2,38%
Empregado doméstico (serviço)	1,32%	4,20%	1,90%	1,83%	0,06%	0,88%	0,27%	0,46%	0,04%	0,45%	0,31%
Enlatado, conserva, alimento preparado e industrializado	0,71%	1,14%	0,80%	0,98%	1,13%	0,60%	1,05%	0,73%	0,78%	2,64%	2,51%
Fumo	0,71%	0,31%	0,85%	0,86%	0,94%	1,10%	1,27%	3,10%	2,22%	2,26%	1,24%
Hospedagem	0,16%	0,75%	0,08%	0,07%	0,10%	0,04%	0,02%	0,03%	0,00%	0,06%	0,21%
Jóia e bijuteria	0,20%	0,41%	0,30%	0,21%	0,23%	0,20%	0,18%	0,03%	0,19%	0,20%	0,33%
Leguminosa, hortaliça e fruta	1,02%	1,46%	1,79%	1,57%	2,32%	2,39%	2,09%	0,67%	1,56%	2,93%	2,15%
Leite, queijo e derivado	1,72%	2,11%	2,14%	1,71%	1,76%	2,54%	2,22%	1,34%	1,46%	4,24%	4,25%
Massa e farinha	0,31%	0,44%	0,31%	0,38%	0,51%	0,33%	0,32%	0,16%	0,58%	1,73%	0,71%
Material de construção	0,45%	1,31%	0,26%	0,92%	0,86%	1,36%	0,37%	2,08%	0,41%	1,54%	0,89%
Material e artigo escolar e serviço de papelaria	0,21%	0,48%	0,33%	0,48%	0,50%	0,61%	0,49%	0,35%	0,59%	0,45%	0,62%
Mobiliário	1,32%	1,35%	2,71%	1,68%	1,14%	0,97%	1,34%	0,53%	0,81%	1,99%	1,70%
Óculos, lentes e material oftalmológico	0,56%	0,46%	0,11%	0,21%	0,49%	0,14%	0,11%	0,02%	0,73%	0,27%	0,44%
Outra bebida alcoólica (exceto cerveja e chopp)	0,14%	0,09%	0,20%	0,06%	0,00%	0,07%	0,09%	0,00%	0,12%	0,46%	0,51%
Outra despesa de consumo	0,64%	0,94%	0,51%	0,46%	0,23%	0,51%	0,34%	0,11%	0,24%	0,57%	0,06%
Outro produto de alimentação (óleo, gordura, margarina, sal, etc.)	0,40%	0,38%	0,18%	0,43%	0,51%	0,47%	0,84%	0,12%	0,40%	2,13%	1,11%
Pão e bolo de padaria	1,06%	1,29%	1,99%	1,74%	2,65%	3,72%	1,65%	4,84%	3,14%	3,53%	2,99%
Peixe e fruto do mar	0,43%	0,37%	0,24%	0,44%	0,72%	0,21%	0,24%	0,93%	0,24%	0,52%	0,56%
Plano de saúde e odontológico	2,18%	5,26%	2,08%	1,82%	1,13%	0,53%	0,69%	0,24%	0,36%	0,47%	0,84%
Refeição pronta, alimento preparado para viagem	0,39%	0,73%	0,02%	0,44%	0,24%	0,22%	0,48%	0,38%	0,35%	0,50%	0,41%
Remédio e material de tratamento	1,99%	3,37%	2,62%	2,92%	4,83%	24,48%	2,94%	2,57%	3,08%	2,67%	0,74%
Roupa feminina e masculina	2,25%	2,90%	2,82%	3,04%	2,82%	2,34%	2,05%	1,85%	2,86%	2,79%	3,40%
Roupa, calçado e acessório para criança	0,83%	0,83%	0,66%	1,09%	1,20%	1,43%	0,93%	1,25%	1,01%	1,54%	1,33%
Seguro em geral (exceto de saúde e odontológico)	0,97%	0,91%	0,30%	0,53%	0,03%	0,06%	0,06%	0,05%	0,01%	0,11%	0,07%
Serviço de reforma e manutenção do imóvel	0,38%	1,25%	0,16%	0,83%	0,50%	1,48%	0,31%	0,36%	0,15%	0,38%	0,00%
Serviço e aparelho para comunicação (exceto internet)	4,45%	5,66%	5,24%	6,05%	20,80%	5,56%	5,47%	1,79%	3,93%	2,84%	13,28%
Serviço e tratamento médico e odontológico	1,23%	3,43%	0,50%	1,21%	0,71%	1,31%	0,62%	0,80%	0,80%	1,17%	0,24%
Serviço para cuidado e higiene pessoal	0,99%	1,53%	1,48%	1,49%	1,57%	1,42%	1,18%	1,69%	1,54%	1,18%	1,61%
Serviço, taxa e outro gasto com habitação	1,45%	4,44%	3,32%	1,02%	0,61%	0,58%	1,38%	0,10%	0,34%	0,49%	0,09%
Transporte com viagem e outro gasto relacionado	0,92%	1,73%	1,04%	0,80%	0,94%	0,51%	0,28%	0,09%	0,45%	0,34%	1,45%
Transporte urbano	2,54%	4,66%	8,98%	3,28%	8,13%	4,23%	9,63%	4,83%	34,74%	5,25%	7,60%
Veículo (aquisição)	32,63%	0,84%	0,00%	2,37%	0,50%	0,16%	0,15%	0,00%	0,09%	0,45%	0,28%
Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

ANEXO IV

Porcentagem média de cada grupo de despesa em cada tipo de família – Método com desenho

Descrição	Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8	Grupo9	Grupo10
Açúcar, bala, doce, biscoito e sobremesa	1,37%	1,57%	1,07%	1,90%	1,50%	1,16%	3,98%	1,52%	1,62%	1,31%
Alimentação fora de casa - almoço e jantar	2,82%	2,17%	2,44%	1,12%	1,42%	1,93%	0,67%	1,65%	0,08%	0,58%
Alimentação fora de casa (exceto almoço e jantar)	1,85%	0,92%	1,36%	2,82%	1,28%	1,27%	2,24%	1,74%	0,57%	1,52%
Aluguel de habitação	1,12%	1,62%	2,00%	0,18%	34,28%	1,54%	0,75%	0,48%	0,00%	0,26%
Animal doméstico (aquisição, produto, serviço e tratamento)	0,37%	0,27%	0,28%	0,24%	0,09%	0,15%	0,55%	0,39%	0,51%	1,01%
Artigo de limpeza	0,44%	0,81%	0,57%	0,47%	0,29%	0,49%	1,70%	0,79%	0,13%	0,41%
Artigo doméstico	0,59%	0,88%	0,43%	0,43%	0,59%	0,53%	0,99%	0,49%	0,40%	1,18%
Artigo e equipamento para cuidado e higiene pessoal	1,90%	2,11%	1,73%	2,08%	2,66%	1,67%	3,47%	2,50%	1,18%	1,71%
Auto Peça, acessório e outro gasto com veículo (exceto combustível)	5,65%	1,56%	6,06%	0,13%	0,09%	2,24%	0,77%	0,67%	0,15%	0,36%
Aves e ovos	1,00%	0,98%	0,69%	1,12%	1,08%	0,72%	3,10%	1,53%	1,16%	1,27%
Bebida não alcoólica (refrigerante, suco, chá, mate, café, etc.)	1,57%	1,46%	1,03%	2,25%	0,92%	1,10%	3,13%	1,75%	2,29%	1,22%
Brinquedo e jogo	0,50%	0,68%	0,65%	0,55%	0,30%	0,53%	0,62%	0,61%	0,35%	0,56%
Calçado e acessório adulto	1,23%	1,15%	0,95%	1,15%	0,77%	0,93%	1,31%	1,06%	0,87%	0,96%
Carne vermelha	1,83%	1,47%	0,60%	1,53%	1,02%	1,07%	3,95%	1,87%	0,78%	2,20%
Cereal (arroz, feijão, grão, etc.)	0,72%	0,49%	0,39%	1,17%	0,65%	0,57%	3,88%	0,92%	1,33%	0,78%
Cerveja e chopp	1,23%	0,59%	0,20%	0,96%	0,04%	0,93%	2,12%	0,85%	0,51%	0,63%
Combustível	20,19%	2,57%	10,92%	0,21%	0,28%	4,52%	0,86%	1,34%	0,12%	0,61%
Computador e acessórios, serviço de acesso à internet	0,35%	0,64%	0,53%	0,04%	0,00%	0,60%	0,12%	0,35%	0,00%	0,56%
Conserto e manutenção de artigo doméstico e pessoal	0,14%	0,64%	0,21%	0,48%	0,27%	0,52%	0,73%	0,49%	0,30%	0,38%
Conta de água	1,31%	0,54%	0,34%	2,27%	1,98%	0,70%	2,06%	2,91%	3,34%	3,15%
Conta de energia elétrica	5,05%	3,49%	4,38%	7,60%	6,52%	3,15%	5,35%	9,91%	39,94%	9,43%
Conta de gás, aquisição de bujão de gás	1,69%	0,90%	0,80%	3,75%	2,48%	0,99%	3,83%	2,55%	10,16%	4,92%
Curso de formação regular	1,65%	9,44%	14,27%	0,31%	0,44%	2,88%	0,22%	0,82%	0,18%	0,12%
Curso não regular	0,49%	1,79%	1,31%	0,10%	0,14%	0,59%	0,29%	0,41%	0,06%	0,08%
Diversão, esporte e cultura	1,52%	1,89%	1,45%	1,11%	0,61%	1,30%	1,23%	1,36%	0,60%	1,22%
Educação, leitura e gasto escolar	0,95%	1,77%	1,67%	0,68%	0,93%	1,05%	0,71%	1,29%	0,13%	0,32%
Eletrodoméstico de grande porte	0,76%	0,56%	0,43%	1,08%	0,93%	0,57%	1,88%	1,32%	0,29%	1,87%
Eletrodoméstico de pequeno e médio porte, equipamentos domésticos	0,20%	0,21%	0,13%	0,55%	0,08%	0,22%	0,48%	0,32%	0,21%	0,13%
Eletroeletrônico - Equipamento de som, fotografia e tv	0,85%	1,33%	1,29%	1,53%	0,51%	1,11%	1,95%	1,72%	1,45%	0,51%
Empregado doméstico (serviço)	1,77%	4,88%	2,30%	0,14%	0,30%	1,44%	0,55%	0,58%	0,47%	1,21%
Enlatado, conserva, alimento preparado e industrializado	0,94%	1,31%	0,82%	0,77%	0,98%	0,63%	2,53%	1,27%	0,74%	0,59%
Fumo	0,84%	0,26%	0,64%	2,19%	1,30%	0,75%	2,21%	0,89%	3,07%	1,64%
Hospedagem	0,07%	0,88%	0,93%	0,00%	0,03%	0,16%	0,09%	0,15%	0,03%	0,00%
Jóia e bijuteria	0,22%	0,50%	0,26%	0,19%	0,16%	0,21%	0,21%	0,21%	0,03%	0,19%
Leguminosa, hortaliça e fruta	1,54%	1,51%	1,02%	1,54%	2,09%	1,03%	2,88%	2,34%	0,63%	2,40%
Leite, queijo e derivado	1,67%	2,43%	1,53%	1,54%	2,07%	1,63%	3,99%	2,09%	1,29%	2,52%
Massa e farinha	0,35%	0,48%	0,43%	0,51%	0,35%	0,30%	1,60%	0,45%	0,16%	0,26%
Material de construção	0,90%	1,23%	0,54%	0,60%	0,51%	0,29%	1,64%	0,91%	2,12%	1,20%
Material e artigo escolar e serviço de papelaria	0,47%	0,52%	0,53%	0,51%	0,52%	0,20%	0,45%	0,46%	0,35%	0,45%
Mobiliário	1,64%	1,52%	0,78%	0,84%	1,21%	1,35%	2,02%	1,34%	0,54%	0,84%
Óculos, lentes e material oftalmológico	0,25%	0,54%	0,25%	0,64%	0,03%	0,60%	0,26%	0,44%	0,02%	0,19%
Outra bebida alcoólica (exceto cerveja e chopp)	0,10%	0,08%	0,12%	0,13%	0,00%	0,14%	0,42%	0,11%	0,00%	0,03%
Outra despesa de consumo	0,44%	1,14%	0,31%	0,25%	0,39%	0,56%	0,58%	0,20%	0,11%	0,41%
Outro produto de alimentação (óleo, gordura, margarina, sal, etc.)	0,40%	0,43%	0,31%	0,43%	0,76%	0,36%	1,98%	0,49%	0,13%	0,39%
Pão e bolo de padaria	1,70%	1,08%	1,14%	3,06%	1,60%	1,06%	3,27%	2,76%	4,90%	4,37%
Peixe e fruto do mar	0,43%	0,39%	0,67%	0,23%	0,24%	0,35%	0,48%	0,55%	0,95%	0,21%
Plano de saúde e odontológico	1,80%	5,42%	4,19%	0,45%	0,79%	2,24%	0,54%	1,62%	0,24%	1,34%
Refeição pronta, alimento preparado para viagem	0,43%	0,32%	2,17%	0,27%	0,25%	0,37%	0,54%	0,27%	0,39%	0,20%
Remédio e material de tratamento	2,90%	3,27%	2,33%	2,97%	2,83%	2,03%	2,78%	3,24%	2,61%	23,10%
Roupa feminina e masculina	2,99%	2,83%	2,53%	3,05%	1,63%	2,36%	2,88%	2,79%	1,83%	2,53%
Roupa, calçado e acessório para criança	1,06%	0,89%	0,76%	1,13%	0,98%	0,83%	1,53%	1,21%	1,27%	1,32%
Seguro em geral (exceto de saúde e odontológico)	0,60%	0,66%	1,33%	0,01%	0,01%	0,99%	0,21%	0,22%	0,04%	0,05%
Serviço de reforma e manutenção do imóvel	0,80%	0,92%	0,04%	0,11%	0,42%	0,31%	0,64%	0,44%	0,36%	1,29%
Serviço e aparelho para comunicação (exceto internet)	6,07%	5,17%	5,03%	4,57%	6,04%	4,44%	3,13%	18,34%	1,66%	4,61%
Serviço e tratamento médico e odontológico	1,26%	3,66%	1,71%	0,75%	0,47%	1,60%	1,37%	0,67%	0,81%	2,50%
Serviço para cuidado e higiene pessoal	1,46%	1,54%	1,46%	1,56%	0,96%	0,98%	1,28%	1,60%	1,71%	1,51%
Serviço, taxa e outro gasto com habitação	1,07%	4,47%	3,04%	0,32%	1,33%	1,60%	0,55%	0,85%	0,11%	0,76%
Transporte com viagem e outro gasto relacionado	1,02%	1,74%	1,78%	0,41%	0,31%	0,99%	0,44%	0,75%	0,09%	0,41%
Transporte urbano	3,11%	4,26%	2,56%	32,94%	8,82%	2,45%	5,54%	8,47%	4,65%	4,10%
Veículo (aquisição)	2,38%	1,16%	0,31%	0,07%	0,50%	32,72%	0,43%	0,65%	0,00%	0,13%
Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)