

UMA PROPOSTA DE IDENTIFICAÇÃO DE PERFIS DE CONDIÇÕES DE VIDA DOS DOMICÍLIOS DE MINAS GERAIS ATRAVÉS DA PESQUISA DE ORÇAMENTOS FAMILIARES 2002/2003

Helena Cruz Castanheira
Cedeplar/UFMG

Ana Maria Hermeto Camilo de Oliveira
Cedeplar/UFMG

Resumo

Trata-se de identificação de perfis socioeconômicos dos domicílios de Minas Gerais. É utilizada, para isso, a metodologia de análise de aglomerados, ou análise de *cluster*. A base de dados é a POF 2002/2003 do IBGE, da qual são utilizadas variáveis de pessoas, características estruturais do domicílio, despesas e quantidade de bens duráveis. São identificados três diferentes perfis domiciliares, de acordo com o grau de condições de vida favoráveis. Com a análise, as principais deficiências e necessidades dos domicílios da amostra são identificadas. Pretende-se, com isso, contribuir na orientação de diretrizes para futuras políticas públicas.

Palavras-chave: perfis socioeconômicos, domicílios, Minas Gerais, POF 2002/2003, análise de aglomerados

INTRODUÇÃO

O propósito deste trabalho é identificar perfis socioeconômicos dos domicílios brasileiros, baseados nas despesas familiares, com o objetivo de avançar na análise das deficiências e necessidades básicas da população como um todo. Importante, não somente para um maior conhecimento das condições de vida da população de Minas Gerais, mas também para direcionamento de políticas públicas. As despesas são consideradas atualmente de grande importância no estudo de padrões de vida da população. Segundo Deaton e Zaidi (2002, p.4), por exemplo, “Consumption measures are limited in their scope, but are nevertheless a central component of any assessment of living standards¹”.

A metodologia de análise utilizada é a de aglomerados que tem como objetivo classificar uma mostra de entidades em um pequeno número de grupos mutuamente excludentes, de acordo com as similaridades e diferenças entre eles. Utilizaremos, para isso, os dados da Pesquisa de Orçamentos Familiares (POF) 2002-2003, do IBGE (Instituto Brasileiro de Geografia e Estatística), sobre despesas da família², características demográficas, aspectos físicos dos domicílios e a quantidade de bens duráveis. Assim, tem-se desde características físicas do imóvel a características socioeconômicas das pessoas estudadas.

Há uma extensa literatura sobre bem-estar ou condições de vida no Brasil. Descrevem-se níveis e dinâmicas dos resultados da pobreza, disparidades regionais e setoriais, estudos sobre a relação entre mercado de trabalho, capital humano e padrões de gasto público e assim por diante. A maioria dos trabalhos possuem a renda como unidade básica de investigação. Neste trabalho, optou-se por utilizar como variável chave o consumo e não a renda. Outras variáveis (educação, condições domiciliares, inventário de bens duráveis, dentre outras) também foram utilizadas.

O motivo da utilização dessas variáveis e não a renda para a análise das condições de vida da população é porque considera-se que a renda é reflexo, ou consequência dessas outras variáveis e elas estão amplamente relacionadas. Veja, por exemplo, o que diz Deaton e Zaidi (2002) sobre o uso do consumo como análise de condições de vida:

Among economic measures of living standards, the main competitor to consumption-based measure is a measure based on income. In most industrialized countries, including the U.S., living standards and poverty are assessed with reference to income, not consumption. This tradition is followed in much Latin America, where many household surveys make no attempt to collect consumption data[...] There are both theoretical and practical reasons that must be considered when making the choice to use income or consumption to measure living standards. (DEATON e ZAIDI 2002, p.11)

Já sobre a POF 2002-2003, Deaton (2007, p.1) afirma que “*This new POF is both nationally representative and extensive in its questionnaire coverage of income, consumption and expenditures. It thus provides a new and unique opportunity for a thorough analysis of the*

¹ As medidas de consumo são limitadas em seus alcances, mas são um componente central em qualquer avaliação de padrões de vida. (tradução nossa)

² Para efeito de divulgação, na POF 2002-2003 o termo "família" tem sido utilizado para representar o conceito Unidade de Consumo. Uma unidade de consumo é um único morador ou conjunto de moradores que compartilham da mesma fonte de alimentação ou compartilham as despesas com moradia. Na amostra da POF 2002-2003, 99,8% dos domicílios possuem apenas uma unidade de consumo. Com isso, neste trabalho, foram considerados apenas os domicílios com uma unidade de consumo. Assim, os termos, família, pessoas do domicílio e unidade de consumo possuem os mesmos significados.

³. (tradução nossa)

measurement of poverty and inequality in Brazil”.

Deaton et alli (2007) menciona quatro razões para a utilização do consumo em detrimento da renda. O primeiro é que o consumo oscila menos que a renda no curto prazo. O segundo, é que a renda comumente sofre de inúmeros erros de medida devido à renda proveniente do trabalho informal, trabalhos sazonais e de pagamentos em espécie, o que é muito comum no Brasil. O terceiro fator mencionado é a variação da renda ao longo do ciclo de vida da pessoa, tipicamente aumenta e diminui ao longo da vida, uma ilustração disso é a Teoria do Ciclo de Vida⁴. O último fator mencionado é que a renda pode sofrer de perceptíveis incentivos à um sub-relato (*under-reporting*) dos respondentes.

Assim, em resumo, o que impulsionou este trabalho foram os estudos do Deaton mencionados acima, a falta de trabalhos que utiliza principalmente a despesa na análise socioeconômica dos domicílios Brasileiros⁵ e, de certa forma também, a tentativa de dialogar com o trabalho de Duesenberry e seus seguidores sobre consumo e padrão de vida.

Dentre os trabalhos internacionais que utilizaram o consumo como medida de análise do padrão de vida da população vale mencionar Lindert et alli (2005), que faz uma comparação do resultado de políticas sociais de acordo com diferentes medidas de bem-estar, baseados em consumo e renda. Os autores encontraram divergências nas medidas de bem-estar estimada para a renda em relação à estimada para consumo. Para a Guatemala, por exemplo, essa divergência foi a maior (p.41).

Outro trabalho que utiliza consumo como uma medida de bem-estar é Crossley e Pendakur (2002). Eles fazem uma análise de coorte (1969-1999) da desigualdade baseada no consumo para o Canadá. Concluem que renda e rendimentos são fatores importantes na mensuração do bem-estar material. Porém, devido ao fato das unidade de consumo terem acesso a pegar emprestado ou poupar, a mensuração da desigualdade baseada no consumo é uma melhor proxy da desigualdade. Para eles,

It is natural, at least for economists, to think of the material well-being of individuals as being determined by the goods and services which they actually consume in any period. Of course, wages, earnings and income are important determinants of this. We might think of a chain as follows: wages → Earnings → Income → Consumption → Material Well-being. (CROSSLEY e PENDAKUR 2002, p.1)

⁴ MODIGLIANI, Franco; COHEN, Kalman J. (1961) *The role of anticipations and plans in economic behavior and their use in economic analysis and forecasting*. Urbana.

⁵ Dos trabalhos que utilizam a Pesquisa de Orçamentos Familiares vale mencionar Barros et alli (2007), que faz uma interessante comparação entre a desigualdade de renda baseada na PNAD e na POF. Ele encontra que na PNAD⁵ há uma subestimação da renda das famílias, devido à dificuldade em capturar adequadamente todas as fontes de renda, o que gera impacto direto na renda média. É feita ainda, uma comparação entre renda e consumo com os dados da POF 2002/2003. O autor encontra que as diferenças entre eles são muito elevadas, principalmente no décimo mais pobre da população, com o consumo chegando a quase o dobro da renda. Ele afirma que essas diferenças são muito elevadas e não podem ser explicadas pela disponibilidade de ativos de endividamento desses grupos, porque esses ativos são limitados. Dessa forma, conclui haver evidências que, mesmo na POF, a renda dos mais pobres encontra-se subestimada, o que pode estar levando a uma superestimação do grau de desigualdade calculado quando utiliza-se apenas a variável renda.

Outro trabalho a ser mencionado é Quitanes et alli (2006) que, baseados em dados sobre renda na POF 2002-2003, traça o quadro de pobreza e desigualdade no Brasil (utilizam Índice de Theil, Coeficiente de Gini, Consumo médio, proporção de pobres, hiato médio de pobreza, dentre outros).

Silveira e Bertasso (2004), também utilizando a POF, porém 95-96, fizeram uma tipologia socioeconômica das famílias brasileiras, eles utilizaram 24 variáveis. Com isso encontraram 10 tipos diferentes de famílias no Brasil. Concluem que a renda, o tamanho da família, a composição etária do domicílio e a idade do chefe são variáveis chave.

Ferreira e Lanjouw (2001) baseados no consumo, elaboraram um indicador de desigualdade e pobreza. Para que isso fosse possível eles fizeram uma imputação dos dados da PPV (Pesquisa sobre padrões de vida) de 1996, que possui informações sobre consumo, nos da PNAD 1996-97 (Pesquisa Nacional por Amostra de Domicílios). Eles chegaram à conclusão que os índices de pobreza e desigualdade estimadas na PNAD utilizando a renda como base são superestimados se comparados aos que tem o consumo como base.

E afirmam que, a ligação entre salários e rendimentos está mediada pela oferta de trabalho. Já a relação entre rendimentos e renda é baseada no sistema de governo de impostos e benefícios, e por decisões de poupança no passado. Similarmente, renda e consumo se dão através das decisões de poupança e empréstimo que são determinadas por necessidades passadas e futuras, riscos e condições do mercado de crédito. E, por último, afirma que a relação entre consumo e bem-estar material é mediada por vários fatores psicológicos, segundo eles, pouco entendidos pelos economistas.

Sobre classe social e medidas de consumo e renda temos o estudo de Schaninger (1981) que afirma:

Social class proponents have contended that social class is linked more closely to consumption patterns than income (Coleman 1960⁶; Martineau1958⁷). Levy (1966)⁸ argued that 'social class variation are variation in lifestyle', and concluded that resulting consumption, media and store patronage differences are 'not simply the result of variations in income'. Myers and Guttman (1974)⁹ also regarded lifestyle as an intervening variable closely related to consumption, contending that social class is a valuable segmentation base because it captures lifestyle differences that income ignores (SCHANINGER, 1981 p.192)

Como se pode observar, já foram feitos vários trabalhos empíricos que validam a utilização de outras variáveis, além da renda, para a análise de perfis socioeconômicos de domicílios. Além disso, considera-se que o consumo pode ser uma melhor medida para a análise de unidades domiciliares que a renda, se o consumo for bem coletado. Especificamente para a POF 2002-2003 Deaton et alli (2007) afirma ser uma pesquisa confiável, de ampla abrangência e de questionários com extensiva cobertura no que diz respeito a consumo, renda e gastos em geral.

FONTE DE DADOS E METODOLOGIA

Fonte de Dados

A fonte de dados utilizada neste trabalho é a Pesquisa de Orçamentos Familiares (POF) 2002-2003, do Instituto Brasileiro de Geografia e Estatística (IBGE) a última pesquisa mais abrangente no que diz respeito às despesas dos domicílios brasileiros.

A pesquisa foi feita no período de julho de 2002 a junho de 2003. São investigados domicílios particulares permanentes “o domicílio particular permanente é aquele construído para servir exclusivamente à habitação, e no momento da coleta, tinha finalidade de servir de moradia a uma ou mais pessoa” (IBGE 2004, p.23).

No domicílio é identificada a unidade básica da pesquisa, a Unidade de Consumo. De acordo com IBGE 2004, uma unidade de consumo é um único morador ou conjunto de moradores que compartilham da mesma fonte de alimentação ou compartilham as despesas com moradia. Para efeito de divulgação, o termo "família" tem sido utilizado para representar o

⁶ Coleman, R.P. (1960), "The Significance of Social Stratification in Selling" in Marketing: A Maturing Discipline, Martin L. Bell, ed. Chicago: American Marketing Association, 171-84.

⁷ Martineau, Pierre (1958), "Social class and Spending Behavior", Journal of Marketing, 23, 121-41

⁸ Levy, S.J. (1966), Social Class and Consumer Behavior", in On Knowing the Consumer. J.W.Newman, ed. New York: John Wiley & Sons, Inc. 146-160.

⁹Myers, J.H. and Guttman, J. (1974) "Life Style: The Essence of Social Class" in Lifestyle and Psychographics, William D. Wells, ed. Chicago: American Marketing Association, chapter 10.

conceito Unidade de Consumo. Nesta pesquisa, serão analisadas apenas as unidades de consumo principais do domicílio, que representam 99,8% das uc's pesquisadas. Assim, para esta pesquisa, domicílio e unidade de consumo possuem o mesmo significado.

O objetivo da POF 2002-2003 é mensurar as estruturas de consumo, dos gastos e dos rendimentos das famílias brasileiras. A amostra foi estruturada de tal forma que possibilita a publicação de resultados para o Brasil, para as Grandes Regiões (Norte, Nordeste, Sudeste, Sul e Centro-Oeste) e por situação urbana e rural. Em Minas Gerais foram entrevistados 3.004 domicílios.

Indicadores Selecionados para a construção dos perfis domiciliares brasileiros

1. Pessoas

Na base de dados com características das pessoas pertencentes à unidade de consumo (uc) foram extraídas as seguintes variáveis:

QUADRO 1	
Variáveis: Pessoas	
Ntotal:	número total de pessoas na u.c.
Prop_adultos:	porcentagem de pessoas entre 15 e 64 anos na u.c.
Prop_crianças:	porcentagem de pessoas entre 0 e 7 anos na u.c.
Prop_jovens:	porcentagem de pessoas entre 8 e 14 anos na u.c.
Prop_idosos:	porcentagem de pessoas maiores de 65 anos na u.c.
Densmor:	proporção do número total de pessoas na unidade de consumo pela quantidade de cômodos no domicílio.
Ndependentes:	número de idosos, jovens e crianças na uc.
Prop_depend:	proporção entre número de dependentes no domicílio e o número de adultos.
grupidade_ref:	indica a qual grupo de idade pertence a pessoa nomeada como responsável pela uc (o chamado "chefe de família", ou pessoa referência). São 8 grupos de idade, 1: de 13 a 19 anos, 2: 20 a 29, 3: 30 a 39, 4: 40 a 49, 5: 50 a 59, 6: 60 a 69, 7: 70 a 79, 8: 80 ou mais.
ganosestud_ref:	indica a qual grupo de quantidade de anos de estudo a pessoa de referência pertence. Grupo 0: pessoas com 0 anos de estudo. Grupo 1: 1 a 4 anos de estudo, grupo 2: 5 a 8 anos de estudo, grupo 3: 9 a 11 anos de estudo, grupo 4 : 12 a 15 anos de estudo e grupo 5: 16 a 17 anos de estudos.
grupraca_ref:	raça da pessoa referência da u.c. (1: branca, 2: preta, 3:parda, 4:outras);
grupreli_ref:	religião da pessoa referência na família (0: sem religião, 1: Católica Apostólica Romana, 2:Igreja Evangélica Assembléia de Deus, 3: Outras Religiões).
cc_referencia:	pessoa referência possui cartão de crédito ou não;
chespecial_ref:	pessoa referência possui cheque especial ou não;
plsaude_ref:	pessoa referência possui plano de saúde ou não;
sexo_ref:	sexo pessoa referência;
presen_conjuje:	existência de cônjuge na unidade de consumo ou não;
Fonte: Elaboração Própria	

2 Domicílio

QUADRO 2	
Variáveis: Domicílio	
qdemor:	quantidade total de moradores no domicílio
tipodom:	tipo de domicílio (1: casa não rústica, 2: casa rústica, 3: apartamento, 4: cômodo)
condocup:	condição de ocupação do domicílio. (1: próprio já pago, 2: próprio em aquisição, 3: alugado, 4:cedido por empregador, 5: cedido por particular, 6: outra)
qdcom:	total de cômodos existentes no domicílio
comdorm:	total de cômodos servindo de dormitório
qdbanh:	total de banheiros localizados dentro da propriedade
abastagua:	forma de abastecimento de água no domicílio (1:rede geral com canalização interna, 2: poço ou nascente com canalização interna, 3: outra forma com canalização interna, 4: rede geral sem canalização interna, 5: poço ou nascente sem canalização interna, 6: outra forma sem canalização interna)
escsan:	tipo de Escoadouro Sanitário do banheiro sanitário do domicílio. (1:rede geral de esgoto ou pluvial, 2: fossa séptica, 3: fossa rudimentar, 4: vala, 5: rio, lago ou mar, 6: outro escoadouro, 7: não tem)
origemee:	procedência da energia elétrica utilizada pelo domicílio. (1:companhia de energia elétrica, 2: outros produtores de energia, 3:próprios , 4: não tem)
matpiso:	tipo de material que predomina no piso do domicílio. (1: carpete, 2: cerâmica/lajota/pedra, 3: madeira aparelhada, 4: cimento, 5: madeira aproveitada, 6: terra, 7: outros)
pavrua:	existe ou não pavimentação na rua onde se localiza o domicílio.

Fonte: Elaboração Própria

3 Despesas

Para todas as despesas de cada domicílio foi obtida a relação entre despesa da categoria correspondente divididas por despesas totais do domicílio. Ou seja, foi preferido adotar nesta pesquisa a proporção de gastos da despesa analisada em relação à despesa total do que a despesa em si. Isso porque o valor da despesa varia de região pra região. No nosso caso é mais interessante analisar a “importância” de cada tipo de gasto em relação ao gasto total. Neste sentido, todas as despesas a seguir são em proporções e anualizadas e deflacionadas.

Quadro 3	
Variáveis: despesas, em proporções, anualizadas e deflacionadas	
cons_total	soma de todas as despesas do domicílio
conspercapita	cons_total dividido pela quantidade de moradores
des_acesso-s	acessórios e manutenção de veículos
des_alim_a-a	açúcares e produtos de confeitaria
des_alim~es	carnes e vísceras
des_alim_a-c	aves, ovos e laticínios
des_a~salcoo	bebidas alcoólicas
des_a~nalcoo	bebidas não-alcoólicas e infusões
des_alim_c-b	cereais, leguminosas e hortaliças tuberosas
des_alim~tos	sais e condimentos, despesas anualizadas e deflacionadas
des_alim_e-s	enlatados e conservas
des_alim_f-a	farinhas, féculas, massas, cocos, castanhas e nozes
des_alimfo-a	proporção das despesas totais com alimentação fora de casa
des_alim_f-s	frutas
des_alim_h-s	hortaliças folhosas, frutosas e outras
des_alim_o-p	óleos, gorduras e alimentos preparados ou semi-preparados
des_a~icados	panificados e carnes industrializadas
des_a~scados	pescados marinhos, pescados de água-doce
des_prodlis	produtos de limpeza doméstica e outros
des_alutidom	aluguel de aparelhos e utilidades de uso doméstico
des_aquisi-s	aquisição de móveis
des_aquisi-o	aquisição de veículos no período de referência de 12 meses
des_artigo~r	artigos de toucador
des_bolsac~o	bolsas, calçados e cintos
des_brinqu~s	brinquedos e material de recreação
des_cabele-a	serviços pessoais (cabeleireiro, barbeiro, manicuro e pedicuro, costureira, dentre outros)
des_cerifa~i	cerimônias familiares e práticas religiosas
des_compra-s	aquisição de aparelhos e máquinas de uso doméstico
des_comuni-o	comunicações (telefone público, correio, telemensagem, telegrama fonado, dentre outros)
des_constr~o	construção, reforma de habitação e jazigo
des_divers~r	diversões, esportes e uso de celulares
des_educacão	educação
des_eetelint	serviços públicos ou similares (energia elétrica, esgoto, gás encanado, telefone residencial, acesso Internet, tv assinatura, gás de bujão, aquisição de água lítro, querosene, lenha e carvão)
des_encarg~s	contribuições, transferência e encargos financeiro, mesadas e dinheiro para outras uc's
des_fumo	fumo
des_habita~o	habitação
des_instrm-s	aquisição de instrumentos e acessórios musicais fotográficos, artigos para acampamentos, máquinas de escritório (para o domicílio) e utilidades de uso doméstico
des_jogose-s	jogos e apostas
des_joiase-s	jóias, relógios, aparelhos e acessórios de telefonia celular
des_leitura	leitura (jornal, revista infantil, outras revistas, álbum e figurinhas, dentre outros)

des_manut	manutenção e pequenos reparos com habitação, jazigo e jardinagem
des_manutm~s	consertos e manutenção de móveis, aparelhos, máquinas e utensílios de uso doméstico
des_outras~v	outras despesas individuais (bijuteria, óculos sem grau, bateria de relógio, dentre outras)
des_outro~os	documentação, seguro e outros gastos com veículos
des_outros~t	outros serviços de utilidade pública e habitação
des_outro~is	outros imóveis
des_papela~a	artigos de papelaria, livros não-didáticos e assinatura de periódicos
des_remedios	produtos farmacêuticos
des_roupas~a	roupas de criança até 14 anos
des_roupas~o	artigos de armarinho tecidos e roupas de cama, mesa e banho
des_roupas~m	roupas de homem
des_roupas~r	roupas de mulher
des_servdo~o	serviços domésticos
des_serviç~e	serviços de assistência à saúde
des_serviç~s	serviços bancários, de cartório e profissionais, despesas anualizadas e deflacionadas
des_transp~e	transportes
des_utensi~a	utensílios avulsos e artigos de banheiro, copa e cozinha
des_viagens	viagens
des_decora~o	decoração e forração

Fonte: Elaboração Própria

4 Inventário bens duráveis

Os bens duráveis foram agregados em grupos de acordo com a frequência desses itens nos domicílios em geral e suas características em relação à elasticidade-renda (bem normal, inferior ou de luxo). É importante mencionar, porém, que esta elasticidade-renda não foi calculada matematicamente, apenas o conceito foi utilizado.

Quadro 4	
Variáveis: Quantidade de Bens Duráveis no Domicílio	
d_torraaspiencer:	torradeira elétrica, apirador de pó, enceradeira
d_fogão:	fogão
d_geladeira:	geladeira
d_batliqferro:	batedeira de bolo, liquidificador, ferro elétrico no domicílio
d_tvsembicicl:	televisão em cores, conjunto de som acoplado, bicicleta
d_maqluxo:	freezer, máquina de lavar louça, máquina de lavar roupas, máquina de secar roupas, ar condicionado
d_semiluxo:	máquina de costura, forno de microondas, antena parabólica
d_videodvdventil:	vídeo cassete, DVD, ventilador e circulador de ar, secador de cabelos, toca disco a laser
d_beminferior:	televisão em preto e branco, gravador e toca fitas, rádio de mesa, rádio portátil
d_motocicletacomputador:	motocicleta, microcomputador
d_automovel:	automóvel
d_filtroagua:	filtro de água a vela, purificador de água tipo ionizador

Fonte: Elaboração Própria

Metodologia

Neste trabalho será utilizada a análise fatorial e a análise multivariada de agrupamentos. A análise de agrupamentos, que é uma técnica multivariada

[...]tem como objetivo classificar uma amostra de entidades (objetos ou indivíduos) em um pequeno número de grupos mutuamente excludentes, com base nas similaridades entre as entidades [...] Na análise de grupamentos, diferentemente da análise discriminante, os grupos não são predefinidos. Ao invés disso, a técnica é usada para identificar grupos. (HAIR et alli 2005, p. 33)

Esse objetivo está em total consonância ao que almejamos nesta pesquisa, a identificação de grupos nos domicílios da Pesquisa de Orçamentos Familiares 2002-2003, para uma análise sucinta dos diferentes padrões de despesa e domicílios no Brasil.

Optou-se por utilizar essa técnica de análise multivariada devido ao fato de nenhuma variável ter que ser definida como dependente ou independente e pelo fato de podermos utilizar variáveis contínuas e categóricas ao mesmo tempo. E ainda, com a análise de agrupamentos, o modelo que fornece o número de agrupamentos na amostra e possibilita a utilização de uma grande base de dados.

A medida de similaridade utilizada pelo método é a de log-verossimilhança que tem a distância entre as observações como critério de similaridade e estabelece uma medida de distribuição de probabilidades às variáveis. Assume-se que as variáveis contínuas possuem uma distribuição normal (*Gaussiana*) e as categóricas seguem uma distribuição multinomial. Para definir qual é a melhor quantidade de clusters, o modelo utiliza o Critério Bayesiano de Schwarz (BIC) que fornece o número adequado de agrupamentos.

A análise fatorial foi utilizada para ver se era possível reduzir o número de variáveis utilizadas sem interferir na qualidade dos resultados. Essa análise identifica variáveis, ou fatores, que explicam um padrão de correlações do grupo de variáveis. Na análise fatorial utilizada, só é possível a imputação de variáveis contínuas. O resultado em geral é uma quantidade reduzida de variáveis que explicam a maior parte do grupo de variáveis.

ANÁLISE DOS RESULTADOS

Análise Fatorial das variáveis e seus graus de significância

O primeiro passo adotado em relação às variáveis foi uma análise fatorial das variáveis para ver a possibilidade de redução da quantidade das mesmas para a análise de *cluster*. É importante mencionar que essa análise é válida apenas para as variáveis contínuas.

Para o grupo sobre despesas foi obtido um resultado na qual não era possível eliminar nenhuma das variáveis. As variáveis de despesa totalizam 58, sendo que 21 explicavam apenas 50,09% da variância total dos dados. Ou seja, se reduzíssemos o número de variáveis de despesas, perderíamos 49,91% das informações. Além disso, as variações deste grupo eram amplamente correlacionadas entre si. Para que a redução fosse possível, o número de variáveis selecionadas pelo modelo, neste caso 21, deveriam ser responsáveis por pelo menos 70 % das variações totais do banco de dados e ser pouco correlacionadas entre si.

O passo seguinte foi fazer a análise fatorial para as variáveis contínuas do grupo de pessoas. Especificamente, as variáveis: `prop_nadultos`, `prop_ncriança`, `prop_njovens`,

prop_nidosos, densmor, prop_depend e n_dependentes. Elas falam sobre o ciclo de vida da família e pensamos que poderiam ter comportamentos parecidos na base de dados. O resultado que obtivemos com a análise fatorial foi de uma redução amplamente possível dessas variáveis. Das sete acima, três conseguiam explicar 89,77% da variação total dos dados, um resultado bastante significativo.

A primeira variável, responsável pela maior parte das variações neste grupo foi prop_depend (proporção total de crianças, jovens e idosos sobre o número total de adultos) responsável por 47,68 % das variações. A variável seguinte foi densmor, quantidade de moradores em relação ao número total de cômodos do domicílio, que foi responsável por 26,83% das variações. E, por último, prop_ncrianças que é a proporção da quantidade total de crianças (indivíduos até 7 anos) no domicílio sobre o número total de pessoas, responsável por 15,25% das variações.

Foi feita ainda uma análise fatorial para o grupo de bens duráveis, mas nenhum apresentou possibilidade de redução da quantidade de variáveis dos mesmos.

Características gerais dos *clusters* obtidos (variáveis de região, pessoas e características do domicílio)

Para estimar os grupos utilizamos todas as variáveis do grupo despesa, pessoas, bens duráveis e domicílio mencionadas anteriormente, com exceção das variáveis eliminadas pelo processo de análise fatorial.

A fim de que os domicílios analisados possam ter significância estatística para todo o estado de Minas Gerais, as tabelas abaixo foram geradas com o “peso final” fornecido e estimado na POF 2002-2003. Dez domicílios da amostra possuíam seus pesos “missing”, ou seja, eram sem dados para essa variável. Por isso nossa amostra, ao fazer as tabelas, passou de 3004 domicílios em Minas Gerais para 2994 domicílios. O resultado do exercício de agrupamento está na TAB 1. Como se pode observar, o primeiro *cluster* agrupou 32,43% dos domicílios, o segundo 19,86%, o terceiro é o maior grupo com 47,71% dos domicílios da amostra.

TABELA 1
Distribuição domicílios por *cluster*

<i>Cluster</i>	Freqüência Simples	Freqüência Relativa
1	971	32.43
2	595	19.86
3	1428	47.71
Total	2994	100

Fonte: Elaboração Própria

Nota: Dados extraídos da POF 2002/2003

A TAB 2 nos mostra a distribuição dos domicílios por estrato geográfico. O *cluster* número 2 é o que possui a maior quantidade de domicílios em área rural (58,41%). O *cluster* 1 possui mais domicílios na região metropolitana (61,76%) e o *cluster* 3 é o que possui mais domicílios em outras áreas urbanas da UF (71,24%). As diferenças das médias, pelo teste t-estatístico, foram significativas estatisticamente.

TABELA 2
Distribuição percentual dos domicílios de Minas Gerais por estrato geográfico segundo *cluster*

<i>Cluster</i>	RURAL		URBANO		Total
	domicílios em área rural	domicílios da capital e região metropolitana	domicílios urbanos em "outras" áreas da UF		
1	3.74	34.5	61.76		100 (n=971)
2	58.41	6.77	34.82		100 (n=595)
3	3.67	25.09	71.24		100 (n=1428)
Total	14.56	24.5	60.94		100 (n=2994)

Fonte: Elaboração Própria

Nota: Dados extraídos da POF 2002/2003

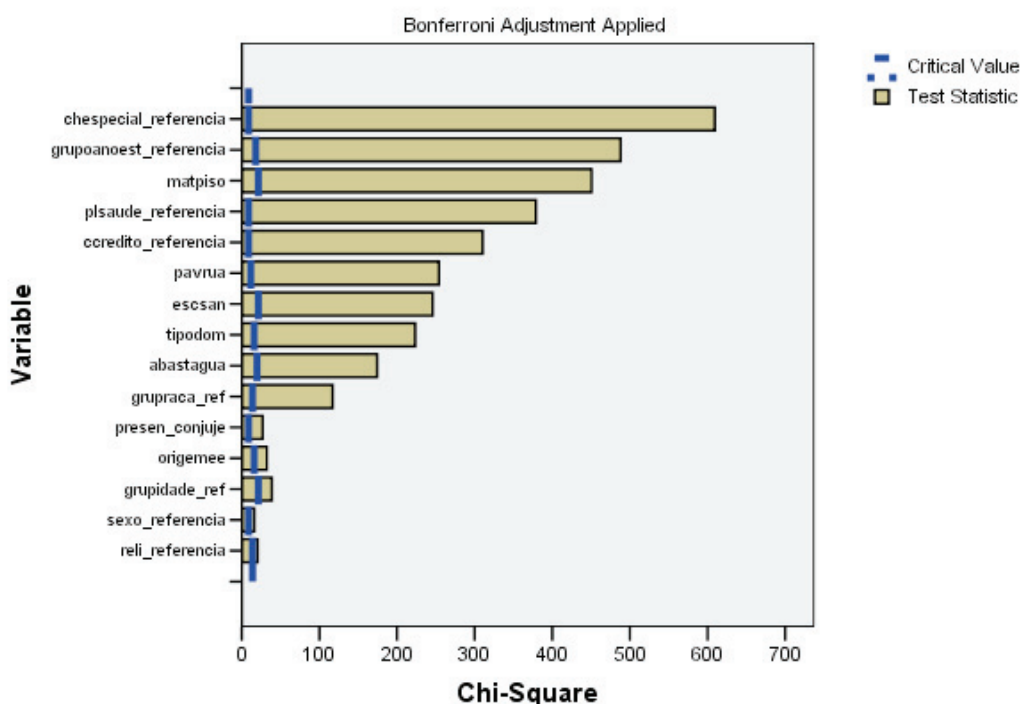
Nota2: A classificação de "Rural" e "Urbano" utilizada neste trabalho é a divulgada pelo IBGE na POF 2002-2003, variável "estrato geográfico".

Será feito agora um breve resumo, por aglomerado, das variáveis que definiram cada *cluster* e suas estatísticas descritivas. As variáveis categóricas serão as primeiras a serem analisadas, em seguida as contínuas.

Grupo de domicílios com menor demanda de políticas públicas, ou seja, com favoráveis condições de vida (*cluster* 1)

A FIG 1 nos mostra as variáveis categóricas por ordem de importância para o aglomerado 1. As linhas verticais são os valores críticos (t-estatísticos) que determinaram a significância de cada variável. Para a variável ser considerada significativa ela precisa estar além da linha azul pontilhada que aparece na figura (valor crítico). Como se pode observar, no aglomerado 1 todas as variáveis categóricas foram consideradas importantes e significativas na definição do grupo.

Figura 1
TwoStep Cluster Number = 1



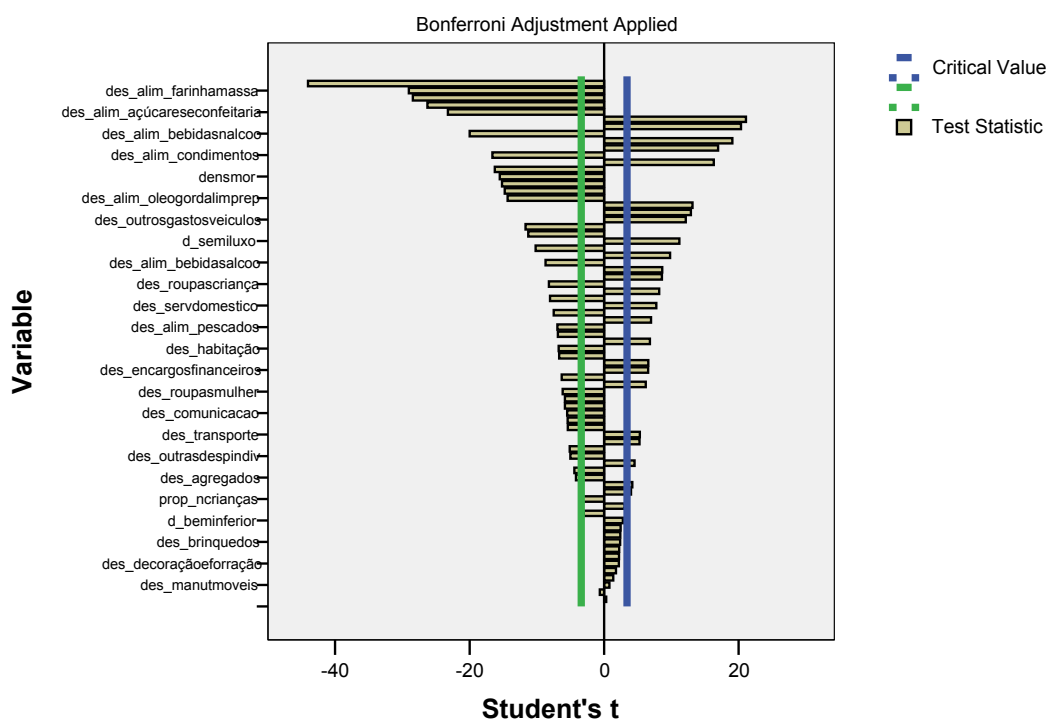
Em linhas gerais, e seguindo a ordem das variáveis da FIG 1, este grupo possui 87,13% das pessoas com cheque especial, 93,33% dos chefes de família com 12 a 15 anos de estudo e 100% dos com 16 ou mais anos de estudo. Além disso, 85,26% dos domicílios possuem piso de cerâmica, lajota, pedra ou madeira aparelhada. Possui também a maior quantidade de domicílios com chefes de família com plano de saúde (70%) e cartão de crédito (65,93%). A menor quantidade de domicílios sem pavimentação nas ruas é deste grupo, apenas 10,86% e eles não há nenhum domicílio com escoadouro sanitário considerado rudimentar, são na grande maioria rede geral de esgoto.

Em relação ao tipo de domicílio, 89,33% dos apartamentos da amostra são deste *cluster*. Mas a porcentagem de apartamento na amostra total é pequena, apenas 6,72% dos domicílios, a maioria dos domicílios em Minas Gerais são casas não rústicas (90,55%). A raça da pessoa referência também foi uma variável significativa para a formação deste agrupamento. Ao analisarmos os dados percebemos que 69,77% dos chefes de família são brancos, o que está muito acima da média do *cluster 2* (38,75%) e do *cluster 3* (45,33%). Além disso, a menor porcentagem de negros também é deste grupo, apenas 4,77% dos domicílios.

Em relação ao comportamento das variáveis contínuas tem-se a tabela em anexo que nos mostra todas as variáveis contínuas do modelo, algumas estatísticas descritivas e as características de forma geral. Ele possui informações detalhadas da FIG 2. Como se pode observar esta figura não apresenta todas as 73 variáveis do modelo listadas segundo seus graus de importância, apenas algumas. Ela foi incluída como exemplo ilustrativo de como os t-estatísticos deste *cluster* para essas variáveis se distribuem ao longo do valor 0. Serão descritas aqui apenas

as variáveis significativas na definição do modelo (que seus valores extrapolam as linhas azuis e verdes da FIG 2, ou seja, os valores críticos), para o total de 73 variáveis contínuas 58 foram estatisticamente definidoras deste perfil.

Figura 2
TwoStep Cluster Number = 1



Este grupo de domicílios possui uma proporção de gastos totais com alimentos geralmente menor que a média da amostra como um todo (t-estatístico positivo), o que significa que os alimentos possuem um menor peso no orçamento total destes domicílios que para os domicílios da amostra como um todo. Ou seja, as proporções entre valor total de gastos com alimentos sobre os gastos totais do domicílio são menores para este *cluster*. Em relação aos bens duráveis este agrupamento possui todos os sinais do teste-t negativo, o que significa que os domicílios deste agrupamento possuem, em média, mais bens duráveis do que os outros domicílios da amostra.

Para os outros tipos de despesa que não alimentos, os valores das proporções sobre despesa total geralmente são menores que a média da amostra, o que reflete um menor peso no orçamento total. Porém, para 12 tipos de despesas diferentes esse comportamento não se aplica. São elas (por ordem de importância na definição do *cluster*): documentação, seguro e outros gastos com veículos (*des_outro~os*), despesa com educação (*des_educação*), despesa com aquisição de veículos no período de referência de 12 meses (*des_aquisi~o*), despesa com serviços domésticos (*des_servdo~o*), despesas com serviços bancários, de cartório e profissionais (*des_serviç~s*), despesas com acessórios e manutenção de veículos (*des_acesso~s*), despesas com serviços de assistência à saúde (*des_serviç~e*), despesas com contribuições, transferência e encargos financeiro, mesadas e dinheiro para outros domicílios (*des_encarg~s*), despesas com outros imóveis (*des_outro~is*), despesa com transporte (*des_transp~e*), despesas com outros

serviços de utilidade pública e habitação (des_outros~t), despesas com leitura (jornal, revista infantil, outras revistas, álbum e figurinhas, dentre outros) (des_leitura).

No que diz respeito às variáveis que nos indicam a densidade de moradores no domicílio (densmor¹⁰) e a proporção de dependentes no domicílio (propdepend¹¹), os valores do t-estatísticos são significativamente diferentes que os valores dos outros clusters e são positivos. Ou seja, em geral, a densidade de moradores e a proporção de dependentes deste cluster são menores que a média total da amostra.

Grupo de domicílios com maior demanda de políticas públicas, ou seja, com desfavoráveis condições de vida (*cluster 2*)

Em relação às variáveis categóricas dos domicílios do grupo 2, temos a FIG 3 abaixo que tem a mesma informação da figura do *cluster* anterior. É importante lembrar da característica, mencionada anteriormente, que os domicílios deste *cluster* se localizam principalmente em área rural, diferentemente dos outros dois *clusters* que se concentram em áreas urbanas, como mencionado anteriormente.

Os domicílios deste grupo possuem a maior porcentagem de domicílios sem canalização interna de água (31,21%), contra 6,78% da amostra como um todo. Além disso, as maiores porcentagens de tipos de escoadouro sanitário precários¹² (83,85%), a maior porcentagem de domicílios em ruas sem pavimentação (83,47%) e a maior porcentagem de domicílios com piso de cimento ou terra (77,56%). A outra variável categórica mais representativa estatisticamente nesse perfil é a anos de estudo do chefe do domicílio, que também apresenta valores explicitamente diferente dos outros *clusters* pois 82,13% dos chefes de família possuem escolaridade de 0 a 4 anos de estudo, esse valor para o grupo 1 é 32,03%.

Este *cluster* possui 87,16% dos domicílios da amostra que se encontram em casas rústicas e 69,59% dos domicílios tipo cômodos (os outros 30,41% são do grupo 3). Além disso, possui 96,41% dos domicílios da amostra que não possuem energia elétrica e 100% dos que a procedência da energia elétrica é de “outros produtores de energia”. Sobre plano de saúde, cartão de crédito e cheque especial, os valores de domicílios neste grupo que o chefe de família possui esses benefícios são, respectivamente, 1,3% , 3,78% e 1,3% .

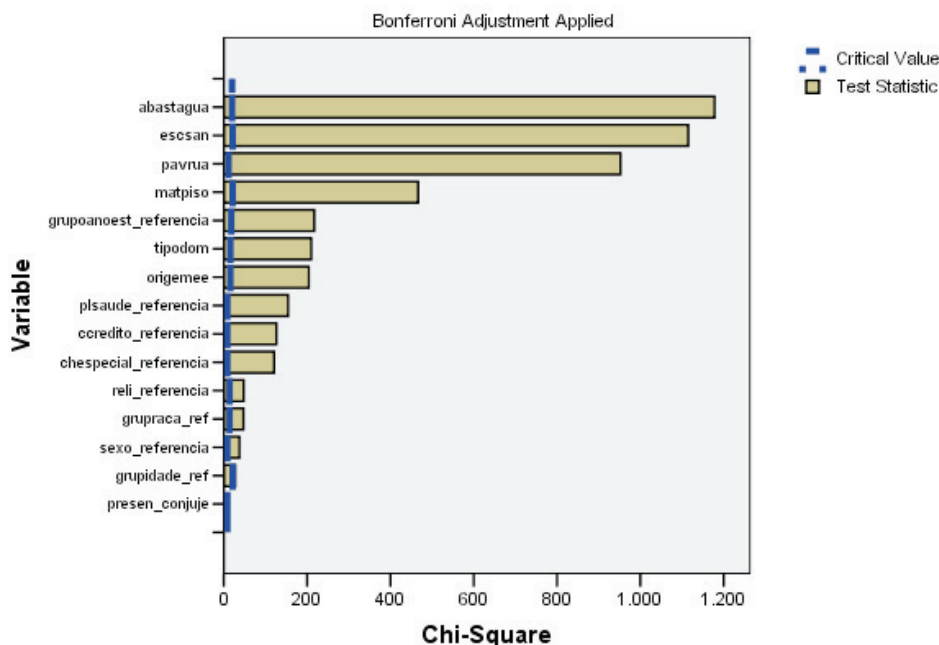
Em relação à religião, este grupo possui a porcentagem mais alta de domicílios da religião Católica Apostólica Romana (86,31% dos domicílios) e a menor porcentagem de pessoas referência sem religião (2,61%). No que diz respeito a raça dos chefes de família este grupo é o de menor proporção de chefes de família brancos (38,75%). O comportamento da variável “sexo do chefe de família” é o mesmo do grupo 1, 81,41% são homens e 18,59% são mulheres.

¹⁰ Proporção do número total de pessoas na unidade de consumo pela quantidade de cômodos no domicílio.

¹¹ Proporção entre número de dependentes (crianças e idosos) no domicílio e o total de adultos.

¹² Os tipos de escoadouro sanitários considerados precários aqui são os das categorias “outra forma com canalização interna”, “rede geral sem canalização interna”, “poço ou nascente sem canalização interna”, “outra forma sem canalização interna”.

Figura 3
TwoStep Cluster Number = 2

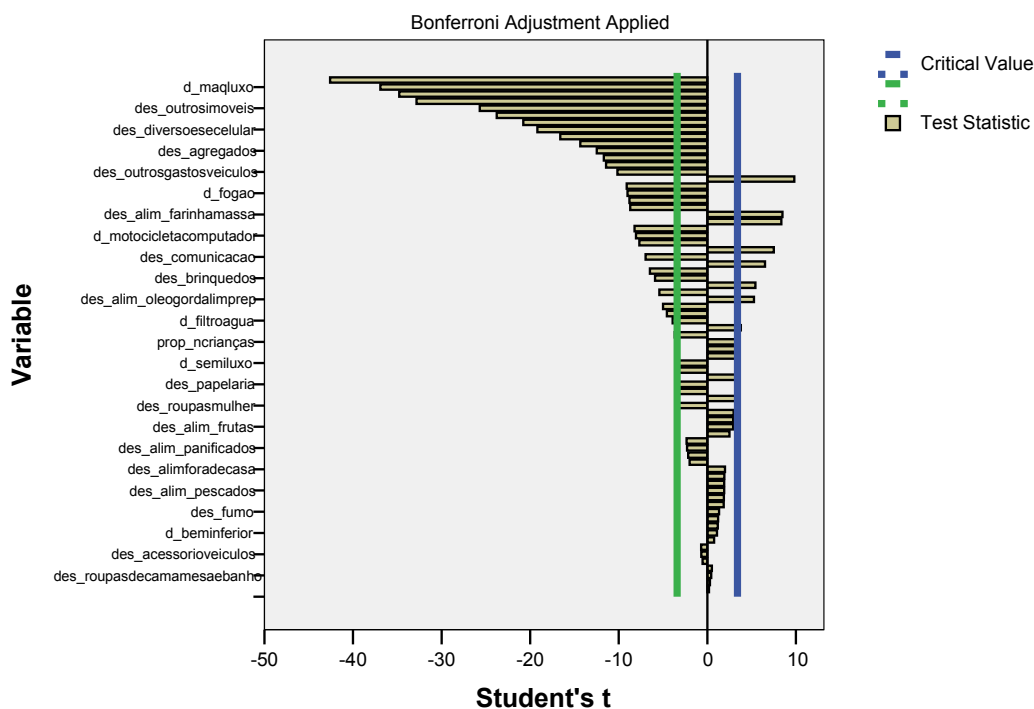


Em relação às variáveis contínuas do modelo, além da tabela em anexo, tem-se a FIG 4, que foi incluída para exemplo ilustrativo do comportamento dos valores dos t-estatísticos ao longo do eixo de valor 0. Serão descritas aqui apenas as variáveis contínuas importantes na definição do modelo (variáveis em que seus valores extrapolam as linhas azuis e verdes da FIG 4, ou seja, os valores críticos) ou também, na tabela em anexo, $Dif\ c2 = 1$. Para o total de 73 variáveis contínuas 40 foram estatisticamente definidoras deste perfil.

As de gastos com alimentos são, como esperado, todas com o valor do t-estatístico negativas, ao contrário do grupo 1. Ou seja, a despesa com alimentos possuem um alto peso na proporção de gastos totais, responsáveis por grande parte dos mesmos. Apenas para um tipo de alimento esse valor é positivo, proporção de despesas com alimentos enlatados e em conservas (*des_alim_e~s*), o que, acredita-se, reflete um menor consumo destes tipos de produto para os domicílios deste *cluster*.

Para as outras despesas, todas são menores que para a média geral da amostra como um todo. Apenas para despesa com utensílios avulsos e artigos de banheiro, copa e cozinha (*des_utensi~a*) e despesas com roupas de homem (*des_roupas~m*), a proporção de gastos deste *cluster* é maior que a média total das proporções de despesas da amostra como um todo.

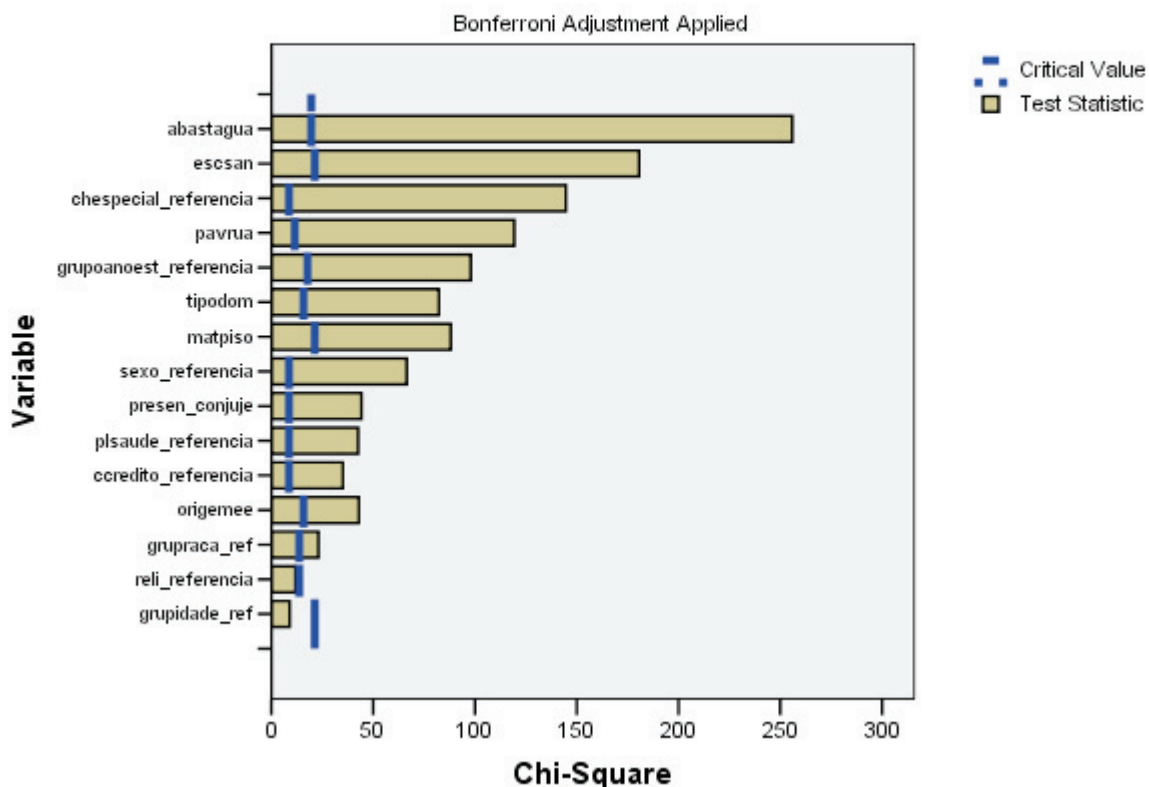
Como esperado, a quantidade de bens duráveis é significativamente menor neste grupo que na amostra como um todo. Além disso, as variáveis contínuas significativas que dizem respeito à estrutura etária e à quantidade de pessoas residentes no domicílio são maiores que a média da amostra total. São elas densidade de moradores (*densmor*), proporção de crianças (*prop_ncria~s*) e proporção de dependentes (*prop_depend*) no domicílio.

Figura 4**TwoStep Cluster Number = 2**

Grupo de domicílios com nível intermediário de demanda por políticas públicas, ou seja, com condições de vida intermediárias em relação aos *clusters* anteriores (*cluster 3*)

O grupo 3 seria um meio termo entre o grupo 1 e 2. Pela análise das variáveis categóricas até então pode-se perceber que o grupo 1 é o que possui menores demandas em termos de políticas sociais no que diz respeito a políticas públicas, e o grupo 2 é o que possui maiores demandas, pois apresentam as piores condições de vida da amostra analisada.

Será seguida a ordem de apresentação das variáveis categóricas por importância de acordo com a FIG 5 . Em relação ao abastecimento de água dos domicílios deste *cluster*, 99,06% possuem canalização interna, 97,26% possuem um escoadouro do banheiro sanitário não precário (rede geral de esgoto ou pluvial e fossa séptica) , poucos possuem cheque especial (4,14%), 21,06% dos domicílios não possuem pavimentação na rua, 88,65% dos chefes de família possuem de 0 a 8 anos de estudo. Este grupo possui a maior porcentagem de mulheres como chefe de família do domicílio, 34,64% , bem como a maior porcentagem de domicílios sem cônjuge (39,58%) . Além disso, possui um valor intermediário entre os cluster 1 e 2 de domicílios que possuem plano de saúde e cartão de crédito, com 12,28% e 13,41% respectivamente. A maior parte dos chefes de família são considerados negros ou pardos (53,57% contra 45,33% de brancos).

Figura 5**TwoStep Cluster Number = 3**

Em relação às variáveis contínuas do modelo, além da tabela em anexo, tem-se a FIG 6, que foi incluída para exemplo ilustrativo do comportamento dos valores dos t-estatísticos ao longo do eixo de valor 0. Serão descritas aqui apenas as variáveis contínuas importantes na definição do modelo (variáveis em que seus valores extrapolam as linhas azuis e verdes da FIG 4, ou seja, os valores críticos) ou também, pela tabela em anexo, $Dif\ c2 = 1$. Para o total de 73 variáveis contínuas apenas 33 foram estatisticamente definidoras deste perfil.

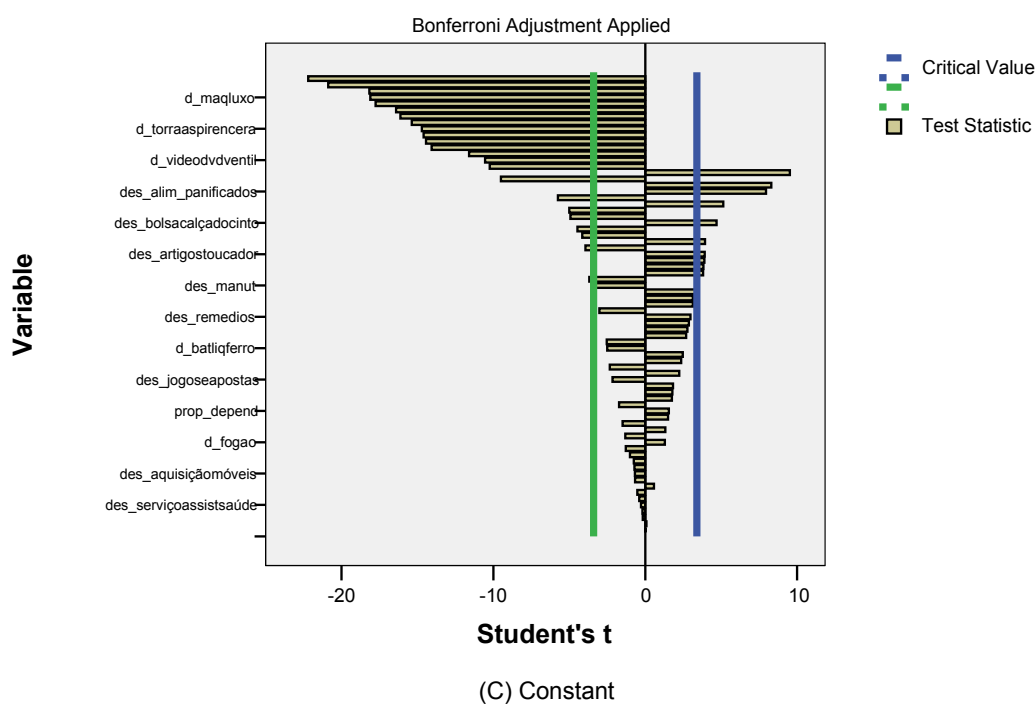
Diferente dos *clusters* anteriores, nenhuma das variáveis contínuas que dizem respeito à estrutura etária do domicílio foram importantes na definição deste grupo. Todas as variáveis de bens duráveis apresentaram um sinal positivo que nos diz que a quantidade total de bens neste domicílio é menor que a média da amostra.

As variáveis de despesa apresentaram comportamentos diferentes entre si. No que diz respeito às despesas com alimentos, apenas seis variáveis foram definidoras deste perfil. Destas, três apresentaram sinais negativos e três positivos. Para despesas com alimentos panificados e carnes industrializadas (*des_a~icados*), sais e condimentos (*des_alim~tos*) e despesas com bebidas alcólicas (*des_a~salcoo*) as proporções sobre os gastos totais foram significativamente maiores que a média da amostra com um todo. Ao contrário, para as despesas com hortaliças folhosas, frutosas e outras (*des_alim_h~s*), cereais, leguminosas e hortaliças tuberosas (*des_alim_c~b*) e farinhas, féculas, massas, cocos, castanhas e nozes (*des_alim_f~a*) os gastos foram menores que a média.

Sobre gastos em geral, excetuando os com alimentos mencionados acima, 18 variáveis foram estatisticamente significantes ao modelo. Em geral elas apresentaram proporções menores que a amostra total. Porém, para 7 variáveis as proporções dos domicílios deste *cluster* foram maiores que as dos outros *clusters*. São elas: despesas com serviços públicos ou similares (energia elétrica, esgoto, gás encanado, telefone residencial, acesso Internet, tv assinatura, gás de bujão, aquisição de água litro, querosene, lenha e carvão) (*des_eetelint*), despesa com habitação (*des_habita~o*), despesas com roupas de mulher (*des_roupas~r*), despesas com bolsas, calçados e cintos (*des_bolsac~o*), despesas com comunicações (telefone público, correio, telemensagem, telegrama fonado, dentre outros) (*des_comuni~o*), artigos de toucador (*des_artigo~r*) e despesas com serviços pessoais (cabeleireiro, barbeiro, manicuro e pedicuro, costureira, dentre outros) (*des_cabele~a*).

Figura 6

TwoStep Cluster Number = 3



Comparação medidas de renda *versus* medidas de despesa

Como foi observado na descrição dos grupos, o grupo 1 seria o que requer menos demanda em termos de políticas públicas pois possui melhor condição de vida. O grupo 2 seria o que requer maiores demandas em termos de políticas públicas, pois possui piores condições de vida. E o grupo 3 seria um intermediário entre esses dois grupos.

Neste momento será feita uma comparação entre despesa total e renda total por *cluster* para ver se há realmente uma diferença estatisticamente significativa entre as duas variáveis. Pelo quadro 4, observa-se que o p-valor é 0,5054 o que nos informa que há uma diferença estatisticamente significativa entre os valores das duas variáveis.

QUADRO 4						
Teste t variável despesa total versus renda total do domicílio						
. ttest despesatotal==rend_total						
Paired t test						

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
despes~l	3004	20753.29	652.4568	35760.35	19473.98	22032.6
rend_t~l	3004	20403.62	758.1694	41554.32	18917.04	21890.21

+-----						
diff	3004	349.6656	524.9187	28770.14	-679.5709	1378.902

mean(diff) = mean(despesatotal - rend_total)				t = 0.6661		
Ho: mean(diff) = 0				degrees of freedom = 3003		
Ha: mean(diff) < 0		Ha: mean(diff) != 0		Ha: mean(diff) > 0		
Pr(T < t) = 0.7473		Pr(T > t) = 0.5054		Pr(T > t) = 0.2527		

Fonte: Elaboração Própria

Essa diferença de médias no que diz respeito a cada *cluster* separadamente é significativa dependendo do *cluster* e do p-valor analisado, ver TAB 3 . Para o *cluster* 1 ela é significativa para todos os p-valores, para os grupos 2 e 3 apenas em Pr(T<t).

TABELA 3
Difereça dos valores médio das variáveis "Despesa Anual Total" e "Renda Anual Total" por cluster

Variáveis	Despesa Total	Renda Total	Diferença	Pr(T<t)	Pr(T > t)	Pr(T>t)
cluster 1	44515.85	46130.07	1614.22	0.1666	0.3332	0.8334
cluster 2	8737.313	7648.812	-1088.501	0.9973	0.0054	0.0027
cluster 3	11522.09	10252.57	-1269.52	1.0000	0.0000	0.0000
total	20753.29	20403.62	-349.67	0.7473	0.5054	0.2527

Fonte: Elaboração Própria

Nota: Dados extraídos da POF 2002/2003

Agora será feita uma comparação da tipologia de *clusters* adotada neste trabalho e a tipologia que identifica os domicílios precários apenas através da variável renda *per capita* dos mesmos. Essa comparação é feita na TAB 4, em que pode-se perceber que ao utilizar as linhas de miséria e pobreza do IPEA para a POF 2002-2003 (baseadas na renda *per capita*) foram encontrados domicílios “pobres” e “indigentes” até mesmo no *cluster* 1, que pela nossa análise possui as melhores condições de vida de todos os outros domicílios (3,98 % dos domicílios se encontram no estado de pobreza e 1,36% de miséria). Para os domicílios do *cluster* 2, que pela nossa análise são os que possuem as maiores demandas no que diz respeito a políticas públicas pois possuem se encontram em piores condições de vida, com a tipologia de renda O que quer dizer que, ao utilizarmos a renda total declarada na POF e compararmos que utiliza renda não são totalmente identificados, apenas a metade deles (50,56%) são identificados.

TABELA 4
Distribuição percentual dos domicílios pobres* e indigentes**
por agrupamentos de domicílios da amostra da POF 2002/2003

Variáveis	Pobre*	Indigente**	Não Pobre	Total
<i>cluster 1</i>	3.98	1.36	96.02	100
<i>cluster 2</i>	50.56	26.43	49.44	100
<i>cluster 3</i>	25.3	8.04	74.7	100
total	23.41	9.53	76.59	100

Fonte: Elaboração Própria

Nota1: Dados extraídos da POF 2002/2003

Nota2: pobres* = Renda familiar per capita é menor que o necessário para adquirir a cesta de produtos alimentares que supre o mínimo per capita de calorias diárias recomendada pela FAO (2100 calorias). A renda per capita utilizada como linha de pobreza varia de estado para estado.

Indigentes** = A linha de indigência é construída endogeneamente e é metade da linha de pobreza.

Nota3: As linhas de pobreza e indigência foram extraídas de IPEA
 <<http://www.ipea.gov.br>>

Acredita-se, com isso, que o modelo gerado neste trabalho, no qual utilizou variáveis de consumo, quantidade de bens duráveis, características pessoais do chefe de família e características do domicílio, apresenta um “retrato” diferente da situação socioeconômica das famílias de Minas Gerais do que as medidas que utilizam apenas a renda como variável

5 COMENTÁRIOS CONCLUSIVOS

O objetivo deste trabalho era identificar perfis socioeconômicos dos domicílios do estado de Minas Gerais através da utilização de outras variáveis, que não a renda, com o propósito de avançar na análise das deficiências e necessidades básicas da população como um todo.

Os resultados obtidos foram três diferentes *clusters* para o estado de Minas Gerais. O *cluster 1* é composto pelo grupo de domicílios com menor demanda de políticas públicas, ou seja, com favoráveis condições de vida. O *cluster 2* é o oposto do anterior, grupo com maior demanda de políticas públicas, ou seja, que apresentam piores condições de vida. E o *cluster 3* é o intermediário entre os dois.

Em relação a características sociais e demográficas dos domicílios é importante destacar aqui a forte conexão entre desigualdade racial e desigualdade de recursos historicamente presente no Brasil. Não por coincidência, os domicílios com as melhores condições de vida têm a maioria dos chefes de família da raça branca e, ao contrário, os com piores condições de vida são em sua maioria das raças preta e parda.

Além disso, é importante destacar ainda a precária condição de vida dos domicílios em área rural. Uma atenção especial no que diz respeito a políticas públicas focalizadas nesses domicílios é emergencial. Este trabalho forneceu um perfil detalhado desses domicílios. Dentre outras coisas, pôde-se perceber que esse grupo de domicílios é o que possui maior quantidade de crianças (*prop_ncrianças*) e que o maior peso no orçamento dessas famílias é com alimentos em geral. A estrutura física dos domicílios é bastante precária, com muitos possuindo piso de terra no interior das residências, sem escoamento sanitário apropriado e sem canalização interna de água. Sabe-se que políticas públicas de infra-estrutura em áreas rurais são muito onerosas, mas são consideradas fundamentais para o desenvolvimento humano dos indivíduos desses domicílios.

Uma política pública, focalizada nessas famílias, no que diz respeito a benefícios para compra de alimentos também é sugerida.

Depois da análise das condições de vida desses domicílios e seus principais tipos de gastos, foi feita uma comparação entre as medidas que utilizam a renda para classificar os domicílios e a forma como eles foram classificados neste trabalho. Encontrou-se que, as medidas geram resultados distintos. Ao se utilizar a linha de miséria do IPEA, por exemplo, foi encontrado que apenas 50,56% dos domicílios do *cluster 2* são pobres. Além disso, até mesmo no *cluster 1* foram encontrados domicílios considerados pobres (3,98%), e eles são os que possuem as melhores condições de vida da população.

Acredita-se que os resultados atingidos nesta pesquisa de fato avancem no que diz respeito à análise das condições de vida da população do estado de Minas Gerais. No que diz respeito a políticas públicas e melhora das condições de vida das populações mais necessitadas sugere-se amplo investimento em infra-estrutura de áreas rurais, visando a melhoria de escoadouros sanitários, incentivos para a exterminação do piso de terra dos domicílios e até mesmo pavimentação de vias públicas. Além disso, subsídio em alimentos em geral que são responsáveis pela maior parte dos gastos desses domicílios. Ou, ao invés de subsídio, ajuda para financeira para a compra. Acredita-se que essas medidas são fundamentais porque os domicílios de maior necessidade de melhoria nas condições de vida são os que possuem maiores proporções de crianças e jovens e boas condições de vida são fundamentais para o desenvolvimento humano desses indivíduos.

Para avanços em pesquisas futuras sugere-se uma análise da relação dos *clusters* obtidos com variáveis de saúde, tanto antropométricas quanto de acesso a serviços.

6 REFERÊNCIAS

- ALMEIDA, A. N.; FREITAS, R. E.; IPEA. **Renda e despesa familiar no Brasil segundo a pesquisa de orçamentos familiares (POF) 2002-2003**. Brasília: IPEA, 2006. 43 p.
- BARROS, Ricardo Paes de. et al. **A desigualdade de renda no Brasil encontra-se subestimada?: uma análise comparativa com base na PNAD, na POF e nas contas nacionais**. Rio de Janeiro: IPEA, 2007. 36 p.
- CROSSLEY, T. F.; PENDAKUR, K. Consumption inequality .In: GREEN, D.; KESSELMAN, J. (eds). **Dimensions of Inequality in Canada**. Canadá: UBC Press, 2002. p.1-34. Disponível em : (http://www.sfu.ca/~pendakur/pendakur_crossley.pdf). Acesso em: 24 mar. 2007.
- DEATON, A.; ZAIDI, S. **Guidelines for constructing consumption aggregates fr welfare analysis**. Washington D.C: World Bank, 2002. Working Paper n. 135.
- DEATON, A. et al. **Brazil measuring poverty using household consumption**. Washington D.C: World Bank., 2007. Report n. 38358 Disponível em: <http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2007/02/06/000090341_20070206093204/Rendered/PDF/363580BR.pdf> Acesso em : 24 de março 2007
- DUESENBERY, James S. **Income, saving and the theory of consumer behavior**. New York: Oxford University, 1967. 128p.
- FERREIRA, F.H.G., LANJOUW, P. A New Poverty profile for Brazil using PPV, PNAD and Census Data. Rio de Janeiro: Departamento de Economia PC-RIO, mar. 2001. (Texto para Discussão, 418).
- FRANK, Robert H. The demand for unobservable and other nonpositional goods. **American Economic Review**, v.75, n. 1, p.101-117, mar. 1985.
- GARCIA, F.; CASTELO, A.M. **Evolução do déficit habitacional**. São Paulo: Fundação Getúlio Vargas, 2006. 10p. Disponível em: <<http://www.sindusconsp.com.br/downloads/Relat%C3%B3rio%20Deficit%202004.pdf>> Acesso: 17 jun. 2007.
- GRUSKY, D. B. The Past, Present, and Future os Social Inequality. In: GRUSKY, D. B. **Social stratification**, 2. ed. Boulder: Westview Press, 2001. p. 3-51.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa de orçamentos familiares 2002-2003**: primeiros resultados, Brasil e grandes regiões. 2. ed. Rio de Janeiro: IBGE, 2004.
- LINDERT, K.; SKOUFIAS,E.; SHAPIRO, J. How effectively do public transfers in Latin America redistribute income? In: LATIN AMERICA AND THE CARIBBEAN ECONOMIC ASSOCIATION (LACEA),. ANNUAL MEETINGS,. apr. 29 2005. **Working papers...** [s.l.]: LACEA, 2005.
- MILLER, R.L. et al. **Economics today: the macro view**. 2nd Canadian Edition. Totonto: Longman Press, 2002.
- NORUSIS, M. **SPSS 13.0 statistical procedures companion**. Upper Saddle-River, N.J.: Prentice Hall, Inc, 2004.
- RIBEIRO, M. N. O; BARBOSA, L. M. Avaliação das condições de vida da população do Brasil,

- Nordeste e Rio Grande do Norte. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, ANEP, 15., set. 2006. **Trabalhos apresentados...** [s.l.]: [s.n.], 2006.
- SCALON, Maria Celi. **Mobilidade social no Brasil** : padrões e tendências. Rio de Janeiro: IUPERJ/UCAM; Revan, 1999. 190p.
- SILVEIRA, B. F.; BERTASSO, B. Tipologia socioeconômica das famílias das grandes regiões urbanas brasileiras e seu perfil de gastos. In: CONGRESSO DA ASSOCIAÇÃO LATINO AMERICANA DE POPULAÇÃO, 1., 2004., Caxambu, MG. **Trabalhos apresentados...** [s.l.]: ALAC, 2004.
- HAIR, J.F. et al. **Análise multivariada de dados**. 5. ed. Porto Alegre: Bookman, 2005.
- VEBLER, Thorstein Bunde. **A teoria da classe ociosa**: um estudo econômico das instituições. São Paulo: Abril Cultural, 1983. (Os Economistas).

APÊNDICE

Estatísticas descritivas e características gerais das 73 variáveis contínuas do modelo

Variable	P c1	M c1	D. P. c1	S c1	Dif c1	P c2	M c2	D. P. c2	S c2	Dif c2	P c3	M c3	D. P. c3	Sc3	Dif c3	
d_automovel	9	0.8161	0.7544687	-	sim	10	0.1400	0.3912689	+	sim	1	0.1265	0.344002	+	sim	
d_battliqfe~o	6	2.7181	1.042548	-	sim	7	1.2625	0.9885926	+	sim	44	1.9259	0.9025442	+	não	
d_beminfer~r	62	0.8018	0.9818177	-	não	65	0.7388	0.7201713	-	não	28	0.6296	0.7416054	+	sim	
d_filtroagua	57	0.8117	0.5603553	-	sim	35	0.6613	0.5237761	+	sim	63	0.7238	0.5071748	+	não	
d_fogao	30	1.1289	0.366641	-	sim	17	0.9013	0.4021201	+	sim	59	1.0386	0.2578063	-	não	
d_geladeira	51	1.1090	0.3290525	-	sim	69	1.0475	0.2128388	-	não	13	1.0139	0.117075	+	sim	
d_maqluxo	12	1.0055	1.113698	-	sim	2	0.0600	0.2626523	+	sim	4	0.1921	0.4188232	+	sim	
d_motocicl~r	19	0.4879	0.6394481	-	sim	23	0.1150	0.3491811	+	sim	6	0.0841	0.2858741	+	sim	
d_semiluxo	23	1.0308	0.9766494	-	sim	41	0.5788	0.724192	+	não	15	0.4699	0.6979431	+	sim	
d_torraasp~a	18	0.6619	0.9209084	-	sim	3	0.0288	0.1883289	+	sim	9	0.1219	0.3389031	+	sim	
d_tvsonbic~l	10	3.4703	1.711946	-	sim	9	1.7213	1.339138	+	sim	23	2.3179	1.383584	+	sim	
d_videodvd~l	7	2.3899	1.793651	-	sim	1	0.2813	0.5958962	+	sim	14	0.8835	1.007082	+	sim	
densmor	14	0.4938	0.2330968	+	sim	30	0.7120	0.5111458	-	sim	45	0.6376	0.3442197	0	-	não
des_a~icados	16	0.0157	0.0155252	+	sim	53	0.0208	0.0313039	+	não	19	0.0302	0.0311929	-	sim	
des_a~nalcoo	8	0.0060	0.0079215	+	sim	46	0.0151	0.0331658	-	não	42	0.0125	0.0170412	-	não	
des_a~salcoo	26	0.0073	0.0131582	+	sim	73	0.0113	0.0308912	-	não	31	0.0136	0.0238492	-	sim	
des_a~scados	35	0.0006	0.0033952	+	sim	59	0.0024	0.0141456	-	não	71	0.0014	0.0084783	+	não	
des_acesso~s	37	0.0231	0.051465	-	sim	68	0.0106	0.0357559	+	não	5	0.0039	0.0154066	+	sim	
des_agrega~s	56	0.0042	0.0174459	+	sim	11	0.0014	0.011655	+	sim	35	0.0115	0.0532904	-	não	
des_alim_~es	13	0.0120	0.0186573	+	sim	43	0.0290	0.056889	-	não	46	0.0249	0.0425872	-	não	
des_alim_a~a	5	0.0053	0.0106224	+	sim	27	0.0227	0.0399948	-	sim	72	0.0135	0.0223471	-	não	
des_alim_a~c	3	0.0199	0.0210313	+	sim	21	0.0635	0.0806586	-	sim	65	0.0390	0.0419755	+	não	
des_alim_c~b	1	0.0092	0.0182334	+	sim	15	0.0750	0.1127888	-	sim	25	0.0303	0.0443879	+	sim	
des_alim_e~s	55	0.0005	0.0016397	+	sim	31	0.0003	0.0020363	+	sim	41	0.0011	0.0053018	-	não	
des_alim_f~a	2	0.0035	0.007482	+	sim	20	0.0216	0.0361086	-	sim	33	0.0091	0.0160282	+	sim	
des_alim_f~s	52	0.0055	0.0100769	+	sim	50	0.0097	0.0242553	-	não	62	0.0069	0.0163538	+	não	
des_alim_h~s	4	0.0030	0.0050667	+	sim	25	0.0146	0.0269104	-	sim	22	0.0061	0.0094375	+	sim	
des_alim_o~p	17	0.0063	0.013762	+	sim	32	0.0231	0.0552028	-	sim	38	0.0111	0.0208004	+	não	
des_alim~tos	11	0.0018	0.0041333	+	sim	58	0.0048	0.0106285	-	não	30	0.0052	0.0106571	-	sim	
des_alimfo~a	39	0.0357	0.0470785	+	sim	56	0.0523	0.0858979	-	não	51	0.0498	0.0720467	-	não	
des_alutidom	45	0.0000	0.000071	+	sim	63	0.0001	0.001001	-	não	73	0.0000	0	-	não	
des_aquisi~o	28	0.0719	0.1401459	-	sim	28	0.0177	0.0622423	+	sim	10	0.0127	0.0473211	+	sim	
des_aquisi~s	48	0.0097	0.0190973	+	sim	48	0.0178	0.0442575	-	não	64	0.0127	0.0241626	+	não	
des_artigo~r	61	0.0087	0.0135104	+	não	42	0.0079	0.0177998	+	não	29	0.0124	0.021184	-	sim	
des_bolsac~o	31	0.0123	0.0140874	+	sim	67	0.0155	0.0230104	+	não	24	0.0191	0.0229416	-	sim	
des_brinqu~s	65	0.0041	0.0114362	-	não	29	0.0020	0.0059462	+	sim	66	0.0033	0.0083761	-	não	
des_cabele~a	67	0.0093	0.0102823	-	não	12	0.0052	0.008065	+	sim	32	0.0101	0.0144339	-	sim	
des_cerifa~i	66	0.0041	0.0124745	-	não	16	0.0012	0.0062246	+	sim	57	0.0038	0.0160359	-	não	
des_compra~s	49	0.0174	0.0305503	+	sim	49	0.0294	0.0634821	-	não	70	0.0226	0.0400393	+	não	
des_comuni~o	47	0.0011	0.0037626	+	sim	26	0.0007	0.0042068	+	sim	27	0.0029	0.0102368	-	sim	
des_constr~o	64	0.0282	0.0855207	-	não	54	0.0172	0.0553163	+	não	60	0.0193	0.060118	+	não	
des_decora~o	68	0.0016	0.0042879	-	não	66	0.0015	0.0074016	-	não	26	0.0009	0.0029792	+	sim	
des_divers~r	25	0.0178	0.0233933	-	sim	8	0.0030	0.0105666	+	sim	53	0.0093	0.0188975	+	não	
des_educaçã	27	0.0407	0.06751	-	sim	52	0.0176	0.0444164	+	não	2	0.0100	0.0194771	+	sim	
des_eetelint	21	0.0695	0.0472297	+	sim	34	0.0774	0.0643854	+	sim	16	0.1071	0.0728245	-	sim	
des_encarg~s	41	0.0406	0.0712794	-	sim	13	0.0126	0.0309094	+	sim	43	0.0219	0.0445981	+	não	
des_fumo	24	0.0041	0.0112564	+	sim	62	0.0092	0.0271224	-	não	40	0.0098	0.0236045	-	não	
des_habita~o	38	0.1569	0.1133029	+	sim	33	0.1595	0.1295085	+	sim	18	0.2143	0.1383225	-	sim	
des_instrm~s	73	0.0024	0.0063663	-	não	70	0.0026	0.0130903	-	não	58	0.0021	0.0054039	+	não	
des_jogose~s	69	0.0035	0.0282435	-	não	19	0.0007	0.0039333	+	sim	49	0.0015	0.0068159	+	não	

des_joiase~s	63	0.0039	0.0084637	-	não	24	0.0015	0.0065437	+	sim	50	0.0039	0.0120993	-	não
des_leitura	58	0.0019	0.0091016	-	sim	6	0.0001	0.0007509	+	sim	17	0.0002	0.0017261	+	sim
des_manut	70	0.0212	0.0819086	-	não	64	0.0204	0.0668517	-	não	34	0.0133	0.0462511	+	não
des_manutm~s	71	0.0031	0.0081741	-	não	72	0.0031	0.0174874	-	não	61	0.0026	0.0094687	+	não
des_outras~v	53	0.0016	0.0028742	+	sim	37	0.0017	0.0035772	+	sim	37	0.0027	0.0071163	-	não
des_outro~is	43	0.0469	0.1474237	-	sim	5	0.0024	0.0157392	+	sim	8	0.0043	0.0289308	+	sim
des_outro~os	20	0.0117	0.0166499	-	sim	14	0.0022	0.0075864	+	sim	3	0.0019	0.0059728	+	sim
des_outros~t	54	0.0288	0.1270936	-	sim	4	0.0007	0.0077691	+	sim	7	0.0020	0.0173486	+	sim
des_papela~a	72	0.0045	0.0080624	+	não	44	0.0037	0.0078288	+	não	52	0.0053	0.0143204	-	não
des_procli~s	15	0.0065	0.0122497	+	sim	51	0.0152	0.0276754	-	não	36	0.0155	0.0323651	-	não
des_remedios	22	0.0248	0.0329796	+	sim	61	0.0412	0.0620029	-	não	39	0.0433	0.0744076	-	não
des_roupas~a	29	0.0068	0.0119449	+	sim	57	0.0125	0.0353862	-	não	55	0.0109	0.0190066	-	não
des_roupas~m	42	0.0125	0.0174054	+	sim	40	0.0209	0.0379298	-	sim	67	0.0158	0.0243323	+	não
des_roupas~o	36	0.0032	0.0072998	+	sim	71	0.0051	0.0140095	-	não	48	0.0059	0.0165919	-	não
des_roupas~r	44	0.0123	0.0167154	+	sim	47	0.0128	0.025145	+	não	21	0.0200	0.0294021	-	sim
des_servdo~o	32	0.0177	0.0373011	-	sim	55	0.0054	0.0375031	+	não	11	0.0030	0.0127168	+	sim
des_serviç~e	40	0.0361	0.0485052	-	sim	22	0.0143	0.0388791	+	sim	69	0.0251	0.0562356	+	não
des_serviç~s	34	0.0164	0.0401032	-	sim	18	0.0027	0.0142579	+	sim	12	0.0033	0.0096756	+	sim
des_transp~e	50	0.0555	0.0663275	-	sim	45	0.0353	0.0722126	+	não	56	0.0410	0.0693925	+	não
des_utensi~a	33	0.0009	0.0027098	+	sim	36	0.0027	0.0084563	-	sim	47	0.0014	0.0034849	+	não
des_viagens	60	0.0200	0.0336143	-	não	60	0.0194	0.042178	-	não	20	0.0125	0.0256386	+	sim
prop_depend	46	0.2776	0.2632416	+	sim	39	0.3659	0.2989059	-	sim	54	0.3412	0.2933926	-	não
prop_ncria~s	59	0.0721	0.1354036	+	não	38	0.1100	0.1706037	-	sim	68	0.0869	0.1543022	+	não

Fonte: Elaboração Própria

Nota1: Dados extraídos da POF2002-2003

Nota2: P = posição da variável por ordem de importância no *cluster*; M = Média da variável; D.P. = Desvio Padrão; S = sinal do t-estatístico (*Student's t*) do *cluster*. Para valores negativos (-), a média geral da amostra total é geralmente menor que a média geral do *cluster*. Para valores positivos (+) a média geral da amostra total é maior que a média geral do *cluster*. Dif= a variável contribui na formação do *cluster*? (ou seja, o t-estatístico, *Student's t*, é significante ou não?); c1 = *cluster 1*; c2 = *cluster 2*; c3 = *cluster 3*.