

# AVALIAÇÃO DO EMPREGO DAS TÉCNICAS DE ANÁLISE DE CORRESPONDÊNCIA E ANÁLISE DE CONGLOMERADOS EM TESES E EM DISSERTAÇÕES DE ALGUMAS INSTITUIÇÕES DE ENSINO SUPERIOR

*Evaluating the use of the techniques of correspondence analysis and cluster analysis in theses and dissertations of some higher education institutions*

*Maria Aparecida Gouvêa*

Professor do Curso de Graduação e Pós-Graduação da FEA/USP – São Paulo – SP, Brasil. E-mail: [magouvea@usp.br](mailto:magouvea@usp.br)

*Leandro Campi Prearo*

Professor e Pesquisador da Universidade Municipal de São Caetano do Sul – São Caetano do Sul – SP, Brasil.

E-mail: [leandro.prearo@uscs.edu.br](mailto:leandro.prearo@uscs.edu.br)

*Maria do Carmo Romeiro*

Professora e Pesquisadora da Universidade Municipal de São Caetano do Sul – São Caetano do Sul – SP, Brasil.

E-mail: [mromeiro@uscs.edu.br](mailto:mromeiro@uscs.edu.br)

## Resumo

Este artigo faz parte de um amplo estudo de avaliação da adequação no uso de técnicas estatísticas multivariadas em teses e em dissertações de algumas instituições de ensino superior na área de marketing na temática do comportamento do consumidor entre 1997 e 2006. Neste artigo são focalizadas as técnicas de análise de correspondência e análise de conglomerados, duas técnicas com grande potencial de uso em estudos de marketing. Foi objetivo deste trabalho a análise da adequação do emprego dessas técnicas às necessidades dos problemas de pesquisa apresentados nas teses e nas dissertações e a aferição do nível de acerto no atendimento de suas premissas. De forma geral, os resultados sugerem a necessidade de um aumento do comprometimento dos pesquisadores na verificação de todos os preceitos teóricos de aplicação dessas técnicas classificadas na categoria de agrupamento.

**Palavras-chave:** Métodos Quantitativos. Análise de Correspondência. Análise de Conglomerados. Comportamento do Consumidor.

## Abstract

This paper is part of a large study to assess the adequacy of the use of multivariate statistical techniques in theses and dissertations of some higher education institutions in the area of marketing in the theme of consumer behavior from 1997 to 2006. The correspondence analysis and cluster analysis are focused on in this paper, two techniques with great potential of using in marketing studies. The objective of this study was to analyze whether the employment of these techniques suits the needs of the research problem presented in these theses and dissertations as well as to evaluate the level of meeting of their assumptions. Overall, the results suggest the need for more involvement of researchers in the verification of all the theoretical precepts of application of these techniques classified in grouping category.

**Key words:** Quantitative Methods. Correspondence Analysis. Analysis of Conglomerates. Consumer Behavior.



## 1 INTRODUÇÃO

A partir da década de 1990 tem sido intensificada a atenção de acadêmicos em relação à avaliação da qualidade da produção científica da área de Administração. Nesse sentido, destacam-se algumas referências de autores com abordagem dessa questão em campos específicos da Administração, a saber: Administração Pública (MACHADO-DA-SILVA *et al.*, 1989); Finanças (LEAL *et al.*, 2003); Marketing (POWERS *et al.*, 1998; BOTELHO e MACERA, 2001); Métodos Quantitativos (BREI; LIBERALI, 2004); Operações (ARKADER, 2003), Organizações (MACHADO-DA-SILVA *et al.*, 1990; BERTERO e KEINERT, 1994; VERGARA e CARVALHO, 1995; BERTERO; CALDAS; WOOD Jr., 1999; RODRIGUES FILHO, 2002), Pesquisa em Administração (MARTINS, 1994; TORRES, 2000; PERIN, 2002); Recursos Humanos (ROESCH *et al.*, 1997; CALDAS *et al.* 2002; TONELLI *et al.*, 2003), Sistemas de Informação. (HOPPEN, *et al.*, 1998)

Esses estudos estão, em geral, norteados por duas linhas de abordagem: as discussões sobre aspectos epistemológicos e as discussões sobre aspectos metodológicos, esta segunda com ênfase na adoção de critérios de qualidade e consistência da produção. Entretanto, na literatura pesquisada, a segunda linha parece ainda contar com um número reduzido de estudos.

Com relação ao uso de métodos quantitativos, Martins (1994), evidenciou, em estudo sobre a Epistemologia da Pesquisa em Administração, que as dissertações e teses apresentadas na FEA/USP, FEA/PUC e EASP/FGV, entre os anos de 1980 e 1993, abusavam da utilização deste instrumental de forma bastante superficial, sem consideração do nível de complexidade e de sofisticação dessas técnicas. O autor alertou ainda para a despreocupação para com as premissas teóricas de aplicação da maior parte dos métodos.

Para Gamboa (1987), a importância da análise da produção científica, amparada nas competências metodológicas, justifica-se, visto que a sua questão é de fundamental e decisiva importância para o desenvolvimento e para os resultados da pesquisa científica.

Conforme o número de variáveis focalizadas simultaneamente nas análises quantitativas, têm-se os seguintes contextos: univariado (enfoque analítico de apenas uma variável por vez), bivariado (enfoque simultâneo de duas variáveis) e multivariado (enfoque

simultâneo de três ou mais variáveis). O uso crescente de técnicas de análise multivariada decorre da contribuição que esse contexto pode oferecer à análise quantitativa dos dados para a produção científica e do fato de que, nas últimas décadas, vários pacotes computacionais estatísticos se aperfeiçoaram sobremaneira no sentido de tornar seus conteúdos distantes das complexidades matemáticas, próprias deles.

A evolução tecnológica notória nos *softwares* estatísticos tem proporcionado maior facilidade na operacionalização das ferramentas de análise multivariada, contribuindo para a diminuição de barreiras para os usuários e atendendo adequadamente à demanda dos cientistas das ciências sociais aplicadas, entre elas a Administração. Entretanto, tem-se constatado a ocorrência de erros de aplicação, seja na inadequação dos objetivos de uso das ferramentas com os objetivos propostos na pesquisa empírica, seja na violação de premissas de aplicação das técnicas. Há situações em que o pesquisador apenas exercita o emprego de uma técnica e distancia-se de seu problema de pesquisa e do alcance dos objetivos inicialmente traçados.

Martins (1994) afirma que os métodos quantitativos são, em Administração, mais aplicados pelos autores de pesquisa nas áreas de Marketing, Produção e Finanças, com uso mais intenso pela primeira, especialmente da análise multivariada em decorrência da necessidade de se conhecer o mercado consumidor, por meio da mensuração das opiniões, atitudes, preferências, perfil e outras características dos consumidores. (MALHOTRA, 2001)

Nesse sentido, os estudos em Marketing, especialmente na temática do Comportamento do Consumidor, vêm se utilizando sobremaneira das ferramentas de análise multivariada. Para Milagre (2001), o uso da técnica multivariada tornou-se mais comum a partir do momento que os acadêmicos e profissionais de Marketing passaram a aplicá-la em estudos sobre a preferência e satisfação do consumidor, bem como o seu perfil e comportamento de compras.

Assim, neste estudo pretende-se aprofundar a discussão e oferecer subsídios à reflexão metodológica na temática dos Métodos Quantitativos, especialmente sob o recorte de suas aplicações nos estudos em Marketing sobre o Comportamento do Consumidor.

Entre as técnicas estatísticas multivariadas, a análise de correspondência e a análise de conglomerados

dos destacam-se pela ampla possibilidade de aplicação em Marketing e por pertencerem à mesma categoria de técnicas: agrupamento de objetos ou variáveis (JOHNSON; WICHERN, 1998). A análise de correspondência é uma técnica com grande potencial de aplicação na área de Marketing, sobretudo pelo fato de possibilitar a realização de estudos de posicionamentos de marcas (produtos ou serviços) associados a características do mercado efetivo e/ou potencial. Em um ambiente altamente competitivo, como por exemplo, o setor bancário, o uso dessa técnica pode facilitar a visualização da posição relativa de bancos concorrentes e as respectivas correspondências a diferentes perfis de mercado, podendo tal análise dar suporte para a identificação de estratégias de posicionamento. A análise de conglomerados (*cluster analysis*), por sua vez, é intensamente empregada em Marketing, sobretudo pelo fato de possibilitar a atividade de segmentação de mercado, um dos principais alicerces para a compreensão do comportamento do consumidor.

Devido à importância dessas técnicas no campo de marketing, decidiu-se selecioná-las como foco de análise neste trabalho, com os seguintes objetivos a serem alcançados:

- a) Identificar a intensidade de uso das técnicas multivariadas na área de Marketing, temática do Comportamento do Consumidor, tendo como unidade de análise as dissertações e as teses, do período entre 1997 e 2006, de duas universidades públicas: Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo e Escola de Administração da Universidade Federal do Rio Grande do Sul.
- b) Identificar o nível de adequação do uso das técnicas de análise de correspondência e de análise de conglomerados aos problemas empíricos propostos nessas dissertações e nas teses selecionadas.
- c) Identificar fontes de erros da aplicação das técnicas de análise de correspondência e análise de conglomerados, a partir do não atendimento às suas premissas básicas, no conjunto da produção científica examinada.

A seleção dos Programas de Pós-Graduação, alvo deste estudo, deve-se à necessidade de delimitação, devido à dificuldade operacional de um levantamento amostral representativo da produção nacional, principalmente pela indisponibilidade de material em base

de dados *on-line* por uma parcela importante das Instituições do País.

Nesse sentido, optou-se inicialmente por um recorte focado nos Programas de Pós-Graduação de alta *performance* na última avaliação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Registre-se que a opção por esse critério não se motivou pela suposição de que os programas com avaliação menos positiva, abaixo da nota 6, tratariam as técnicas em estudo de forma mais ou menos corretas. A avaliação da CAPES aponta três Instituições com nota 6: a FEA/USP, a EA/UFRGS e a EASP/FGV.

Um segundo critério deu-se da necessidade de garantir a maior homogeneidade possível da população-alvo, já que não é objetivo específico desse estudo a comparação entre Instituições. Nesse sentido, a seleção levou em conta a natureza da Instituição: pública ou privada. Considerando que entre as três citadas, duas são públicas, optou-se pela análise das dissertações e teses das Instituições públicas: FEA/USP e EA/UFRGS.

Destacam-se alguns fatores restritivos na abrangência dessa investigação:

- a) delimitação do público-alvo: os resultados desse estudo são apenas válidos para as dissertações e teses da FEA/USP e da EA/UFRGS, de 1997 a 2006; e
- b) as informações sobre o atendimento às premissas da análise de correspondência e da análise de conglomerados são baseadas nos relatos dos autores de cada dissertação ou tese sob análise, que evidentemente podem conter uma descrição subestimada ou superestimada do que foi realizado.

## 2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção será realizada uma introdução sobre análise multivariada, seguida por uma apresentação geral de técnicas multivariadas, finalizando com considerações sobre as técnicas de análise de correspondência e análise de conglomerados.

### 2.1 Análise multivariada

A análise multivariada refere-se a um conjunto de métodos estatísticos que torna possível a análise

simultânea de medidas múltiplas para cada indivíduo, objeto ou fenômeno observado (CORRAR *et al.*, 2007). Além disso, a análise multivariada permite estudar e evidenciar as ligações, as semelhanças e as diferenças existentes entre todas as variáveis envolvidas no processo. (BOUROCHE; SAPORTA, *apud* TRIVELLONI; HOCHEIM, 1998)

Segundo Steiner (1995), a necessidade de entender a relação entre diversas variáveis aleatórias faz da análise multivariada uma metodologia com grande potencial de uso.

Para Lourenço e Matias (2001), as técnicas estatísticas multivariadas são mais complexas do que aquelas da estatística univariada. Apesar dessa complexidade teórica fundamentada na matemática, as técnicas multivariadas, por permitirem o tratamento de diversas variáveis simultaneamente, podem oferecer ao pesquisador um material bastante robusto para a análise dos dados da pesquisa.

Hair *et al.* (2005) corrobora essa abordagem, enfatizando que a análise multivariada auxilia na formulação de questões relativamente complexas de forma específica e precisa, possibilitando a condução de pesquisas teoricamente significativas.

Nesse sentido, conforme Pereira (2001), a multiplicidade de conceitos estatísticos e matemáticos inerentes na análise multivariada representa um grande desafio aos pesquisadores que não pertencem ao campo de matemática e estatística. Para não incorrer em má utilização das técnicas multivariadas, é necessário que o pesquisador tenha os conhecimentos essenciais que o tornem apto a usar produtivamente os *softwares* estatísticos e a interpretar os resultados com propriedade.

## 2.2 Técnicas Estatísticas de Análise Multivariada

A escolha dos métodos e dos tipos de análises empregadas nos trabalhos científicos deve ser determinada pelo problema de pesquisa. Nesse sentido, Johnson e Wichern (1998) propõem uma classificação dos objetivos para atendimento do problema em cinco categorias, as quais determinam as técnicas a serem utilizadas, a saber: investigação da dependência entre as variáveis, predição, construção de hipóteses e de

testes, redução dos dados ou simplificação estrutural e agrupamento de objetos ou variáveis.

CLASSIFICAÇÃO	TÉCNICAS RELACIONADAS
<b>Investigação da dependência entre as variáveis</b> Todas as variáveis são mutuamente independentes ou uma ou mais variáveis são dependentes de outras.	Análise Discriminante Análise de Regressão Correlação Canônica Regressão Logística Análise Conjunta MANOVA
<b>Predição</b> As relações entre as variáveis devem ser determinadas com o objetivo de prever o valor de uma ou mais variáveis com base nas observações de outras variáveis.	Análise Discriminante Análise de Regressão Análise de Regressão Logística
<b>Construção de hipóteses e testes</b> Hipóteses estatísticas específicas, formuladas em termos de parâmetros da população multivariada, são testadas. Isso pode ser feito para validar premissas ou para reforçar convicções prévias.	Modelagem de Equações Estruturais Análise Fatorial Confirmatória
<b>Redução dos dados ou simplificação estrutural</b> O fenômeno em estudo é representado de um modo tão simples quanto possível sem sacrificar informações importantes.	Análise Fatorial Exploratória
<b>Agrupamento de objetos ou variáveis</b> Grupos de objetivos ou variáveis "similares" são criados com base nas medidas características.	Análise de Conglomerados Análise de Correspondência Escalonamento Multidimensional

Quadro 1: Categorias dos objetivos das técnicas estatísticas de análise multivariada

Fonte: Adaptado de Johnson e Wichern (1998)

Após considerar os objetivos do problema de pesquisa, o passo seguinte para a escolha da técnica de análise multivariada é verificar o tipo de relação examinada, número de variáveis dependentes e tipo de escala utilizada.

Sobre o tipo de relação, as técnicas são classificadas como de dependência ou de interdependência. Na primeira situação, uma ou mais variáveis (variáveis dependentes) podem ser explicadas ou preditas por outras (variáveis independentes). Na segunda, todas as variáveis são analisadas simultaneamente, sem a orientação de dependência ou independência.

Sobre o tipo de escala utilizada, pode-se generalizar a classificação teórica dessas escalas de mensuração em dois grandes grupos: variáveis métricas e variáveis não métricas.

Do cruzamento entre o tipo de relação examinada, número de variáveis dependentes e tipo de escala utilizada, tem-se um esquema de classificação para o conjunto de técnicas, exibido nos Quadros 2 e 3.

		VARIÁVEL DEPENDENTE OU PREDITA			
		UMA VARIÁVEL		DUAS OU MAIS VARIÁVEIS	
		MÉTRICA	NÃO MÉTRICA	MÉTRICA	NÃO MÉTRICA
VARIÁVEL INDEPENDENTE OU PREDITOR	UMA VARIÁVEL				
	MÉTRICA	Regressão Simples	Análise Discriminante Regressão Logística	Correlação Canônica	Correlação Canônica Não-Paramétrica
	NÃO-MÉTRICA	Regressão Simples com variável <i>dummy</i>	Regressão Logística ANOVA <sup>(1)</sup>	Análise da Variância Multivariada (MANOVA)	Correlação Canônica Não-Paramétrica
	DUAS OU MAIS VARIÁVEIS				
	MÉTRICA	Regressão Múltipla	Análise Discriminante Regressão Logística	Correlação Canônica Modelagem de Equações Estruturais (SEM)	Correlação Canônica Não Paramétrica
	NÃO-MÉTRICA	Análise Conjunta Regressão Múltipla com variável <i>dummy</i>	Regressão Logística Análise Conjunta	Análise da Variância Multivariada (MANOVA) Modelagem de Equações Estruturais (SEM)	Correlação Canônica Não Paramétrica

<sup>(1)</sup> A ANOVA é considerada uma técnica estatística de análise univariada, não fazendo parte, portanto, do escopo deste estudo.

Quadro 2: Métodos de dependência  
Fonte: Adaptado de Sharma (1996)

VARIÁVEIS MÉTRICAS	VARIÁVEIS NÃO MÉTRICAS
Análise Fatorial Exploratória Análise Fatorial Confirmatória Análise de Conglomerados Escalonamento Multidimensional	Análise de Correspondência Modelo Loglinear Escalonamento Multidimensional Análise de Conglomerados

Quadro 3: Métodos de interdependência  
Fonte: Adaptado de Sharma (1996)

### 2.3 Análise de Correspondência e Análise de Conglomerados

Esse tópico apresenta sucintamente os principais conceitos e premissas das técnicas focalizadas neste estudo, observando-se a relevância do respeito às premissas como uma condição necessária para uma análise credível dos resultados.

#### 2.3.1 Análise de Correspondência

As primeiras considerações matemáticas a respeito da Análise de Correspondência foram feitas por Hirschfeld (1935). A partir daí, os procedimentos numéricos e algébricos foram aplicados em diferentes contextos, notadamente em ecologia e psicologia. A técnica foi redescoberta na França no início da década de 1960 e tem sido extensivamente usada naquele País como um método gráfico de análise de dados. Desde 1975, a técnica vem sendo utilizada em diversas áreas do conhecimento, inclusive Marketing, em publicações de diversos idiomas.

Segundo Hair *et al.* (2005), a Análise de Correspondência é uma técnica de interdependência que facilita tanto a redução dimensional da classificação dos objetos em um conjunto de atributos quanto o mapeamento perceptual de objetos relativo a esses atributos.

Nesse sentido, a técnica de Análise de Correspondência se enquadra na categoria de técnica de agrupamento de casos, sendo seus objetivos principais apresentados no Quadro 4.

CATEGORIA	PRINCIPAL OBJETIVO INDICADO NO ESTUDO	FONTES
Agrupamento	Encontrar e representar as relações existentes entre as variáveis em um espaço multidimensional, reduzindo a dimensão de variáveis com dados qualitativos para obter um pequeno número de fatores.	Pestana e Gageiro (2000)
	Descobrir relações existentes entre variáveis nominais a partir de uma tabela de correspondências colocadas em um espaço de poucas dimensões.	SPSS (2001)
	Retratar a correspondência de categorias de variáveis, por meio de mapas perceptuais.	Hair <i>et al.</i> (2005)
	Fatorar variáveis categóricas e as exibir em um espaço de propriedades que apresente a associação dessas variáveis em duas ou mais dimensões.	Garson (2007a)

Quadro 4: Objetivo de aplicação da técnica de Análise de Correspondência

Fonte; Elaborado pelos autores deste artigo

Sobre as premissas subjacentes ao uso da Análise de Correspondência, dado o caráter exclusivamente exploratório e sem apoio de testes de significância, não foi encontrada nenhuma premissa fundamental na literatura selecionada. Apenas algumas considerações operacionais são citadas, entre elas:

- A necessidade de visualização da relação das variáveis em uma tabela de dupla entrada, o que conduz a um número restrito de categorias por variável. No caso de variáveis contínuas é necessário um procedimento de categorização, apesar de haver uma inevitável perda de informação nesse processo. (GARSON, 2007a)
- O número mínimo de categorias por variável é 3, já que nos *softwares* estatísticos são produzidos mapeamentos cujo número de dimensões é o mínimo entre o número de linhas e colunas em uma tabela de correspondência menos 1. Assim, se uma variável tiver duas categorias, o número de dimensões será 1, não sendo

produzido nenhum mapeamento. Aliás, o uso da análise de correspondência para variáveis com poucas categorias parece não agregar mais informação do que a inspeção descritiva dessas variáveis. (GARSON, 2007a)

- O tamanho da amostra deve garantir que as frequências esperadas, calculadas no teste qui-quadrado, que precede esta técnica, sejam maiores ou iguais a 5. (HAIR *et al.*, 2005)

O Quadro 5 sintetiza as premissas da técnica de análise de correspondência.

PREMISSAS SUBJACENTES	FONTES
Número reduzido de categorias por variável e categorização de variáveis contínuas.	Garson (2007a)
Número de categorias por variável deve ser maior ou igual a 3.	Garson2(007a)
Tamanho da amostra suficiente para que as frequências esperadas sejam maiores ou iguais a 5.	Hair <i>et al.</i> (2005)

Quadro 5: Premissas subjacentes ao uso da Análise de Correspondência

Fonte: Elaborado pelos autores deste artigo

### 2.3.2 Análise de Conglomerados

A Análise de Conglomerados, também chamada de Análise de *Cluster*, Análise de Agrupamentos, Análise de Taxionomia e Análise de Segmentação, foi primeiramente usada por Tyron (1939) e abriga uma variedade de algoritmos de classificação diferentes, todos voltados para a questão de como organizar dados observados em estruturas que façam sentido, ou como desenvolver taxonomias capazes de classificar dados observados em diferentes grupos.

Quanto aos objetivos de aplicação, a análise de conglomerados enquadra-se na categoria de técnica de agrupamento de objetos ou variáveis.

Os principais algoritmos de classificação disponíveis nos *softwares* estatísticos comerciais são o método hierárquico, o método *K-means* e o método *Two Step Cluster*. Considerando que os três algoritmos atendem aos mesmos objetivos de aplicação, a opção por um deles é feita a partir, principalmente, do atendimento a premissas específicas de cada técnica.

OBJETIVO	DESCRIÇÃO	FONTES
Agrupamento	Detectar grupos homogêneos nos dados, podendo os grupos ser constituídos por variáveis ou casos.	Pestana e Gageiro (2000)
	Agrupar sujeitos ou variáveis em grupos homogêneos a partir de uma ou mais características comuns.	Maroco (2003)
	Agregar objetos com base nas características que eles possuem.	Hair <i>et al.</i> (2005)
	Agrupar indivíduos semelhantes de acordo com a similaridade das variáveis em estudo.	Mingoti (2005)
	Identificar um conjunto de grupos que minimize a variação dentro do grupo e maximize a variação entre os grupos.	Garson (2007b)

Quadro 6: Objetivo de aplicação da técnica de Análise de Conglomerados

Fonte; Elaborado pelos autores deste artigo

O método de agrupamento hierárquico é o mais apropriado para amostras inferiores a 250 casos (GARSON, 2007b). Nesse algoritmo, é o pesquisador quem define como a semelhança ou distância será estabelecida, como se agregam os agrupamentos e quantos agrupamentos são necessários (de acordo com o esquema de aglomeração), com aporte ou não da teoria.

O método *K-means* é mais apropriado para amostras grandes (a partir de 250 casos), ainda que também possa ser aplicado em amostras menores. Utiliza como método de agrupamento a Distância Euclideana e exige que o pesquisador especifique *a priori* o número desejado de agrupamentos.

O método *Two Step Cluster* é o mais apropriado para amostras com uma ou mais variáveis categóricas e também é recomendado para grandes bases de dados.

Sobre as premissas para aplicação da técnica e, considerando o fato de que a Análise de Conglomerados é construída fundamentalmente a partir de medidas matemáticas de semelhança ou de medidas de distância (dissimilaridade), tem-se essa técnica como uma das mais elementares do ponto de vista da robustez analítica. (MAROCO, 2003)

Nesse sentido, as premissas subjacentes ao uso da técnica parecem bastante elementares do ponto de vista da análise estatística:

- a) Variáveis métricas: os métodos, hierárquico e *K-Means*, exigem escala métrica ou admitem a escala não métrica somente se esta for apresentada na forma binária; o método *Two Step Cluster* admite variáveis em qualquer nível: métricas ou não métricas. (HAIR *et al.*, 2005; GARSON, 2007b)
- b) Padronização das variáveis: variáveis com magnitudes diferentes precisam ser padronizadas antes do cálculo de distância entre objetos a serem posicionados no mapeamento gerado pela técnica. (HAIR *et al.*, 2005; GARSON, 2007b)
- c) Ausência de observações atípicas: as observações atípicas ou extremas (*outliers*) são aquelas substancialmente diferentes das outras. Podem ser identificadas no nível univariado, bivariado ou multivariado.
- d) Alguns fatores podem causar a presença de observações atípicas univariadas, alterando o resultado de análise e causando violações de normalidade: erro de entrada de dados ou na codificação; evento extraordinário; observações extraordinárias para os quais o pesquisador não tem explicação; observações que representam uma combinação única de valores das variáveis (HAIR *et al.*, 2005; GARSON, 2007c). Sobre os *outliers* multivariados (casos com valores extremos no conjunto de múltiplas variáveis), há três medidas mais comumente usadas para identificá-los: Distância de Cook, Distância de Mahalanobis e Distância de Leverage.
- e) Ausência de multicolinearidade: segundo Gujarati (2000, p. 318), o termo multicolinearidade foi cunhado por Ragnar Frish em 1934. Originalmente, significa a existência de uma “perfeita” (ou exata) relação linear entre algumas ou todas as variáveis explicativas de um modelo de regressão. Formalmente, o termo multicolinearidade refere-se à existência de mais de uma relação linear exata, ao passo que o termo colinearidade refere-se à existência de uma única relação linear.
- f) Algumas sugestões para tratamento de multicolinearidade: ignorar, se o objetivo é prever;

eliminar variáveis; reformular o modelo – usando, por exemplo, a razão entre variáveis; aumentar o tamanho da amostra (TABACHNICK; FIDELL, 1996; GUJARATI, 2000; PESTANA; GAGEIRO, 2000; HAIR *et al.*, 2005; GARSON, 2007c). Testes estatísticos para avaliação da multicolinearidade: MSA (Measure of Sampling Adequacy), KMO (Kaiser-Meyer-Olkin), VIF (Variance inflation factor), Índice de condição, Índice de tolerância (HAIR *et al.*, 2005; GARSON, 2007d).

- g) Sensibilidade ao tamanho da amostra: aliado a um correto procedimento de amostragem, o tamanho da amostra deve ser adequado para permitir a generalização dos resultados, os quais podem ser verificados quanto à significância estatística dos testes. No modelo hierárquico, o tamanho da amostra não pode superar 250 casos; em contrapartida, este é o limite mínimo para os modelos *K-Means* e *Two Step Cluster*. (GARSON, 2007b)

O Quadro 7 sintetiza as premissas da técnica de análise de conglomerados.

PREMISSAS SUBJACENTES	FONTES
Variáveis métricas	Hair et al. (2005); Garson (2007b)
Padronização das variáveis	Hair et al. (2005); Garson (2007b)
Ausência de observações atípicas	Hair et al. (2005); Garson (2007b)
Ausência de multicolinearidade	Hair et al. (2005)
Tamanho da amostra	Hair et al. (2005); Garson (2007b)

Quadro 7: Premissas subjacentes ao uso da Análise de Conglomerados

Fonte: Elaborado pelos autores deste artigo

### 3 METODOLOGIA DE PESQUISA

Nessa seção serão apresentados os tópicos referentes: à caracterização da população, ao desenho metodológico da pesquisa empírica, à construção do critério de adequação da aplicação das técnicas análise de correspondência e da análise de conglomerados, ao instrumento de coleta de dados e às variáveis de resultados.

#### 3.1 Caracterização da População-Alvo

Retomando as considerações iniciais sobre o público-alvo, a população sob a análise contempla as dissertações e teses em Administração, área de Marketing, temática do Comportamento do Consumidor, apresentadas nos Programas de Pós-Graduação da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo e da Escola de Administração da Universidade Federal do Rio Grande do Sul, no período entre 1997 e 2006.

O universo de interesse foi definido como a dissertação ou a tese apresentada aos PPGA das instituições públicas com alta avaliação pela CAPES na área de Administração, com foco em Marketing, especificamente comportamento do consumidor, com utilização das técnicas estatísticas: análise de correspondência e/ou análise de conglomerados como instrumento de solução do problema de pesquisa.

A unidade populacional, no entanto, refere-se à aplicação de técnica estatística multivariada nesse universo de interesse, podendo ser mais do que uma aplicação em cada estudo integrante do universo de interesse.

A identificação da população-alvo foi realizada em três etapas:

- Seleção dos trabalhos que apresentavam o termo comportamento do consumidor como uma das palavras-chaves.
- Leitura dos resumos dos trabalhos não selecionados na primeira etapa, com o objetivo de identificar aqueles que tratavam do comportamento do consumidor, ainda que não apresentassem o termo como palavra-chave e, assim, selecioná-los.
- Exame do conteúdo dos trabalhos selecionados na primeira e na segunda etapa, identificando aqueles com aplicação de alguma técnica de análise multivariada.

A identificação da população-alvo desse estudo registrou 196 dissertações e teses sobre Comportamento do Consumidor, 56 dissertações e teses com aplicação de técnicas estatísticas multivariadas e 99 aplicações de técnicas estatísticas de análise multivariada (população-alvo do estudo).

### 3.2 Desenho Metodológico da Pesquisa

A construção metodológica da pesquisa empírica foi orientada, inicialmente, pela avaliação do processo de solução do problema gerador de dissertações e de teses com aplicação de técnica estatística de análise multivariada, na temática aqui selecionada.

Considerou-se, ainda, o fato de que esta avaliação exige ser orientada por um critério, o qual foi construído dentro desse próprio estudo, não tendo sido submetido a um processo de validação anterior.

Essas duas condições, por si só, remetem esse estudo ao âmbito da pesquisa exploratória, visto que o processo de aprofundamento do entendimento do problema é uma etapa aqui cumprida para subsidiar a construção do critério de avaliação da adequação da aplicação da estatística multivariada aos trabalhos selecionados do período 1997-2006.

Nesse sentido, este estudo insere-se na abordagem quantitativa, utilizando a análise de conteúdo.

O delineamento exploratório para este estudo vai ao encontro da abordagem de Selltiz (1974) que evidencia ser uma das finalidades desse método a apresentação de um recenseamento de problemas considerados urgentes por pessoas que trabalham em determinado campo de relações sociais. Exemplifica essa convergência o alerta feito em outros estudos de que a rigidez teórica, explicitada nas premissas para aplicações das técnicas estatísticas de análise multivariada, parece não ser acompanhada de rigidez empírica, explicitada nas concessões feitas pelos pesquisadores, o que, por vezes, pode resultar em conclusões não precisas sobre a solução dos problemas.

A opção pelo uso da técnica de análise de conteúdo foi orientada especialmente pela necessidade primária desse estudo de interpretar a situação problema das dissertações e das teses selecionadas com o propósito de identificar o processo de sua solução por meio da seleção de uma técnica estatística de análise multivariada pertinente.

Da mesma forma como encontrado na abordagem de vários autores, entre eles Richardson (1999) e Rocha e Deusdará (2005), também aqui, a identificação precisa da natureza da técnica de análise oscila entre a discussão quantitativa e a qualitativa.

Assim, de um lado é buscada uma objetividade bastante intensa (RICHARDSON, 1999) para cate-

gorização dos problemas das dissertações e teses, mediante a definição de critérios rígidos de julgamento da solução desses problemas. De outro, o processo geral de avaliações a ser implementado às unidades de análise (dissertações e teses) está contaminado por julgamentos, na medida em que a análise do conteúdo supõe também o exame das características ausentes ou registros parciais do atendimento às premissas da técnica estatística utilizada.

Nesse sentido, algumas definições de análise de conteúdo parecem contemplar o caminho metodológico da coleta de dados, de acordo com propostas encontradas em Janis *et al.* (*apud* RICHARDSON, 1999, p. 222):

Assim, a análise de conteúdo pode ser definida como qualquer técnica: na base de regras explicitamente formuladas e sempre quando os juízos do analista sejam considerados como relatórios de um observador científico.

Entretanto, a definição encontrada em Bardin (1979) propicia a convivência com as duas abordagens, na medida em que inclui em sua definição de análise de conteúdo, a geração de indicadores quantitativos ou não no processo analítico.

Análise de conteúdo é um conjunto de técnicas de análise das comunicações, visando obter, através de procedimentos sistemáticos e objetivos de descrição dos conteúdos das mensagens, indicadores (quantitativos ou não), que permitam inferir conhecimentos relativos às condições de produção dessas mensagens. (BARDIN, 1979, p. 31)

### 3.3 A Construção do Critério de Avaliação da Aplicação das Técnicas Estatísticas: análise de correspondência e/ou análise de conglomerados

Dois critérios são apresentados a seguir com a finalidade de avaliar o grau de acerto na aplicação das duas técnicas focalizadas neste trabalho.

#### 3.3.1 Procedimento de Categorização do Problema de Pesquisa da Unidade de Análise: Critério 1

Embora várias técnicas de análise multivariada tenham sido utilizadas no material analisado, neste artigo serão destacados os resultados correspondentes à pertinência e adequação do emprego da análise de correspondência e/ou da análise de conglomerados. A orientação para categorizar o problema de pesquisa de cada unidade de análise (unidade “i” de análise) foi dada pela finalidade ou objetivo teórico de aplicação de diferentes técnicas estatísticas de análise multivariada. Segundo Johnson e Wichern (1998, p. 2), as categorias são sintetizadas em cinco modalidades:

- a) redução dos dados ou simplificação estrutural;
- b) agrupamento;
- c) dependência entre variáveis;
- d) predição; e
- e) formulação de hipóteses e testes.

No caso da análise de correspondência e da análise de conglomerados, o objetivo alcançado em seu uso é de agrupamento de observações.

A categoria identificada na unidade “i” de análise foi confrontada com a categoria-objetivo da técnica análise de correspondência e/ou análise de conglomerados, na condição de tratamento estatístico aplicado para solução do problema. A avaliação feita por meio desse confronto tem um caráter dicotômico, ou seja, foi considerada como uma aplicação adequada das referidas técnicas quando a categoria-objetivo de aplicação das técnicas ajustou-se à categoria do problema de pesquisa da unidade “i” de análise.

### 3.3.2 Procedimento de Avaliação do Atendimento às Premissas das Técnicas Estatísticas: análise de correspondência e/ou análise de conglomerados: Critério 2

O atendimento às premissas das técnicas aqui focalizadas foi decorrente da confirmação de que todas as premissas postuladas pela teoria foram atendidas. A violação de pelo menos uma das premissas, explicitada no conteúdo apresentado na unidade “i” de análise, bem como a não explicitação da situação de cada premissa (atendimento ou não atendimento) resultou

na categoria “não atendimento às premissas”. Dessa forma, o critério de avaliação do atendimento às premissas é dicotômico: atendimento a todas as premissas das técnicas e não atendimento a pelo menos uma das premissas dessas técnicas.

### 3.3.3 Avaliação Final do Nível de Adequação do Uso da Análise de Correspondência e/ou da Análise de Conglomerados

O registro do nível de adequação do uso destas técnicas foi feito em três níveis (Figura 1):

- a) Nível 1 (Não adequação do uso da técnica): não ajuste da categoria-objetivo de aplicação da técnica à categoria do problema de pesquisa da unidade “i” de análise, independentemente do atendimento ou não atendimento às premissas desta técnica (Não atendimento ao critério “1”).
- b) Nível 2: ajuste da categoria-objetivo de aplicação da técnica à categoria do problema de pesquisa da unidade “i” de análise e não atendimento a pelo menos uma das premissas da técnica ou à não explicitação da situação de cada premissa (atendimento ou não atendimento) no documento da unidade “i” (Atendimento ao critério “1” e não atendimento ao critério “2”).
- c) Nível 3: ajuste da categoria-objetivo de aplicação da técnica à categoria do problema de pesquisa da unidade “i” de análise e atendimento de todas as premissas da técnica, conforme documento da unidade “i”. (Atendimento ao Critério 1 e ao Critério 2).

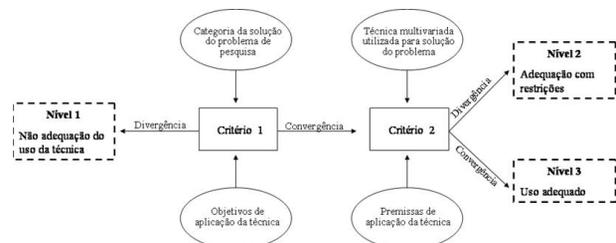


Figura 1: Avaliação do nível de adequação da técnica

Fonte: Elaborada pelos autores deste artigo

### 3.4 Instrumento de Coleta de Dados e Resultados

Para a coleta de dados foi desenvolvido um instrumento para registro das ocorrências encontradas em cada unidade “i” de análise.

Esses registros referem-se:

- a) à identificação numérica das unidades de análise;
- b) ao ano de defesa do trabalho;
- c) à descrição do problema de pesquisa;
- d) à categorização do problema de pesquisa a partir dos objetivos de aplicação das técnicas de análise multivariada;
- e) à identificação da técnica de análise multivariada utilizada no trabalho;
- f) à aplicação do Critério 1: categorização do problema de pesquisa da unidade de análise;
- g) à identificação e contagem das premissas atendidas na aplicação da análise de correspondência e/ou da análise de conglomerados, se pelo menos uma dessas foi utilizada;
- h) à aplicação do Critério 2: procedimento de avaliação do atendimento às premissas dessas técnicas, se alguma foi utilizada; e
- i) à avaliação final do nível de adequação do uso dessas técnicas.

Os resultados são identificados por meio da ocorrência evidenciada em cada critério. A partir desses resultados, a complementação do plano analítico pressupõe:

- a) evidenciar quais técnicas estatísticas de análise multivariada foram mais utilizadas no período de estudo, independentemente da categoria dos problemas de pesquisa;

- b) evidenciar as categorias dos problemas de pesquisa que estariam apresentando maior número de ocorrências de aplicação; e
- c) evidenciar quais premissas são mais negligenciadas no processo de aplicação da análise de correspondência e/ou da análise de conglomerados.

## 4 ANÁLISE DOS RESULTADOS

A primeira etapa de análise dos resultados tratou de apresentar as evidências quanto à intensidade de uso das técnicas multivariadas de modo geral, conforme o objetivo “a”.

A segunda parte, conforme o objetivo “b”, tratou do nível de adequação do uso da análise de correspondência e/ou da análise de conglomerados aos problemas de pesquisa das dissertações e teses sob a análise.

A terceira parte identificou, a partir do não atendimento, as premissas básicas de aplicação dessas técnicas, fontes potenciais de erro.

### 4.1 Resultados Relativos ao Objetivo “a”

Entre 1997 e 2006, as dissertações e as teses sobre a temática “Comportamento do Consumidor”, defendidas nos PPGA da FEA/USP e EA/UFRGS, apresentaram maior intensidade de uso de técnicas estatísticas multivariadas para atender ao objetivo de reduzir ou simplificar a estrutura de dados coletada (62,5%), conforme apresentado na Figura 2, com a aplicação da técnica de Análise Fatorial Exploratória, única representante dessa categoria. A categoria de investigação de dependência entre variáveis foi a

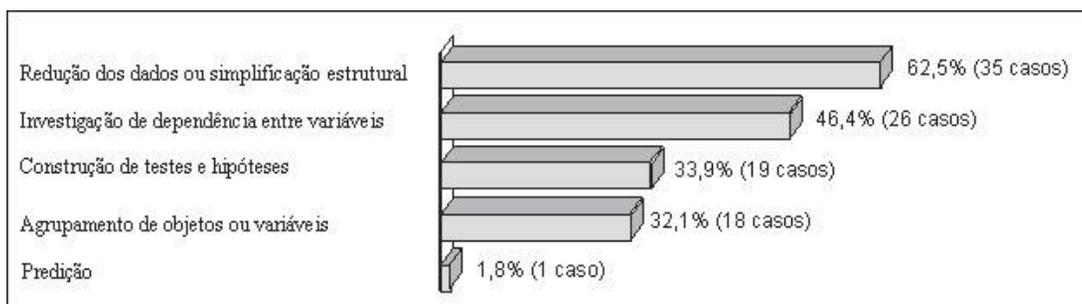


Figura 2: Categoria dos objetivos de aplicação das técnicas estatísticas multivariadas (base: universo de interesse)

Fonte: Elaborada pelos autores deste artigo

segunda mais utilizada (46,4%). Nesse contexto, seis técnicas de análise multivariada apresentam essa categoria como um dos objetivos de aplicação: Análise de Regressão (15 casos), Análise Discriminante (3 casos), Análise Multivariada da Variância (3 casos), Análise Conjunta (3 casos), Regressão Logística (1 caso) e Correlação Canônica (1 caso).

A Construção de Hipóteses e Testes, exclusivamente representada, neste estudo, pela técnica estatística multivariada de Modelagem de Equações Estruturais somou 18 casos (32,1%).

Já a categoria de agrupamento de objetos ou variáveis foi utilizada em 18 estudos (33,9%), com maior número de casos com uso da técnica de Análise de Conglomerados (12 casos), seguido da técnica de Análise de Correspondência (5 casos) e da técnica de Escalonamento Multidimensional (1 caso).

Chama a atenção que apenas um dos estudos selecionados utilizou-se de técnica de análise multivariada (regressão logística) com o objetivo de predição de algum fenômeno ou fato, o que possivelmente ocorre mais pela ausência de interesse por problemas de pesquisa na temática de Comportamento do Consumidor do que pela restrição do número de técnicas disponíveis, visto que essa categoria é representada, nesse estudo, por três técnicas de análise multivariada: Análise de Regressão, Análise Discriminante e Regressão Logística.

Ainda, a fim de ilustrar o ambiente de uso das técnicas estatísticas multivariadas, registre-se que em 64,3% dos estudos selecionados, pelo menos duas técnicas são utilizadas para atendimento à solução dos problemas de pesquisa. O uso mais intenso das técnicas ocorre, em termos relativos, nas teses, com

81,2% dos casos com aplicação de pelo menos duas técnicas, contra 57,5% nas dissertações, sugerindo maior sofisticação dos estudos no primeiro grupo.

Tabela 1: Quantidade de técnicas diferentes de análise multivariada usada nas dissertações e teses

TÉCNICA ESTATÍSTICA DE ANÁLISE MULTIVARIADA	NÍVEL DO ESTUDO		TOTAL DE CASOS
	DISSERTAÇÃO	TESE	
Uma técnica	42,5% (17 casos)	18,8% (3 casos)	35,7% (20 casos)
Dois técnicas	47,5% (19 casos)	68,7% (11 casos)	53,6% (30 casos)
Três técnicas	7,5% (3 casos)	12,5% (2 casos)	8,9% (5 casos)
Quatro técnicas	2,5% (1 caso)	Nenhum caso	1,8% (1 caso)
Média de técnicas	1,7	1,9	1,8
Mediana de técnicas	2,0	2,0	2,0
Moda de técnicas	2,0	2,0	2,0

Fonte: Elaborada pelos autores deste artigo

Quanto à categoria de aplicação das técnicas: análise de correspondência e análise de conglomerados ao longo do período 1997-2006, é preciso registrar que a categoria de agrupamento de observações apresentou tendência de crescimento no período estudado.

A Tabela 2 apresenta a evolução do uso das técnicas estatísticas a partir do objetivo de agrupamento e a distribuição específica do emprego da análise de correspondência e da análise de conglomerados. Cumpre destacar que nos anos de 1997 a 1999 não se detectou aplicação de técnicas de agrupamento nas Instituições de ensino, que foram o alvo deste estudo.

Tabela 2: Distribuição do uso das técnicas com objetivo de agrupamento ao longo do tempo (período 1997-2006)

	2000	2001	2002	2003	2004	2005	2006
Agrupamento de objetos ou variáveis	10,0%	8,0%	20,0%	0,0%	20,0%	25,0%	31,0%
Análise de correspondência	10,0%	0,0%	6,7%	0,0%	5,0%	10,0%	0,0%
Análise de conglomerados	0,0%	7,7%	13,3%	0,0%	15,0%	15,0%	23,1%
<b>Total</b>	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fonte: Elaborada pelos autores deste artigo

## 4.2 Resultados Relativos ao Objetivo “b”

Como resultado positivo deste estudo pode-se observar a coerência nas dissertações e nas teses selecionadas quanto à adequação das duas técnicas ao problema de pesquisa.

Assim, pelo Critério 1, todas as aplicações das duas técnicas são “adequadas”.

## 4.3 Resultados Relativos ao Objetivo “c”

Nos cinco casos em que se detectou a aplicação da análise de correspondência foram respeitadas duas das três premissas encontradas na literatura pesquisada. A condição não atendida em alguns trabalhos ocorreu para a premissa de tamanho suficiente da amostra para garantia de frequências esperadas maiores ou iguais a 5.

A técnica de análise de conglomerados apresentou um número expressivo de premissas atendidas (75% de atendimento a mais de 50% das premissas).

Não obstante, registre-se que as três premissas mais atendidas são as que exigem menor grau de envolvimento do pesquisador quando do uso da técnica, ou

seja, entre os 11 estudos que atenderam à premissa de padronização de variáveis, apenas um aplicou alguma técnica de padronização de dados. Quanto ao tipo de variável e tamanho da amostra, não se pode afirmar se a determinação do tamanho da amostra e a definição do tipo de medição utilizada foram realizadas levando em conta a técnica multivariada que seria empregada.

De forma oposta, as duas premissas que exigiram maior participação do pesquisador quando da aplicação da técnica (multicolinearidade e tratamento de observações atípicas – *outliers*) são as que obtiveram menor taxa de uso (25%, cada).

## 4.4 Critério “2” de Avaliação dos Níveis de Adequação do Emprego das duas Técnicas

O resultado final da avaliação das aplicações encontradas na população-alvo sugere um frágil ambiente operacional de aplicação das técnicas análise de correspondência e análise de conglomerados, no que se refere especificamente ao atendimento às suas premissas.

Tabela 3: Atendimento às premissas subjacentes ao uso das técnicas

PREMISSAS SUBJACENTES	CORRESPONDÊNCIA		CONGLOMERADOS	
	NÚMERO DE CASOS (TOTAL DE 5 CASOS)	% SOBRE O TOTAL DE CASOS QUE UTILIZARAM A TÉCNICA	NÚMERO DE CASOS (TOTAL: 12 CASOS)	% SOBRE O TOTAL DE CASOS QUE UTILIZARAM A TÉCNICA
Número reduzido de categorias por variável e categorização de variáveis contínuas	5	100,0%	-	-
Número de categorias por variável maior ou igual a 3	5	100,0%	-	-
Padronização das variáveis	-	-	11	91,7%
Tamanho da amostra	3	60,0%	8	66,7%
Tipo de variáveis	-	-	11	91,7%
Ausência de multicolinearidade	-	-	3	25,0%
Ausência de observações atípicas	-	-	3	25,0%
Atendimento a <b>nenhuma</b> premissa	0	0,0%	0	0,0%
Atendimento a <b>até 50%</b> das premissas	0	0,0%	3	25,0%
Atendimento a <b>mais de 50%</b> das premissas	5	100,0%	9	75,0%
Atendimento a <b>todas</b> as premissas	0	0,0%	1	8,3%

Fonte: Elaborada pelos autores deste artigo

O emprego do Critério 2 do processo de avaliação revela que, respectivamente, 0% e somente 8,3% das aplicações de análise de correspondência e análise de conglomerados atendem plenamente às exigências de verificação de suas premissas.

#### 4.5 Avaliação Final do Nível de Adequação do Uso das duas Técnicas

A avaliação final do nível de adequação do uso de análise de correspondência e de análise de conglomerados pode ser sintetizada pelos índices percentuais obtidos nos três níveis retratados na Figura 1:

- a) Nível 1: 100% das aplicações das duas técnicas apresentaram convergência entre a categoria-objetivo de aplicação com a categoria do problema da pesquisa;
- b) Nível 2: 100% (correspondência) e 91,7% (conglomerados) das aplicações da técnica atenderam parcialmente as suas premissas; e
- c) Nível 3: 0% (correspondência) e 8,3% (conglomerados) das aplicações da técnica apresentaram uso adequado quanto ao atendimento de suas premissas.

## 5 CONSIDERAÇÕES FINAIS

Nos trabalhos avaliados, as soluções para os problemas de pesquisa estudados concentraram-se no emprego de técnicas de redução ou simplificação estrutural dos dados (62,5% dos trabalhos examinados), seguido de técnicas de investigação de dependência entre variáveis (46,4% dos trabalhos examinados).

No conjunto dos trabalhos focalizados, as técnicas de agrupamento somente começaram a ser aplicadas em 2000, com trajetória ascendente em seu uso, com destaque para a Análise de Conglomerados, utilizada em mais de 15% das teses e dissertações avaliadas entre os anos de 2004 e 2006.

Um resultado satisfatório encontrado sobre a aplicação das técnicas desta categoria é que 100% das suas aplicações apresentaram-se adequadas à resolução dos problemas de pesquisa das dissertações e teses-alvo desse estudo.

Quanto à qualidade de aplicação dessas técnicas, destaque-se que a verificação de todas as premissas

ocorreu, respectivamente, em 0% e em 8,3% das aplicações de correspondência e da análise de conglomerados. As premissas menos verificadas foram ausência de multicolinearidade e ausência de observações atípicas em Análise de Conglomerados (25%) e tamanho da amostra em Análise de Correspondência (60,0%).

Assim, ainda que os achados desse estudo não possam ser extrapolados para a produção acadêmica de outras Instituições, outros cursos e mesmo de outros períodos de tempo, eles sugerem a necessidade de maior cuidado conceitual nas aplicações das técnicas de agrupamento.

A partir dos resultados aqui obtidos pode-se ressaltar a importância das bancas de qualificação no sentido de alertar os pesquisadores sobre a necessidade desses cuidados, incentivando a prática do atendimento às premissas subjacentes destas técnicas.

Espera-se que o resultado apurado nesse estudo, relativo ao elevado nível de não adequação da implementação das técnicas de agrupamento na população-alvo, em função da não verificação de premissas exigidas, seja um estímulo aos pesquisadores para a busca de um entendimento mais detalhado e aprofundado das técnicas a serem empregadas como parte do processo de solução do problema de pesquisa, de forma a minimizar eventuais erros decorrentes da aplicação não adequada.

## REFERÊNCIAS

- ARKADER, R. A Pesquisa Científica em Gerência de Operações No Brasil. **Revista de Administração de Empresas – RAE**, São Paulo, v. 43, n. 1, p. 70-79, 2003.
- BARDIN, L. **Análise de conteúdo**. Lisboa: Edições 70, 1979.
- BERTERO, C. O.; KEINERT, T. M. M. A evolução da análise organizacional no Brasil (1961-93). **Revista de Administração de Empresas – RAE**, São Paulo, v. 34, n. 3, p. 81-90, 1994.
- BERTERO, Carlos Osmar; CALDAS, Miguel Pinto; WOOD Jr., Thomaz. Produção científica em administração de empresas: provocações, insinuações e contribuições para um debate local. **Revista de Administração Contemporânea – RAC**, v. 3, n. 1, jan.-abr., p. 147-178, 1999.

BOTELHO, D.; MACERA. Análise meteórica de teses e dissertações da área de Marketing apresentadas na FGV-EAESP (1974-1999). In: ENCONTRO ANUAL DA ASSOCIAÇÃO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO E PESQUISA EM ADMINISTRAÇÃO, 25., Campinas. **Anais...** Campinas: ANPAD, Marketing, 2001.

BREI, V. A.; LIBERALI, G. O uso de modelagem em equações estruturais na área de marketing no Brasil. In: I ENCONTRO DE MARKETING DA ASSOCIAÇÃO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO (EMA), **Anais...** Porto Alegre, RS: 2004.

CALDAS, *et al.* Espelho, espelho meu: Meta-estudo da Produção científica em Recursos Humanos nos ENANPADs da década de 90. In: XXVI, ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO (EnANPAD) Salvador, **Anais...** Salvador: BA, 2002.

CORRAR, Luiz J. *et al.* **Análise multivariada.** São Paulo: Editora Atlas, 2007.

GAMBOA, S. S. Epistemologia da Pesquisa em educação. Tese de Doutorado. Unicamp. Campinas, 1987.

GARSON, G. D. **Correspondence analysis, from statnotes:** topics im multivariate analysis. 2007a. Disponível em: <<http://www2.chass.ncsu.edu/garson/pa765.statnote.html>>. Acesso em: 20 jul. 2007.

\_\_\_\_\_. **Cluster analysis, from statnotes:** topics im multivariate analysis 2007b. Disponível em: <<http://www2.chass.ncsu.edu/garson/pa765.statnote.html>>, Acesso em: 20 jul. 2007.

\_\_\_\_\_. **Testing of assumption, from statnotes:** topics im multivariate analysis. 2007c. Disponível em: <<http://www2.chass.ncsu.edu/garson/pa765.statnote.html>>. Acesso em: 20 jul. 2007.

\_\_\_\_\_. **Fatorial Analysis, from Statnotes:** topics im multivariate analysis. 2007d. Disponível em: <<http://www2.chass.ncsu.edu/garson/pa765.statnote.html>>. Acesso em: 20 jul. 2007,

GUJARATI, D. N. **Econometria básica.** São Paulo – SP: Makron Books, 2000.

HAIR, J. F. *et al.* **Análise multivariada de dados.** 5. ed. Porto Alegre: Editora Bookman, 2005.

HIRSCHFELD, H. O. A connection between correlation and contingency. **Proc. of the Camb. Phil. Soc.** Massachusetts, v. 31, p. 520-524, 1935.

HOPPEN, N. *et al.* Sistemas de informação no Brasil: uma análise dos artigos científicos dos anos 90. In: XXII ENCONTRO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO. EnANPAD, Foz do Iguaçu, **Anais...** Foz do Iguaçu: PR, 1998. p. 36.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis.** Fourth Edition. New Jersey: Printice Hall, 1998.

LEAL, R. P. C. *et al.* Perfil da pesquisa em finanças no Brasil. **Revista de Administração de Empresas,** São Paulo, v. 43, n. 1, p. 91-103, 2003.

LOURENÇO, A.; MATIAS, R. P. **Estatística multivariada.** Porto: Instituto Superior de Engenharia do Porto, 2001.

MACHADO-DA-SILVA, C. *et al.* Produção Acadêmica em Administração Pública: período 1983-1988. In: ENCONTRO ANUAL DA ASSOCIAÇÃO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO E PESQUISA M ADMINISTRAÇÃO, Belo Horizonte: ANPAD, **Anais...** 1989, p. 1.599-1.618, ANPAD, Belo Horizonte, 1989.

\_\_\_\_\_. Organizações: O estado da arte da produção acadêmica no Brasil. In: XIV ENCONTRO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO (EnANPAD). 1990. **Anais...** 1990. EnANPAD, 1990.

MALHOTRA, N. K. **Pesquisa de marketing.** Porto Alegre: Bookman, 2001.

MAROCO, J. **Análise Estatística:** com utilização do SPSS. 1. ed. Lisboa: Edições Silabo. 2003.

MARTINS, G. A. **Epistemologia da pesquisa em administração.** Tese de Livre-Docência. Faculdade de Administração, Economia e Contabilidade da Universidade de São Paulo. São Paulo, 1994.



- MILAGRE, R. A. **Estatística**: uma proposta de ensino para os cursos de administração de empresas. 160 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2001.
- MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada**: uma abordagem aplicada. Belo Horizonte: UFMG, 2005.
- PEREIRA, J. C. R. **Análise de dados qualitativos**. 3. ed. Edusp, São Paulo, 2001.
- PERIN, M. G. *et al.* A pesquisa Survey em artigos de Marketing nos ENANPADs da Década de 90. **Revista Interdisciplinar de Marketing**, Maringá, v. 1, n.1, jan.-abr., p. 44-59, 2002.
- PESTANA, M. H.; GAGEIRO, J. N. **Análise de dados para ciências sociais**: a complementaridade do SPSS. 2. ed. Lisboa: Editora Silabo, 2000.
- POWERS, T. *et al.* Career research productivity patterns of marketing academicians. **Journal of Business Research**, New York, v. 42, n. 1, p. 75-86, 1998.
- RICHARDSON, R. J. *et al.* **Pesquisa social**. 3. ed. São Paulo: Atlas, 1999.
- ROCHA, D.; DEUSDARÁ, B. Análise de Conteúdo e Análise do Discurso: Aproximações e afastamentos na re(construção) de uma trajetória. **ALEA**, Rio de Janeiro, v. 7, n. 2, p. 305-322, 2005.
- RODRIGUES FILHO, J. **Estudos críticos em administração no Brasil**: classificação da produção de conhecimento sob a ótica da teoria crítica de Jurgen Habermas. 262 p. Tese. Universidade Federal da Paraíba, 2002.
- ROESCH, S. M. *et al.* Tendências da Pesquisa em recursos Humanos e Organizações: uma análise das dissertações de mestrado. In: XXII ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO (EnANPAD), 1997, Rio de Janeiro. **Anais...** Rio das Pedras: Anpad, 1997.
- SHARMA, S. **Applied Multivariate Techniques**. New York: Wiley, 1996.
- STEINER, M. T. A. **Uma Metodologia para o Reconhecimento de Padrões Multivariados com Resposta Dicotômica**. 285 p. 1995. Tese (Doutorado em Engenharia), UFSC, Florianópolis. Disponível em: <[http://www.eps.ufsc.br/teses/steiner/capit\\_2/cap2\\_ste.htm](http://www.eps.ufsc.br/teses/steiner/capit_2/cap2_ste.htm)>. Acesso em: 28 out. 2007.
- SPSS –Statistical Package for the Social Sciences. **Categories 11.0**. Chicago: SPSS, 2001.
- TABACHNICK, B. G.; FIDELL, L. S. **Using multivariate statistics**. 3. ed. New York: Harper Collins, 1996.
- TONELLI, M. *et al.* Produção Acadêmica em Recursos Humanos no Brasil: 1991-2000. **Revista de Administração de Empresas – RAE**, São Paulo, v. 43, n. 1, p. 105-122, 2003.
- TORRES, R. R. **Estudo sobre os planos amostrais das dissertações e teses em Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo e da Escola de Administração da Universidade Federal do Rio Grande do Sul**: uma contribuição crítica. 246 p. Dissertação de Mestrado. FEA/USP, São Paulo, 2000.
- TRIVELLONI, Carlos A. P.; HOCHHEIM, Norberto. Avaliação de Imóveis com Técnicas de Análise Multivariada. In: CONGRESSO BRASILEIRO DE CADASTRO TÉCNICO MULTIFINALITÁRIO, 1998. **Anais...** Florianópolis, UFSC, 1998.
- TYRON, R. C. **Cluster analysis**. Ann Arbor: Edwards Brothers, 1939.
- VERGARA, S. C.; CARVALHO Jr., D. S. Nacionalidade dos Autores Referenciados na Literatura Brasileira sobre Organizações. In: ANPAD 19º, 1995, João Pessoa, 1995. **Anais...** João Pessoa, 1995.