

Avaliação do emprego das técnicas de análise de regressão e correlação canônica em teses e dissertações de algumas instituições de ensino superior

Evaluation of different techniques of regression analysis and canonical correlation in theses and dissertations of some institutions of higher education

Leandro Campi Prearo¹

Maria Aparecida Gouvêa²

Maria do Carmo Romeiro³

Resumo

Este artigo faz parte de um amplo estudo de avaliação da adequação no uso de técnicas estatísticas multivariadas em teses e dissertações de algumas instituições de ensino superior na área de *marketing* na temática do comportamento do consumidor entre 1997 e 2006. Focalizam-se as técnicas de análise de regressão e correlação canônica, duas técnicas com grande potencial de uso em estudos de *marketing*. Foi objetivo deste trabalho a análise da adequação do emprego destas técnicas às necessidades dos problemas de pesquisa apresentados nas teses e 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

¹ Doutorando e Mestre em Administração – Área de Métodos Quantitativos – FEA/USP. Docente e pesquisador da Universidade Municipal de São Caetano do Sul (USCS). São Caetano do Sul – São Paulo. E-mail: leandro.prearo@uscs.edu.br. Av. Goiás, 3400 – Bairro Barcelona – São Caetano do Sul – São Paulo.

² Livre-Docência em Administração – Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo – FEA/USP. Docente de Estatística e Metodologia de Pesquisa – Curso de Graduação e Pós-Graduação da FEA/USP no Departamento de Administração da FEA/USP. E-mail: mgouvea@usp.br. Av. Prof. Luciano Gualberto, 908 – FEA 1 – Sala E110 – Cidade Universitária – CEP 05508-900 – São Paulo- São Paulo.

³ Doutora e Mestre em Administração – FEA/USP. Docente e pesquisador da Universidade Municipal de São Caetano do Sul (USCS). São Caetano do Sul – São Paulo. e-mail: mromeiro@uscs.edu.br. Av. Goiás, 3400 – Bairro Barcelona – São Caetano do Sul – São Paulo.

pesquisadores na verificação de todos os preceitos teóricos de aplicação destas técnicas classificadas na categoria de investigação de dependência entre variáveis.

Palavras-chave: Métodos Quantitativos. Análise de Regressão. Correlação Canônica. *Marketin*. 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 regression analysis and canonical correlation 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 the category of investigation of dependence among variables.

Keywords: *Quantitative Methods. Regression Analysis. Canonical Correlation. Marketin. Consumer Behavior.*

Introdução

A ênfase na exigência de qualidade da produção científica nacional em Administração tem ocorrido mais intensamente a partir da década de 1990: Administração Pública (MACHADO DA SILVA *et al.*, 1989); Finanças (LEAL *et al.*, 2003); Marketing (POWERS *et al.*, 1998; VIEIRA, 1998; BOTELHO; MACERA; 2001); Métodos Quantitativos (BREI; LIBERALI, 2004); Operações (ARKADER, 2003), Organizações (MACHADO-DA-SILVA *et al.*, 1990; BERTERO; KEINERT, 1994; VERGARA; CARVALHO, 1995; BERTERO *et al.*, 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).

Há duas vertentes na avaliação crítica da qualidade da produção acadêmica nacional: aspectos epistemológicos e aspectos metodológicos. Essa segunda linha parece ainda contar com um número reduzido de estudos.

Em estudo sobre a Epistemologia da Pesquisa em Administração, Martins (1994, p. 65-66) evidenciou que as dissertações e teses apresentadas na FEA/USP, FEA/PUC e EASP/FGV, entre os anos de 1980 e 1993, abusavam da utilização do instrumental dos métodos quantitativos de forma bastante superficial, dado o nível 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, p.17), 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 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 este 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 a 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, 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.

Uma das áreas da Administração que mais utiliza o método quantitativo, especialmente a análise multivariada, é a área de Marketing. Principalmente pela necessidade de se conhecer o mercado consumidor busca-se a mensuração das opiniões, atitudes, preferências, perfil e outras características dos consumidores (MALHOTRA, 2001).

Os estudos em *Marketing*, especialmente na temática do Comportamento do Consumidor, vêm se utilizando sobremaneira das ferramentas de análise multivariada. Para Milagre (2001, p. 74), 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. Martins (1994, p. 66) afirma que os métodos quantitativos são, em Administração, mais aplicados pelos autores de pesquisa nas áreas de *Marketing*, Produção e Finanças.

Nesse sentido, esse estudo pretende aprofundar a discussão e oferecer subsídios à reflexão 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 regressão e a análise de correlação canônica destacam-se pela ampla possibilidade de aplicação em Marketing e por pertencerem à mesma categoria de técnicas: investigação da dependência entre variáveis (JOHNSON; WICHERN, 1998, p. 2).

A análise de regressão é uma técnica com grande aplicação na área de *Marketing*, sobretudo pelo fato de possibilitar a identificação da relação entre variáveis e o dimensionamento do potencial de mercados. A correlação canônica, por sua vez, é uma extensão da regressão múltipla e tem a proposta de facilitar o estudo das interrelações entre um conjunto

de variáveis dependentes múltiplas e um conjunto de independentes múltiplas. Assim, por exemplo, pode ser interessante no contexto estratégico relacionar e fazer previsões a respeito do desempenho de uma marca (vendas, participação de mercado, rentabilidade - conjunto de variáveis dependentes) e avaliações dos clientes sobre variáveis do composto de *marketing* (preço, produto, comunicação, distribuição - conjunto de variáveis independentes).

Tendo em vista a relevância destas técnicas no campo de marketing, decidiu-se selecioná-las para foco de análise neste trabalho.

Neste contexto, os objetivos a serem alcançados neste estudo são:

- a) Identificar a intensidade de uso das técnicas multivariadas na área de Marketing, temática do Comportamento do Consumidor, tendo como unidade de análise dissertações e teses, do período 1997-2006, 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 regressão e correlação canônica aos problemas de pesquisa apresentados nestas dissertações e teses selecionadas.
- c) Identificar fontes de erros da aplicação das técnicas de análise de regressão e correlação canônica, a partir do não atendimento às suas premissas básicas, no conjunto da produção científica examinada.

A seleção destes Programas de Pós-Graduação como público-alvo desse estudo deveu-se à necessidade de delimitação, dada a 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:

- ▶ 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.
- ▶ As informações sobre o atendimento às premissas da análise de regressão e correlação canônica 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.

Fundamentação Teórica

Nesta seção, haverá uma introdução sobre análise multivariada, seguida por uma apresentação geral de técnicas multivariadas, finalizando com considerações sobre a análise de regressão e correlação canônica.

Análise multivariada

A análise multivariada permite estudar e evidenciar as ligações, as semelhanças e 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), por um lado, as técnicas estatísticas multivariadas são mais complexas do que aquelas da estatística univariada. Por outro lado, apesar de uma razoável complexidade teórica fundamentada na matemática, as técnicas multivariadas, por permitirem o tratamento de diversas variáveis ao mesmo tempo, podem oferecer ao pesquisador um material bastante robusto para a análise dos dados da pesquisa.

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

Técnicas estatísticas de análise multivariada

A escolha dos métodos e tipos de análises empregadas nos trabalhos científicos deve ser determinada pelo problema de pesquisa. Nesse sentido, Johnson e Wichern (1998, p. 2) propõem uma classificação dos objetivos para atendimento do problema em cinco categorias.

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

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

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

Fonte: Adaptado de JOHNSON; WICHERN (1998, p. 2).

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.

Análise de regressão e análise de correlação canônica

A seguir, apresentam-se os principais conceitos e premissas das técnicas focalizadas neste estudo.

Análise de regressão

A terminologia “regressão” foi proposta pela primeira vez por Francis Galton, 1885, em estudo onde demonstrou que a altura dos filhos não tende a refletir a altura dos pais, mas tende a regredir para a média da população (MAROCO, 2003).

O modelo clássico de regressão linear baseia-se na associação entre uma variável dependente Y e uma coleção de variáveis preditoras X_1, X_2, \dots, X_r .

Sobre o objetivo de aplicação da técnica de análise de regressão, em função de inúmeros objetivos “secundários” possíveis, faz-se necessário sublinhar que o objetivo mais geral dessa técnica é a avaliação do relacionamento entre variáveis (Quadro 4).

- ▶ 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.
- ▶ 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,

Quadro 2 – Métodos de dependência

			Variável dependente ou predita				
			Uma variável		Duas ou mais variáveis		
			Métrica	Não-métrica	Métrica	Não-métrica	
Variável Independente ou preditor	Uma variável		Métrica	<ul style="list-style-type: none"> ▶ Regressão Simples 	<ul style="list-style-type: none"> ▶ Análise Discriminante ▶ Regressão Logística 	<ul style="list-style-type: none"> ▶ Correlação Canônica 	<ul style="list-style-type: none"> ▶ Correlação Canônica Não-Paramétrica
	Uma variável		Não-métrica	<ul style="list-style-type: none"> ▶ Regressão Simples com variável <i>dummy</i> 	<ul style="list-style-type: none"> ▶ Regressão Logística ▶ ANOVA⁽¹⁾ 	<ul style="list-style-type: none"> ▶ Análise da Variância Multivariada (MANOVA) 	<ul style="list-style-type: none"> ▶ Correlação Canônica Não-Paramétrica
	Duas ou mais variáveis		Métrica	<ul style="list-style-type: none"> ▶ Regressão Múltipla 	<ul style="list-style-type: none"> ▶ Análise Discriminante ▶ Regressão Logística 	<ul style="list-style-type: none"> ▶ Correlação Canônica ▶ Modelagem de Equações Estruturais (SEM) 	<ul style="list-style-type: none"> ▶ Correlação Canônica Não-Paramétrica
	Duas ou mais variáveis		Não-métrica	<ul style="list-style-type: none"> ▶ Análise Conjunta ▶ Regressão Múltipla com variável <i>dummy</i> 	<ul style="list-style-type: none"> ▶ Regressão Logística ▶ Análise Conjunta 	<ul style="list-style-type: none"> ▶ Análise da Variância Multivariada (MANOVA) ▶ Modelagem de Equações Estruturais (SEM) 	<ul style="list-style-type: none"> ▶ Correlação Canônica Não-Paramétrica

(1) A ANOVA é considerada uma técnica estatística de análise univariada, não fazendo parte, portanto, do escopo desse estudo.

Fonte: Adaptado de SHARMA (1996, p. 6)

Quadro 3 – Métodos de interdependência

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

Fonte: Adaptado de SHARMA (1996, p. 6)

Quadro 4 - Objetivos de aplicação da técnica de Análise de Regressão

Categoria	Principal objetivo indicado no estudo	Fontes
<p>► Investigação da dependência entre as variáveis</p> <p>► Predição</p>	Avaliação do relacionamento de uma variável dependente com diversas variáveis independentes	Tabachnick e Fidell (1996)
	Predição de valores de uma variável dependente a partir de uma coleção de variáveis independentes	Johnson e Wichern (1998)
	<p>Identificar se as variáveis independentes explicam uma variação significativa da variável dependente</p> <p>Predizer os valores da variável dependente</p> <p>Determinar a estrutura da relação a partir da equação matemática que relaciona as variáveis independentes e dependente</p> <p>Identificar quanto da variação na variável dependente pode ser explicado pelas variáveis independentes: intensidade da relação</p>	Malhotra (2001)
	<p>Estabelecer que um grupo de variáveis independentes explica uma proporção da variação de uma variável dependente, a um nível de significância.</p> <p>Estabelecer a importância de predição relativa das variáveis independentes (comparando os coeficientes padronizados)</p>	Garson (2007a)

Fonte: Autoria própria.

As premissas subjacentes ao uso da análise de regressão são as seguintes:

2007a). 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.

- ▶ 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.
- ▶ 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 e GAGEIRO (2000), HAIR *et al* (2005), GARSON (2007b). 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, 2007c).
- ▶ 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; na literatura pesquisada encontram-se os seguintes critérios: 104 casos + m, sendo m o número de variáveis independentes (TABACHNICK; FIDELL, 1996); pelo menos 5 casos para cada variável independente (GARSON, 2007a); pelo menos 40 casos para cada variável independente quando da utilização do método *stepwise* (GARSON, 2007a).

- ▶ Homoscedasticidade: quando a variância dos termos de erro (e) parece constante ao longo do domínio da variável preditora, tem-se homoscedasticidade (HAIR et al., 2005); a homoscedasticidade pode ser verificada por meio de gráficos de resíduos (erros) contra os valores reais e contra os valores calculados pela equação; há também alguns testes estatísticos para avaliação da homoscedasticidade, entre eles (HAIR et al., 2005; GARSON, 2007d): Teste de Goldfeld-Quandt; Teste de Park; Teste de Breusch-Pagan-Godfrey; Teste de White's; Teste de Levene para Homogeneidade das Variâncias; Teste de Bartlett's para Homogeneidade das Variâncias; Teste F-max; Teste Box's M
- ▶ Normalidade multivariada: Johnson e Wichern (1998) alertam que, para dados reais, a presença de variáveis com distribuição normal multivariada exata dificilmente ocorre. Nesse sentido, a densidade normal é freqüentemente uma aproximação útil à verdadeira distribuição da população; há poucos métodos disponíveis para testar a normalidade multivariada (SHARMA, 1996, p. 380); nos softwares estatísticos LISREL e EQS é possível obter o índice de Mardia, que é baseado nas funções de Skewness e Kurtosis; valores menores que três dessa estatística sugerem a normalidade multivariada
- ▶ Linearidade: os modelos lineares prevêem valores que se ajustam a uma linha reta, que tem uma mudança com unidade constante da variável dependente em relação a uma mudança constante na variável independente (HAIR et al., 2005); a inspeção simples de gráficos como o diagrama de dispersão (*scatterplot*), por exemplo, é um método simples e comum para a verificação de linearidade; além disso, a análise de correlação e outros testes mais complexos são disponíveis para a detecção dessa premissa; Eisenbeis (1977) e Huberty (1994) admitem o relaxamento dessa premissa quando a normalidade multivariada é atendida e quando a amostra é razoavelmente grande

- ▶ Ausência de erros correlacionados: para Gujarati (2000, p.413), a implicação do não atendimento a essa premissa ocorre principalmente nos testes de hipóteses utilizados na aplicação da técnica; o autor afirma que há o risco de se considerar um coeficiente como estatisticamente significativo, ainda que ele possa não ser significativo; alguns testes para a detecção da presença de autocorrelação: Teste de Geary (ou Teste das Carreiras), Teste de Breush-Godfrey, Teste M de Durbin e Teste de Durbin-Watson (teste mais largamente utilizado) (GUJARATI, 2000; HAIR *et al.*, 2005).

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

Quadro 5 - Premissas subjacentes ao uso da Análise de Regressão

Premissas	Fontes
Ausência de dados atípicos	Tabachnick e Fidell (1996); Hair <i>et al</i> (2005); Garson (2007a)
Ausência de multicolinearidade	Tabachnick e Fidell (1996); Gujarati (2000); Malhotra (2001); Hair <i>et al</i> (2005); Garson (2007a)
Tamanho da amostra	Tabachnick e Fidell (1996); Garson (2007a)
Homoscedasticidade	Tabachnick e Fidell (1996); Gujarati (2000); Malhotra (2001); Hair <i>et al</i> (2005); Garson (2007a)
Normalidade Multivariada	Tabachnick e Fidell (1996); Gujarati (2000); Hair <i>et al</i> (2005); Garson (2007a)
Linearidade	Tabachnick e Fidell (1996); Gujarati (2000); Malhotra (2001); Hair <i>et al</i> (2005); Garson (2007a)
Ausência de erros correlacionados	Tabachnick e Fidell (1996); Gujarati (2000); Malhotra (2001); Hair <i>et al</i> (2005); Garson (2007a)

Fonte: Autoria própria.

Análise de correlação canônica

A Análise de Correlação Canônica é a técnica mais geral de análise multivariada e, conforme Hair *et al.* (2005), pode ser entendida como uma extensão lógica da Análise de Regressão.

Para Johnson e Wichern (1998), a Correlação Canônica é um procedimento estatístico que permite o exame da estrutura de relações existentes entre dois grupos ou conjunto de variáveis.

Sobre os objetivos de aplicação, a Correlação Canônica enquadra-se na categoria investigação da dependência entre as variáveis (Quadro 6).

Quadro 6 - Objetivo de aplicação da técnica de Correlação Canônica

Categoria	Principal objetivo indicado no estudo	Fontes
Investigação da dependência entre as variáveis	Encontrar pares de combinações lineares das variáveis de cada conjunto de variáveis, de modo que a correlação entre as combinações lineares seja máxima	Johnson e Wichern (1998)
	Maximizar a relação entre as dimensões associadas às variáveis dependentes e independentes	Hair <i>et al.</i> (2005)
	Maximizar a correlação linear entre o grupo de variáveis dependentes e o grupo de variáveis independentes	Garson (2007e)

Fonte: Autoria própria.

Sobre as premissas de aplicação de técnica, assim como na maioria das outras técnicas de dependência, a Correlação Canônica é bastante rigorosa em suas exigências:

- ▶ Ausência de observações atípicas;
- ▶ Ausência de multicolinearidade;

- ▶ Sensibilidade ao tamanho da amostra: critérios - 10 vezes o número de variáveis na análise (HAIR *et al.*, 2005); 20 vezes o número de variáveis na análise (STEVENS, 1996, *apud* GARSON, (2007e); 40 a 60 vezes o número de variáveis na análise (BARCIKOWSKI and STEVENS, 1975 *apud* GARSON, 2007e);
- ▶ Homoscedasticidade.
- ▶ Normalidade multivariada;
- ▶ Linearidade: com exceção do procedimento não-linear de Correlação Canônica (GARSON, 2007 e).

O Quadro 7 apresenta as premissas da análise de correlação canônica.

Quadro 7 - Premissas subjacentes ao uso da Correlação Canônica

Premissas	Fontes
Ausência de observações atípicas	Garson (2007 e).
Ausência de multicolinearidade	Dillon e Goldstein (1984); Hair <i>et al</i> (2005); Garson (2007e).
Tamanho da amostra	Hair <i>et al</i> (2005); Stevens, 1996 <i>apud</i> Garson (2007e); Barcikowski and Stevens, 1975 <i>apud</i> Garson (2007 e).
Homoscedasticidade	Hair <i>et al.</i> (2005); Garson (2007 e).
Normalidade Multivariada	Hair <i>et al</i> (2005); Garson (2007 e).
Linearidade	Hair <i>et al</i> (2005); Stevens, 1996 <i>apud</i> Garson (2007e); Barcikowski and Stevens, 1975 <i>apud</i> Garson (2007 e).

Fonte: Autoria própria.

Metodologia de Pesquisa

Nessa seção, serão apresentados os tópicos referentes a: caracterização da população, desenho metodológico da pesquisa empírica, a construção do critério de adequação da aplicação das técnicas de análise de regressão e correlação canônica, instrumento de coleta de dados e variáveis de resultados.

Caracterização da população-alvo

A população sob 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 tese, com foco em Marketing, especificamente comportamento do consumidor, e utilização das técnicas estatísticas de análise de regressão e/ou correlação canônica como instrumento de solução do problema de pesquisa, apresentada aos PPGA¹s das instituições públicas com mais alta avaliação pela CAPES na área de Administração.

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 Tabela 1 apresenta o ambiente de identificação da população-alvo desse estudo, registrando 196 dissertações e teses sobre Comportamento do Consumidor (universo – U), 56 dissertações e teses com aplicação de técnicas estatísticas multivariadas (universo de interesse – UI) e 99 aplicações de técnicas estatísticas de análise multivariada (população-alvo – PA).

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 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 Selltitz (1974, p. 60) 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.

Tabela 1 – Dissertações e teses por ano de defesa

ANO	FEA/USP						EA/UFRGS						Total								
	Dissertações			Teses			Dissertações			Teses						Total					
	U	UI	PA	U	UI	PA	U	UI	PA	U	UI	PA	U	UI	PA						
	Total			Total			Total			Total			Total								
2006	2	0	0	6	5	10	8	5	10	8	2	3	1	0	0	9	2	3	17	7	13
2005	12	6	11	3	1	2	15	7	13	4	3	5	5	1	2	9	4	7	24	11	20
2004	10	5	7	4	2	5	14	7	12	10	4	8	1	0	0	11	4	8	25	11	20
2003	9	0	0	3	1	2	12	1	2	9	3	4	0	0	0	9	3	4	21	4	6
2002	2	1	4	4	3	5	6	4	9	17	3	6	1	0	0	18	3	6	24	7	15
2001	8	1	1	1	1	3	9	2	4	19	5	8	2	1	1	21	6	9	30	8	13
2000	12	0	0	5	1	1	17	1	1	11	5	9	0	0	0	11	5	9	28	6	10
1999	3	1	1	1	0	0	4	1	1	5	1	1	1	0	0	6	1	1	10	2	2
1998	2	0	0	2	0	0	4	0	0	4	0	0	0	0	0	4	0	0	8	0	0
1997	5	0	0	0	0	0	5	0	0	3	0	0	1	0	0	4	0	0	9	0	0
Total	65	14	24	29	14	28	94	28	52	90	26	44	12	2	3	102	28	44	196	56	99

Fonte: Autoria própria.

U – Universo.

UI – Universo de interesse.

PA – População-alvo.

Exemplifica essa convergência o alerta feito em outros estudos de que a rigidez teórica, explicitadas nas premissas para aplicações das técnicas estatísticas de análise multivariada, parece não ser acompanhada de rigidez empírica, explicitadas 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 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 na abordagem de vários autores, entre eles Richardson (1999, p.221-222) e Rocha e Deusdará (2005, p. 309), 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, p. 221) para categorização 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, p. 31) 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).

A construção do critério de avaliação da aplicação das técnicas estatísticas: análise de regressão e/ou correlação canônica

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

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 regressão e/ou da correlação canônica. 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:

- ▶ Redução dos dados ou simplificação estrutural
- ▶ Agrupamento
- ▶ Dependência entre variáveis
- ▶ Predição
- ▶ Formulação de hipóteses e testes.

No caso da análise de regressão e da correlação canônica, o objetivo comum em ambas é a investigação da dependência entre variáveis.

A categoria identificada na unidade “i” de análise foi confrontada com a categoria-objetivo da técnica de análise de regressão e/ou correlação canônica, 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.

Procedimento de avaliação do atendimento às premissas das técnicas estatísticas: análise de regressão e/ou correlação canônica: 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 destas técnicas.

Avaliação final do nível de adequação do uso da análise de regressão e/ou da correlação canônica

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

- ▶ 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”).
- ▶ 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”).
- ▶ 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).

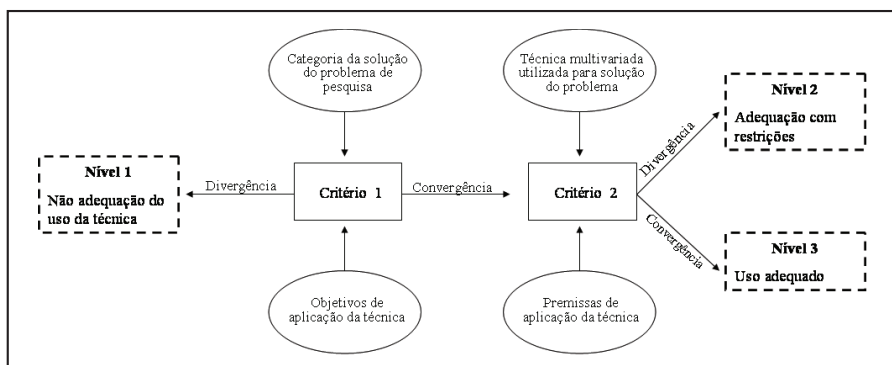


Figura 1– Avaliação do nível de adequação da técnica

Fonte: Autoria própria.

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:

- ▶ Identificação numérica das unidades de análise.
- ▶ Ano de defesa do trabalho.
- ▶ Descrição do problema de pesquisa.
- ▶ Categorização do problema de pesquisa a partir dos objetivos de aplicação das técnicas de análise multivariada.
- ▶ Identificação da técnica de análise multivariada utilizada no trabalho.
- ▶ Aplicação do Critério 1: categorização do problema de pesquisa da unidade de análise.
- ▶ Identificação e contagem das premissas atendidas na aplicação da análise de regressão e/ou da correlação canônica, se pelo menos uma destas foi utilizada.
- ▶ Aplicação do Critério 2: procedimento de avaliação do atendimento às premissas destas técnicas.
- ▶ Avaliação final do nível de adequação do uso destas técnicas.

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

- ▶ Evidenciar quais técnicas estatísticas de análise multivariada foram mais utilizadas no período de estudo, independentemente da categoria dos problemas de pesquisa.
- ▶ Evidenciar as categorias dos problemas de pesquisa que estariam apresentando maior número de ocorrências de aplicação.

- ▶ Evidenciar quais premissas são mais negligenciadas no processo de aplicação da análise de regressão e/ou da correlação canônica.

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 regressão e/ou da correlação canônica aos problemas de pesquisa das dissertações e teses sob análise.

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

Resultados relativos ao objetivo “a”

Entre 1997 e 2006, as dissertações e teses sobre a temática Comportamento do Consumidor, defendidas nos PPGA's 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 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 (Figura 2).

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 2).

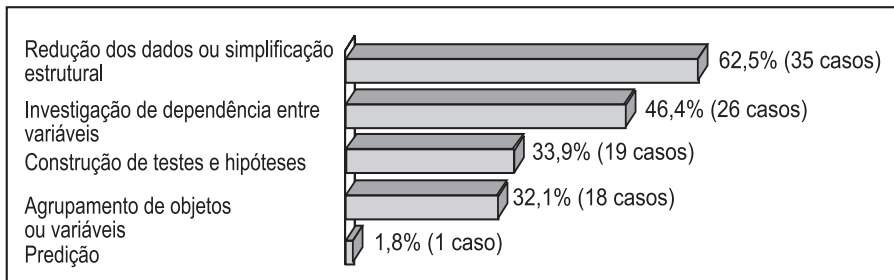


Figura 2 – Categoria dos objetivos de aplicação das técnicas estatísticas multivariadas (base: universo de interesse).

Fonte: Autoria própria.

Tabela 2 – Quantidade de técnicas de análise multivariada diferentes usadas 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)
Duas 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: Autoria própria.

Quanto às categorias de aplicação da Análise de Regressão e Correlação Canônica ao longo do período 1997-2006, deve-se registrar que a categoria de investigação da dependência entre variáveis apresentou tendência de queda no período 1999-2003, voltando a crescer a partir de 2004. Nos anos de 1997 e 1998 não se detectou aplicação de técnica estatística multivariada nas Instituições de ensino alvo desse estudo. A tabela 03 apresenta a evolução do uso das técnicas estatísticas a partir dos objetivos de investigação da dependência entre variáveis e a distribuição específica do emprego da Análise de Regressão e Correlação Canônica (Tabela 3).

Resultados relativos ao objetivo “b”

Um resultado positivo deste estudo é a coerência nas dissertações e teses selecionadas, quanto à adequação das duas técnicas ao problema de pesquisa.

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

Tabela 3 – Distribuição do uso das técnicas com objetivo de investigação da dependência entre variáveis, ao longo do tempo (período 1997-2006)

	1999	2000	2001	2002	2003	2004	2005	2006
Investigação da dependência entre variáveis	50%	30%	38%	20%	0%	25%	25%	31%
Análise de Regressão	0%	30%	39%	7%	0%	10%	15%	8%
Correlação Canônica	0%	0%	0%	7%	0%	0%	0%	0%
Total	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fonte: Autoria própria.

Resultados relativos ao objetivo “c”

Dentre as técnicas utilizadas nos estudos selecionados, a técnica de Análise de Regressão foi a que apresentou o menor número de premissas atendidas. Mais da metade das dissertações e teses (66,7%) não atendeu a nenhuma das premissas exigidas.

Com exceção do tamanho da amostra, nenhuma das premissas apresentou mais de 30% de atendimento, com destaque negativo para homoscedasticidade e linearidade, com atendimento por 13,3% dos estudos em cada uma dessas premissas.

Apenas um dos estudos que aplicaram a técnica de Análise de Regressão atendeu a todas as premissas exigidas.

Algumas particularidades na aplicação dessa técnica merecem destaque:

- ▶ Dois entre os quatro casos de atendimento à premissa de multicolinearidade utilizaram como variáveis independentes os escores fatoriais resultantes da aplicação da técnica estatística multivariada de Análise Fatorial Exploratória. Os outros dois casos fizeram uso do método *stepwise*, o qual não exige atendimento à premissa de multicolinearidade;

- ▶ Cinco entre os dez casos de não atendimento de nenhuma das premissas aplicaram a técnica de forma bastante secundária para o atendimento de seus objetivos. Mais especificamente, esses estudos fizeram uso da técnica com o único intuito de hierarquizar variáveis.

Quanto à correlação canônica, tem-se o resultado de apenas um caso de aplicação:

- ▶ Apenas duas premissas foram examinadas: multicolinearidade e sensibilidade ao tamanho da amostra;
- ▶ Sobre a multicolinearidade, registre-se que a sua verificação deu-se por meio de um processo de inércia, visto que as variáveis de entrada da técnica são variáveis resultantes da aplicação de Análise Fatorial Exploratória (escores fatoriais) (Tabela 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 regressão e correlação canônica, no que se refere especificamente ao atendimento às suas premissas.

O emprego do critério “2” do processo de avaliação revela que, respectivamente, apenas 6,7% e 0% das aplicações de análise de regressão e correlação canônica atendem plenamente às suas premissas.

Tabela 4 – Atendimento às premissas subjacentes ao uso das técnicas

Premissas subjacentes		Regressão múltipla		Correlação canônica	
		Número de casos (total de 15 casos)	% sobre o total de casos que utilizaram a técnica	Número de casos (total de 1 caso)	% sobre o total de casos que utilizaram a técnica
Ausência de dados atípicos		3	20,0%	0	0%
Multicolinearidade		4	26,7%	1	100%
Tamanho da amostra		15	100,0%	1	100%
Homoscedasticidade		2	13,3%	0	0%
Normalidade	multivariada	0	0,0%	0	0%
	univariada	3	20,0%	0	0%
Linearidade		2	13,3%	0	0%
Ausência de erros correlacionados		3	20,0%	0	0%
Atendimento a nenhuma premissa		10	66,7%	0	0%
Atendimento a até 50% das premissas		13	86,7%	1	100%
Atendimento a mais de 50% das premissas		2	13,3%	0	0%
Atendimento a todas as premissas		1	6,7%	0	0%

Fonte: Autoria própria.

Avaliação final do nível de adequação do uso das duas técnicas

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

- Nível 1: 100% das aplicações das duas técnicas apresentaram convergência entre a categoria-objetivo de aplicação com a categoria do problema da pesquisa.
- Nível 2: 93,3% (regressão) e 100% (correlação canônica) das aplicações das técnicas atenderam parcialmente as suas premissas.
- Nível 3: 6,7% (regressão) e 0% (correlação canônica) das aplicações das técnicas apresentaram uso adequado quanto ao atendimento de suas premissas.

Conclusões

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

No conjunto dos trabalhos focalizados, o uso de técnicas de investigação da dependência entre variáveis apresentou tendência de queda no período 1999-2003, voltando a crescer a partir de 2004.

Um fato bastante positivo sobre o emprego destas técnicas é que 100% das suas aplicações apresentaram-se adequadas à resolução dos problemas de pesquisa das dissertações e teses focalizadas nesse estudo.

Quanto à qualidade de aplicação destas técnicas, destaque-se que a verificação de todas as premissas ocorreu, respectivamente, em 6,7% e em 0% das Análise de Regressão e de Correlação Canônica.

Mesmo que os resultados desse estudo não possam ser extrapolados para a produção acadêmica de outras Instituições ou de outros períodos de tempo, recomenda-se maior cuidado conceitual nas aplicações das técnicas de Análise de Regressão e de Correlação Canônica.

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 inerentes a estas técnicas.

Espera-se que o resultado apurado nesse estudo, relativo ao elevado nível de não adequação do uso da análise de regressão e da correlação canônica na população-alvo, em função da não verificação de suas premissas, seja um estímulo aos pesquisadores para a busca de um entendimento mais profundo das técnicas a serem empregadas como parte do processo de solução do problema de pesquisa, de forma a minimizar a ocorrência de erros de aplicação.

Referências

- ARKADER, R. A pesquisa científica em gerência de operações no Brasil. *Revista de Administração de Empresas*, São Paulo, v. 43, n. 1, p. 70-79, 2003.
- BARDIN, L. *Análise de conteúdo*. Lisboa: Edições 70, 1979.
- BERTERO, C. O.; CALDAS, M. P.; WOOD JUNIOR, T. Produção científica em administração de empresas: provocações, insinuações e contribuições para um debate local. *Revista de Administração Contemporânea*, Curitiba, v. 3, n. 1, p. 147-178, jan./abr. 1999.
- BERTERO, C. O.; KEINERT, T. M.M. A evolução da análise organizacional no Brasil (1961-93). *Revista de Administração de Empresas*, São Paulo, v. 34, n. 3, p. 81-90, 1994.
- 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, SP. *Anais...* Campinas, SP: ANPAD, 2001. 1 CD-ROM.

BREI, V. A.; LIBERALI, G.. O uso de modelagem em equações estruturais na área de marketing no Brasil. In: ENCONTRO DE MARKETING DA ASSOCIAÇÃO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO (EMA), 2004, Porto Alegre. *Anais...* Porto Alegre: ANPAD, 2004. 1 CD-ROM.

CALDAS, M.; TONELLI, M.; LACOMBE, B. Espelho, espelho meu: meta-estudo da produção científica em recursos humanos nos EnANPADs da década de 90. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, 26., 2002, Salvador. *Anais...* Salvador, BA: ANPAD, 2002. CD-ROM.

DILLON, W. R.; GOLDSTEIN, M. *Multivariate analysis: methods and applications*. New York: Wiley, 1984.

EISENBEIS, R. Pitfalls in the application of discriminant analysis in business, finance and economics. *The Journal of Finance*, New York, v. 32, n. 3, p. 875-900, 1977.

GAMBOA, S. S. *Epistemologia da Pesquisa em educação*. 1987. 229 f. Tese (Doutorado em Educação)- Universidade de Campinas, Campinas, SP, 1987.

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

_____. *Fatorial analysis, from statnotes: topics in multivariate analysis*. 2007c. Disponível em: <<http://www2.chass.ncsu.edu/garson/pa765.statnote.html>>. Acesso em: 20 jul. 2007.

_____. *Regression analysis, from statnotes: topics in multivariate analysis*. 2007a. Disponível em: <<http://www2.chass.ncsu.edu/garson/pa765.statnote.html>>. Acesso em: 20 jul. 2007.

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

_____. *Testing of assumptions: topics in 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: Makron, 2000.

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

HOPPEN, N. et al. Sistemas de informação no Brasil: uma análise dos artigos científicos dos anos 90. In: ENCONTRO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, 22., 1998, Foz do Iguaçu, PR. *Anais ...Foz do Iguaçu, PR: ANPAD, 1998. p. 36.*

HUBERTY, C. J. *Applied discriminant analysis*. New York: Wiley, 1994.

JOHNSON, R. A.; WICHERN, D. W. *Applied multivariate statistical analysis*. 4th. ed. New Jersey: Prentice 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, 1989, Belo Horizonte. *Anais...Belo Horizonte: ANPAD, 1989. p. 1599-1618.*

_____. et al. Organizações: o estado da arte da produção acadêmica no Brasil. In: XIV ENCONTRO ANUAL DA ANPAD, 14., 1990, Florianópolis. *Anais... Florianópolis: ANPAD, 1990. v. 6, p. 11-28.*

- MALHOTRA, N. K. *Pesquisa de marketing*. Porto Alegre: Bookman, 2001.
- MAROCO, J. *Análise estatística: com utilização do SPSS*. Lisboa: Silabo, 2003.
- MARTINS, G. A. *Epistemologia da pesquisa em administração*. Tese (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*. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Florianópolis, 2001.
- PERIN, M. G. et al. A pesquisa survey em artigos de marketing nos EnANPADs da década de 90. *Revista Interdisciplinar de Marketing*, São Paulo, v. 1, n.1, p. 44-59, jan./abr. 2002.
- PESTANA, M. H.; GAGEIRO, J. N. *Análise de dados para ciências sociais: a complementaridade do SPSS*. 2. ed. Lisboa: 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.
- RODRIGUES FILHO, J. Um estudo da produção acadêmica em administração estratégica no Brasil na terminologia de Habermas. *RAE Eletrônica*, São Paulo, v. 3, n. 2, jul./dez. 2004.
- ROESCH, S. M. et al. Tendências da pesquisa em recursos humanos e organizações: uma análise das dissertações de mestrado. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, 22., 1997, Rio de Janeiro. *Anais...* Rio das Pedras: ANPAD, 1997. 1 CD-ROM.

RUDIO, F. V. *Introdução ao projeto de pesquisa científica*. 22. ed. Petrópolis: Vozes, 1998.

SHARMA, S. *Applied multivariate techniques*. New York: Wiley, 1996.

STEINER, M. T. A. *Uma metodologia para o reconhecimento de padrões multivariados com resposta dicotômica*. 1995. Tese (Doutorado em Engenharia)-Universidade Federal de Santa Catarina, Florianópolis, 1995. Disponível em: <http://www.eps.ufsc.br/teses/steiner/capit_2/cap2_ste.htm>. Acesso em: 28 out. 2007.

STEVENS, J. *Applied multivariate for the social sciences*. 3rd. ed. Mahwah, NJ: Lawrence Erlbaum, 1996.

TABACHNICK, B.G.; FIDELL, L.S. *Using multivariate statistics*. 3rd. ed. New York: Harper Collins, 1996.

TONELLI, M. et al. Produção acadêmica em recursos humanos no Brasil: 1991-2000. *Revista de Administração de Empresas*, 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 USP e da Escola de Administração da UFGRS: uma contribuição crítica*. Dissertação (Mestrado em Administração)-Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, 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, Florianópolis. *Anais...* Florianópolis: UFSC, 1998.

VERGARA, S. C.; CARVALHO JÚNIOR, D. S. Nacionalidade dos autores referenciados na literatura brasileira sobre organizações. In: ENCONTRO NACIONAL DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO DE EMPRESAS, 19., 1995, João Pessoa. *Anais...* João Pessoa: ANPAD, 1995.

VIEIRA, G. D. Por quem os sinos dobram? Uma análise da publicação científica na área de marketing do EnANPAD. In: ENCONTRO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, 22., Foz do Iguaçu, 1998. *Anais...* Foz do Iguaçu: ANPAD, 1998. p. 113.

Recebido em: 6/4/2011.

Aprovado para publicação em: 10/5/2011.