

**AVALIAÇÃO DO EMPREGO DA TÉCNICA DE ANÁLISE MULTIVARIADA DE
VARIÂNCIA EM TESES E DISSERTAÇÕES DOS PROGRAMAS DE PÓS
GRADUAÇÃO EM ADMINISTRAÇÃO DA UNIVERSIDADE DE SÃO PAULO E DA
UNIVERSIDADE FEDERAL DO GRANDE ABC**

Maria Aparecida Gouvea¹

Leandro Campi Prearo²

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. Neste artigo é focalizada a análise multivariada de variância, uma técnica com grande potencial de uso em estudos de marketing. Foi objetivo deste trabalho a análise da adequação do emprego desta técnica 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 pesquisadores na verificação de todos os preceitos teóricos de aplicação da técnica de análise multivariada de variância.

Palavras-chave: Análise Multivariada, Análise Multivariada da Variância, Análise Quantitativa, Comportamento do Consumidor

1 INTRODUÇÃO

A qualidade da produção científica nacional em Administração tem recebido atenção de muitos acadêmicos, sobretudo 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 e MACERA, 2001); Métodos Quantitativos (BREI e LIBERALI, 2004); Operações (ARKADER, 2003), Organizações (BERTERO *et al*, 1999; RODRIGUES FILHO, 2002), Pesquisa em Administração (MARTINS, 1994; TORRES, 2000; PERIN, 2002); Recursos Humanos (CALDAS *et al* 2002; TONELLI *et al*, 2003), Sistemas de Informação (HOPPEN, *et al*, 1998).

Em geral, essa análise crítica da produção acadêmica nacional se divide em duas linhas: a linha das discussões sobre aspectos epistemológicos e a linha com ênfase na adoção de critérios de qualidade e consistência da produção em Administração, ou seja, os aspectos metodológicos. Essa segunda linha parece ainda contar com um número reduzido de estudos.

Martins (1994, p.65-66), 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 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.

A evidente contribuição que a análise multivariada pode oferecer à análise quantitativa dos dados para a produção científica alia-se ao fato de que, nas últimas décadas, vários pacotes computacionais estatísticos se aperfeiçoaram sobremaneira no sentido de tornar seus conteúdos distantes das complexidades matemáticas, próprias das técnicas de Análise Multivariada.

As técnicas de análise multivariada têm sofrido sofisticações para atender a demanda dos cientistas das ciências sociais aplicadas, entre elas a Administração. Essas sofisticações das ferramentas de análise, que atraem os pesquisadores pelas facilidades oferecidas, podem levá-los a 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, notadamente 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, neste estudo pretende-se 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.

A análise multivariada de variância (*multivariate analysis of variance - MANOVA*) é uma técnica com grande potencial de aplicação na área de Marketing, sobretudo pelo fato de possibilitar a comparação de médias de variáveis ao longo de estratos amostrais de interesse na pesquisa (por exemplo, comparação das médias das vendas e das notas de satisfação com uma marca de perfume por estratos formados pela combinação de cidades e de faixas de renda familiar).

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 da técnica de análise multivariada de variância aos problemas de pesquisa apresentados nestas dissertações e teses selecionadas.

- c) Identificar fontes de erros da aplicação da técnica de análise multivariada de variância, 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, em particular 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 avaliação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – triênio 2004-2007, ou seja, aqueles programas pontuados pela CAPES com as maiores avaliações numa escala entre 1 e 7. Registre-se que a opção por esse critério não se motivou pela suposição de que os programas com avaliação menos positiva, tratariam as técnicas em estudo de forma mais ou menos corretas. A avaliação da CAPES aponta três Instituições com as maiores avaliações, 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 multivariada de variância 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 haverá uma introdução sobre análise multivariada, seguida por uma apresentação geral de técnicas multivariadas, finalizando com considerações sobre a análise multivariada de variância.

2.1 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; HOICHEIM, 2001).

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.

2.2 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.

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

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

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

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.

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.

2.3 Análise multivariada de variância (MANOVA)

Quando há o interesse em se verificar se as variáveis categóricas independentes afetam as variáveis dependentes métricas, a Análise Multivariada de Variância (MANOVA) é a técnica estatística de análise multivariada mais adequada.

Segundo Sharma (1996, p.342), a MANOVA é uma extensão da ANOVA (extensão do "test t", que permite verificar qual o efeito sobre uma variável dependente categórica - PESTANA e GAGEIRO, 2000), com a única diferença que a primeira envolve um conjunto de variáveis dependentes.

Sobre os objetivos de aplicação, a análise multivariada de variância enquadra-se na categoria de investigação da dependência entre as variáveis.

Categoria	Principal objetivo indicado no estudo	Fontes
▪ Investigação da dependência entre as variáveis	Verificar se uma ou mais variáveis independentes têm influência sobre um conjunto de variáveis dependentes	HAIR <i>et al</i> (2005)
	Examinar o efeito das variáveis independentes em um conjunto de variáveis dependentes de natureza quantitativa	PESTANA E GAGEIRO (2000)
	Examinar as diferenças entre grupos simultaneamente ao longo de múltiplas variáveis dependentes	MALHOTRA (2001)
	Verificar a principal interação das variáveis categóricas em variáveis métricas dependentes múltiplas Identificar as variáveis independentes que diferenciam um grupo de variáveis dependentes	GARSON (2007h)

Figura 2 – Objetivos de aplicação da técnica de Análise Multivariada de Variância (MANOVA)

2.3.1 Premissas da técnica estatística de análise multivariada de variância

Ainda que a técnica de análise multivariada de variância exija o atendimento a um grande número de premissas subjacentes ao uso, segundo Lewis (1995, p.19), muitos estudos empíricos têm demonstrado que os efeitos negativos do não atendimento a essas premissas parecem não ser tão comprometedores. Ainda assim, na falta de uma conclusão científica que desconstrua a teorização em cima dessa técnica, deve-se considerar o conjunto de premissas referenciado pela maior parte dos autores da área.

A seguir, são comentadas as premissas presentes na análise multivariada de variância.

2.3.1.1 Sensibilidade ao tamanho da amostra

Aliado a um correto procedimento de amostragem, o tamanho da amostra (número de casos, indivíduos, observações, entrevistas) deve ser adequado para permitir a generalização dos resultados, os quais podem ser verificados quanto à significância estatística dos testes.

As facilidades oferecidas pela maioria dos pacotes computacionais de aplicação estatística podem comprometer a observância do tamanho mínimo da amostra pelo pesquisador, fato esse que pode causar uma série de problemas importantes. Para Hair *et al* (2005), essa omissão pode resultar em baixíssimo poder estatístico dos testes de significância ou a um ajuste muito adequado dos dados, tornando os resultados artificialmente bons.

2.3.1.2 Normalidade multivariada

Nas técnicas de análise multivariada que se utilizam de variáveis métricas e testes estatísticos, a normalidade multivariada é a condição mais fundamental de aplicação.

A distribuição normal multivariada tem a forma de sinus tridimensionais simétricos quando o eixo de x apresentar os valores de uma determinada variável, o eixo y apresentar a contagem para cada valor da variável de x e, o eixo de z, apresentar os valores de qualquer outra variável em consideração.

Entretanto, 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 frequentemente uma aproximação útil à verdadeira distribuição da população.

Para Sharma (1996, p.380), há poucos métodos disponíveis para testar a normalidade multivariada.

O índice de Mardia parece ser o teste para normalidade multivariada mais disponível para os usuários de pacotes estatísticos. Baseado nas funções de *Skewness* e *Kurtosis*, o índice de Mardia só é disponível no pacote estatístico LISREL e no pacote estatístico EQS.

2.3.1.3 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, a partir da leitura de TABACHNICK e FIDELL (1996), GUJARATI (2000), PESTANA e GAGEIRO (2000), HAIR *et al* (2005), GARSON (2007a): ignorar, se o objetivo é prever; eliminar variáveis; reformular o modelo – usando, por exemplo, a razão entre variáveis; aumentar o tamanho da amostra.

2.3.1.4 Influência de dados extremos univariados e multivariados

As observações atípicas ou extremas (*outliers*) são aquelas substancialmente diferentes das outras. Podem ser identificadas no nível univariado, bivariado ou multivariado.

Segundo Hair *et al* (2005) e Garson (2007a), alguns fatores podem causar a presença de observações atípicas univariadas nos dados coletados, alterando o resultado de análise e também causando violações de normalidade: erro de entrada de dados ou na codificação; evento extraordinário: o pesquisador deve decidir se mantém ou não o dado extremo na análise; observações extraordinárias para os quais o pesquisador não tem explicação; observações que representam uma combinação única de valores das variáveis: o pesquisador deve manter o dado a menos que haja evidência específica para desconsiderá-lo.

Sobre os *outliers* multivariados, esses são casos com valores extremos no conjunto de múltiplas variáveis e são operacionalmente definidos por meio de três medidas mais comumente usadas e disponíveis na maioria dos pacotes estatísticos comerciais: Distância de Cook, Distância de Mahalanobis e Distância de Leverage.

2.3.1.5 Homoscedasticidade

Diz-se que quando a variância dos termos de erro (ε) parece constante ao longo do domínio da variável preditora, tem-se homoscedasticidade (HAIR *et al*, 2005). Esta é uma propriedade fundamental que deve ser garantida, sob pena de invalidar toda a análise estatística.

Espera-se que os erros sejam aleatórios e, se isto não ocorre, há heteroscedasticidade. Significa dizer que há chances de ocorrerem erros grandes (ou pequenos). Há tendências nos erros. Por exemplo, se na avaliação de imóveis residenciais a equação obtida indica erros maiores para os imóveis mais caros, progressivamente (quanto maior o imóvel, maior o erro), não há variância constante.

A homoscedasticidade pode inicialmente ser verificada por meio de gráficos de resíduos (erros). Os gráficos dos erros contra os valores reais e contra os valores calculados pela equação são importantes. Se os pontos estão distribuídos aleatoriamente, sem demonstrar um comportamento definido, há homoscedasticidade. Mas se existe alguma tendência, então há heteroscedasticidade. Havendo heteroscedasticidade, podem ser tentadas transformações nas variáveis (geralmente transformações logarítmicas) ou outras soluções mais complexas.

A literatura consultada apresenta uma série de testes estatísticos para avaliação da homoscedasticidade, entre eles (Hair *et al*, 2005; Garson, 2007): 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*.

2.3.1.6 Linearidade

Conforme Hair *et al* (2005), de forma geral, os modelos lineares preveem 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.

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.

Ainda sobre a linearidade das relações, Eisenbeis (1977) e Huberty (1994) admitem o relaxamento dessa premissa quando a normalidade multivariada é atendida e quando a amostra é razoavelmente grande. A figura 3 sintetiza as premissas da técnica de análise multivariada de variância.

.Premissas subjacentes	Considerações
Tipo de variáveis ^{(1) (2) (3)}	Variáveis independentes categóricas ⁽³⁾ Variáveis dependentes contínuas ou intervalares ⁽³⁾
Tamanho da amostra ^{(1) (2) (3)}	20 casos por variável ⁽²⁾
Normalidade Multivariada ^{(1) (2) (3)}	Violação tem pouco efeito nos testes estatísticos ^{(1) (2) (3)}
Multicolinearidade ^{(1) (2) (3)}	
Ausência de observações atípicas ^{(1) (2) (3)}	
Homoscedasticidade ^{(1) (2) (3)}	
Linearidade ^{(1) (2) (3)}	
Moderado grau de correlação entre as variáveis preditas ^{(1) (2) (3)}	

Figura 3 - Premissas subjacentes ao uso da Análise Multivariada de Variância (MANOVA)

Fonte: Elaboração própria a partir da abordagem de (1) Sharma (1996) (2) Hair *et al* (2005), (3) Garson (2007h)

3 METODOLOGIA DE PESQUISA

Nessa seção serão apresentados os tópicos referentes à: caracterização da população, desenho metodológico da pesquisa empírica, a construção do critério de adequação da aplicação da técnica de análise multivariada de variância, instrumento de coleta de dados e variáveis de resultados.

3.1 Caracterização da população-alvo

O universo de interesse foi definido como a dissertação ou tese, com foco em Marketing, especificamente comportamento do consumidor, e utilização de técnica estatística de análise multivariada como instrumento de solução do problema de pesquisa, apresentada aos PPGA's 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 identificação da população-alvo desse estudo registrou 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).

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 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 Seltiz (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. 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 encontrado 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, uma vez 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”.

3.3. A construção do critério de avaliação da aplicação da técnica estatística de análise multivariada de variância

Dois critérios são apresentados a seguir com a finalidade de avaliar o grau de acerto na aplicação da análise multivariada de variância.

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 multivariada de variância. 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 multivariada de variância, o objetivo alcançado em seu uso 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 multivariada de variância, 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 da técnica análise multivariada de variância quando a categoria-objetivo de aplicação da técnica ajustou-se à categoria do problema de pesquisa da unidade “i” de análise.

3.3.2 Procedimento de avaliação do atendimento às premissas da técnica estatística de análise multivariada de variância: Critério 2

O atendimento às premissas da análise multivariada de variância 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 da análise multivariada de variância e não atendimento a pelo menos uma das premissas desta técnica.

3.3.3 Avaliação final do nível de adequação do uso da análise multivariada de variância

O registro do nível de adequação do uso da análise multivariada de variância foi feito em três níveis (Figura 4):

- Nível 1 (Não adequação do uso da técnica): não ajuste da categoria-objetivo de aplicação da análise multivariada de variância 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 análise multivariada de variância à categoria do problema de pesquisa da unidade “i” de análise e não atendimento a pelo menos uma das premissas da técnica ou a 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).
-

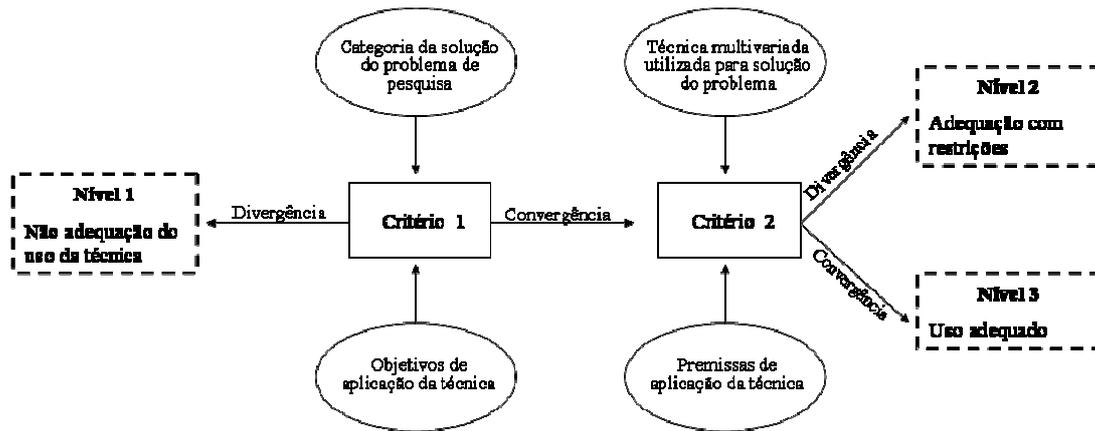


Figura 4- Avaliação do nível de adequação da técnica

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 multivariada de variância aos problemas de pesquisa das dissertações e teses sob análise.

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

4.1 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 5, 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.

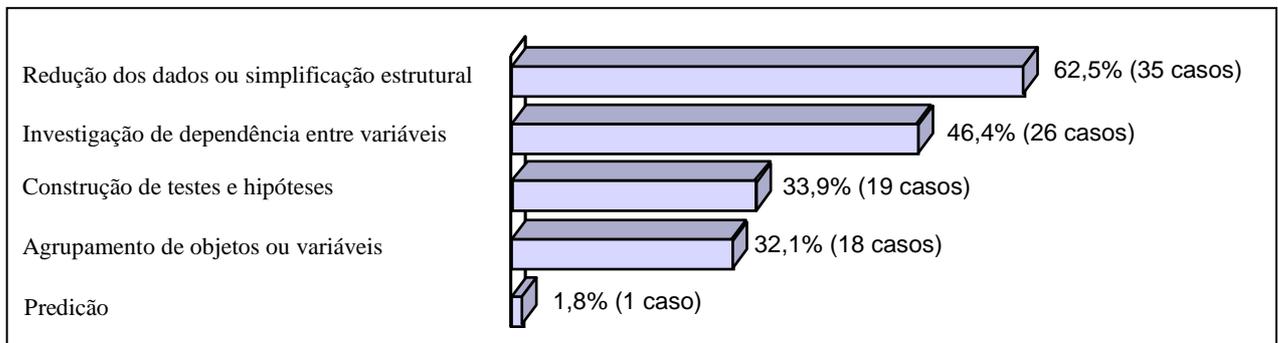


Figura 5 – Categoria dos objetivos de aplicação das técnicas estatísticas multivariadas (base: universo de interesse)

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 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

Quanto à categoria de aplicação da análise multivariada de variância 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 02 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 multivariada de variância.

Tabela 2 – Distribuição do uso das técnicas com objetivo de investigação da dependência entre variáveis e predição e uso da análise multivariada de variância, 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 multivariada de variância	0%	0%	0%	0%	0%	0%	5%	15,4%
Total	100,0 %							

4.2 Resultados relativos ao objetivo “b”

Um resultado positivo deste estudo é a coerência nas dissertações e teses selecionadas quanto à adequação da técnica escolhida de análise multivariada de variância ao problema de pesquisa.

Assim, pelo critério “1”, todas as aplicações da análise multivariada de variância são “adequadas”.

4.3 Resultados relativos ao objetivo “c”

De modo geral, a análise multivariada de variância apresentou baixos níveis de atendimento às premissas subjacentes ao seu uso nestas dissertações e teses.

Com três casos de aplicação examinados, é reforçada a ausência da prática de verificação da linearidade e multicolinearidade (sob o conceito de verificação direta) no uso das técnicas estatísticas multivariadas. Em nenhum dos casos verificou-se a premissa de linearidade e em apenas um foi verificada a premissa de multicolinearidade. O tratamento de observações atípicas (*outliers*) deixou de ser verificado em apenas uma das aplicações, embora o procedimento adotado tenha sido univariado.

Tabela 3 – Atendimento às premissas subjacentes ao uso da análise multivariada de variância

Premissas subjacentes	Número de casos (total de 3 casos)	% sobre o total de casos que utilizaram a técnica
Tamanho da amostra	3	100%
Normalidade univariada	3	100%
Normalidade multivariada	0	0,0%
Multicolinearidade	1	33,3%
Ausência de observações atípicas	2	66,7%
Homoscedasticidade	3	100%
Linearidade	0	0,0%
Correlação entre as variáveis preditas	3	100%
Atendimento a <u>nenhuma</u> premissa	0	0,0%
Atendimento a <u>até 50%</u> das premissas	0	0,0%
Atendimento a <u>mais de 50%</u> das premissas	3	100%
Atendimento a <u>todas</u> as premissas	0	0,0%

A tabela 4 ilustra alguns dimensionamentos quanto ao tamanho das amostras e à relação entre casos e variáveis no emprego da análise multivariada de variância nos estudos selecionados.

Tabela 4: Dimensionamentos no geral e segundo a análise multivariada de variância

	Número de aplicações	Tamanho da amostra		Número de variáveis		Número de casos por variável	
		Número médio de casos	Número médio de casos	Número médio de variáveis	Número mediano de variáveis	Número médio de casos	Número mediano de casos
Resultado Geral	99	482	300	23	19	30	14
Análise multivariada de variância	3	429	478	14	10	39,8	30

4.4 Critério “2” de avaliação dos níveis de adequação da análise multivariada de variância

O resultado final da avaliação das aplicações encontradas na população-alvo sugere um frágil ambiente operacional de aplicação da análise multivariada de variância, no que se refere ao seu uso e especificamente ao atendimento às premissas subjacentes à técnica.

O emprego do critério “2” do processo de avaliação revela que nenhuma das aplicações de análise multivariada de variância atendeu plenamente às suas premissas.

4.5 Avaliação final do nível de adequação do uso da análise multivariada de variância

A avaliação final do nível de adequação do uso da análise multivariada de variância pode ser sintetizada pelos índices percentuais obtidos nos três níveis retratados na figura 1:

- Nível 1: 100% das aplicações da técnica de análise multivariada de variância apresentaram convergência entre a categoria-objetivo de aplicação desta técnica com a categoria do problema da pesquisa;
- Nível 2: 100% das aplicações da análise multivariada de variância atenderam parcialmente as suas premissas;
- Nível 3: 0% das aplicações da análise multivariada de variância apresentou uso adequado quanto ao atendimento de suas premissas.

5 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, a análise multivariada de variância só passou a ser aplicada em 2005, com trajetória ascendente em seu uso.

Um fato bastante positivo sobre a aplicação da análise multivariada de variância 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 desta técnica, destaque-se que a verificação de todas as suas premissas não foi constatada em nenhuma das aplicações. Todo esse ambiente de evidências, revela, a partir da aplicação do critério “2”, que nenhuma, entre as 3 aplicações de análise multivariada de variância estudadas, foi totalmente adequada.

Não obstante, os achados desse estudo devem ser avaliados com cautela em função do pequeno número casos encontrados de aplicações da técnica de análise multivariada de variância.

No entanto, ainda que os achados desse estudo não possam ser extrapolados para a produção acadêmica de outras Instituições ou de outros períodos de tempo, eles sugerem a necessidade de maior cuidado conceitual nas aplicações da técnica de análise multivariada de variância.

A partir dos resultados aqui obtidos pode-se ressaltar a importância das bancas de qualificação para alertar os pesquisadores sobre a necessidade desses cuidados, incentivando a prática do atendimento às premissas subjacentes desta técnica.

Nesse sentido, espera-se que o resultado apurado nesse estudo, relativo ao elevado nível de não adequação do uso da análise multivariada de variância na população-alvo, em função da não verificação de suas premissas, seja um estímulo aos pesquisadores para a busca de um entendimento mais detalhado e aprofundado da técnica a ser empregada como parte do processo de solução do problema de pesquisa, de forma a minimizar eventuais erros decorrentes da aplicação não adequada.

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 multivariate analysis of variance is focused on in this paper, a technique with great potential of using in marketing studies. The objective of this study was to analyze whether the employment of this technique suits the needs of the research problem presented in these theses and dissertations as well as to evaluate the level of meeting of its assumptions. Overall, the results suggest the need for more involvement of researchers in the verification of all the theoretical precepts of application of the technique of multivariate analysis of variance.

Key words: Multivariate Analysis, Multivariate Analysis of Variance, Quantitative Analysis, Consumer Behavior

NOTAS

¹ Doutora em Administração pela FEA/USP. Docente e pesquisadora da Universidade de São Paulo.

² Doutorando em Mestre em Métodos Quantitativos pela Universidade de São Paulo. Docente e pesquisador da Universidade Municipal de São Caetano do Sul.

³ Doutora em Administração pela Universidade de São Paulo. Docente e pesquisadora da Universidade Municipal de São Caetano do Sul.

REFERÊNCIAS

ARKADER, R.. A Pesquisa Científica em Gerência de Operações No Brasil. *Revista de Administração de Empresas - RAE*, v. 43, n. 1, p. 70-79, 2003.

_____. *et al.* 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 tesese 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...* ampinas: ANPAD, 2001. Marketing.

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.

EISENBEIS, R. Pitfalls in the Application of Discriminant Analysis in Business, Finance and Economics. *The Journal of Finance*, 1977, Vol. XXXII, n. 3, 875-900

GARSON, G. D. *Testing of Assumption, from Statnotes: Topics im Multivariate Analysis*. Acessado em 20/07/2007.

_____. *MANOVA, from Statnotes: Topics im Multivariate Analysis*. Acessado em 20/07/2007.

_____. *Testing of Assumptions: Topics im Multivariate Analysis*. Acessado em 20/07/2007.

GUJARATI, D. N. – *Econometria Básica*, Makron Books, 2000,

HAIR, J.F *et al. Análise Multivariada de Dados*. 5ª Edição. Editora Bookman. Porto Alegre, 2005.

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

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

JOHNSON, R. A. e 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*, v. 43, n. 1, p. 91-103, 2003.

LEWIS, D.G. *Análise de Variância*. São Paulo: Harbra, 1995.

LOURENÇO, A.; MATIAS, R. P. *Estatística Multivariada*. 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, 1989, p. 1599-1618.

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

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*. 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*, 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*. 2ª edição. Editora Silabo. Lisboa, 2000.

POWERS, T. *et al.* Career research productivity patterns of marketing academicians. *Journal of Business Research*, New York, v.42, nº 1, p.75-86, 1998.

RICHARDSON, R. J. *et al.* *Pesquisa Social*. 3ª edição. 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*, v.7, nº 2, 2005. p.305-322.

RODRIGUES FILHO, J. *Estudos Críticos em Administração no Brasil – Classificação da Produção de Conhecimento Sob a Ótica da Teoria Crítica de Jurgen Habermas*. Tese. Universidade Federal da Paraíba, 2002.

SELLTIZ, C. *et al. Métodos de pesquisa nas relações sociais*. São Paulo: Herder, 1974.

SHARMA, S. *Applied Multivariate Techniques*, Wiley, New York, 1996.

STEINER, M. T. A. *Uma Metodologia para o Reconhecimento de Padrões Multivariados com Resposta Dicotômica*. 1995. Tese (Doutorado em Engenharia), UFSC, Florianópolis. Disponível em <http://www.eps.ufsc.br/teses/steiner/capit_2/cap2_ste.htm> Acesso em 28/10/2007.

TABACHNICK, B.G.; FIDELL, L.S. *Using Multivariate Statistics*. 3^o edição. Harper Collins. New York, 1996.

TONELLI, M. *et al. Produção Acadêmica em Recursos Humanos no Brasil: 1991-2000*. *Revista de Administração de Empresas - RAE*, 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*. Dissertação de Mestrado. FEA/USP, São Paulo, 2000.

TRIVELLONI, Carlos A. P. ; HOCHHEIM, Norbeto. Avaliação de Imóveis com Técnicas de Análise Multivariada. In: *CONGRESSO BRASILEIRO DE CADASTRO TÉCNICO MULTIFINALITÁRIO*, 1998, Florianópolis, UFSC.

VIEIRA, G. D. Por quem os sinos dobram? Uma análise da publicação científica na área de marketing do ENANPAD, In: *XXII ENCONTRO NACIONAL DOS*

PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO (EnANPAD), 1998:
Foz do Iguaçu. *Anais...Foz do Iguaçu*, 1998, p.113.