

UNIVERSIDADE FEDERAL DO PARANÁ

ANGELO DA SILVA CABRAL

**AVALIAÇÃO DE DESEMPENHO UTILIZANDO MÉTODOS ESTATÍSTICOS: UM  
ESTUDO DE CASO NA ÁREA DA SAÚDE**

CURITIBA

2017

ANGELO DA SILVA CABRAL

**AVALIAÇÃO DE DESEMPENHO UTILIZANDO MÉTODOS ESTATÍSTICOS: UM  
ESTUDO DE CASO NA ÁREA DA SAÚDE**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Produção, na Área de Concentração em Pesquisa Operacional e na Linha de Pesquisa em Métodos Estatísticos Aplicados à Engenharia, do Departamento de Engenharia de Produção, Setor de Tecnologia, Universidade Federal do Paraná, como parte das exigências para a obtenção do título de Mestre em Engenharia de Produção

Orientadora: Profa. Dra. Sonia Isoldi Marty Gama Müller

CURITIBA

2017

---

C117a

Cabral, Angelo da Silva

Avaliação de desempenho utilizando métodos estatísticos: um estudo de caso na área da saúde / Angelo da Silva Cabral. – Curitiba, 2017.

87 f. : il. color. ; 30 cm.

Dissertação - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-Graduação em Engenharia de Produção, 2017.

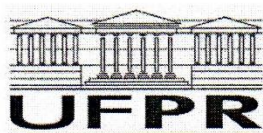
Orientador: Sonia Isoldi Marty Gama Müller.

Bibliografia: p. 72-74.

1. Engenharia de produção. 2. Avaliação de desempenho. 3. Análise fatorial. I. Universidade Federal do Paraná. II. Müller, Sonia Isoldi Marty Gama. III. Título.

CDD: 658.306

---



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
Setor TECNOLOGIA  
Programa de Pós-Graduação ENGENHARIA DE PRODUÇÃO


## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA DE PRODUÇÃO da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **ANGELO DA SILVA CABRAL** intitulada: **AVALIAÇÃO DE DESEMPENHO UTILIZANDO METODOS ESTATISTICOS: UM ESTUDO DE CASO NA AREA DE SAUDE**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO.

Curitiba, 16 de Fevereiro de 2017.

  
SONIA ISOLDI MARTY GAMA MULLER  
Presidente da Banca Examinadora (UFPR)

  
SACHIKO ARAKI LIRA  
Avaliador Externo ( )

  
ARINEI CARLOS LINDBECK DA SILVA  
Avaliador Interno (UFPR)

  
IZABEL CRISTINA ZATTAR  
Avaliador Interno (UFPR)

## **AGRADECIMENTOS**

Ao meu pai, pelo incentivo e apoio incondicional.

À Profa. Dra Sonia Isoldi Marty Gama Müller pela orientação, apoio e confiança.

Ao corpo docente do Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Paraná, que direta ou indiretamente colaboraram com este estudo.

À equipe da empresa 2iM, em especial ao Dr. César Abicalaffe e a Jovaldo Savian por fornecerem os dados que possibilitaram a realização das análises.

## RESUMO

A avaliação de desempenho dos prestadores de serviços de saúde tornou-se uma temática constante e relevante no fomento de discussões sobre a melhoria da qualidade dos sistemas de saúde, com um crescente número de publicações sobre o tema nos últimos anos. O objetivo desse estudo é aplicar métodos estatísticos para avaliar a qualidade do serviço prestado por um hospital, por meio da avaliação de desempenho dos médicos dessa instituição, e comparar os resultados obtidos com os índices propostos por modelo desenvolvido por uma empresa que presta serviços de gestão em saúde. Primeiramente, estimativas dos escores de desempenho dos médicos foram gerados de forma univariada para cada indicador de desempenho original fornecido pela empresa, baseados em Modelos Lineares Mistos. Os escores estimados dos indicadores foram consolidados em uma base de dados multivariada em que cada médico possui um escore por indicador. Foi utilizado o modelo de Análise Fatorial para analisar a intercorrelação dos escores e identificar quais os indicadores mais relevantes e escores fatoriais para gerar a classificação final dos avaliados. Os dois modelos foram comparados por meio da análise da correlação entre as classificações fornecidas pelos escores fatoriais e os índices de *performance* fornecidos pela empresa. Como resultado, verificou-se que a Análise Fatorial é uma metodologia eficiente na avaliação de desempenho, permitindo identificar as especialidades médicas que se destacaram das demais, tanto positiva como negativamente.

Palavras-chave: Avaliação de desempenho, Pagamento por performance, Saúde, Análise fatorial, Escores fatoriais.

## **ABSTRACT**

The performance evaluation of health service providers has become a constant and relevant thematic in fomenting discussions on improving the quality of health systems, with an increasing number of publications on the subject in the last years. The objective of this study is to apply statistical methods to assess the quality of service provided by a hospital, through the medical performance evaluation of this institution, and compare the results obtained with the indexes proposed by a model developed by a company that provides health management. First, physicians' scores estimates were generated univariate for each original performance indicator provided by the company, based on Mixed Linear Models. The scores estimates of the indicators were consolidated in a multivariate database in which each physician has a score by indicator. The Factor Analysis model was used to analyze the intercorrelation of the scores and to identify the most relevant indicators and factorial scores to generate the final classification of the evaluated ones. The two models were compared by analyzing the correlation between classifications provided by the factor scores and the performance indexes provided by the company. As a result, it was verified that Factor Analysis is an efficient methodology in the performance evaluation, allowing to identify the medical specialties that stood out from the others, both positively and negatively.

Key-words: Performance evaluation, Pay for performance, Health, Factor analysis, Factor scores.

## LISTA DE FIGURAS

FIGURA 1 – EXEMPLO ILUSTRATIVO DO MODELO GPS.2IM©.....	21
FIGURA 2 – DIAGRAMA DE AGRUPAMENTO COM A VARIAÇÃO ENTRE E DENTRO DOS GRUPOS.....	23
FIGURA 3 – COMPARAÇÃO DOS MÉTODOS DE DISTÂNCIA PARA A LIGAÇÃO SIMPLES E LIGAÇÃO COMPLETA.....	25
FIGURA 4 – REPRESENTAÇÃO GRÁFICA DO MÉTODO DE FISHER PARA DUAS POPULAÇÕES.....	32
FIGURA 5 – ROTAÇÃO FATORIAL ORTOGONAL.....	44
FIGURA 6 – ROTAÇÃO FATORIAL OBLÍQUA.....	44
FIGURA 7 – PROCEDIMENTOS METODOLÓGICOS UTILIZADOS NO ESTUDO.....	49
FIGURA 8 – MATRIZ DE CORRELAÇÃO DOS INDICADORES PARA O PRIMEIRO GRUPO DE ESPECIALIDADES.....	62
FIGURA 9 – MATRIZ DE CORRELAÇÃO DOS INDICADORES PARA O SEGUNDO GRUPO DE ESPECIALIDADES.....	65

## LISTA DE GRÁFICOS

GRÁFICO 1 – BANDAS IDEAIS DO MODELO GPS.2IM©. ....	19
GRÁFICO 2 – RESULTADOS MENSIS DOS INDICADORES DE DESEMPENHO DO HOSPITAL EM 2015. ....	53
GRÁFICO 3 – DENDROGRAMA PARA OS MÉDICOS DA ESPECIALIDADE 905 PELO MÉTODO DE WARD. ....	56
GRÁFICO 4 – DISTRIBUIÇÃO DOS ESCORES FATORIAIS. ....	66
GRÁFICO 5 – DISTRIBUIÇÃO DOS ESCORES FATORIAIS POR ESPECIALIDADE. ....	67
GRÁFICO 6 – DIAGRAMA DE DISPERSÃO ENTRE OS ESCORES FATORIAIS E O ÍNDICE DE PERFORMANCE DO MODELO GPS.2IM©. ....	68
GRÁFICO 7 – GRÁFICO DE RESÍDUOS VERSUS VALORES PREDITOS DO MODELO LINEAR MISTO. ....	82
GRÁFICO 8 – GRÁFICO DE RESÍDUOS VERSUS VALORES PREDITOS DO MODELO LINEAR MISTO – CONTINUAÇÃO. ....	83
GRÁFICO 9 – GRÁFICO DE RESÍDUOS VERSUS PERÍODO DO MODELO LINEAR MISTO. ....	84
GRÁFICO 10 – GRÁFICO DE RESÍDUOS VERSUS PERÍODO DO MODELO LINEAR MISTO – CONTINUAÇÃO. ....	85

## LISTA DE QUADROS

QUADRO 1 – INDICADORES AVALIADOS PELO MODELO GPS.2IM© NO HOSPITAL.....	22
QUADRO 2 – MATRIZ DAS SOMAS DOS QUADRADOS E PRODUTOS CRUZADOS DA MANOVA. ....	28
QUADRO 3 – DISTRIBUIÇÃO AMOSTRAL DO LAMBDA DE WILKS. ....	28
QUADRO 4 – MATRIZ DE CONFUSÃO PARA A VALIDAÇÃO DOS RESULTADOS DA ANÁLISE DISCRIMINANTE.....	36
QUADRO 5 – INDICADORES AVALIADOS PELO MODELO GPS.2IM© NO HOSPITAL POR ESPECIALIDADE. ....	60

## LISTA DE TABELAS

TABELA 1 – MÉDICOS POR ESPECIALIDADE PRESENTES NA BASE DE DADOS. .....	50
TABELA 2 – ESTATÍSTICAS DESCRITIVAS DOS INDICADORES DE DESEMPENHO DO HOSPITAL EM 2015.....	52
TABELA 3 – COMPARAÇÃO DAS MÉDIAS MENCIAIS DOS INDICADORES DE DESEMPENHO DO HOSPITAL EM 2015.....	54
TABELA 4 – COMPARAÇÃO DAS MÉDIAS DOS INDICADORES DE DESEMPENHO DO HOSPITAL POR ESPECIALIDADE EM 2015. ....	55
TABELA 5 – MANOVA PARA OS GRUPOS GERADOS PELO MÉTODO DE WARD PARA A ESPECIALIDADE 905.....	57
TABELA 6 – ANÁLISE DISCRIMINANTE – MATRIZ DE CONFUSÃO COM VALIDAÇÃO CRUZADA PARA OS MÉDICOS DA ESPECIALIDADE 905. ....	57
TABELA 7 – EFEITOS PADRONIZADOS DOS INDICADORES DE DESEMPENHO DO HOSPITAL.....	58
TABELA 8 – INDICADORES AVALIADOS PELO MODELO GPS.2IM© NO HOSPITAL POR GRUPO DE ESPECIALIDADES.....	61
TABELA 9 – MATRIZ DE CARREGAMENTOS PARA O GRUPO 1 – ROTAÇÃO VARIMAX.....	63
TABELA 10 – MATRIZ DE CARREGAMENTOS PARA O GRUPO 2 – ROTAÇÃO VARIMAX.....	64
TABELA 11 – TABELA P PARA O COEFICIENTE DE CORRELAÇÃO DE SPEARMAN.....	87

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>12</b>
1.1 OBJETIVOS.....	13
1.1.1 Objetivo geral.....	13
1.1.2 Objetivos específicos .....	13
1.2 JUSTIFICATIVA.....	13
1.3 ESTRUTURA DO TRABALHO.....	14
<b>2 REVISÃO DE LITERATURA</b> .....	<b>15</b>
2.1 QUALIDADE EM SAÚDE .....	15
2.2 DIMENSÕES DA QUALIDADE .....	15
2.3 INDICADORES DE DESEMPENHO .....	17
2.4 AJUSTE DE RISCO .....	18
2.5 MODELO GPS.2IM© .....	18
2.6 MÉTODOS ESTATÍSTICOS .....	23
2.6.1 Análise de agrupamentos.....	23
2.6.2 Análise multivariada de variância (MANOVA) .....	27
2.6.3 Análise Discriminante.....	29
2.6.4 Modelo linear misto .....	36
2.6.5 Análise Fatorial .....	40
2.6.6 Coeficiente de correlação de Spearman.....	47
<b>3 METODOLOGIA</b> .....	<b>48</b>
3.1 ESTRATÉGIA DA PESQUISA .....	48
3.2 MATERIAIS E MÉTODOS .....	48
3.3 PREPARAÇÃO DA BASE DE DADOS .....	51
3.3.1 Seleção dos indicadores .....	51
3.3.2 Padronização da base de dados .....	53
<b>4 RESULTADOS</b> .....	<b>60</b>
4.1 ANÁLISE FATORIAL .....	60
4.1.1 Análise fatorial: grupo 1.....	61
4.1.2 Análise fatorial: grupo 2.....	64
4.2 CLASSIFICAÇÃO DOS AVALIADOS.....	66
4.3 COMPARAÇÃO DOS MODELOS.....	67

<b>5 CONSIDERAÇÕES FINAIS.....</b>	<b>70</b>
<b>REFERÊNCIAS .....</b>	<b>72</b>
<b>APÊNDICE 1 – COMANDOS DO R UTILIZADOS PARA A ANÁLISE DE AGRUPAMENTOS .....</b>	<b>76</b>
<b>APÊNDICE 2 – COMANDOS DO R UTILIZADOS PARA A ANÁLISE DISCRIMINANTE .....</b>	<b>77</b>
<b>APÊNDICE 3 – COMANDOS DO R UTILIZADOS PARA O MODELO LINEAR MISTO .....</b>	<b>79</b>
<b>APÊNDICE 4 – COMANDOS DO R UTILIZADOS PARA A ANÁLISE FATORIAL.</b>	<b>80</b>
<b>APÊNDICE 5 – RESULTADOS SUPLEMENTARES.....</b>	<b>82</b>
<b>ANEXO 1 – TABELA P DE VALORES CRÍTICOS DE <math>r_s</math> DO COEFICIENTE DE CORRELAÇÃO DE SPEARMAN .....</b>	<b>87</b>

## 1 INTRODUÇÃO

Uma das principais formas de auxiliar profissionais a crescer é por meio da avaliação de desempenho, em que é possível identificar competências e comportamentos que merecem ser aprimorados, além de qualidades com potencial de serem reforçadas (RANGEL, 2015). A avaliação de desempenho é uma ferramenta que facilita a análise dos pontos fortes e fracos de uma pessoa, permitindo que se verifique, sobretudo, qual é a sua contribuição atual e potencial para a empresa (ABRAPP, 2013).

A avaliação de desempenho pode ser definida como uma apreciação sistemática do desempenho de cada pessoa, em função das atividades que ela desempenha, das metas e objetivos a serem alcançados e do seu potencial de desenvolvimento (CHIAVENATO, 2010). Por meio dessa ferramenta, avalia-se o desempenho para saber a forma pela qual as atividades estão sendo realizadas, corrigir desvios e melhorar produtividade (ALEXANDRE, 2011).

Qualquer que seja a técnica de avaliação escolhida pelo empregador, seu propósito deve ser a fixação de objetivos claros que possam ser mensurados de forma simples e assertiva (RANGEL, 2015). Desempenho e objetivos devem ser tópicos inseparáveis da avaliação de desempenho. A avaliação do desempenho deve ser usada para melhorar a produtividade do funcionário dentro da organização, tornando-o mais bem equipado para produzir com eficácia e eficiência (CHIAVENATO, 2010). O reconhecimento por parte do funcionário é baseado em expectativas de amadurecimento profissional, desenvolvimento de carreira e conquista de recompensas. A avaliação de desempenho funciona como estimuladora dessas expectativas, na medida em que cria oportunidade para o funcionário se mostrar no desenvolvimento de sua competência e habilidades (SEBRAE, 2015).

Para uma avaliação justa, o funcionário precisa compreender o método e cobiçar por sua aplicação. Para isto, a empresa deve cuidar para que a ferramenta seja utilizada de forma plena. O primeiro passo é tornar o funcionário conhecedor da sistemática de avaliação existente, seus objetivos, suas regras, periodicidade, aplicações e consequências. O sistema de avaliação deve ser elaborado buscando-se trabalhar com requisitos cujos indicadores ofereçam a maior objetividade possível (SEBRAE, 2015). Para garantir melhores resultados, as avaliações devem ser

periódicas e com padrões definidos para todos os cargos existentes na empresa. Deve-se monitorar continuamente a evolução dos avaliados (ALEXANDRE, 2011).

Há uma grande variedade de métodos para avaliar o desempenho laboral. Avaliar o desempenho de um grande número de pessoas dentro das organizações utilizando-se dos critérios de equidade e de justiça e, ao mesmo tempo, estimulando as pessoas não é tarefa fácil. Assim, muitas organizações constroem seus próprios sistemas de avaliação ajustados às suas características (CHIAVENATO, 2010).

Desse modo, apresenta-se o problema da pesquisa a ser investigado: como avaliar o desempenho dos médicos de um hospital? Esse estudo propõe a utilização de métodos estatísticos em todo o processo de avaliação.

## 1.1 OBJETIVOS

### 1.1.1 Objetivo geral

Avaliar o desempenho dos médicos de um hospital por meio de métodos estatísticos e comparar com o modelo utilizado atualmente na empresa.

### 1.1.2 Objetivos específicos

- Obter estimativas dos escores de desempenho baseados no ajuste de Modelos Lineares Mistos, para cada indicador de desempenho original.
- Analisar a intercorrelação das estimativas dos escores de desempenho por meio de Análise Fatorial.
- Classificar os médicos avaliados de acordo com seu desempenho individual utilizando escores fatoriais.
- Comparar os resultados do modelo proposto com o método utilizado atualmente na empresa por meio da análise de correlação.

## 1.2 JUSTIFICATIVA

Diversos sistemas de avaliação de desempenho são utilizados atualmente pelas empresas. De acordo com Rangel (2015), os três sistemas mais utilizados são:

avaliação pelo superior hierárquico, auto avaliação e avaliação a 360 graus. Na avaliação pelo superior, o funcionário pode ser prejudicado caso a relação da dupla não seja boa ou o avaliador tenha uma percepção equivocada do desempenho de seu subordinado. Na auto avaliação, os resultados podem ser condicionados pelos pontos de vista e interesses individuais do avaliado. Na avaliação a 360º, em que todos na empresa são simultaneamente avaliadores e avaliados, o processo requer um nível de maturidade profissional mais elevado, sobretudo da chefia, que deverá estar aberta a críticas.

As desvantagens citadas nos três sistemas têm em comum a figura do avaliador que, por mais capacitado que seja, irá realizar sua avaliação de forma subjetiva, podendo incorrer em erro. A Estatística dispõe de métodos capazes de classificar o desempenho de indivíduos sem a interferência de um avaliador, apenas utilizando os resultados dos indicadores de desempenho das empresas. Por meio dos métodos estatísticos a avaliação passa a ter validade científica.

Este trabalho pretende fomentar pesquisas voltadas à avaliação de desempenho em Programas de Pós-Graduação em Engenharia de Produção, com utilização de métodos estatísticos.

### 1.3 ESTRUTURA DO TRABALHO

No Capítulo 2, é apresentada a revisão de literatura e a descrição de técnicas de Análise Multivariada, tais como: Análise Fatorial, Análise Discriminante, Análise de Agrupamentos e Análise Multivariada de Variância (MANOVA). Apresenta-se também o Coeficiente de Correlação de Pearson e os modelos lineares mistos.

No Capítulo 3 são descritos os materiais e métodos utilizados no desenvolvimento do estudo.

No Capítulo 4 são apresentados os resultados da avaliação de desempenho e a comparação com os resultados do modelo atualmente utilizado na empresa.

A conclusão do trabalho é apresentada no Capítulo 5.

## 2 REVISÃO DE LITERATURA

Neste capítulo, apresenta-se um resumo sobre temas referentes à avaliação de desempenho médico por meio da revisão de trabalhos correlatos e a descrição do modelo Gestão da *Performance* em Saúde (GPS), modelo utilizado pela empresa 2iM, que serve como base de comparação neste estudo. Também são descritos os métodos estatísticos utilizados nas análises.

### 2.1 QUALIDADE EM SAÚDE

Qualidade em saúde pode ser definida como a obtenção dos maiores benefícios com os menores riscos (e custos) para os pacientes (DONABEDIAN, 1990). Problemas de qualidade decorrem do erro humano, o que não significa a pressuposição de culpa. Os erros ocorrem em virtude de sistemas falhos e mal desenhados, estando o problema no sistema, e não nos profissionais (KENNEY, 2008).

Dentro desse contexto, a avaliação de desempenho em saúde tornou-se uma temática constante e relevante no fomento de discussões sobre a melhoria da qualidade do cuidado (VAN HERCK et al., 2010). De forma geral, avaliar o desempenho implica analisar sistematicamente a atuação do profissional em função das atividades que realiza, das metas estabelecidas e dos resultados alcançados (CROMWELL et al., 2011).

### 2.2 DIMENSÕES DA QUALIDADE

A primeira etapa da avaliação do desempenho médico é a definição dos domínios ou dimensões de desempenho em que os profissionais serão avaliados. Essa definição orienta a escolha de indicadores relevantes ao modelo de avaliação (CROMWELL et al., 2011).

Donabedian (1990) propôs que a qualidade em saúde seja dividida em sete dimensões:

- **eficácia:** é o melhor que pode ser feito sob as condições de trabalho mais favoráveis, considerando as condições de saúde do paciente e circunstâncias inalteráveis;
- **efetividade:** em contraste com a eficácia, é a melhoria na saúde que se espera atingir, dada as circunstâncias ordinárias do dia a dia;
- **eficiência:** é a habilidade de obter a melhoria em saúde com o mais baixo custo;
- **otimalidade:** balanceamento entre custo e benefícios;
- **aceitabilidade:** adaptação do cuidado com o paciente aos desejos, expectativas e valores dele e seus familiares;
- **legitimidade:** preocupação com o bem-estar do indivíduo e da coletividade;
- **equidade:** princípio que determina o que é razoável na distribuição do cuidado e seus benefícios entre os membros de uma população (DONABEDIAN, 1990).

Em 2001, o Institute of Medicine publicou um estudo sobre qualidade em saúde em que os autores reorganizaram alguns desses conceitos (INSTITUTE OF MEDICINE, 2001). As dimensões de qualidade propostas são as seguintes:

- **segurança:** evitar lesões aos pacientes a partir do cuidado destinado a eles;
- **eficiência:** prover serviços para todos que poderiam se beneficiar e abster-se de prover serviços para aqueles que não estão propensos aos benefícios, evitando desperdícios;
- **centrado no paciente:** garante que os valores do paciente orientem todas as decisões clínicas;
- **oportuno:** redução de esperas, tanto para os que recebem cuidados como para os que os disponibilizam;
- **equitativo:** disponibilizar serviços que não tenham variação em qualidade por causa de características pessoais do paciente.

Outro estudo produzido pelo mesmo instituto, em 2007, simplifica os grandes objetivos de melhoria na saúde ao afirmar que focando em três dimensões, as demais serão conquistadas. Essas dimensões são: eficiência, efetividade e centralidade no paciente (INSTITUTE OF MEDICINE, 2007).

A Secretaria de Saúde do Estado de São Paulo instituiu, por meio da Lei Complementar 1193 de 2013, o programa denominado Prêmio de Produtividade Médica (PPM) que avalia o desempenho dos médicos do SUS no Estado. Os médicos são avaliados nas seguintes dimensões: qualidade dos trabalhos prestados, grau de resolutividade, responsabilidade e eficiência na execução de atividades, assiduidade e produtividade (SES/SP, 2017).

O Quality and Outcomes Framework (QOF), programa de recompensa anual e voluntária para os médicos generalistas (GP) da Inglaterra, utiliza três domínios em sua avaliação: Clínica; Saúde Pública e Saúde Pública – Serviços Adicionais (HEALTH & SOCIAL CARE INFORMATION CENTRE, 2017).

Em um estudo publicado em 2009, realizado a partir de dados coletados de 21 grandes hospitais públicos da China, foram propostas apenas duas dimensões da qualidade: despesa médica e média de permanência dos pacientes (JIAN et al., 2009).

### 2.3 INDICADORES DE DESEMPENHO

A etapa mais importante da implementação da avaliação de desempenho é a escolha dos indicadores do modelo que devem ser escolhidos seguindo as dimensões da qualidade previamente determinadas. Em termos gerais, um bom indicador de desempenho deve ter as seguintes características: aceitabilidade (ser aceitável por quem está sendo avaliado), factibilidade (os dados devem estar disponíveis, serem válidos, confiáveis e consistentes) e confiabilidade (que seja representativo) (ABICALAFFE, 2015).

A edição de 2014/15 do QOF contou com 81 indicadores, em que os médicos podiam obter até 559 pontos. O domínio Clínica foi composto por 69 indicadores, em que os médicos podem atingir até 435 pontos. O domínio Saúde Pública foi composto por sete indicadores, em que se pode atingir até 97 pontos. O domínio Saúde Pública – Serviços Adicionais foi composto por cinco indicadores, que somados atingiam até 27 pontos (HEALTH & SOCIAL CARE INFORMATION CENTRE, 2016).

O modelo testado com os hospitais chineses é mais simples, possuindo apenas dois indicadores, um para cada dimensão: Índice de Eficiência de Custo e Índice de Eficiência de Tempo (JIAN et al., 2009).

O problema da criação de um índice de desempenho composto, combinando diversos indicadores de qualidade, é analisado em um artigo publicado em 2012, no qual os autores afirmam que a grande controvérsia está no estabelecimento de pesos adequados para ponderar indicadores distribuídos em múltiplas dimensões. A proposta para a solução do problema foi a obtenção de escores a partir de uma generalização do método de Análise de Componentes Principais, que considera todos os pesos do modelo como não negativos (LIU; JAIN; SHI, 2012).

## 2.4 AJUSTE DE RISCO

Ajuste de Risco na medicina é um processo estatístico usado para identificar e ajustar variações nos tipos de desfechos que decorrem de diferenças nas características do paciente (fatores de risco). Dependendo da presença de fatores de risco durante os atendimentos aos pacientes, estes podem apresentar diferentes resultados, independentemente da qualidade do cuidado oferecido pelos médicos (ABICALAFFE, 2015).

No modelo que estudou os hospitais públicos chineses foi criado um escore de risco de morte em que se verificou que, nos casos de alto risco, a mortalidade estava mais relacionada com a doença do que a evolução clínica. Para casos de baixo risco, o inverso se aplica (JIAN et al., 2009).

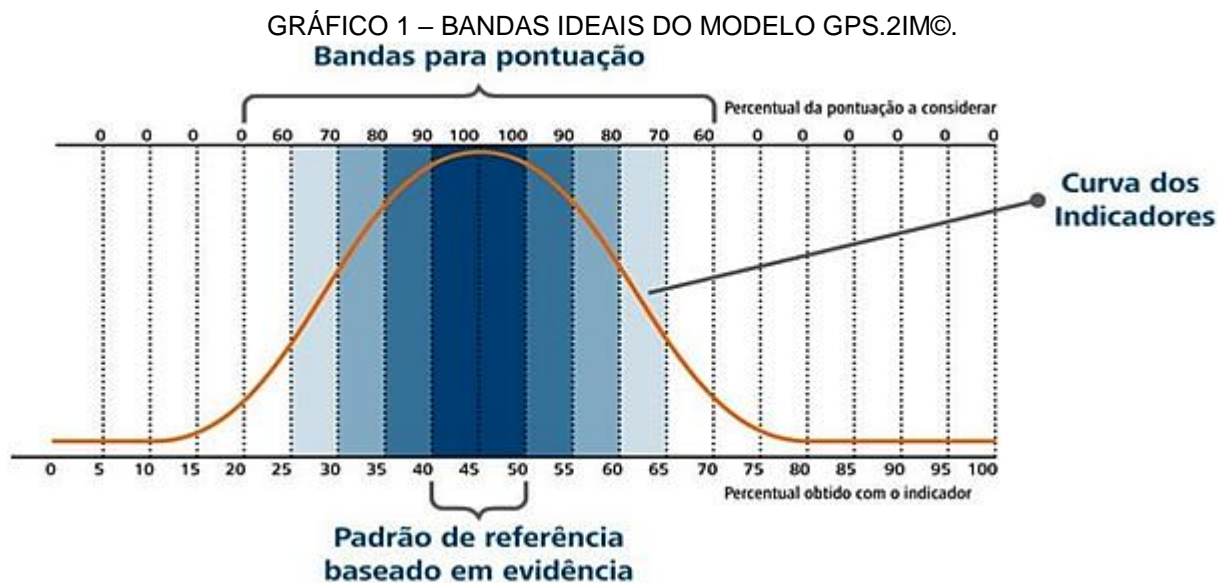
Em um estudo de 2012 foram considerados para análise todos os pacientes com quatro das condições pediátricas mais comuns (gastroenterite aguda, febre, ferimento na cabeça e doença respiratória). Foi utilizado um sistema de triagem com o objetivo de solucionar o problema do ajuste de risco, selecionando para a análise apenas os pacientes com gravidade mediana. Assim, foram incluídas no estudo cerca de um terço das visitas anuais dos dois hospitais analisados (LIU; JAIN; SHI, 2012).

## 2.5 MODELO GPS.2iM©

O modelo Gestão de *Performance* em Saúde da empresa 2iM (GPS.2iM©) é organizado em seis dimensões, sendo quatro domínios – estrutura, eficiência, efetividade e satisfação ou experiência do cliente; e duas perspectivas transversais – segurança e acesso. No domínio Estrutura estão contemplados indicadores

relacionados aos recursos humanos e ao uso de informação. O domínio Eficiência possui indicadores essencialmente relacionados a processos, custos e utilização de recursos do sistema. No domínio Efetividade estão relacionados indicadores de resultados ou desfechos intermediários e finais decorrentes do cuidado à saúde. O domínio Satisfação ou Experiência do paciente se refere à percepção do paciente quanto ao atendimento disponibilizado (ABICALAFFE, 2015).

No modelo GPS.2iM©, os indicadores são categorizados em faixas de pontuação, denominadas bandas ideais, em que se atribui ao médico 100% dos pontos do indicador quando seus valores estiverem dentro da banda de referência e, à medida que se afastam, o percentual é reduzido gradativamente, de acordo com uma curva normal de probabilidade (Gráfico 1).



FONTE: Adaptado de Abicalaffe (2015).

As bandas ideais são definidas por indicador e podem ser encontradas tanto com base em evidências publicadas como em metas contratuais. Uma característica das bandas ideais é a possibilidade de serem ajustadas para cada perfil dentro de uma especialidade. Por perfil, entende-se a classificação de um profissional em função de características dos casos que atende, tais como faixa etária, diagnóstico principal, gravidade da doença, entre outras. Assim, uma mesma especialidade pode ter diferentes perfis. Para cada perfil poderão existir bandas ideais diferentes. Como cada especialidade ou perfil tem uma banda ideal específica para cada indicador, é

possível comparar determinado indicador para especialidades e perfis distintos. No modelo GPS.2iM©, o Ajuste de Risco é realizado separando os médicos de uma mesma especialidade em perfis, considerando os fatores de risco dos pacientes atendidos. O objetivo é que os médicos de um mesmo perfil sejam semelhantes, possibilitando que seus resultados possam ser comparados diretamente. Esse ajuste tem efeito direto nas bandas ideais dos indicadores no modelo, que são definidas de forma separada em cada perfil (ABICALAFFE, 2015).

O modelo GPS.2iM© considera que os escores obtidos nas múltiplas dimensões devem ser unificados em um índice de *performance* composto e disponibiliza ao avaliado, via *internet*, todas as componentes de sua avaliação mediante o acesso a um *software* que possui uma planilha visual do índice de *performance*, denominado *scorecard*, medido na escala de 0 a 100, que determina o desempenho do avaliado com base nos indicadores aferidos (ABICALAFFE, 2015).

A Figura 1 apresenta um exemplo ilustrativo do modelo GPS.2iM© em sua íntegra. O primeiro indicador – “média de permanência por especialidade” – pertence ao domínio Efetividade e tem um peso de 12 dos 100 pontos disponíveis. O médico apresentou média de permanência de 7,556 dias, muito acima do resultado ideal para esse indicador que é de 0 a 1,19 dias. Pelo critério definido, esse médico ficou com 0% do peso desse indicador, portanto, o escore do indicador ficou em zero. A soma dos escores de cada indicador traduz-se no índice de performance, ou seja, uma nota de zero a 100 (no exemplo acima, o índice foi de 85).

A escolha dos indicadores de desempenho a serem utilizados pelo modelo GPS.2iM© é definida de acordo com as características de cada hospital, ou seja, cada hospital apresenta seu conjunto de indicadores a serem utilizados na avaliação. Os indicadores analisados nesse estudo são apresentados no Quadro 1. Indicadores com código 6XXX foram fornecidos diretamente pelo hospital, enquanto os com código 7XXX foram calculados pela 2iM com as fórmulas informadas. O quadro também informa em quais domínios do modelo GPS.2iM© os indicadores foram alocados.

FIGURA 1 – EXEMPLO ILUSTRATIVO DO MODELO GPS.2IM©.

Domínios / Indicador - Dezembro de 2014	Pontos MAX	Valor Indicador	Banda Inferior			Banda Ideal				Banda Superior					Score			
			0%	50%	60%	70%	80%	90%	100%	90%	80%	70%	60%	50%		0%		
<b>EFETIVIDADE</b>																		
(7026) CALC - Média de permanência por especialidade - Dias	12	7,556							0,000	1,190	1,191 a 1,890	1,891 a 2,590	2,591 a 3,290	3,291 a 3,990	3,991 a 4,690	Acima de 4,691	0,00	
(7027) CALC - Taxa de pacientes de longa permanência (acima de 20 dias) - %	10	11,111							0,000	10,950	10,591 a 14,260	14,261 a 17,570	17,571 a 20,880	20,881 a 24,190	24,191 a 27,500	27,501 a 100,000	9,00	
(7078) CALC - Taxa de mortalidade geral	12	5,556							0,000	6,250						6,251 a 100,000	12,00	
<b>TOTAL DO DOMÍNIO 34,00 EFICIÊNCIA</b>																		21,00
			0%	50%	60%	70%	80%	90%	100%	90%	80%	70%	60%	50%	0%			
(6077) FORN - História clínica - qtde	4	N.D	0,000 a 49,999	50,000 a 60,999		61,000 a 70,999	71,000 a 80,999		81,000	100,000							4,00	
(6078) FORN - Resumo de alta - qtde	4	N.D	0,000 a 49,999	50,000 a 60,999		61,000 a 70,999	71,000 a 80,999		81,000	100,000							4,00	
(7024) CALC - Ticket médio (geral) - R\$	8	R\$ 14.889,482	0,000 a 1.091,939						1.091,940	Acima de 1.091,940							8,00	
(7066) CALC - Faturamento por médico por mês	13	R\$ 268.010,670	0,000 a 12.084,989					12.084,990 a 37.611,389	37.611,390	Acima de 37.611,390							13,00	
(7071) CALC - Número de internações / mês por médico - %	6	18,000	0,000 a 4,849					4,850 a 23,039	23,040	Acima de 23,040							5,40	
(7074) CALC - Alta dada pelo médico até às 11h - %	7	12,500	0,000 a 5,139				5,140 a 13,369	13,370 a 21,559	21,600	Acima de 21,600							5,60	
(7079) CALC - Taxa de internação - %	8	62,069	0,000 a 0,979					0,990 a 4,279	4,280	Acima de 4,280							8,00	
(7095) CALC - Margem de contribuição - %	0	70,840	0,000 a 49,999	50,000 a 60,999		61,000 a 70,999	71,000 a 80,999		81,000	100,000							0,00	
<b>TOTAL DO DOMÍNIO 50,00 ESTRUTURA</b>																		48,00
			0%	50%	60%	70%	80%	90%	100%	90%	80%	70%	60%	50%	0%			
(6101) FORN - Fazer parte voluntariamente da diretoria, comissões, conselho e coordenar áreas do hospital - qtde	5	N.D	0,000 a 0,000						0,000	1,000							5,00	
(6103) FORN - Atividade científica: participação em congressos, coordenação e publicação científica com o nome do hospital - qtde	5	N.D	0,000 a 0,000						0,000	1,000							5,00	
<b>TOTAL DO DOMÍNIO 10,00 SATISFAÇÃO</b>																		10,00
(6106) FORN - Número de reclamações dos clientes - qtde	6	N.D							0,000	0,000							6,00	
<b>TOTAL DO DOMÍNIO 6,00 TOTAL GERAL 100,00</b>																		6,00
<b>TOTAL DA PERFORMANCE</b>																<b>85,00</b>		

FONTE: Adaptado de Abicalaffe (2015).

QUADRO 1 – INDICADORES AVALIADOS PELO MODELO GPS.2IM© NO HOSPITAL.

CÓDIGO	INDICADOR	DOMÍNIOS 2IM	FÓRMULA
6070	Taxa de não conformidade de auditoria de prontuários	Eficiência	
6102	Taxa de infecção de sitio cirúrgico	Efetividade	
6103	Atividade científica: participação em congressos, coordenação, palestrante de atividades pelo ICEP, publicação científica com nome do hospital	Estrutura	
6112	Atividade científica: participação de atividades pelo ICEP	Estrutura	
6114	Taxa de satisfação do cliente – tempo para atendimento	Satisfação	
6115	Taxa de satisfação do cliente – cordialidade	Satisfação	
6116	Taxa de satisfação do cliente – orientações	Satisfação	
7049	Ticket médio	Eficiência	$\frac{\text{faturamento total}}{\text{quantidade de cirurgias do médico}}$
7050	Ticket médio (a)	Eficiência	$\frac{\text{faturamento total}}{\text{número de saídas}}$
7051	Taxa de movimento cirúrgico	Eficiência	$\frac{\text{total de procedimentos cirúrgicos}}{\text{total de atendimentos}}$
7052	Número de procedimentos cirúrgicos / mês por médico	Eficiência	$\frac{\text{número de procedimentos cirurgicos}}{\text{quantidade de altas até às 11 horas}}$
7053	Alta dada pelo médico até às 11h	Eficiência	$\frac{\text{quantidade de altas até às 11 horas}}{\text{número de pacientes por dia}}$
7054	Média de permanência por especialidade	Efetividade	$\frac{\text{número de saídas}}{\text{número de saídas}}$
7055	Número de internação / mês por médico	Eficiência	$\frac{\text{número de saídas}}{\text{número de saídas}}$
7056	Taxa de internação da especialidade	Eficiência	$\frac{\text{número de saídas}}{\text{número de saídas por especialidade}}$
7057	Taxa de mortalidade geral	Efetividade	$\text{taxa de mortalidade}$
7058	Taxa de pacientes de longa permanência	Efetividade	$\frac{\text{número de pacientes com permanência } \geq \text{ a 20 dias}}{\text{número de saídas}}$
7072	Taxa de mortalidade cirúrgica até 48 horas	Efetividade	$\text{taxa de mortalidade cirúrgica até 48 horas}$
7073	Faturamento por médico por mês	Eficiência	$\text{faturamento total}$

FONTE: 2iM (2017).

## 2.6 MÉTODOS ESTATÍSTICOS

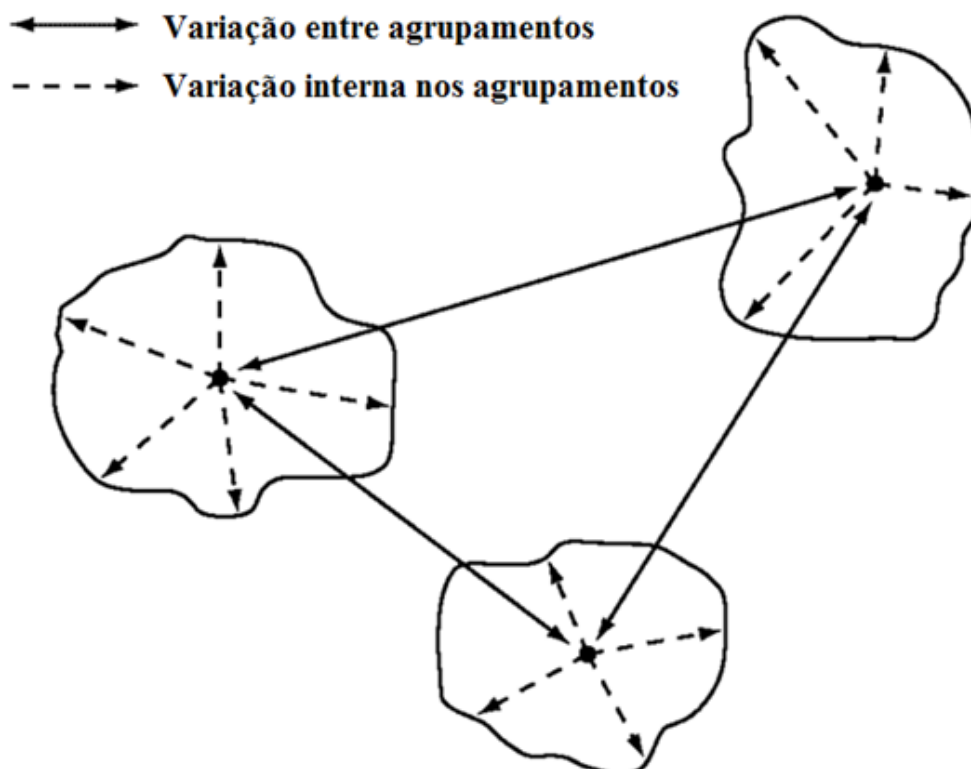
### 2.6.1 Análise de agrupamentos

Análise de Agrupamentos ou Análise de *Clusters* é um método de classificação não supervisionada que classifica objetos em diferentes grupos, cada um dos quais contendo objetos semelhantes segundo alguma função de distância, em geral a Distância Euclidiana. Os grupos resultantes devem exibir elevada homogeneidade interna e elevada heterogeneidade externa.

#### 2.6.1.1 Algoritmos de agrupamento

Os algoritmos de agrupamento têm por objetivo construir grupos de tal forma a maximizar as diferenças entre os grupos, conforme apresentado na Figura 2. Os algoritmos de agrupamento mais comumente usados podem ser classificados em duas categorias gerais: hierárquica e não hierárquica.

FIGURA 2 – DIAGRAMA DE AGRUPAMENTO COM A VARIAÇÃO ENTRE E DENTRO DOS GRUPOS.



FONTE: Adaptado de Hair et al. (2005).

### 2.6.1.2 Métodos de agrupamento hierárquico

Segundo Johnson e Wichern (2007), nesse método, de início existem tantos grupos quanto objetos. Diversos objetos semelhantes são agrupados primeiro, unindo esses grupos iniciais de acordo com as suas similaridades, atenuando o critério de similaridade os subgrupos vão se unindo a outros subgrupos até formar um grupo único.

O procedimento de agrupamento é o seguinte:

- 1) no início, tem-se  $n$  grupos, cada um formado por um único objeto; calcula-se a matriz simétrica de distâncias  $n \times n$ ,  $D = (d_{ij})$ , em que  $d_{ij}$  é a distância entre o objeto  $i$  e o objeto  $j$ ;

$$D = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix}, \quad (1)$$

sendo:  $d_{11} = d_{22} = \cdots = d_{nn} = 0$ ;

- 2) na matriz  $D$ , identifica-se o par de grupos mais próximo (menor distância) que passa a formar um único grupo;
- 3) o novo grupo formado é denominado, por ex.,  $(A, B)$ , se os grupos primitivos do par forem  $A$  e  $B$ . Nova matriz de distâncias é construída, apagando-se as linhas e colunas correspondentes aos grupos  $A$  e  $B$  e inserindo à matriz a linha e a coluna dadas pelas distâncias entre  $(AB)$  e os grupos remanescentes;
- 4) os passos 2 e 3 são repetidos  $(n-1)$  vezes observando as identidades dos grupos que são agrupados.

### 2.6.1.3 Ligação simples

As ligações se referem ao algoritmo utilizado para agrupar os objetos semelhantes. O procedimento de ligação simples é baseado na distância mínima: os dois objetos separados pela menor distância são identificados e colocados no primeiro agrupamento. Em seguida, a próxima distância mais curta dentre todos os pares de

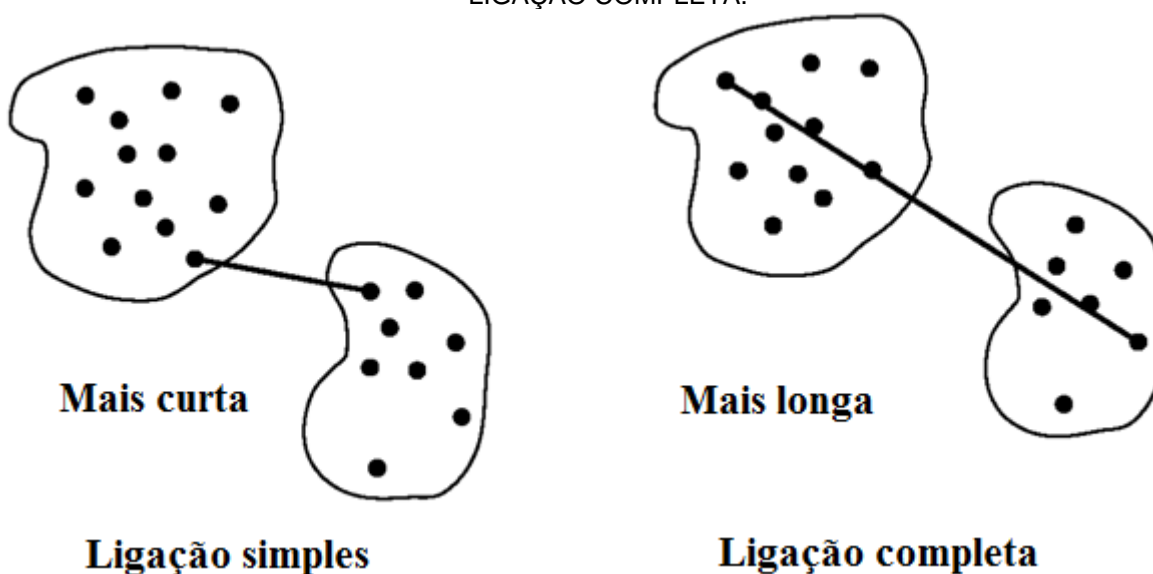
elementos de um mesmo grupo é determinada e um terceiro objeto se junta aos dois primeiros para formar um agregado, ou um novo grupo de dois membros é formado. O processo continua até que todos os objetos formem um só agregado. Esse procedimento também foi chamado de abordagem do vizinho mais próximo (HAIR et al., 2005; EVERITT et al., 2011; RENCHER, 2002).

#### 2.6.1.4 Ligação completa

O procedimento de ligação completa é semelhante ao da ligação simples, exceto que o critério de agrupamento se baseia em distância máxima. Por essa razão, também é conhecido como abordagem do vizinho mais distante. Esse método é chamado de ligação completa porque todos os objetos de um agrupamento são conectados a alguma distância máxima ou similaridade mínima (HAIR et al., 2005).

A Figura 3 mostra como as distâncias menores (ligação simples) e maiores (ligação completa) representam a similaridade entre grupos. O uso da menor distância reflete o par de objetos mais próximos, enquanto a ligação completa reflete o par mais distante.

FIGURA 3 – COMPARAÇÃO DOS MÉTODOS DE DISTÂNCIA PARA A LIGAÇÃO SIMPLES E LIGAÇÃO COMPLETA.



FONTE: Adaptado de Hair et al. (2005).

#### 2.6.1.5 Ligação média

O método da ligação média começa da mesma forma que a ligação simples ou completa, mas o critério de agrupamento é a distância média entre pares de objetos, sendo um em cada grupo. Nessa técnica a partição é baseada em todos os elementos dos agregados, ao invés de um único par de membros extremos. A abordagem da ligação média tende a combinar agregados com pequena variação interna e aproximadamente mesma variância (HAIR et al., 2005; EVERITT et al., 2011; RENCHER, 2002).

#### 2.6.1.6 Método do centroide

No método do centroide, a distância entre os agrupamentos é a distância entre os seus centroides, que são os valores médios dos vetores na variável estatística de agrupamento. Nesse método, toda vez que indivíduos são reunidos, um novo centroide é computado. Os centroides migram quando ocorrem fusões de agregados, ou seja, existe uma mudança de centroide do agrupamento toda vez que um novo indivíduo ou grupo de indivíduos é acrescentado a um grupo já existente. Esse método pode produzir resultados confusos por causa das inversões, ou seja, casos em que a distância entre os centroides de um par pode ser menor do que a distância entre os centroides de outro par unido em uma combinação anterior. A vantagem desse método é que ele é menos afetado por observações atípicas do que outros métodos hierárquicos (HAIR et al., 2005; EVERITT et al., 2011; RENCHER, 2002).

#### 2.6.1.7 Método de Ward

No método de Ward, a distância entre dois agrupamentos é a soma dos quadrados entre os dois agrupamentos sobre todas as variáveis. Em cada estágio do procedimento de agrupamento, a soma interna dos quadrados é minimizada sobre todas as partições que possam ser obtidas pela combinação de dois agregados do estágio anterior. Esse procedimento tende a combinar agrupamentos com um pequeno número de observações, produzindo agregados com o mesmo número de observações (HAIR et al., 2005; EVERITT et al., 2011; RENCHER, 2002).

### 2.6.1.8 Métodos de agrupamento não-hierárquico

Segundo Johnson e Wichern (2007), os métodos de agrupamento não-hierárquicos são técnicas usadas quando se deseja formar um número pré-definido de grupos de itens ou objetos. O método não-hierárquico mais usado é o algoritmo das k-médias, que é composto por 3 etapas:

- 1) partir arbitrariamente os itens em  $k$  grupos iniciais;
- 2) realocar cada item no grupo cuja média esteja mais próxima. Em geral é usada a Distância Euclidiana. O centroide é recalculado para o grupo que recebeu novo item e para o grupo que perdeu algum item;
- 3) repetir a 2ª etapa até que não restem mais realocações a serem feitas.

### 2.6.2 Análise multivariada de variância (MANOVA)

A análise multivariada de variância é a extensão multivariada das técnicas univariadas para avaliar as diferenças entre as médias de grupos. A hipótese nula testada é a de igualdade dos vetores de médias sobre múltiplas variáveis dependentes. O modelo possui as seguintes suposições:

- 1) independência (as amostras aleatórias de cada população são independentes);
- 2) homocedasticidade (todas as populações têm mesma matriz covariância  $\Sigma$ );
- 3) todas as populações são normalmente distribuídas.

A estatística do teste da MANOVA é o Lambda de Wilks, equivalente ao teste F na ANOVA, que é obtido da seguinte forma:

$$\Lambda^* = \frac{|B|}{|B + W|}, \quad (2)$$

sendo os valores de B e W obtidos conforme apresentado no Quadro 2.

QUADRO 2 – MATRIZ DAS SOMAS DOS QUADRADOS E PRODUTOS CRUZADOS DA MANOVA.

FONTE DE VARIAÇÃO	MATRIZ DAS SOMAS DOS QUADRADOS E PRODUTOS CRUZADOS	GRAUS DE LIBERDADE
Tratamentos	$B = \sum_{i=1}^g n_i (\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})'$	$g - 1$
Resíduos	$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(X_{ij} - \bar{X}_i)'$	$\sum_{i=1}^{n_i} n_i - g$
Total	$B + W$	$\sum_{i=1}^{n_i} n_i - 1$

FONTE: Johnson e Wichern (2007).

em que

$X_{ij}$  é a j-ésima observação da i-ésima amostra (ou i-ésimo tratamento);

$\bar{X}_i$  é a média da i-ésima amostra (ou i-ésimo tratamento);

$\bar{X}$  é a média global (todas as amostras).

A distribuição amostral do Lambda de Wilks é apresentada no Quadro 3:

QUADRO 3 – DISTRIBUIÇÃO AMOSTRAL DO LAMBDA DE WILKS.

Nº DE VARIÁVEIS	Nº DE GRUPOS	DISTRIBUIÇÕES AMOSTRAIS
$p = 1$	$g \geq 2$	$\left(\frac{\sum_{i=1}^g n_i - g}{g - 1}\right) \left(\frac{1 - \Lambda^*}{\Lambda^*}\right) \sim F_{g-1, n-g}$
$p = 2$	$g \geq 2$	$\left(\frac{\sum_{i=1}^g n_i - g - 1}{g - 1}\right) \left(\frac{1 - \sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2(g-1), 2(n-g-1)}$
$p \geq 1$	$g = 2$	$\left(\frac{\sum_{i=1}^g n_i - p - 1}{p}\right) \left(\frac{1 - \Lambda^*}{\Lambda^*}\right) \sim F_{p, n-p-1}$
$p \geq 1$	$g = 3$	$\left(\frac{\sum_{i=1}^g n_i - p - 2}{p}\right) \left(\frac{1 - \sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2p, 2(n-p-2)}$

FONTE: Johnson e Wichern (2007).

Se  $n = \sum_{i=1}^g n_i$  é suficientemente grande, a distribuição do  $\Lambda$  passa a ser:

$$-\left(n - 1 - \frac{p + g}{2}\right) \cdot \ln(\Lambda^*) \sim X_{p(g-1)}^2, \quad (3)$$

### 2.6.3 Análise Discriminante

A técnica multivariada conhecida como Análise Discriminante contempla a alocação de novos objetos em conjuntos previamente definidos. A Análise Discriminante é uma técnica estatística apropriada quando a variável dependente é categórica e as variáveis independentes são numéricas. A técnica envolve determinar uma variável estatística, a combinação linear de duas ou mais variáveis independentes que discriminarão melhor entre os grupos *a priori*. A discriminação é conseguida estabelecendo-se os pesos da variável estatística para cada variável para maximizar a variância entre grupos relativa à variância dentro dos grupos. A combinação linear para uma Análise Discriminante, também conhecida como função discriminante, é determinada por uma equação que assume a seguinte forma:

$$Z_{ij} = a + W_1X_{1k} + W_2X_{2k} + \dots + W_nX_{nk}, \quad (4)$$

sendo

$Z_{ij}$  o escore Z discriminante da função discriminante  $j$  para o objeto  $k$ ;

$a$  o intercepto;

$W_{ik}$  o peso discriminante para a variável independente  $i$ .

$X_{ij}$  a variável independente  $i$  para o objeto  $k$  (HAIR et al., 2005).

#### 2.6.3.1 Discriminação e classificação entre duas populações

Segundo Johnson e Wichern (2007), o método de Fisher transforma as observações multivariadas  $\underline{X}$  nas observações univariadas  $Y$ , tal que os  $Y$ 's nas populações  $\pi_1$  e  $\pi_2$  sejam separados tanto quanto possível.

Sejam  $\mu_{1i}$  a média dos  $Y$ 's obtidos dos  $\underline{X}$ 's pertencentes à população  $\pi_1$  e  $\mu_{2i}$  a média dos  $Y$ 's obtidos dos  $\underline{X}$ 's pertencentes à população  $\pi_2$ , Fisher propôs a combinação linear que maximiza a distância quadrática entre  $\mu_{1i}$  e  $\mu_{2i}$  relativamente à variabilidade dos  $Y$ 's. Assim, seja

- $\underline{\mu}_1 = E(\underline{X}/\pi_1)$ : valor esperado de uma observação multivariada de  $\pi_1$ .
- $\underline{\mu}_2 = E(\underline{X}/\pi_2)$ : valor esperado de uma observação multivariada de  $\pi_2$ .

E supondo a matriz de covariância

$$\Sigma = E \left[ (\underline{\mathbf{X}} - \underline{\boldsymbol{\mu}}_i) (\underline{\mathbf{X}} - \underline{\boldsymbol{\mu}}_i)' \right], i = 1, 2, \quad (5)$$

como sendo a mesma para ambas as populações, então considerando a combinação linear  $Y = \underline{\mathbf{C}}' \underline{\mathbf{X}}$ , tem-se

$$\mu_{1Y} = E(Y/\pi_1) = E(\underline{\mathbf{C}}' \underline{\mathbf{X}}/\pi_1) = \underline{\mathbf{C}}' E(\underline{\mathbf{X}}/\pi_1) = \underline{\mathbf{C}}' \underline{\boldsymbol{\mu}}_1. \quad (6)$$

Da mesma forma

$$\mu_{2Y} = E(Y/\pi_2) = E(\underline{\mathbf{C}}' \underline{\mathbf{X}}/\pi_2) = \underline{\mathbf{C}}' E(\underline{\mathbf{X}}/\pi_2) = \underline{\mathbf{C}}' \underline{\boldsymbol{\mu}}_2, \quad (7)$$

e

$$\sigma_Y^2 = V(Y) = V(\underline{\mathbf{C}}' \underline{\mathbf{X}}) = \underline{\mathbf{C}}' V(\underline{\mathbf{X}}) \underline{\mathbf{C}} = \underline{\mathbf{C}}' \Sigma \underline{\mathbf{C}}, \quad (8)$$

que é a mesma para as duas populações.

Segundo Fisher (JOHNSON; WICHERN, 2007), a melhor combinação linear vem da razão entre o quadrado da distância entre as médias e a variância de Y, ou seja,

$$\frac{(\mu_{1Y} - \mu_{2Y})^2}{\sigma_Y^2} = \frac{(\underline{\mathbf{C}}' \underline{\boldsymbol{\mu}}_1 - \underline{\mathbf{C}}' \underline{\boldsymbol{\mu}}_2)^2}{\underline{\mathbf{C}}' \Sigma \underline{\mathbf{C}}} = \frac{(\underline{\mathbf{C}}' \underline{\boldsymbol{\delta}})^2}{\underline{\mathbf{C}}' \Sigma \underline{\mathbf{C}}}, \quad (9)$$

sendo

$$\underline{\boldsymbol{\delta}} = \underline{\boldsymbol{\mu}}_1 - \underline{\boldsymbol{\mu}}_2 \text{ e } V(\underline{\mathbf{X}}) = \Sigma. \quad (10)$$

Seja  $\underline{\boldsymbol{\delta}} = \underline{\boldsymbol{\mu}}_1 - \underline{\boldsymbol{\mu}}_2$  e  $Y = \underline{\mathbf{C}}' \underline{\mathbf{X}}$ , então a razão  $\frac{(\underline{\mathbf{C}}' \underline{\boldsymbol{\delta}})^2}{\underline{\mathbf{C}}' \Sigma \underline{\mathbf{C}}}$  é maximizada por  $\underline{\mathbf{C}} = K \Sigma^{-1} \underline{\boldsymbol{\delta}} = K \Sigma^{-1} (\underline{\boldsymbol{\mu}}_1 - \underline{\boldsymbol{\mu}}_2)$  para qualquer  $K \neq 0$ . Para  $K = 1$  tem-se  $\underline{\mathbf{C}} = \Sigma^{-1} (\underline{\boldsymbol{\mu}}_1 - \underline{\boldsymbol{\mu}}_2)$  e

$$Y = \underline{C}'\underline{X} = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1}\underline{X}, \quad (11)$$

que é conhecida como Função Discriminante Linear de Fisher (F.D.L.)

A F.D.L. de Fisher, que não exige que os dados possuam distribuição Normal (RENCHER, 2002), transforma as populações multivariadas  $\pi_1$  e  $\pi_2$  em populações univariadas tais que as médias das populações univariadas sejam separadas tanto quanto possível relativamente à variância populacional considerada comum. Assim

$$Y_0 = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1}\underline{X}_0, \quad (12)$$

é o valor da F.D.L. de Fisher para uma nova observação  $\underline{X}_0$  e considera-se o ponto médio entre as duas populações univariadas

$$\begin{aligned} m &= \frac{1}{2}(\mu_{1Y} - \mu_{2Y}) = \frac{1}{2}(\underline{C}'\underline{\mu}_1 + \underline{C}'\underline{\mu}_2) = \frac{1}{2}\underline{C}'(\underline{\mu}_1 + \underline{\mu}_2) \\ &= \frac{1}{2}[(\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1}(\underline{\mu}_1 + \underline{\mu}_2)], \end{aligned} \quad (13)$$

e ainda pode-se mostrar que

$$E(Y_0/\pi_1) - m \geq 0 \quad e \quad E\left(\frac{Y_0}{\pi_2}\right) - m < 0, \quad (14)$$

ou seja, se  $\underline{X}_0 \in \pi_1$ , é esperado que  $Y_0 \geq m$ . Por outro lado, se  $\underline{X}_0 \in \pi_2$ , é esperado que  $Y_0 < m$ .

Desta forma, tem-se a regra de classificação:

- alocar  $\underline{X}_0$  em  $\pi_1$  se  $Y_0 \geq m$
- alocar  $\underline{X}_0$  em  $\pi_2$  se  $Y_0 < m$

A F.D.L. de Fisher estimada é dada por:

$$\hat{Y} = \underline{C}'\underline{X} = (\underline{\mu}_1 - \underline{\mu}_2)' S_p^{-1}\underline{X}, \quad (15)$$

sendo o ponto médio  $m$  estimado por:

$$\hat{m} = \frac{1}{2} [(\underline{\mathbf{X}}_1 - \underline{\mathbf{X}}_2)' S_p^{-1} (\underline{\mathbf{X}}_1 + \underline{\mathbf{X}}_2)] = \frac{\bar{Y}_1 + \bar{Y}_2}{2}, \quad (16)$$

em que  $\bar{\mathbf{X}}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{i1}$ ,  $\bar{\mathbf{X}}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{i2}$  e

$$S_p = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2} \quad (17)$$

com

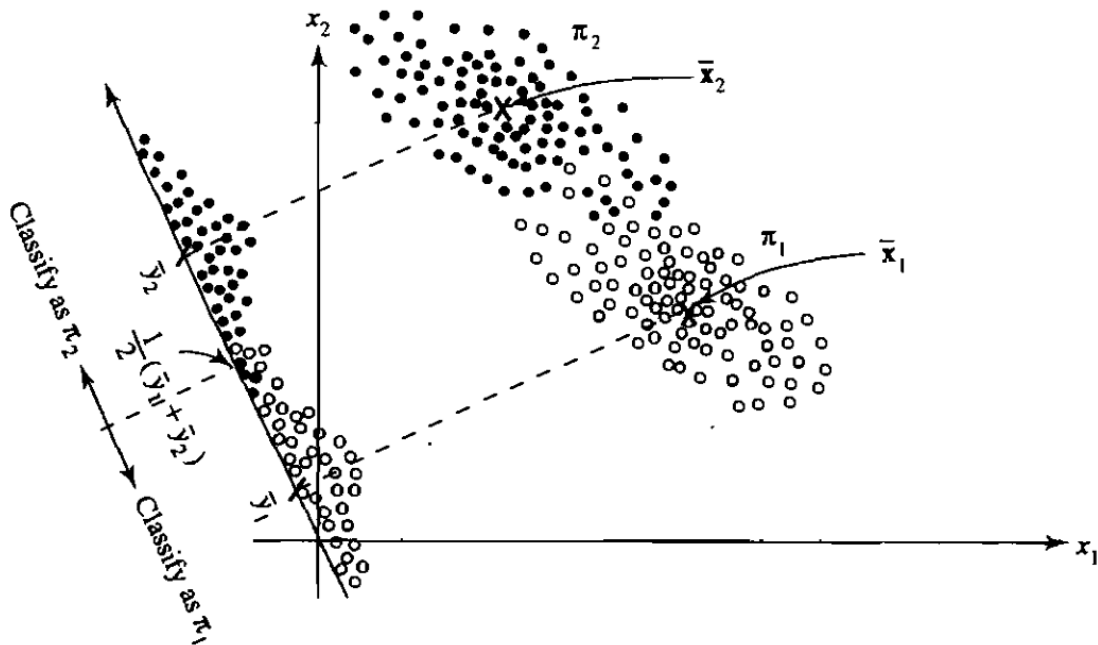
$$S_1 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (\underline{\mathbf{X}}_{i1} - \bar{\mathbf{X}}_1) (\underline{\mathbf{X}}_{i1} - \bar{\mathbf{X}}_1)' \quad e \quad (18)$$

$$S_2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (\underline{\mathbf{X}}_{i2} - \bar{\mathbf{X}}_2) (\underline{\mathbf{X}}_{i2} - \bar{\mathbf{X}}_2)'.$$

Assim, a regra de alocação baseada na F.D.L. de Fisher estimada fica:

- alocar  $\underline{\mathbf{X}}_0$  em  $\pi_1$  se  $\hat{Y}_0 \geq \hat{m}$
- alocar  $\underline{\mathbf{X}}_0$  em  $\pi_2$  se  $\hat{Y}_0 < \hat{m}$

FIGURA 4 – REPRESENTAÇÃO GRÁFICA DO MÉTODO DE FISHER PARA DUAS POPULAÇÕES.



FONTE: Johnson e Wichern (2007).

### 2.6.3.2 Discriminação entre diversas populações

De acordo com Johnson e Wichern (2007), o método de discriminação para duas populações pode ser estendido para diversas populações. Esse método não necessita da suposição de que as diversas populações sejam normais, entretanto é assumido que as matrizes de covariâncias populacionais sejam iguais, isto é,  $\Sigma_1 = \Sigma_2 = \dots = \Sigma_g = \Sigma$ . Assim, seja  $\underline{\bar{\mu}}$  o vetor médio das diversas populações

$$\underline{\bar{\mu}} = \frac{1}{g} \sum_{i=1}^g \underline{\mu}_i, \quad (19)$$

e  $B_0$  a matriz de variação entre grupos populacionais tal que

$$B_0 = \sum_{i=1}^g (\underline{\mu}_i - \underline{\bar{\mu}}) (\underline{\mu}_i - \underline{\bar{\mu}})', \quad (20)$$

a combinação linear  $Y = \underline{C}'\underline{X}$  tem esperança dada por:

$$E(Y) = \mu_{iY} = \underline{C}'E(\underline{X}/\pi_i) = \underline{C}'\underline{\mu}_i, \quad (21)$$

para  $i = 1, 2, \dots, g$ , e variância dada por:

$$\sigma_Y^2 = V(Y) = V(\underline{C}'\underline{X}) = \underline{C}'V(\underline{X})\underline{C} = \underline{C}'\Sigma\underline{C}, \quad (22)$$

para todas as populações. Dessa forma o valor esperado univariado  $\mu_{iY}$  se altera quando a população da qual  $\underline{X}$  é selecionado é outra. Tem-se então uma média global dada por:

$$\bar{\mu}_Y = \frac{1}{g} \sum_{i=1}^g \mu_{iY} = \frac{1}{g} \sum_{i=1}^g \underline{C}'\underline{\mu}_i = \underline{C}'\underline{\bar{\mu}}, \quad (23)$$

e, conseqüentemente, a razão entre a soma dos quadrados das distâncias das populações para a média global de Y em relação à variância de Y é

$$\frac{\sum_{i=1}^g (\mu_{1Y} - \bar{\mu}_Y)^2}{\sigma_Y^2} = \frac{\sum_{i=1}^g (\underline{C}' \underline{\mu}_i - \underline{C}' \bar{\underline{\mu}})^2}{\underline{C}' \underline{\Sigma} \underline{C}} = \frac{\sum_{i=1}^g [\underline{C}' (\underline{\mu}_i - \bar{\underline{\mu}}) (\underline{\mu}_i - \bar{\underline{\mu}})' \underline{C}]}{\underline{C}' \underline{\Sigma} \underline{C}} \quad (24)$$

$$\frac{\underline{C}' \left[ \sum_{i=1}^g (\underline{\mu}_i - \bar{\underline{\mu}}) (\underline{\mu}_i - \bar{\underline{\mu}})' \right] \underline{C}}{\underline{C}' \underline{\Sigma} \underline{C}} = \frac{\underline{C}' B_0 \underline{C}}{\underline{C}' \underline{\Sigma} \underline{C}}.$$

Essa razão mede a variabilidade entre os grupos de valores Y relativamente à variabilidade comum entre os grupos. Analogamente ao problema de duas populações, pode-se selecionar  $\underline{C}$  que maximiza esta razão. É conveniente escalonar  $\underline{C}$  tal que  $\underline{C}' \underline{\Sigma} \underline{C} = 1$ .

Seja  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_s$  os  $s \leq \min(g-1, p)$  autovalores não nulos de  $\Sigma^{-1} B_0$  e  $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_s$  os correspondentes autovetores escalonados tal que  $\underline{e}' \underline{\Sigma} \underline{e} = 1$ . Então o vetor de coeficientes  $\underline{C}$  que maximiza a razão  $\frac{\underline{C}' B_0 \underline{C}}{\underline{C}' \underline{\Sigma} \underline{C}}$  é dado por  $\underline{C}_1 = \underline{e}_1$ . A combinação linear  $\underline{C}'_1 \underline{X} = \underline{e}'_1 \underline{X}$  é chamada de primeira discriminante.

O vetor  $\underline{C}_2 = \underline{e}_2$  que maximiza  $\frac{\underline{C}' B_0 \underline{C}}{\underline{C}' \underline{\Sigma} \underline{C}}$  é sujeito à condição  $Cov(\underline{C}'_1 \underline{X}, \underline{C}'_2 \underline{X}) = 0$ . A combinação linear  $\underline{C}'_2 \underline{X} = \underline{e}'_2 \underline{X}$  é chamada de segunda discriminante. E assim sucessivamente obtemos os vários discriminantes. Ainda se tem que  $V(\underline{C}'_i \underline{X}) = 1$ .

As estimativas de  $\Sigma$  e  $\underline{\mu}_i$  são obtidas a partir de amostras aleatórias de tamanhos  $n_i$  das populações  $\pi_i$  ( $i = 1, 2, \dots, g$ ):

$$\bar{\underline{X}}_i = \frac{1}{n_i} \sum_{j=1}^{n_j} X_{ij}, \quad (25)$$

$$\bar{\underline{X}} = \frac{\sum_{i=1}^g n_i \bar{\underline{X}}_i}{\sum_{i=1}^g n_i} = \frac{\sum_{i=1}^g \sum_{j=1}^{n_i} n_i \underline{X}_{ij}}{\sum_{i=1}^g n_i}, \quad (26)$$

e a matriz soma de produtos cruzados é estimada por:

$$\hat{B}_0 = \sum_{i=1}^g (\bar{\mathbf{X}}_i - \bar{\mathbf{X}}) (\bar{\mathbf{X}}_i - \bar{\mathbf{X}})' . \quad (27)$$

Um estimador da matriz  $\Sigma$  pode ser obtido com base na matriz  $W$

$$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i) (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)' . \quad (28)$$

Consequentemente,

$$S_p = \frac{W}{n_1 + n_2 + \dots + n_g - g} = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2 + \dots + (n_g - 1)S_g}{n_1 + n_2 + \dots + n_g - g} . \quad (29)$$

Para maximizar a razão

$$\frac{\underline{\mathbf{C}}' \hat{B}_0 \underline{\mathbf{C}}}{\underline{\mathbf{C}}' S_p \underline{\mathbf{C}}} \quad \text{ou} \quad \frac{\underline{\mathbf{C}}' \hat{B}_0 \underline{\mathbf{C}}}{\underline{\mathbf{C}}' W \underline{\mathbf{C}}} \quad (30)$$

o vetor  $\underline{\mathbf{C}}$  é dado por  $\hat{\underline{\mathbf{C}}}_1 = \hat{\underline{\mathbf{e}}}_1$ , e a combinação linear  $\hat{\underline{\mathbf{C}}}_1 \mathbf{X}$  é chamada de primeira discriminante amostral. Continuando, tem-se em  $k$ ,  $\hat{\underline{\mathbf{C}}}_k = \hat{\underline{\mathbf{e}}}_k$  ( $k \leq s$ ) e  $\hat{\underline{\mathbf{C}}}_k \mathbf{X}$  o  $k$ -ésimo discriminante amostral, em que:  $\hat{\underline{\mathbf{C}}}_1 = \hat{\underline{\mathbf{e}}}_1$ ,  $\hat{\underline{\mathbf{C}}}_2 = \hat{\underline{\mathbf{e}}}_2$ , ...,  $\hat{\underline{\mathbf{C}}}_k = \hat{\underline{\mathbf{e}}}_k$  são os autovetores correspondentes dos autovalores da matriz  $W^{-1} \hat{B}_0$ .

### 2.6.3.3 Validação dos resultados

O estágio final da Análise Discriminante envolve a validação dos resultados. Considerando a propensão da Análise Discriminante para aumentar a razão de sucesso se avaliada apenas sobre a amostra de análise, a validação cruzada é um passo essencial. A validação cruzada geralmente é feita com a amostra original, mas é possível usar uma amostra adicional como a amostra de teste.

Segundo Johnson e Wichern (2007), para determinar a precisão de classificação pode-se calcular a Taxa Aparente de Erro, definida como sendo a fração das observações no treinamento amostral referente ao reconhecimento errôneo pela função obtida. Essa taxa é calculada pela Matriz de Confusão que apresenta a real

situação das observações nos grupos comparando-a com o reconhecimento apresentado pelo modelo encontrado. Para  $n_1$  observações de  $\pi_1$  e  $n_2$  observações de  $\pi_2$ , a matriz de confusão tem a forma dada no Quadro 4:

QUADRO 4 – MATRIZ DE CONFUSÃO PARA A VALIDAÇÃO DOS RESULTADOS DA ANÁLISE DISCRIMINANTE.

CLASSIFICAÇÃO REAL	CLASSIFICAÇÃO PREDITA	
	$\Pi_1$	$\Pi_2$
$\Pi_1$	$n_{1/1}$	$n_{1/2}$
$\Pi_2$	$n_{2/1}$	$n_{2/2}$

FONTE: Johnson e Wichern (2007).

A taxa aparente de erro é dada por:

$$APER = \frac{n_{1/2} + n_{2/1}}{n_1 + n_2}, \quad (31)$$

devendo ser interpretada como a proporção de observações classificadas incorretamente.

#### 2.6.4 Modelo linear misto

O modelo linear misto, proposto por Laird e Ware (1982), é aquele que apresenta tanto fatores de efeitos fixos como aleatórios, além do erro experimental e um termo constante. O termo adicional de efeitos aleatórios é apropriado para representar agrupamentos e, portanto, dados dependentes. Esses modelos são úteis para a análise de dados coletados hierarquicamente ou quando as observações são obtidas de indivíduos relacionados, como irmãos ou outros tipos de parentesco, ou quando os dados são obtidos repetidas vezes em um mesmo indivíduo ao longo de do tempo. Outro motivo de se adotar um modelo linear misto é a possibilidade de se fazer a predição de efeitos aleatórios, na presença de efeitos fixos (FOX; WEISBERG, 2011).

Um modelo linear misto possui algumas peculiaridades na sua análise de variância, como a composição das esperanças matemáticas dos quadrados médios,

cujo conhecimento permite o estabelecimento correto dos testes de hipóteses (HICKS, 1973). De acordo com Fox e Weisberg (2011), o modelo pode ser descrito da seguinte forma:

$$\begin{aligned} y_{ij} &= \beta_1 x_{1ij} + \dots + \beta_p x_{pij} + \gamma_1 z_{1ij} + \dots + \gamma_q z_{qij} + \epsilon_{ij} \\ \gamma_{ik} &\sim N(0, \psi_k^2), \quad \text{Cov}\{\gamma_{ik}, \gamma_{ik'}\} = \psi_{kk'} \\ \epsilon_{ij} &\sim N(0, \sigma^2 \lambda_{ijj'}), \quad \text{Cov}\{\epsilon_{ij}, \epsilon_{ij'}\} = \sigma^2 \lambda_{ijj'}, \end{aligned} \quad (32)$$

sendo que

- $y_{ij}$  é o valor da variável resposta na  $j$ -ésima de  $n_i$  observações no  $i$ -ésimo agrupamento;
- $\beta_1, \dots, \beta_p$  são os coeficientes dos efeitos fixos, que são idênticos para todos os grupos;
- $x_{1ij}, \dots, x_{pij}$  são os valores das variáveis regressoras dos efeitos fixos para a  $j$ -ésima observação no grupo  $i$ . A primeira dessas regressoras é usualmente a constante,  $x_{1ij} = 1$ ;
- $\gamma_1, \dots, \gamma_q$  são os efeitos aleatórios;
- $\psi_k^2$  é a variância e  $\psi_{kk'}$  é a covariância entre os efeitos aleatórios. Assume-se que é constante para todos os grupos;
- $\epsilon_{ij}$  é o erro para a  $j$ -ésima observação no  $i$ -ésimo grupo. O termo de erro para o grupo  $i$  é considerado com distribuição normal multivariada;
- $\sigma^2 \lambda_{ijj'}$  é a covariância entre os erros no  $i$ -ésimo grupo.

Alternativamente, em forma matricial, o modelo fica representado por

$$\begin{aligned} \mathbf{Y}_i &= \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \boldsymbol{\gamma}_i + \boldsymbol{\epsilon}_i \\ \boldsymbol{\gamma}_i &\sim \mathbf{N}_q(\mathbf{0}, \boldsymbol{\Psi}) \\ \boldsymbol{\epsilon}_i &\sim \mathbf{N}_{n_i}(\mathbf{0}, \sigma^2 \boldsymbol{\Lambda}_i) \end{aligned} \quad (33)$$

sendo

- $\mathbf{Y}_i$  o vetor  $n_i \times 1$  de resposta para as observações no  $i$ -ésimo grupo;
- $\mathbf{X}_i$  a matriz  $n_i \times p$  do modelo para os efeitos fixos das observações no  $i$ -ésimo grupo;

- $\underline{\beta}$  o vetor  $p \times 1$  de coeficientes de efeitos fixos;
- $\underline{Z}_i$  a matriz  $n_i \times q$  do modelo de efeitos aleatórios das observações no  $i$ -ésimo grupo;
- $\underline{\gamma}_i$  vetor  $q \times 1$  de coeficientes de efeitos aleatórios para o grupo  $i$ ;
- $\underline{\epsilon}_i$  o vetor  $n_i \times 1$  de erros das observações no  $i$ -ésimo grupo;
- $\underline{\Psi}$  a matriz  $q \times q$  das covariâncias dos efeitos aleatórios;
- $\sigma^2 \underline{\Lambda}_i$  a matriz  $n_i \times n_i$  das covariâncias dos erros para o grupo  $i$ .

Além disso, de (33) tem-se que  $E(Y) = \underline{X}\underline{\beta}$  e  $Var(Y) = \underline{V} = \underline{Z}\underline{\Psi}\underline{Z}^T + \underline{\Sigma}$ , sendo  $\underline{\Sigma}$  a matriz de covariâncias dos vetores aleatórios de  $\underline{\epsilon}_i$

No contexto da análise de resíduos, este estudo aborda os resíduos condicionais padronizados, resíduo com confundimento mínimo e distância de Mahalanobis que são definidos a seguir.

#### 2.6.4.1 Resíduo condicional padronizado

Os resíduos condicionais são utilizados na identificação de pontos discrepantes e para avaliar a suposição de homocedasticidade do modelo (PINHEIRO; BATES, 2000). Nobre (2004) propôs uma padronização dos resíduos condicionais que os torna mais eficazes na verificação de possível afastamento da suposição de homocedasticidade, dada por:

$$\hat{e}_i^* = \frac{\hat{e}_i}{\sigma \sqrt{q_{ii}}}, \quad (34)$$

em que  $\hat{e}_i$  representa o  $i$ -ésimo elemento de  $\hat{e}$  e  $q_{ii}$  o  $i$ -ésimo elemento da matriz  $\underline{Q}$ , sendo essa definida como  $\underline{Q} = \underline{M} - \underline{M}\underline{X}(\underline{X}^T \underline{M}\underline{X})^{-1} \underline{X}^T \underline{M}$  e  $\underline{M} = \sigma^2 \underline{V}^{-1}$ .

#### 2.6.4.2 Resíduo com confundimento mínimo

O resíduo com confundimento mínimo, proposto por Hilden-Minton (1995), é utilizado para a verificação da normalidade, sendo definido por:

$$CF_i = 1 - \frac{\underline{U}_i^T \underline{Q} \underline{Q} \underline{U}_i}{\underline{U}_i^T \underline{Q} \underline{U}_i}, \quad (35)$$

em que  $\underline{U}_i$  é a  $i$ -ésima coluna de  $I_n$ . A medida  $CF_i$  representa a proporção de variabilidade de  $\hat{e}_i$  quando confundido por  $\hat{b}$ , sendo maior o confundimento quanto mais próximo de 1 for o valor de  $CF_i$ . Assim, o autor propôs utilizar a transformação linear  $\underline{L}^T \hat{e}$  para minimizar o confundimento de  $\underline{L}_i^T \hat{e}$ . O resíduo com confundimento mínimo é o mais eficiente na verificação da normalidade dos resíduos condicionais por meio do gráfico normal de probabilidade com envelope simulado (COSTA et al., 2012).

#### 2.6.4.3 Distância de Mahalanobis

Segundo Pinheiro e Bates (2000), o gráfico de  $\hat{\gamma}_i$  versus o índice dos indivíduos pode ser utilizado na identificação de observações discrepantes. No entanto, Waternaux et al (1989) propõem a utilização da distância de Mahalanobis para a identificação de observações discrepantes. A distância de Mahalanobis é dada por:

$$\zeta_i = \hat{\gamma}_i^T \widehat{Var}(\hat{\gamma}_i - \gamma_i) \hat{\gamma}_i, \quad (36)$$

que, sob a suposição de validade do modelo, tem distribuição qui-quadrado com  $n_i$  graus de liberdade para  $n_i$  suficientemente grande.

### 2.6.5 Análise Fatorial

A Análise Fatorial é uma técnica de análise multivariada cujo objetivo é explicar as correlações entre um conjunto grande de variáveis em termos de um conjunto de poucas variáveis aleatórias não-observáveis chamadas fatores. Nessa técnica, as variáveis aleatórias (V.A.s) são agrupadas de acordo com suas correlações, de forma que dentro dos grupos as V.A.s. são altamente correlacionadas entre si, mas têm correlações relativamente baixas com as V.A.s. de outros grupos (JOHNSON; WICHERN, 2007).

#### 2.6.5.1 Modelo fatorial ortogonal

Seja  $\underline{X} \sim \bullet (\mu, \Sigma)$ . No modelo fatorial  $\underline{X}$  é linearmente dependente sobre algumas v. a. não-observáveis  $F_1, F_2, \dots, F_m$  chamadas fatores comuns e  $p$  fontes de variação aditivas  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ , chamadas erros ou fatores específicos. O modelo fatorial é

$$\begin{aligned} X_1 - \mu_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1 \\ X_2 - \mu_2 &= \ell_{21}F_1 + \ell_{22}F_2 + \dots + \ell_{2m}F_m + \varepsilon_2 \\ &\vdots \\ X_p - \mu_p &= \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pm}F_m + \varepsilon_p, \end{aligned} \quad (37)$$

ou em notação matricial:

$$\underbrace{\underline{X} - \underline{\mu}}_{(p \times 1)} = \underbrace{\underline{L}}_{(p \times m)} \underbrace{\underline{F}}_{(m \times 1)} + \underbrace{\underline{\varepsilon}}_{(p \times 1)}, \quad (38)$$

sendo:

- $\mu_i$  a média da  $i$ -ésima variável;
- $\varepsilon_i$  o  $i$ -ésimo erro ou fator específico;
- $F_j$  o  $j$ -ésimo fator comum;

$\ell_{ij}$  o peso ou carregamento na  $i$ -ésima variável  $X_i$  do  $j$ -ésimo fator  $F_j$ ,  $i = 1, 2, \dots, p$ ,  $j = 1, 2, \dots, m$ .

As seguintes suposições definem o modelo:

$$1. E(\underline{\mathbf{F}}) = \underline{\mathbf{0}}, V(\underline{\mathbf{F}}) = E(\underline{\mathbf{F}}\underline{\mathbf{F}}') = {}_m I_m$$

$$2. E(\underline{\boldsymbol{\varepsilon}}) = {}_p \underline{\mathbf{0}}, V(\underline{\boldsymbol{\varepsilon}}) = E(\underline{\boldsymbol{\varepsilon}}\underline{\boldsymbol{\varepsilon}}') = {}_p \Psi_p = \begin{bmatrix} \Psi_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Psi_p \end{bmatrix}$$

$$3. \text{Cov}(\underline{\boldsymbol{\varepsilon}}, \underline{\mathbf{F}}) = {}_p \underline{\mathbf{0}}$$

### 2.6.5.2 Estrutura da covariância

A estrutura da covariância para o modelo fatorial ortogonal é dada por:

$$1. \text{Cov}(\underline{\mathbf{X}}) = \underline{\mathbf{L}}\underline{\mathbf{L}}' + \underline{\boldsymbol{\Psi}}$$

ou

$$\begin{aligned} \text{Var}(X_i) &= \ell_{i1}^2 + \cdots + \ell_{im}^2 + \psi_i \\ \text{Cov}(X_i, X_k) &= \ell_{i1}\ell_{k1} + \cdots + \ell_{im}\ell_{km} \end{aligned}$$

$$2. \text{Cov}(\underline{\mathbf{X}}, \underline{\mathbf{F}}) = \underline{\mathbf{L}}, \text{ ou}$$

$$\text{Cov}(X_i, F_j) = \ell_{ij},$$

de tal forma que  $\ell_{ij}$  é a covariância da  $i$ -ésima variável  $X_i$  com o  $j$ -ésimo fator comum  $F_j$ . Se a matriz de correlação for usada, então  $\ell_{ij}$  será o coeficiente de correlação entre a  $i$ -ésima variável e o  $j$ -ésimo fator comum (JOHNSON; WICHERN, 2007).

### 2.6.5.3 Comunalidades e variâncias específicas

A porção da variância da  $i$ -ésima v.a.  $X_i$  advinda como contribuição dos  $m$  fatores comuns é chamada comunalidade e a porção da  $V(X_i)$  oriunda do fator específico é a variância específica (JOHNSON; WICHERN, 2007). Assim,

$$\text{Var}(X_i) = \underbrace{\ell_{i1}^2 + \dots + \ell_{im}^2}_{\text{comunalidades}} + \underbrace{\psi_i}_{\text{variância específica}}. \quad (39)$$

O valor da comunalidade é um índice útil para avaliar o quanto de variância em cada variável é explicado pela Análise Fatorial. Comunalidades elevadas indicam que uma grande quantia de variância em uma variável foi extraída pela solução fatorial (HAIR et al., 2005).

#### 2.6.5.4 Estimação do modelo fatorial por componentes principais

A Análise Fatorial por Componentes Principais da matriz de covariância  $S$  é especificada em termos de seus pares de autovalores/autovetores  $(\hat{\lambda}_1, \hat{\underline{e}}_1)$ ,  $(\hat{\lambda}_2, \hat{\underline{e}}_2)$ , ...,  $(\hat{\lambda}_p, \hat{\underline{e}}_p)$ , em que  $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_p \geq 0$  (JOHNSON; WICHERN, 2007). Esse método possui a vantagem de não exigir a normalidade dos dados (RENCHEER, 2002).

Seja  $m < p$  o número de fatores comuns. A matriz dos carregamentos estimados dos fatores  $\hat{\ell}_{ij}$  é dada por  $\hat{L} = \hat{C} \hat{D}_\lambda^{1/2}$ , sendo

$$\hat{C} = \begin{bmatrix} \hat{e}_{11} & \dots & \hat{e}_{1p} \\ \vdots & \ddots & \vdots \\ \hat{e}_{p1} & \dots & \hat{e}_{pp} \end{bmatrix} \quad e \quad \hat{D}_\lambda^{1/2} = \begin{bmatrix} \sqrt{\hat{\lambda}_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{\hat{\lambda}_p} \end{bmatrix}. \quad (40)$$

As variâncias específicas estimadas são dadas pelos elementos da diagonal principal da matriz  $\hat{\Psi} = S - \hat{L} \hat{L}'$ , como apresentado em (41)

$$\hat{\Psi} = \begin{bmatrix} \psi_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \psi_p \end{bmatrix} \quad \text{com } \psi_i = s_{ii} - \sum_{j=1}^m \hat{\ell}_{ij}^2. \quad (41)$$

As comunalidades são estimadas conforme (42):

$$\hat{h}_i^2 = \hat{\ell}_{i1}^2 + \hat{\ell}_{i2}^2 + \dots + \hat{\ell}_{im}^2. \quad (42)$$

Para determinar o número de fatores comuns, o indicado é basear-se na proporção da variância amostral devido a cada fator, que é:

$$\frac{\hat{\lambda}_j}{s_{11} + s_{22} + \dots + s_{pp}} \quad \text{para a análise feita a partir da matriz de covariância;}$$

$$\frac{\hat{\lambda}_j}{p} \quad \text{para a análise feita a partir da matriz de correlação.}$$

#### 2.6.5.5 Seleção dos fatores

Os principais critérios para determinar a quantidade de fatores selecionados em um modelo fatorial são os seguintes:

- experiência no assunto pesquisado (método subjetivo);
- critério de Kaiser: tantos fatores quanto for o número de autovalores maiores que 1;
- escolher o número de fatores que explicam uma proporção especificada da variância total.

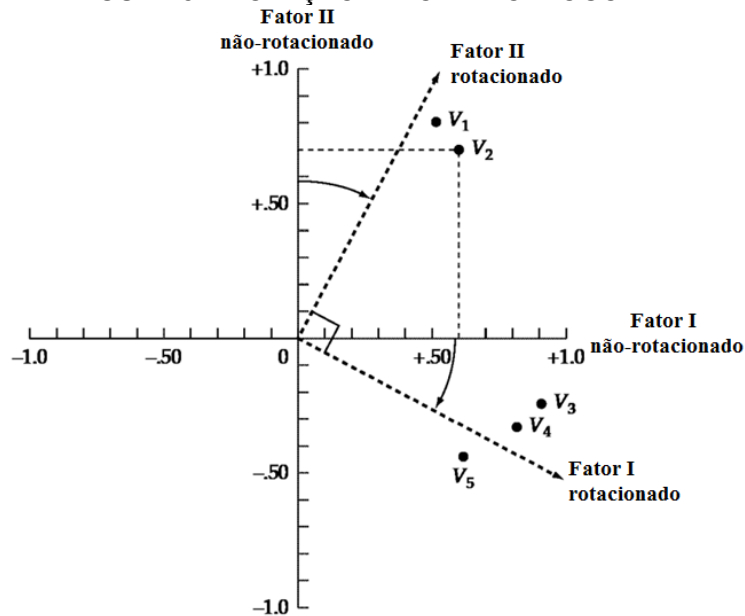
Há consequências negativas na seleção de fatores em excesso ou ao menos para representar os dados. Se poucos fatores são selecionados, a estrutura correta não é revelada e as dimensões importantes podem ser omitidas. Se muitos fatores são mantidos, a interpretação se torna mais difícil quando os resultados são rotacionados (HAIR et al., 2005).

#### 2.6.5.6 Rotação dos fatores

A rotação fatorial é uma ferramenta importante na interpretação dos fatores. Na rotação, os eixos de referência dos fatores são rotacionados em torno da origem até que uma determinada posição seja alcançada. As soluções fatoriais não rotacionadas extraem fatores na sua ordem de importância. O primeiro fator tende a ser um fator geral com quase toda a variável com carga significativa. O segundo fator e os seguintes são baseados na quantia residual de variância. Cada fator explica porções sucessivamente menores de variância. O efeito final de rotacionar a matriz fatorial é redistribuir a variância dos primeiros fatores para os últimos com o objetivo de atingir um padrão fatorial mais simples e teoricamente mais significativo (HAIR et

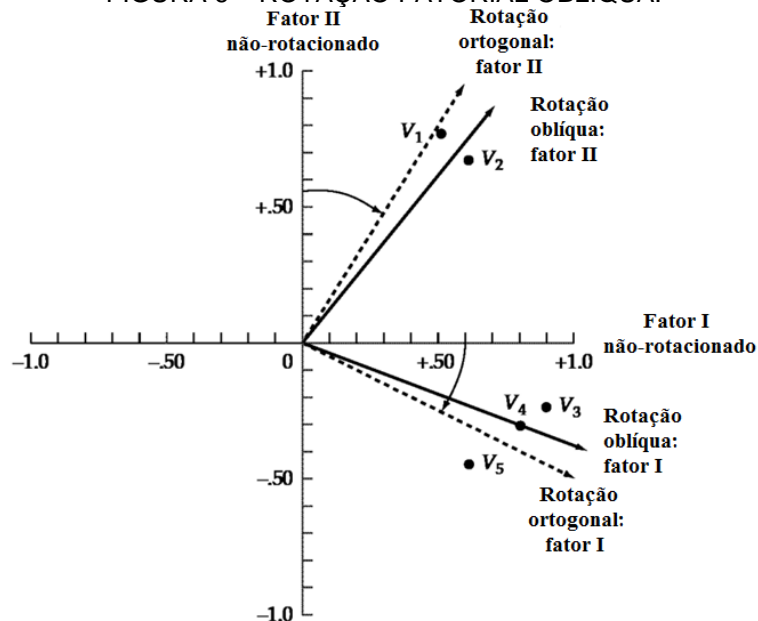
al., 2005). O caso mais simples de rotação é uma rotação ortogonal, na qual os eixos são mantidos a 90 graus. Também é possível rotacionar os eixos sem manter o ângulo de 90 graus entre os eixos de referência. Quando não há a restrição de solução ortogonal, o procedimento de rotação se chama rotação oblíqua. Rotações ortogonais são demonstradas na Figura 5 e rotações oblíquas na Figura 6.

FIGURA 5 – ROTAÇÃO FATORIAL ORTOGONAL.



FONTE: Adaptado de Hair et al. (2005).

FIGURA 6 – ROTAÇÃO FATORIAL OBLÍQUA.



FONTE: Adaptado de Hair et al. (2005).

### 2.6.5.7 Escores fatoriais

Escore fatorial é uma estimativa do valor de cada um dos fatores para uma observação individual. Essas quantidades são frequentemente usadas para diagnosticar propostas da mesma forma que a análise anterior. Pelo método de mínimos quadrados ponderados, os escores fatoriais estimados para variáveis originais são obtidos por:

$$\hat{\mathbf{F}} = (\hat{\mathbf{L}}' \hat{\mathbf{\Psi}}^{-1} \hat{\mathbf{L}})^{-1} \hat{\mathbf{L}} \hat{\mathbf{\Psi}}^{-1} (\mathbf{x} - \bar{\mathbf{x}}). \quad (43)$$

Os escores fatoriais estimados para variáveis padronizadas:

$$\hat{\mathbf{F}} = (\hat{\mathbf{L}}' \hat{\mathbf{L}})^{-1} \hat{\mathbf{L}} \mathbf{z}. \quad (44)$$

Conceitualmente, o escore fatorial representa o grau em que cada indivíduo tem no grupo de itens que têm cargas elevadas em um fator. Assim, valores mais altos nas variáveis com cargas elevadas em um fator resultam em um escore fatorial superior. A desvantagem dos escores fatoriais é que eles não são facilmente repetidos em estudos, pois são baseados na matriz fatorial, a qual é determinada separadamente em cada estudo (HAIR et al., 2005).

Os escores dos fatores podem ser transformados em um escore bruto para cada observação, e assim ranquear as observações de acordo com essa medida (ARTUSO; CHAVES NETO, 2012). Esse escore é ponderado pelos autovalores de acordo com a equação:

$$E_j = \frac{\sum_{i=1}^k \hat{f}_j' \lambda_j}{\sum_{i=1}^k \lambda_j}. \quad (45)$$

O escore bruto obtido em (45) possui como esperança o valor zero. Com isso, espera-se que metade dos escores apresentem valores negativos, o que pode ser uma característica não desejável. A escala dos escores pode ser convertida conforme a equação (46).

$$\frac{x - \text{mín}}{\text{máx} - \text{mín}} = \frac{y - a}{b - a}, \quad (46)$$

sendo  $a$  e  $b$  respectivamente o menor e o maior valor escolhido para a nova escala. Ao isolar  $y$  na equação (46), tem-se o valor na nova escala:

$$y = \frac{a \times \text{máx} - (x \times a - (x - \text{mín}) \times b)}{\text{máx} - \text{mín}}. \quad (47)$$

#### 2.6.5.8 Testes usados em Análise Fatorial

O teste de esfericidade de Bartlett testa a hipótese de que as variáveis não sejam correlacionadas na população. A hipótese nula é a de que a matriz de correlação da população é uma matriz identidade a qual indica que a Análise Fatorial é inapropriada. A estatística do teste é dada por

$$\chi^2 = - \left[ (n - 1) - \frac{2p + 5}{6} \right] \ln|R|, \quad (48)$$

que tem uma distribuição qui-quadrado com  $\frac{p(p-1)}{2}$  graus de liberdade, sendo:

$n$  o tamanho da amostra;

$p$  o número de variáveis;

$|R|$  o determinante da matriz de correlação.

A medida de adequacidade da amostra de Kaiser-Meyer-Olkin (KMO) é representada pelo índice denominado MSA que avalia a adequação da Análise Fatorial, sendo calculada por

$$MSA = \frac{\sum \sum_{j \neq k} r_{jk}^2}{\sum \sum_{j \neq k} r_{jk}^2 + \sum \sum_{j \neq k} q_{jk}^2}, \quad (49)$$

sendo  $r_{jk}^2$  o quadrado dos elementos da matriz de correlação original fora da diagonal e  $q_{jk}^2$  o quadrado da correlação parcial entre as variáveis. Valores altos (entre 0,5 e

1,0) indicam que a análise fatorial é apropriada, enquanto que valores baixos, abaixo de 0,5 indicam que a análise fatorial pode ser inadequada.

#### 2.6.6 Coeficiente de correlação de Spearman

O coeficiente de correlação de Spearman é uma medida de correlação não paramétrica, isto é, ele avalia uma função monótona arbitrária que pode ser a descrição da relação entre duas variáveis, sem fazer nenhuma suposição sobre a distribuição de frequências das variáveis. Ao contrário do coeficiente de correlação de Pearson, não tem a suposição de que a relação entre as variáveis seja linear; pode ser usado para as variáveis medidas no nível ordinal. Assim,

$$r_s = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}, \quad (50)$$

ou seja, o coeficiente de correlação de Spearman se utiliza da expressão do coeficiente de Pearson porém calculado com postos, ao invés dos dados originais (CAMPOS, 1979). A expressão (50) equivale à

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}, \quad (51)$$

sendo  $d_i = x_i - y_i$  a diferença de postos dos escores X e Y.

Para verificar a significância do valor observado de  $r_s$ , deve-se consultar a tabela P (SIEGEL, 1956) para  $n$  entre 4 e 30. Para  $n$  maior ou igual a 10, pode-se usar a distribuição  $t$  de Student

$$t = r_s \sqrt{\frac{n-2}{1-r_s^2}}, \quad (52)$$

em que  $t$  tem  $n - 2$  graus de liberdade sob a hipótese de inexistência de correlação (SIEGEL, 1956).

### 3 METODOLOGIA

Neste capítulo são descritos os procedimentos que foram adotados na realização deste trabalho, apresentados a estratégia de pesquisa, materiais e métodos.

#### 3.1 ESTRATÉGIA DA PESQUISA

Quanto a sua natureza, esta pesquisa se caracteriza como aplicada, pois tem como objetivo gerar um processo com finalidades imediatas, dirigido à solução de problemas específicos.

Quanto aos objetivos, tem-se uma pesquisa explicativa, com o objetivo de identificar fatores que determinam ou que contribuem para a ocorrência de fenômenos.

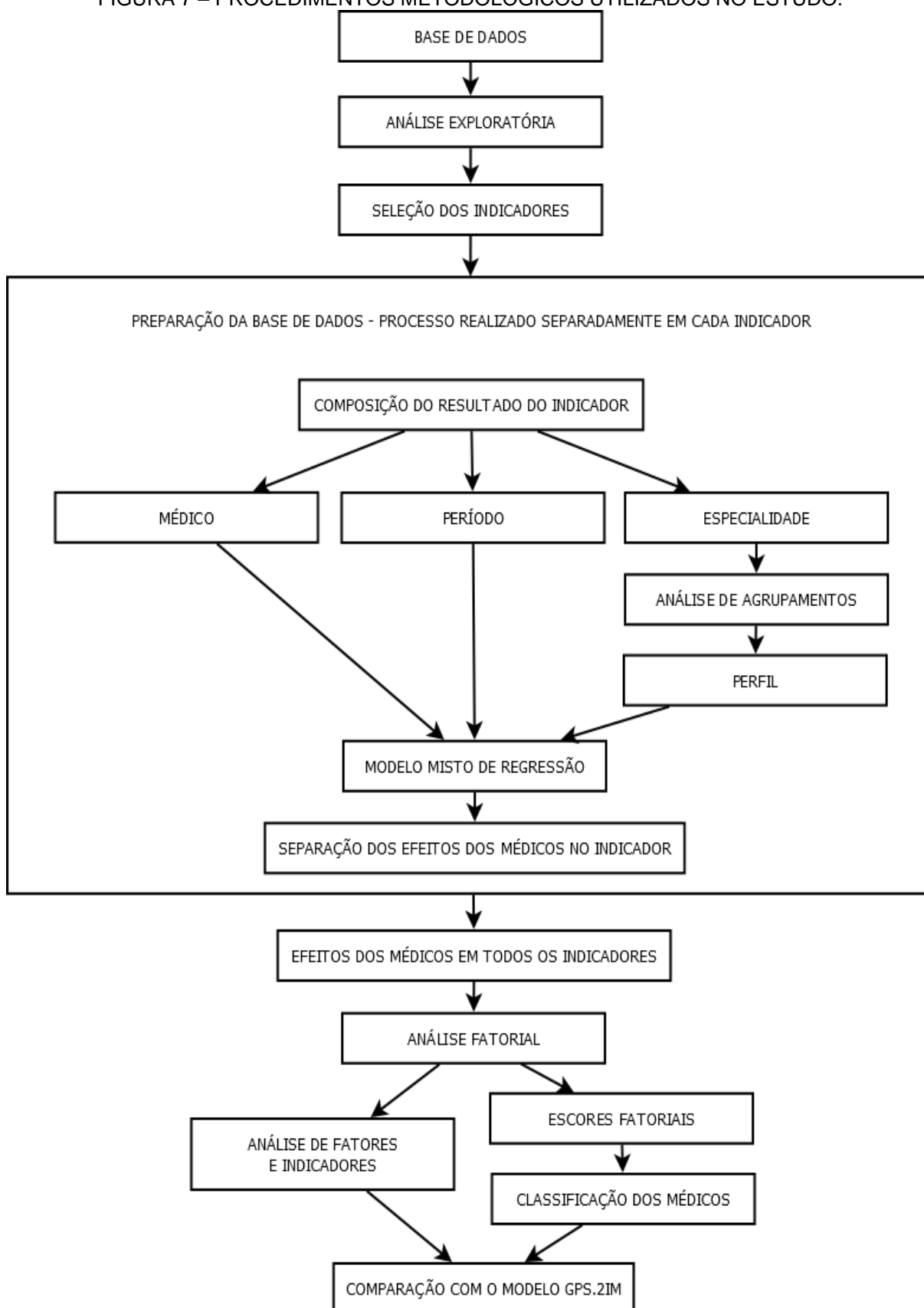
Quanto aos procedimentos, trata-se de um estudo de caso, que consiste no estudo profundo e exaustivo de poucos objetos, de maneira que permita seu amplo e detalhado conhecimento.

#### 3.2 MATERIAIS E MÉTODOS

O estudo de avaliação de desempenho foi realizado com a base de dados de um hospital, fornecida pela empresa 2iM, que faz o acompanhamento do desempenho dos prestadores de serviços de saúde em hospitais privados, públicos e universitários, operadoras de planos de saúde, secretarias de saúde, clínicas especializadas, entre outros. Essa base possui resultados mensais de vinte indicadores de desempenho entre os meses de janeiro e dezembro de 2015. Foram avaliados 150 médicos de 24 especialidades, conforme apresentado na Tabela 1. As informações foram coletadas mensalmente, de forma que cada médico apresenta doze resultados na base de dados.

O procedimento metodológico utilizado neste estudo está representado na Figura 7.

FIGURA 7 – PROCEDIMENTOS METODOLÓGICOS UTILIZADOS NO ESTUDO.



FONTE: O autor (2017).

TABELA 1 – MÉDICOS POR ESPECIALIDADE PRESENTES NA BASE DE DADOS.

CÓDIGO	ESPECIALIDADE	MÉDICOS
405	Angiologia e cirurgia vascular	4
505	Cancerologia clínica	8
506	Cancerologia cirúrgica	4
605	Cardiologia	6
705	Cirurgia cardiovascular	5
805	Cirurgia de cabeça e pescoço	5
905	Cirurgia geral	14
1105	Cirurgia pediátrica	7
1205	Cirurgia plástica	8
1305	Cirurgia torácica	4
1405	Clínica médica	4
1705	Endocrinologia	1
1805	Gastroenterologia	2
2105	Ginecologia e obstetrícia	9
2205	Hematologia e hemoterapia	3
2405	Infectologia	3
3605	Neurocirurgia	12
3705	Neurologia	2
3905	Oftalmologia	2
4005	Ortopedia e traumatologia	21
4105	Otorrinolaringologia	10
4505	Pneumologia	3
5005	Urologia	9
5205	Cirurgia buco-maxilo-facial	4
Total		150

FONTE: O autor (2017).

Primeiramente foi realizada a preparação da base de dados com dois objetivos: verificar a consistência da base de dados (removendo os indicadores com dados faltantes ou inconsistentes) e transformar os doze resultados de cada médico no ano em apenas um, necessário para a Análise Fatorial. Optou-se por aplicar um modelo linear misto em cada indicador, que comporta as medidas repetidas de cada médico como efeito aleatório do modelo. A vantagem de utilizar esse modelo é a possibilidade de isolar o efeito do médico dos demais efeitos fixos (período e especialidade). Esse procedimento está descrito na próxima seção.

Os efeitos dos médicos em cada indicador foram consolidados em uma base de dados multivariada, que foi analisada por meio da Análise Fatorial. Na análise foi utilizado o método das Componentes Principais, que possui a vantagem de não exigir a normalidade dos dados. Para a interpretação das cargas fatoriais foi utilizada a rotação ortogonal varimax. Com o ajuste do modelo de Análise Fatorial foi verificado se os domínios da qualidade do modelo GPS.2iM<sup>®</sup> correspondem aos fatores identificados pelo modelo fatorial. Por meio dessa análise, também foi possível

verificar a relevância de cada indicador de desempenho utilizado no hospital, considerando as comunalidades do modelo.

Por meio da Análise Fatorial foram obtidos os escores fatoriais, que são estimativas do valor de cada um dos fatores para uma observação individual. Os escores fatoriais foram estimados pelo método de Bartlett a partir da matriz de carregamentos fornecida pela Análise Fatorial.

Utilizou-se o coeficiente de correlação de Spearman entre índices de *performance* do modelo GPS.2iM© e os escores fatoriais propostos por este estudo para comparar os dois modelos

Nas análises, foram utilizados os *softwares* Statgraphics Centurion XVI (STATPOINT TECHNOLOGIES, 2009), Minitab 17 (MINITAB, INC, 2010) e R versão 3.3.1 (R CORE TEAM, 2017) com o auxílio das bibliotecas *nlme* (PINHEIRO et al., 2017) e *MASS* (VENABLES; RIPLEY, 2002).

### 3.3 PREPARAÇÃO DA BASE DE DADOS

Nesta seção é apresentado o processo que foi utilizado na seleção dos indicadores e os ajustes prévios que foram realizados na base de dados original.

#### 3.3.1 Seleção dos indicadores

O processo de seleção dos indicadores foi realizado por meio de uma análise exploratória de seus resultados. A Tabela 2 apresenta a média e o desvio padrão dos indicadores presentes na base de dados, além da quantidade de informações não disponíveis, representado pela sigla ND.

Como as informações são mensais, cada médico possui até 12 resultados, um para cada mês, totalizando 1800 informações por indicador. Os resultados dos indicadores 6070, 6102, 6114, 6115 e 6116 ainda não haviam sido informados pelo hospital, e por isso não puderam ser utilizados na análise de desempenho. Além desses resultados, é necessário verificar o comportamento dos indicadores ao longo do ano. O Gráfico 2 apresenta diagramas de dispersão separados por indicador. Cada linha do gráfico representa as observações de um médico ao longo do ano de 2015.

A escala dos valores foi normalizada para permitir a visualização de todos os resultados no mesmo gráfico.

TABELA 2 – ESTATÍSTICAS DESCRITIVAS DOS INDICADORES DE DESEMPENHO DO HOSPITAL EM 2015.

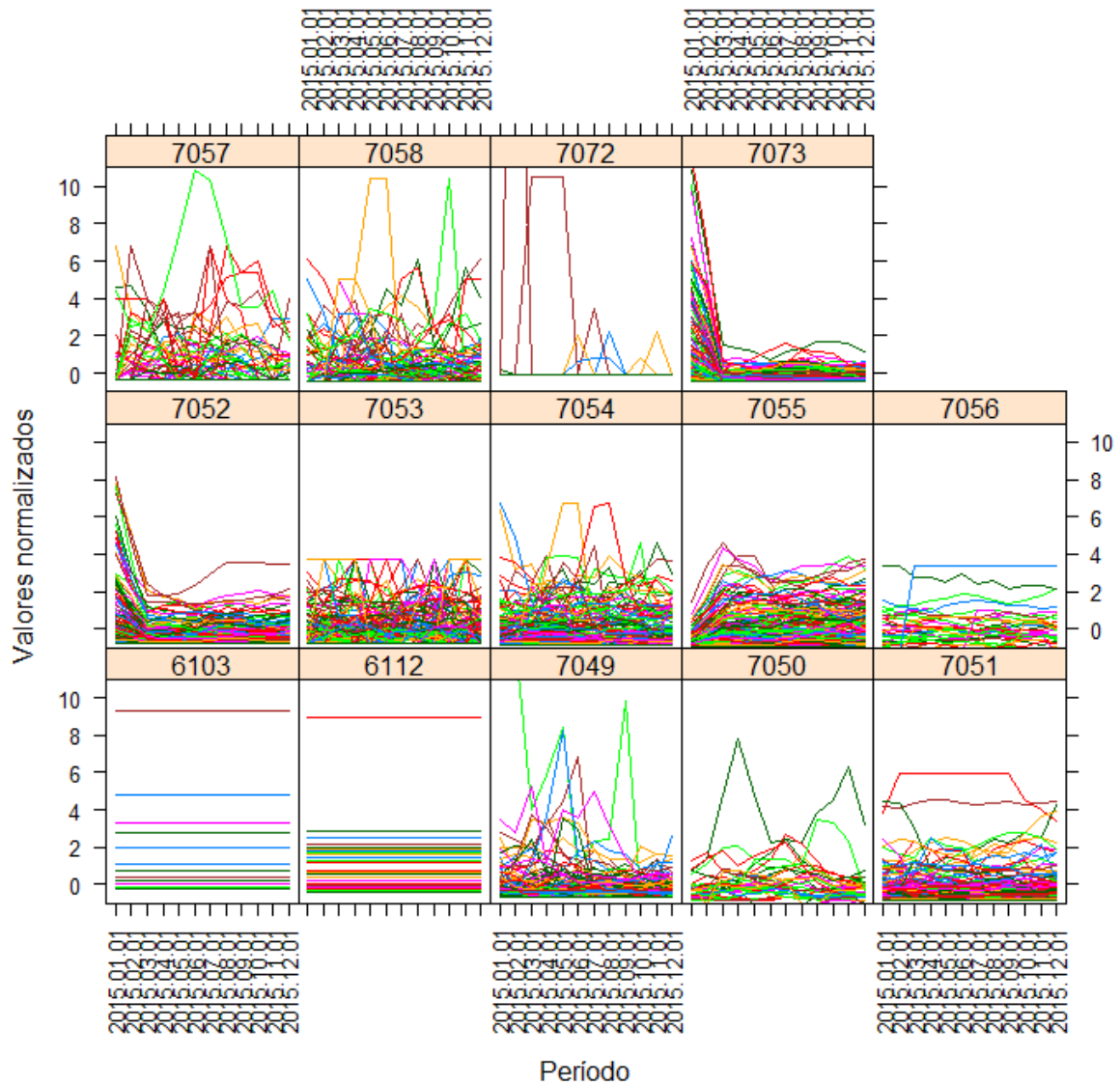
INDICADOR	N	ND	ND (%)
6070	1800	1800	100,00
6102	1368	1368	100,00
6103	1800	24	1,33
6112	1800	48	2,67
6114	1800	1800	100,00
6115	1800	1800	100,00
6116	1800	1800	100,00
7049	1368	0	-
7050	432	0	-
7051	1368	0	-
7052	1368	0	-
7053	1800	0	-
7054	1800	0	-
7055	1800	0	-
7056	432	0	-
7057	1800	0	-
7058	1800	0	-
7072	1368	0	-
7073	1800	0	-

FONTE: O autor (2017).

Verifica-se no Gráfico 2 que os indicadores 6103 e 6112 não apresentam variação ao longo do tempo, indicando que seus resultados foram obtidos apenas uma vez no período e copiados para os demais meses. Essa condição inviabiliza a utilização desses indicadores na Análise Fatorial, e por esse motivo não serão utilizados na avaliação de desempenho. O indicador 7072 apresentou uma grande quantidade de resultados zero, caracterizando a variável como um valor quase constante, o que inviabiliza sua utilização na Análise Fatorial.

Desta forma, somente os indicadores 7049, 7050, 7051, 7052, 7053, 7054, 7055, 7056, 7057, 7058 e 7073 foram utilizados na avaliação de desempenho.

GRÁFICO 2 – RESULTADOS MENSAIS DOS INDICADORES DE DESEMPENHO DO HOSPITAL EM 2015.



FONTE: O autor (2017).

### 3.3.2 Padronização da base de dados

A padronização da base de dados foi necessária nesse estudo para transformar as doze informações de cada médico na base de dados em apenas uma, algo necessário para realizar a Análise Fatorial. A opção mais simples que cumpre esse objetivo é utilizar a média anual dos indicadores, mas essa não é a melhor opção quando os dados apresentam sazonalidade, algo que é possível deduzir ao verificar o Gráfico 2, em que os indicadores 7052 e 7073 apresentaram uma clara alteração na

média ao longo do ano. Para verificar se essa alteração na média é significativa, foi realizada uma Análise de Variância (ANOVA) para medidas repetidas por meio de modelos lineares mistos, considerando o período e a especialidade como efeitos fixos e os médicos como efeitos aleatórios, considerando medidas repetidas (mais de um resultado por indivíduo), para cada indicador selecionado. Os resultados estão na Tabela 3 e a análise de resíduos correspondente se encontra nos apêndices.

TABELA 3 – COMPARAÇÃO DAS MÉDIAS MENSAIS DOS INDICADORES DE DESEMPENHO DO HOSPITAL EM 2015.

Período	7049	7050	7051	7052	7053	7054	7055	7056	7057	7058	7073
Janeiro	16414,46	29190,20	12,28	23,23	12,00	3,91	2,60	25,00	0,99	3,35	1414886,76
Fevereiro	15648,93	31407,07	12,28	14,71	15,87	4,09	5,09	25,04	1,59	3,19	832437,14
Março	15403,36	41452,40	12,28	6,64	16,61	4,44	7,86	28,30	1,66	4,16	169119,49
Abril	15430,21	44024,62	12,28	6,51	19,54	4,72	7,90	28,27	1,87	4,65	167982,63
Mai	18223,35	42196,11	12,28	6,43	17,76	4,87	7,96	27,95	1,59	5,11	180358,41
Junho	15268,12	41292,91	12,28	6,53	18,01	4,75	7,76	28,10	1,75	4,72	187337,49
Julho	13476,35	44960,20	12,28	7,30	18,90	4,52	8,04	29,19	2,23	3,93	208454,67
Agosto	12285,30	38956,69	12,28	7,82	19,62	4,48	8,32	28,76	2,10	3,91	205686,54
Setembro	12901,40	39674,39	12,28	7,62	20,74	4,30	8,21	27,97	1,75	3,20	198142,34
Outubro	10618,98	35986,14	12,28	7,42	18,20	4,38	8,07	28,71	1,44	3,97	180540,07
Novembro	11367,31	37287,87	12,28	7,36	17,34	4,23	7,91	28,69	1,13	3,70	173723,78
Dezembro	11167,65	30140,16	12,28	7,67	17,18	4,22	8,19	27,94	1,02	3,99	169700,93
P-valor	0,0003	0,0027	1,0000	< 0,0001	0,0024	0,2377	< 0,0001	0,5661	0,0048	0,30	< 0,0001

FONTE: O autor (2017).

Na ANOVA, os indicadores 7049, 7050, 7052, 7053, 7055 7057 e 7073 apresentaram p-valor menor que 0,01, sendo possível concluir que houve alteração significativa na média ao longo do ano ao nível de significância de 1%. O indicador 7051 apresentou a mesma média em todos os períodos, embora não seja uma variável constante, como visto no Gráfico 2. Desses resultados, conclui-se que o período do ano em que a avaliação de desempenho será feita deve ser considerado na classificação dos médicos.

Outra possível fonte de variação nos resultados é a especialidade médica. A mesma ANOVA para medidas repetidas da análise anterior foi utilizada para testar se essa hipótese é verdadeira, para cada indicador. Foram considerados o período e a especialidade como efeitos fixos e os médicos como efeitos aleatórios. Os resultados estão na Tabela 4, destacando que alguns indicadores não são avaliados em todas as especialidades, e por isso não apresentam resultados.

TABELA 4 – COMPARAÇÃO DAS MÉDIAS DOS INDICADORES DE DESEMPENHO DO HOSPITAL POR ESPECIALIDADE EM 2015.

Especialidade	7049	7050	7051	7052	7053	7054	7055	7056	7057	7058	7073
405	9913,62	-	25,00	16,34	16,23	3,66	6,51	-	1,39	2,57	256300,44
505	-	58807,59	-	-	14,88	11,01	6,22	12,58	10,32	14,73	688941,01
506	-	24770,93	-	-	6,44	4,32	0,83	29,19	0,00	8,14	69252,57
605	-	38118,08	-	-	12,08	9,86	10,39	16,84	5,34	10,52	767280,71
705	38915,98	-	20,00	1,62	8,00	4,71	0,67	-	2,02	5,00	133469,43
805	8735,16	-	20,00	1,96	10,89	2,00	1,14	-	0,70	0,00	46159,87
905	9366,83	-	7,14	17,22	31,26	3,43	15,30	-	0,60	2,88	422039,36
1105	3963,18	-	14,29	4,02	9,94	1,23	2,22	-	0,00	0,47	40908,22
1205	7671,24	-	12,50	5,36	24,11	1,05	2,90	-	0,00	0,00	66679,56
1305	12073,04	-	25,00	1,99	6,94	4,18	1,10	-	0,00	3,19	52848,12
1405	-	28982,66	-	-	11,73	9,54	11,26	25,00	6,49	9,41	718795,94
1705	-	20651,76	-	-	3,88	7,82	2,86	83,33	0,00	11,83	133204,83
1805	-	17452,00	-	-	9,95	6,38	21,79	50,00	0,87	3,46	883486,66
2105	7752,80	-	11,11	9,25	15,40	2,80	7,04	-	0,00	0,37	118492,20
2205		53277,25	-	-	10,50	11,86	10,93	33,33	6,05	14,07	1138944,99
2405		16564,77	-	-	14,00	7,64	15,69	33,33	0,64	6,41	514540,68
3605	37190,51	-	8,33	6,01	14,11	7,03	3,77	-	2,52	8,06	395900,38
3705		29928,76	-	-	9,98	8,00	22,67	50,00	1,76	7,62	1283828,57
3905	2646,16	-	50,00	1,24	0,00	0,18	0,96	-	0,00	0,00	10187,20
4005	15723,24	-	4,76	10,61	24,15	2,65	7,31	-	0,22	1,25	392659,33
4105	8145,64	-	10,00	9,39	30,64	0,95	7,13	-	0,00	0,00	105714,18
4505		43527,78	-	-	6,33	12,29	6,71	33,33	3,62	10,07	656180,64
5005	8257,46	-	11,11	16,90	18,87	1,73	12,57	-	0,00	0,40	240751,27
5205	11046,53	-	25,00	1,90	17,05	1,50	1,43	-	0,00	0,00	67485,99
P-valor	< 0,0001	0,4502	0,0001	0,0001	0,0009	< 0,0001	< 0,0001	0,0046	< 0,0001	< 0,0001	< 0,0001

FONTE: O autor (2017).

De acordo com os p-valores fornecidos pela ANOVA, há diferença significativa entre os valores médios dos indicadores nas especialidades ao nível de significância de 1%, com exceção do indicador 7050 que não apresentou diferença significativa. Esse resultado indica que os médicos poderiam ser avaliados nesse indicador em um único grupo, sem distinção da especialidade, algo que não é aplicável para os demais indicadores.

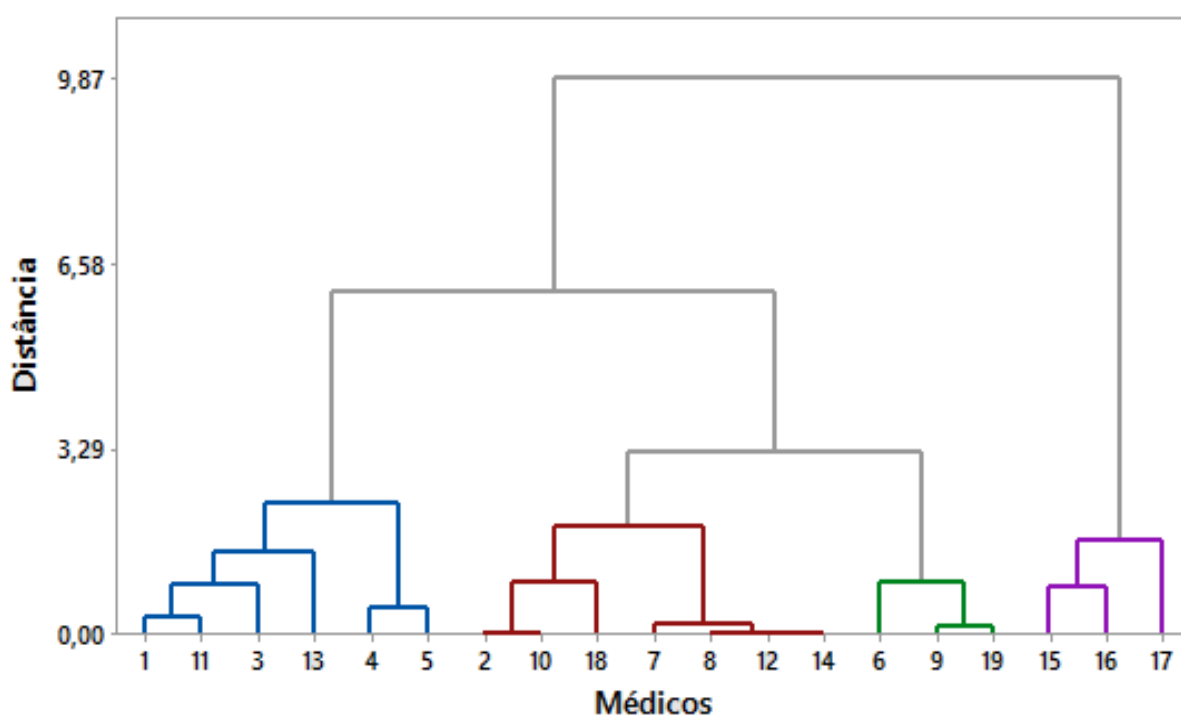
### 3.3.2.1 Análise de agrupamentos

O conteúdo dessa seção só é aplicável quando se tem o objetivo de criar subgrupos de indivíduos a serem avaliados separadamente. Caso os grupos já estejam previamente separados, deve-se desconsiderar a análise de agrupamentos na avaliação.

De acordo com a literatura, além da diferença entre especialidades observada na Tabela 4, pode haver diferença dentro de especialidades, ou seja, dois médicos de uma mesma especialidade poderiam não ser comparáveis devido às diferentes características dos pacientes atendidos por eles. Como a avaliação de desempenho é baseada no desfecho dos atendimentos, médicos que tendem a atender mais pacientes em piores condições de saúde teriam desvantagem nos resultados. Para garantir que a comparação dos desempenhos seja realizada apenas entre médicos com as mesmas características, foi realizada uma análise de agrupamentos com o objetivo de agrupar médicos com características semelhantes, configurando o que se chama na Medicina de Ajuste de Risco. As duas variáveis disponíveis para essa avaliação são a proporção de pacientes com idade superior a 53 anos atendidos pelos médicos e a proporção desses pacientes que também foram atendidos em outras especialidades. O método de agrupamento escolhido foi o de Ward.

O processo de Ajuste de Risco deve ser realizado em cada especialidade, gerando grande quantidade de resultados. Para exemplificar o processo, estão sendo apresentados apenas os resultados da especialidade 905 (cirurgia geral). No Gráfico 3, estão representados os quatro grupos criados pelo método de Ward.

GRÁFICO 3 – DENDROGRAMA PARA OS MÉDICOS DA ESPECIALIDADE 905 PELO MÉTODO DE WARD.



FONTE: O autor (2017).

O método permite a escolha da quantidade de agrupamentos a serem criados, sendo que a solução mais viável para essa especialidade contém quatro grupos. Para verificar se a Análise de Agrupamentos conseguiu cumprir seu objetivo, ou seja, se os grupos diferem significativamente, foi realizada uma Análise Multivariada de Variância (MANOVA), cujos resultados estão representados na Tabela 5.

TABELA 5 – MANOVA PARA OS GRUPOS GERADOS PELO MÉTODO DE WARD PARA A ESPECIALIDADE 905.

	GL	Wilks	F	GL num.	GL den.	P-valor
Grupos	3	0,035748	20	6	28	0,000006073
Resíduos	15					

FONTE: O autor (2017).

O baixo p-valor da MANOVA permite concluir que os grupos são significativamente diferentes. A partir desses resultados será possível classificar novos médicos que entrarem no programa de avaliação sem a necessidade de refazer os agrupamentos. Esse processo de classificação é realizado pela Análise Discriminante, em que o médico fornecerá os seus valores dentro das variáveis e o modelo indicará em qual dos grupos ele deve ser classificado.

O resultado mais interessante da Análise Discriminante para esse estudo é a possibilidade de verificar a eficiência das classificações. Para isso, o modelo utiliza os indivíduos da base de dados que já possuem agrupamento previamente definido e compara com as classificações que seriam feitas pela Análise Discriminante.

Quanto maior a proporção de acertos na classificação, mais bem definidos foram os agrupamentos criados. A proporção de acertos para a especialidade 905, realizada com o método de validação cruzada, está na Tabela 6.

TABELA 6 – ANÁLISE DISCRIMINANTE – MATRIZ DE CONFUSÃO COM VALIDAÇÃO CRUZADA PARA OS MÉDICOS DA ESPECIALIDADE 905.

Grupo Predito	Grupo Verdadeiro			
	1	2	3	4
1	26%	0%	0%	0%
2	5%	37%	0%	0%
3	0%	0%	16%	0%
4	0%	0%	0%	16%

FONTE: O autor (2017).

Os resultados da Tabela 6 demonstram que a Análise Discriminante obteve um poder de classificação de aproximadamente 95%, possibilitando concluir que a Análise de Agrupamentos foi eficiente na subdivisão dos médicos dessa especialidade em subgrupos.

### 3.3.2.2 Modelo linear misto

Com os resultados da Análise de Agrupamentos, a base de dados passou a possuir quatro variáveis explicativas para os resultados dos indicadores: período do ano, a especialidade, o agrupamento e o médico, sendo apenas este último o efeito de interesse nesta análise. O modelo linear misto foi utilizado em cada especialidade com o objetivo de separar esses efeitos, isolando o efeito do médico dos demais. Essa abordagem resolve o problema das medidas repetidas, incompatíveis com o modelo de Análise Fatorial utilizado, pois ao se retirar o efeito dos doze períodos, as informações de cada médico se transformam em apenas um resultado. Como os efeitos dos médicos saem padronizados do modelo misto, todos apresentaram a mesma média, conforme apresentado na Tabela 7.

TABELA 7 – EFEITOS PADRONIZADOS DOS INDICADORES DE DESEMPENHO DO HOSPITAL.

Indicador	Média	Desvio Padrão
7049	0,00	0,67
7050	0,00	0,84
7051	0,00	0,91
7052	0,00	0,72
7053	0,00	0,83
7054	0,00	0,84
7055	0,00	0,85
7056	0,00	0,78
7057	0,00	0,84
7058	0,00	0,81
7073	0,00	0,35

FONTE: O autor (2017).

A análise de resíduos correspondente está disponível nos apêndices, sendo a mesma utilizada na seção 3.3.2, uma vez que os resultados apresentados na Tabela 7 são um subproduto daquela análise.

Caso houvesse apenas um indicador, o modelo linear misto encerraria a análise, sendo o efeito dos médicos no indicador a classificação final, mas como são onze indicadores a serem avaliados, é necessária a utilização de um método multivariado para analisar os indicadores conjuntamente.

## 4 RESULTADOS

Neste capítulo é apresentada a análise da relevância dos indicadores, a classificação dos médicos e a comparação dessa classificação com a do modelo GPS.2iM©.

### 4.1 ANÁLISE FATORIAL

Conforme visto na Tabela 4, alguns indicadores não foram avaliados em todas as especialidades. A relação de indicadores avaliados em cada especialidade é apresentada no Quadro 5.

QUADRO 5 – INDICADORES AVALIADOS PELO MODELO GPS.2iM© NO HOSPITAL POR ESPECIALIDADE.

Especialidade	Indicador										
	7049	7050	7051	7052	7053	7054	7055	7056	7057	7058	7073
405	X		X	X	X	X	X		X	X	X
505		X			X	X	X	X	X	X	X
506		X			X	X	X	X	X	X	X
605		X			X	X	X	X	X	X	X
705	X		X	X	X	X	X		X	X	X
805	X		X	X	X	X	X		X	X	X
905	X		X	X	X	X	X		X	X	X
1105	X		X	X	X	X	X		X	X	X
1205	X		X	X	X	X	X		X	X	X
1305	X		X	X	X	X	X		X	X	X
1405		X			X	X	X	X	X	X	X
1705		X			X	X	X	X	X	X	X
1805		X			X	X	X	X	X	X	X
2105	X		X	X	X	X	X		X	X	X
2205		X			X	X	X	X	X	X	X
2405		X			X	X	X	X	X	X	X
3605	X		X	X	X	X	X		X	X	X
3705		X			X	X	X	X	X	X	X
3905	X		X	X	X	X	X		X	X	X
4005	X		X	X	X	X	X		X	X	X
4105	X		X	X	X	X	X		X	X	X
4505		X			X	X	X	X	X	X	X
5005	X		X	X	X	X	X		X	X	X
5205	X		X	X	X	X	X		X	X	X

FONTE: O autor (2017).

A conclusão que se chega ao analisar o Quadro 5 é a de ser possível separar os médicos em dois grupos distintos de especialidades, de forma que os indicadores avaliados em cada grupo sejam os mesmos, conforme apresentado na Tabela 8. Com isso, será necessário realizar duas avaliações de desempenho, e, conseqüentemente, duas análises fatoriais, sendo uma por grupo.

TABELA 8 – INDICADORES AVALIADOS PELO MODELO GPS.2IM© NO HOSPITAL POR GRUPO DE ESPECIALIDADES.

Grupo	Especialidades	Indicadores	Médicos
Grupo 1	405, 705, 805, 905, 1105, 1205, 1305, 2105, 3605, 3905, 4005, 4105, 5005, 5205.	7049, 7051, 7052, 7053, 7054, 7055, 7057, 7058, 7073.	114
Grupo 2	505, 506, 605, 1405, 1705, 1805, 2205, 2405, 3705, 4505.	7050, 7053, 7054, 7055, 7056, 7057, 7058, 7073.	36

FONTE: O autor (2017).

#### 4.1.1 Análise fatorial: grupo 1

A Figura 8 apresenta a matriz de correlação para o primeiro grupo de especialidades, necessária para os cálculos da Análise Fatorial. Verificam-se fortes correlações em dois agrupamentos de indicadores, sendo o primeiro grupo envolvendo os indicadores 7051, 7052, 7055 e 7073, e o segundo grupo com os indicadores 7054, 7057 e 7058. Os indicadores 7049 e 7073 apresentaram baixas correlações com os demais.

As maiores correlações verificadas foram de 0,90, entre os indicadores 7054 (média de permanência) e 7058 (taxa de pacientes de longa permanência) e 0,89, entre os indicadores 7052 (número de procedimentos cirúrgicos) e 7055 (número de internações). A correlação positiva indica uma forte relação direta de crescimento dos resultados dos médicos nesses indicadores.

FIGURA 8 – MATRIZ DE CORRELAÇÃO DOS INDICADORES PARA O PRIMEIRO GRUPO DE ESPECIALIDADES.



FONTE: O autor (2017).

Os resultados da Análise Fatorial estão na Tabela 9. A determinação da quantidade de fatores escolhidos foi de acordo com o percentual acumulado de variância explicada pelos fatores. Para facilitar a interpretação dos fatores, optou-se por rotacionar as cargas fatoriais. A rotação dos fatores permite obter na maioria dos casos uma estrutura para os pesos tal que cada variável tenha peso alto em um único fator e pesos baixos ou moderados nos demais fatores. Existem diversas rotações disponíveis na literatura, sendo mais utilizada a rotação varimax. Por se tratar de uma rotação ortogonal, os fatores oriundos de uma Análise Fatorial com rotação varimax não são correlacionados entre si.

TABELA 9 – MATRIZ DE CARREGAMENTOS PARA O GRUPO 1 – ROTAÇÃO VARIMAX.

	Fator 1	Fator 2	Fator 3	Fator 4	Comunalidades	Var. específica
Indicador 7052	<b>0,947</b>	0,136	-0,108	0,034	0,929	0,071
Indicador 7055	<b>0,909</b>	0,047	-0,072	0,075	0,838	0,162
Indicador 7051	<b>0,752</b>	0,028	0,003	0,147	0,588	0,412
Indicador 7073	<b>0,724</b>	-0,270	0,310	-0,272	0,767	0,233
Indicador 7054	0,048	<b>0,926</b>	-0,199	0,115	0,913	0,087
Indicador 7058	0,075	<b>0,925</b>	-0,010	0,111	0,874	0,126
Indicador 7057	-0,052	<b>0,843</b>	0,013	-0,054	0,717	0,283
Indicador 7049	-0,032	-0,089	<b>0,974</b>	0,063	0,962	0,038
Indicador 7053	0,094	0,083	0,060	<b>0,965</b>	0,950	0,050
Nº de Itens	4	3	1	1		
Autovalor	2,83	2,53	1,11	1,07	Bartlett	p < 0,0001
% Variância	31,5%	28,2%	12,3%	11,8%	KMO (MSA)	0,66
% Var. acumulada	31,5%	59,6%	71,9%	83,8%		

FONTE: O autor (2017).

O baixo p-valor do teste de Bartlett demonstra que as variáveis são correlacionadas na população e o valor de índice MSA indica que a Análise Fatorial é apropriada para esse conjunto de dados.

Foram escolhidos quatro fatores na Análise Fatorial, que conjuntamente explicam 83,8% da variabilidade dos dados. Dessa forma, é possível resumir os nove indicadores em quatro grupos, conseguindo explicar mais de 80% da informação original.

Como o objetivo da Análise Fatorial é facilitar a explicação das variáveis (indicadores), elas não serão mais analisadas de forma independente, mas em grupos (fatores, também chamados de domínios ou dimensões). Para determinar quais variáveis serão alocadas em cada fator deve-se verificar os valores das cargas fatoriais, que representam a correlação entre a variável e o fator. A rotação facilita esse processo, tornando mais fácil a alocação.

Os indicadores 7051, 7052, 7055 e 7073 apresentam alta carga fatorial com o Fator 1, e este fator explica 31,5% da variabilidade dos dados.

Os indicadores 7054, 7057 e 7058 possuem alta carga fatorial com o Fator 2, e este fator explica 28,2% da variabilidade dos dados.

O indicador 7049 apresenta alta carga fatorial com o Fator 3, e este explica 12,3% da variabilidade dos dados.

O indicador 7053 apresenta alta carga fatorial com o Fator 4, que explica 11,8% da variabilidade dos dados.

Os fatores podem ser nomeados de acordo com as características em comum das variáveis que os compõem. A nomeação dos fatores é um processo subjetivo, geralmente atribuído por um especialista na área em que os dados foram coletados. Dessa forma, comparando com os domínios fornecidos pela 2iM no Quadro 1, o Fator 1 pode ser mencionado como Eficiência e o Fator 2 como Efetividade. Apenas os indicadores 7049 e 7053 não foram classificados conforme as dimensões utilizadas na empresa.

#### 4.1.2 Análise fatorial: grupo 2

A Figura 9 apresenta as correlações entre os indicadores do segundo grupo de especialidades. Novamente, verificam-se dois agrupamentos na matriz de correlação, com um grupo envolvendo os indicadores 7050, 7055, 7056 e 7073 e o segundo grupo envolvendo os indicadores 7054, 7057 e 7058. Os indicadores 7054 (média de permanência) e 7058 (taxa de pacientes de longa permanência) mais uma vez apresentaram uma correlação muito alta, próxima do valor 1, sendo esse um forte indício de que esses dois indicadores fornecem a mesma resposta, estando apenas em escalas de medida diferentes.

Os resultados da Análise Fatorial para o grupo 2 estão na Tabela 10.

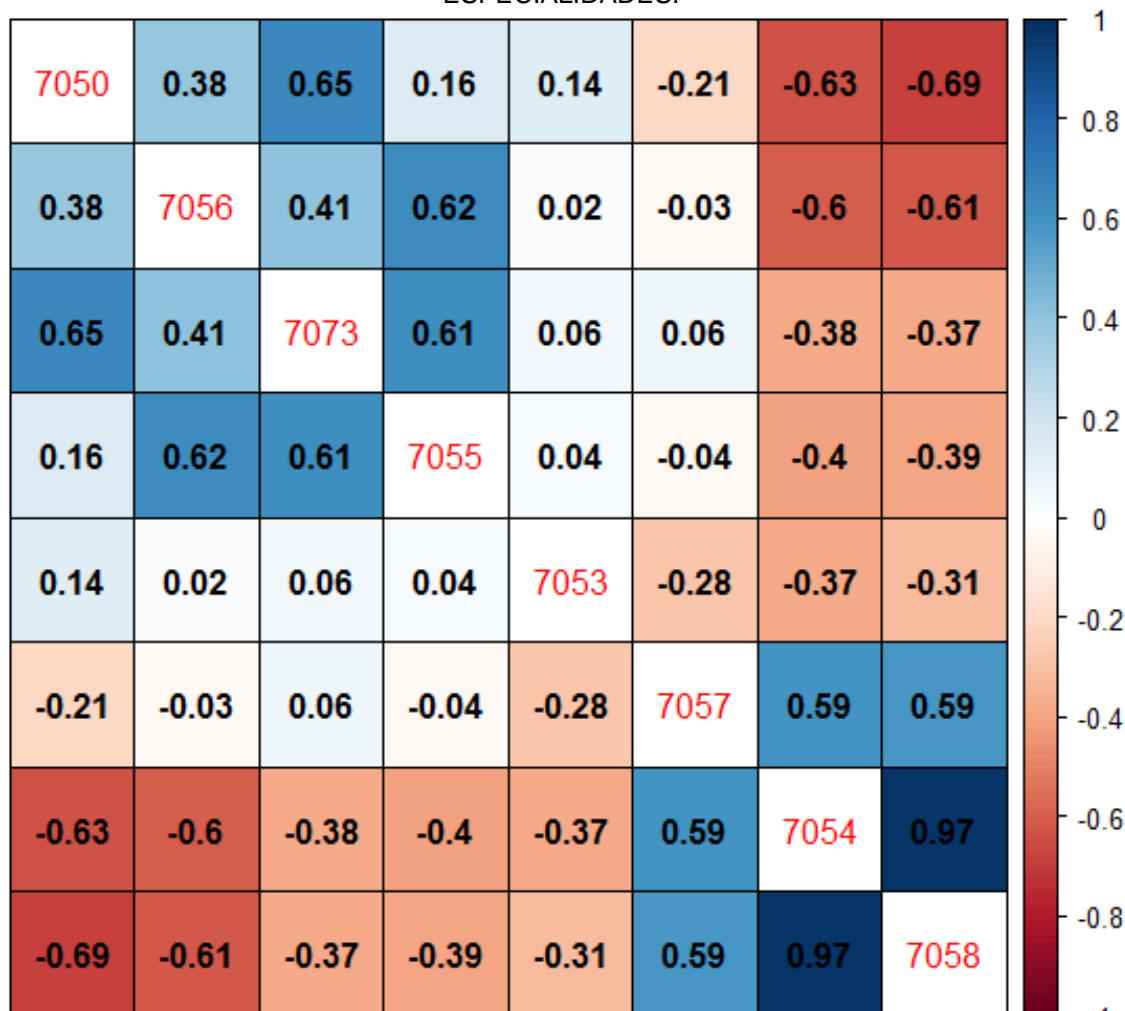
TABELA 10 – MATRIZ DE CARREGAMENTOS PARA O GRUPO 2 – ROTAÇÃO VARIMAX.

	Fator 1	Fator 2	Fator 3	Fator 4	Comunalidades	Var. específica
Indicador 7057	<b>0,877</b>	-0,088	-0,090	0,138	0,804	0,196
Indicador 7058	<b>0,807</b>	0,361	0,409	0,117	0,962	0,038
Indicador 7054	<b>0,795</b>	0,381	0,370	0,190	0,951	0,049
Indicador 7055	0,017	<b>0,927</b>	0,125	0,074	0,880	0,120
Indicador 7056	0,296	<b>0,803</b>	0,217	-0,109	0,791	0,209
Indicador 7050	0,368	0,049	<b>0,909</b>	0,000	0,963	0,037
Indicador 7073	-0,117	0,479	<b>0,789</b>	0,088	0,874	0,126
Indicador 7053	0,215	-0,005	0,049	<b>0,966</b>	0,981	0,019
Nº de Itens	3	2	2	1		
Autovalor	2,33	2,02	1,83	1,03	Bartlett	p < 0,0001
% Variância	29,2%	25,2%	22,8%	12,8%	KMO (MSA)	0,58
% Var. Acumulada	29,2%	54,4%	77,2%	90,1%		

FONTE: O autor (2017).

Os resultados do teste de Bartlett e do KMO indicam que a Análise Fatorial é também é apropriada para o grupo 2. Nesse grupo também foram escolhidos quatro fatores que explicam 90,1% da variabilidade. O Fator 1 desta vez envolve os indicadores de Efetividade, com 29,2% de explicação (contra 28,2% no grupo 1). Os indicadores de Eficiência se dividiram nos Fatores 2 e 3, separando os indicadores relacionados à aspectos financeiros dos demais. O indicador 7053 (média de permanência) ficou outra vez separado dos demais, com 12,8% de variância (contra 11,8% no grupo 1).

FIGURA 9 – MATRIZ DE CORRELAÇÃO DOS INDICADORES PARA O SEGUNDO GRUPO DE ESPECIALIDADES.



FONTE: O autor (2017).

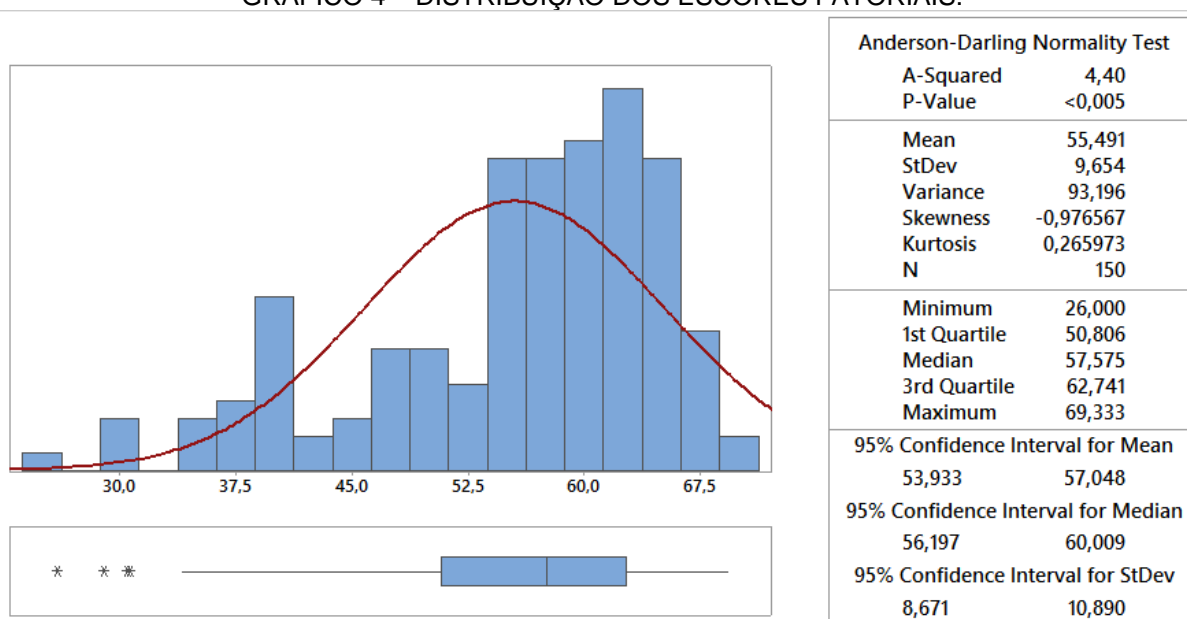
Com os resultados da Análise Fatorial, é possível utilizar as cargas fatoriais para compor escores dos indivíduos avaliados. Com esses escores é possível gerar a classificação dos avaliados, conforme visto na próxima seção.

## 4.2 CLASSIFICAÇÃO DOS AVALIADOS

Uma vez obtida a matriz de carregamentos, é possível gerar os escores fatoriais para todas as observações. Os escores fatoriais estimam o valor de cada um dos fatores para uma observação individual.

Cada avaliado possui quatro escores fatoriais, um em cada fator, visto que foram utilizados quatro fatores na Análise Fatorial. Para unificar os quatro escores em um único resultado foi realizada uma média ponderada, utilizando os autovalores dos fatores na ponderação. No final, os escores tiveram sua escala transformada com o objetivo de facilitar a interpretação. O Gráfico 4 apresenta a distribuição dos escores.

GRÁFICO 4 – DISTRIBUIÇÃO DOS ESCORES FATORIAIS.

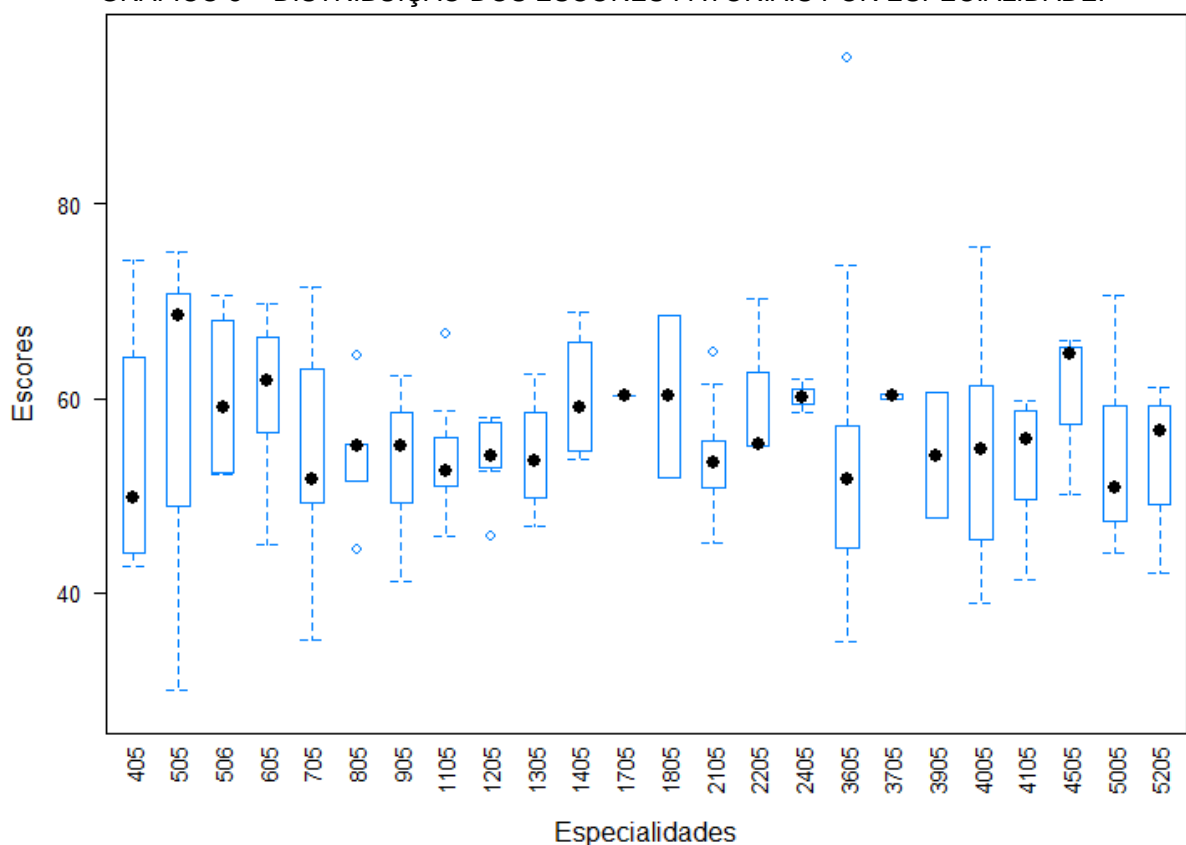


FONTE: O autor (2017).

Os escores fatoriais tiveram sua escala convertida com o objetivo de deixar os resultados em uma escala próxima ao do modelo GPS. Com isso, a média dos escores foi 55,5 e o desvio padrão 9,6. Os quartis demonstram que 50% dos escores estão distribuídos entre as pontuações 50,8 e 62,7. Essa distribuição dos escores se reflete no baixo valor do coeficiente de curtose. O coeficiente de assimetria demonstra que os escores apresentam assimetria à esquerda. O teste de normalidade de Anderson-Darling demonstrou que os escores não possuem distribuição normal.

O Gráfico 5 apresenta a distribuição dos escores por especialidade por meio de gráficos *boxplot*.

GRÁFICO 5 – DISTRIBUIÇÃO DOS ESCORES FATORIAIS POR ESPECIALIDADE.



FONTE: O autor (2017).

O Gráfico 5 mostra que a especialidade que apresentou o maior desempenho mediano no hospital foi a 505 (cancerologia clínica). O menor resultado mediano foi observado na especialidade 405 (angiologia e cirurgia vascular). Chama a atenção a grande variabilidade dos resultados entre os médicos na maioria das especialidades, representada pelo tamanho da caixa dos *boxplots*.

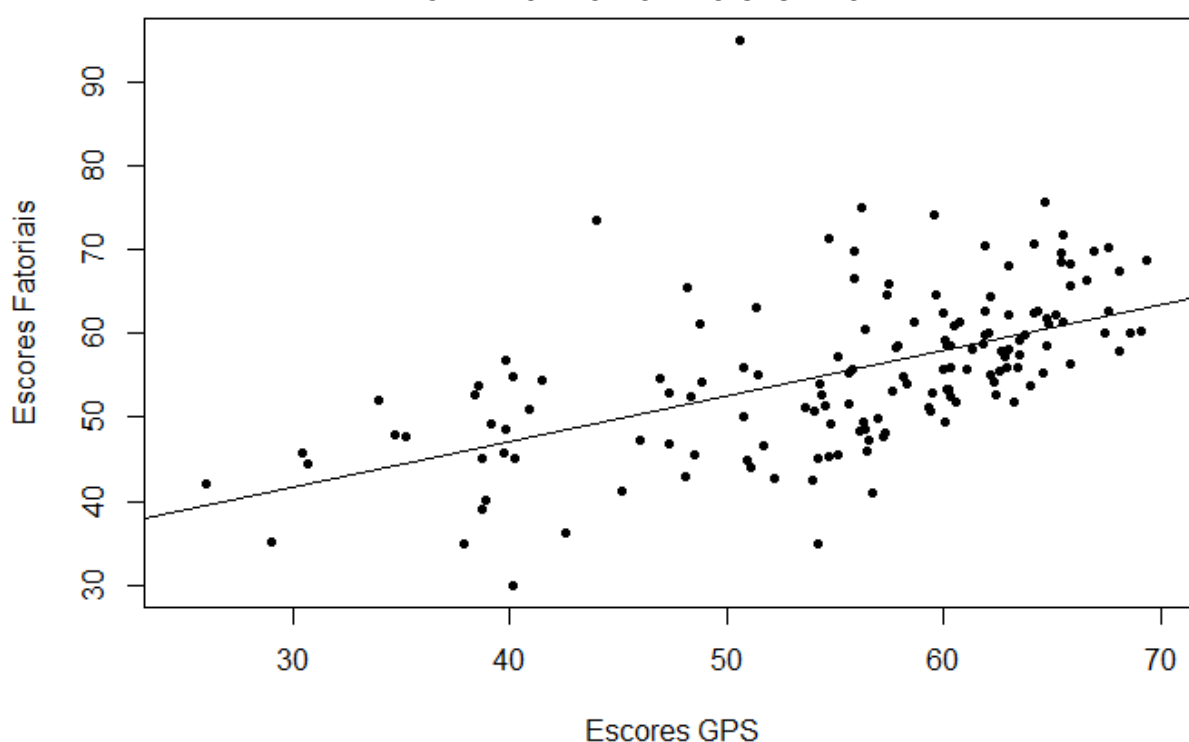
#### 4.3 COMPARAÇÃO DOS MODELOS

Nesta seção estão sendo comparados os escores fatoriais com os escores atribuídos pelo modelo GPS.2iM©. Para os resultados do modelo GPS foram desconsideradas as pontuações dos indicadores não utilizados nesta análise.

A principal diferença entre os dois modelos é que o GPS.2iM© apresenta doze resultados por médico no período avaliado, enquanto os escores fatoriais possui apenas um. Para tornar a comparação possível, foi necessário converter os escores do GPS.2iM© em uma média anual por médico.

O Gráfico 6 apresenta um diagrama de dispersão relacionando os escores dos médicos nos dois modelos. Foi adicionada uma reta de regressão ao gráfico que demonstra uma relação linear positiva entre os escores.

GRÁFICO 6 – DIAGRAMA DE DISPERSÃO ENTRE OS ESCORES FATORIAIS E O ÍNDICE DE PERFORMANCE DO MODELO GPS.2iM©



FONTE: O autor (2017).

Uma forma de comparação é analisar a correlação dos dois métodos de pontuação. Ao se utilizar o coeficiente de correlação de Spearman, o interesse está em comparar os *rankings* de médicos de cada método, e não a pontuação em si. A correlação entre os *rankings* foi de aproximadamente 0,64 ( $p < 0,0001$ ), indicando que os modelos classificam de forma semelhante em grande parte dos casos, embora em poucos casos haja grandes diferenças. Analisando esses casos na base de dados, é possível concluir que as diferenças se acentuam quando há médicos com valores muito extremos em alguns indicadores. O modelo GPS.2iM© possui um sistema

próprio que controla a influência de valores extremos nos indicadores, forçando um maior equilíbrio no resultado final.

## 5 CONSIDERAÇÕES FINAIS

No presente estudo, foi possível concluir que a Análise Fatorial é uma metodologia eficiente para ser utilizada na avaliação de desempenho. Por meio dos resultados é possível avaliar a qualidade dos indicadores de desempenho utilizados e também o desempenho dos indivíduos avaliados, através da geração de escores.

A primeira etapa do estudo exigiu a padronização da base de dados de forma que não houvesse medidas repetidas nos dados a serem utilizados na Análise Fatorial, convertendo as doze informações mensais dos avaliados em um valor anual em cada indicador. A opção por padronizar os resultados utilizando os efeitos aleatórios dos modelos lineares mistos se mostrou mais eficiente que usar a média anual dos avaliados devido aos efeitos de sazonalidade e especialidade presentes nos valores originais. Essa diferença se refletiu na Análise Fatorial, que apresentou melhores resultados quando comparados com os resultados obtidos utilizando somente as médias anuais. Apesar das vantagens, o problema de utilizar modelos lineares mistos é a sua complexidade, o que torna a utilização das médias uma opção mais atrativa na maioria dos casos.

Na Análise Fatorial verificou-se que os indicadores puderam ser melhor explicados em quatro fatores, que conjuntamente explicaram mais de 80% da variabilidade dos dados originais. Todos os indicadores pré-selecionados na análise exploratória se mostraram relevantes quando utilizados quatro fatores, algo que se altera quando a quantidade de fatores é reduzida.

A avaliação de desempenho foi realizada por meio dos escores fatoriais, gerados a partir da Análise Fatorial. Por meio dos resultados foi possível identificar os médicos e as especialidades que mais se destacaram no hospital no período avaliado, embora, de maneira geral, os resultados tenham se apresentado relativamente próximos.

Os escores fatoriais foram comparados com o índice de performance da empresa de forma gráfica e por meio do coeficiente de correlação de Spearman. Nessa análise foi possível verificar que os dois modelos classificam os médicos de forma semelhante na maioria dos casos. Ao se verificar a base de dados, foi possível constatar que as maiores diferenças entre os escores dos dois modelos se dá nos casos em que os médicos apresentam algum resultado muito extremo em pelo menos

um indicador. Essa diferença acontece porque o modelo GPS.2iM© é baseado em bandas no estilo da escala Likert que categoriza os valores dos indicadores, evitando pontuações muito discrepantes. Essa abordagem também pode ser utilizada na Análise Fatorial, se for de interesse que resultados extremos não exerçam grande influência no resultado do desempenho do avaliado, mas com a desvantagem de se obter resultados menos precisos devido à perda de informação na categorização das variáveis.

O presente estudo não abordou o tema da criação de indicadores de desempenho. A metodologia aplicada parte do pressuposto que eles já tenham sido previamente definidos pela empresa. Sendo assim, esse estudo limitou-se apenas a analisar os indicadores disponibilizados pela empresa 2iM, sem a intenção de propor novos indicadores.

Como sugestão para trabalhos futuros, propõe-se a aplicação de uma metodologia estatística na criação de indicadores de desempenho, analisando principalmente o poder de discriminação de cada indicador, que é a capacidade de distinguir os avaliados que têm a proficiência requisitada daqueles quem não a têm, e a consistência interna do conjunto de indicadores criados, identificando indicadores que estão em desacordo com os demais.

## REFERÊNCIAS

- ABICALAFFE, C. L. **Pagamento por performance: o desafio de avaliar o desempenho em saúde**. 1ª. ed. Rio de Janeiro: DOC Content, 2015.
- ABRAPP. **Guia para modelo de avaliação de desempenho**. 2013.
- ALEXANDRE, A. L. A importância da avaliação de desempenho nas organizações. **RH.com.br**, 2011.
- ARTUSO, A. R.; CHAVES NETO, A. Identificação de fatores relevantes e construção de portfólio diversificado no mercado acionário brasileiro. **Produção Online**, p. 1002-1030, 2012.
- CAMPOS, H. **Estatística experimental não-paramétrica**. 3ª. ed. Piracicaba: ESALQ, 1979.
- CHIAVENATO, I. **Gestão de pessoas**. 3ª. ed. Rio de Janeiro: Elsevier, 2010.
- COSTA, T. R. et al. Modelos lineares mistos: uma aplicação na curva de lactação de vacas da raça sindi. **Rev. Bras. Biom.**, São Paulo, v. 30, p. 23-49, 2012.
- CROMWELL, J. et al. **Pay for performance in health care: methods and approaches**. Research Triangle Park, NC: **RTI Press**, 2011.
- DONABEDIAN, A. The seven pillars of quality. **Arch Pathol Lab Med**, p. 1115-1118, 1990.
- EVERITT, B. S. et al. **Cluster analysis**. 5ª. ed. London: John Wiley & Sons, 2011.
- FOX, J.; WEISBERG, S. **An r companion to applied regression**. 2ª. ed. Sage, 2011.
- HAIR, J. F. et al. **Análise multivariada de dados**. 5ª. ed. São Paulo: Bookman, 2005.
- HEALTH & SOCIAL CARE INFORMATION CENTRE. **Quality and Outcomes Framework**. Londres. 2016.
- HICKS, C. R. **Fundamental concepts in the design of experiment**. 2ª. ed. New York: [s.n.], 1973.
- HILDEN-MINTON, J. A. **Multilevel diagnostics for mixed and hierarchical linear models**. Tese de Doutorado - University of California. Los Angeles. 1995.
- INSTITUTE OF MEDICINE. **Crossing the quality chasm: a new health system for the 21st century**. Washington DC. 2001.
- INSTITUTE OF MEDICINE. **Rewarding provider performance: aligning incentives in medicare**. Washington DC. 2007.

JIAN, W. et al. Performance evaluation of inpatient service in Beijing: a horizontal comparison with risk adjustment based on Diagnosis Related Groups. **BMC Health Services Research**, 2009.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 6<sup>a</sup>. ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2007.

KENNEY, C. **The best practice: how the new quality movement is transforming medicine**. Philadelphia. 2008.

LAIRD, N. M.; WARE, J. H. Random-effects models for longitudinal data. **Biometrics**, Oxford, v. 38, p. 963-974, 1982.

LIU, K.; JAIN, S.; SHI, J. Physician performance assessment using a composite quality index. **Statistics in medicine**, p. 2661-2680, 2012.

MINITAB, INC. **Minitab 17 Statistical Software**. State College, PA. 2010.

NOBRE, J. S. **Métodos de diagnóstico para modelos lineares mistos**. Dissertação (Mestrado em Estatística) - Universidade de São Paulo, Instituto de Matemática e Estatística. São Paulo. 2004.

PINHEIRO, J. C.; BATES, D. M. **Mixed effects models in s and s-plus**. New York: Springer, 2000.

PINHEIRO, J. et al. **Nlme: linear and nonlinear mixed effects models**. 2017.

R CORE TEAM. **R: a language and environment for statistical computing**. Viena. 2017.

RANGEL, A. **3 métodos para avaliar o desempenho**. Exame, 2015.

RENCHER, A. C. **Methods of multivariate analysis**. 2<sup>a</sup>. ed. New York: John Wiley & Sons, 2002.

SEBRAE. **Avaliação de desempenho como forma de reconhecimento ao funcionário**. 2015.

SES/SP. **Prêmio de produtividade médica**. São Paulo. 2017.

SIEGEL, S. **Nonparametric statistics for the behavioral sciences**. 1<sup>a</sup>. ed. New York: McGraw-Hill, 1956.

STATPOINT TECHNOLOGIES. **Statgraphics centurion, release 16.1.1**. 2009.

VAN HERCK, et al. Systematic review: effects, design choices, and context of pay-for-performance in health care. **BMC Health Services Research**, p. 10-247, 2010.

VENABLES, W. N.; RIPLEY, B. D. **Modern applied statistics with s**. 4<sup>a</sup>. ed. New York: Springer, 2002.

WATERNAUX, C.; LAIRD, N. M.; WARE, J. H. Methods for analysis of longitudinal data: blood-lead concentrations and cognitive development. **Journal of the American Statistical Association**, v. 84, p. 33-41, 1989.

## APÊNDICES

**APÊNDICE 1 – Comandos do R utilizados para a Análise de Agrupamentos**

```
d <- as.dist(dados)

cl <- hclust(d, method = "ward.D")
cof <- cophenetic(cl) # matriz cofenética
plot(cl)

# Ligações

tab <- data.frame(t(combn(nrow(dados),2)),dist=as.vector(cof))
tab <- tab[order(tab$dist),] # ordena a tabela
tab

### Tabela

tabela <- data.frame(as.vector(d), as.vector(cof))
names(tabela) <- c("Distâncias", "ward")
rownames(tabela) <- paste("[", apply(combn(nrow(dados),2), 2,
paste, collapse=" "), "]")
tabela
```

**APÊNDICE 2 – Comandos do R utilizados para a Análise Discriminante**

```
require(MASS)
analise <- lda(grupo ~ x1 + x2)
analise

### Classificação dos indivíduos

pred <- predict(analise)
pred$class

### Tabela de acertos

tabela <- table(grupo, pred$class)
tabela

### Proporção de acertos

sum(diag(tabela)) / sum(tabela)

### Classificação de um novo indivíduo

new <- data.frame(x1=1, x2=3) # registrando indivíduo
predict(analise, new)$class # classificando o indivíduo

### Gráfico

plot(analise)

### Manova

x <- cbind(x1, x2)
av <- manova(x ~ grupo)
summary(av, test="wilks")
```

```
### Método com validação cruzada

analise2 <- lda(grupo ~ x1 + x2, CV=TRUE)

### Classificação dos indivíduos}

analise2$class

### Tabela de acertos

tabela2 <- table(grupo, analise2$class)
tabela2

### Proporção de acertos

sum(diag(tabela2)) / sum(tabela2)
```

**APÊNDICE 3** – Comandos do R utilizados para o Modelo Linear Misto

```
require(nlme)
```

```
mod1 <- lme(val_indicador ~ periodo + perfil, random =  
~1|medico, data=dados)
```

```
mod2 <- update(mod1, correlation = corARMA(p=2, q=1)) #  
adiciona estrutura de correlação
```

```
summary(mod2) # efeitos fixos  
anova(mod2)
```

```
ranef(mod2, standard=TRUE) # efeitos aleatórios
```

**APÊNDICE 4 – Comandos do R utilizados para a Análise Fatorial**

```
### Matriz de correlação
# Necessária para calcular os autovalores e autovetores.

c <- cor(dados)
c

### Autovalores e autovetores normalizados

eigen(c)

### Componentes principais

# Entrando com a função de componentes principais. Definindo
# "cor=TRUE", o R irá calcular baseado nos autovalores e
# autovetores normalizados.

cp <- princomp(dados, cor=TRUE)
summary(cp) # saída com os desvios, ou seja, a raiz dos
autovalores

### Gráfico

screeplot(cp, type="lines")
abline(h=1) # coloca uma linha horizontal no gráfico de acordo
com o critério Kaiser.

### Matriz de carregamentos

# É necessário neste ponto informar a quantidade de fatores
escolhidos.

n <- 2 # quantidade de fatores escolhidos
L <- t(cp$sdev[1:n] * t(cp$loadings[,1:n])) # multiplica a
raiz dos autovalores pelos autovetores
```

L

```
### Matriz de carregamentos rotacionada
```

```
# Aplicando a rotação varimax a matriz de carregamentos.  
# Atenção para o fato de que os fatores podem aparecer fora de  
# ordem após a rotação. Verifique se os fatores estão  
# ordenados pela "Proportion Var".
```

```
Lr <- varimax(L)
```

```
Lr
```

```
### Comunalidades e variância específica
```

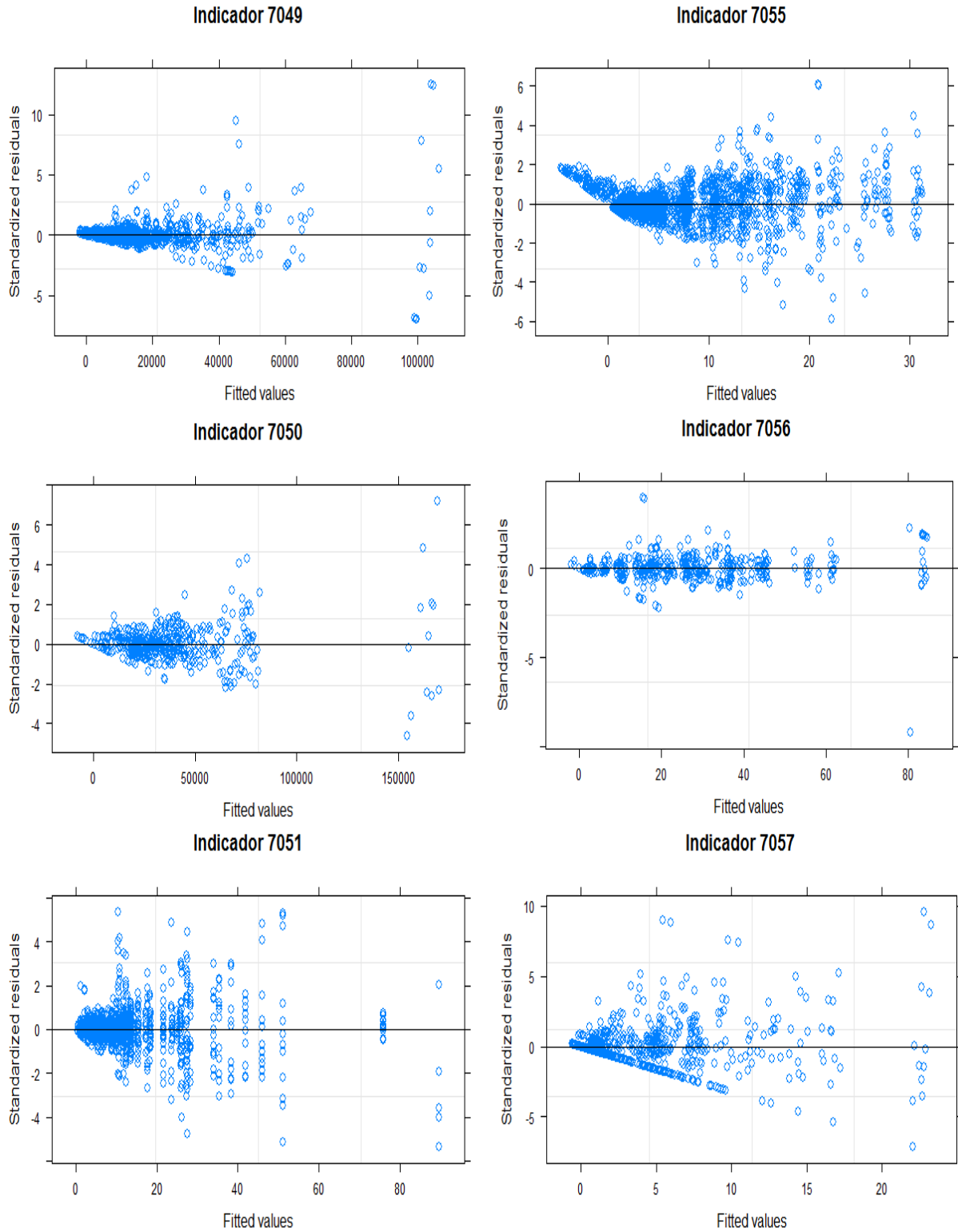
```
Comunalidades <- apply(Lr$loadings[,1:n]^2, 1, sum)
```

```
Var.especifica <- 1 - Comunalidades
```

```
cbind(Comunalidades, Var.especifica)
```

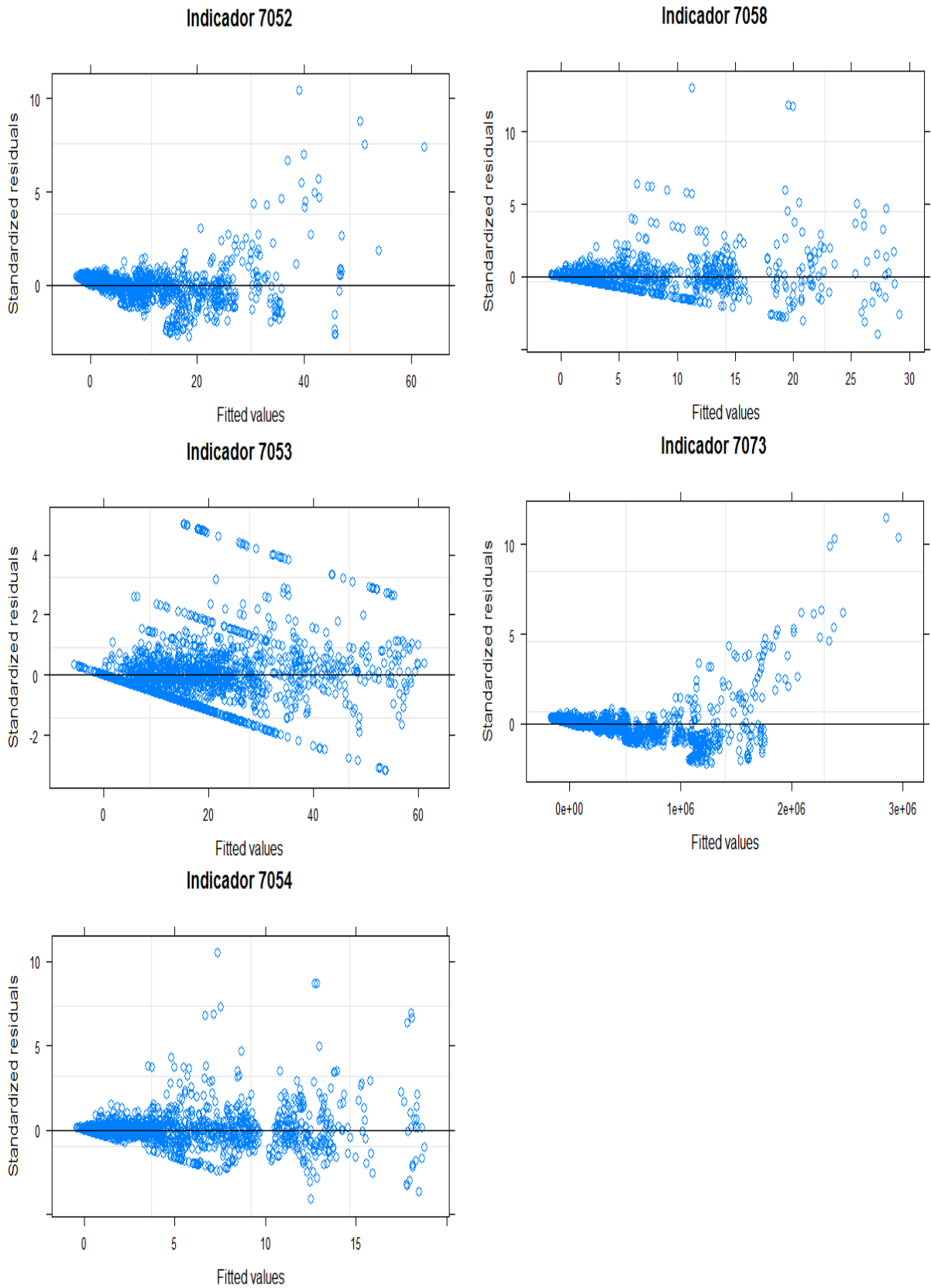
## APÊNDICE 5 – Resultados suplementares

GRÁFICO 7 – GRÁFICO DE RESÍDUOS VERSUS VALORES PREDITOS DO MODELO LINEAR MISTO.



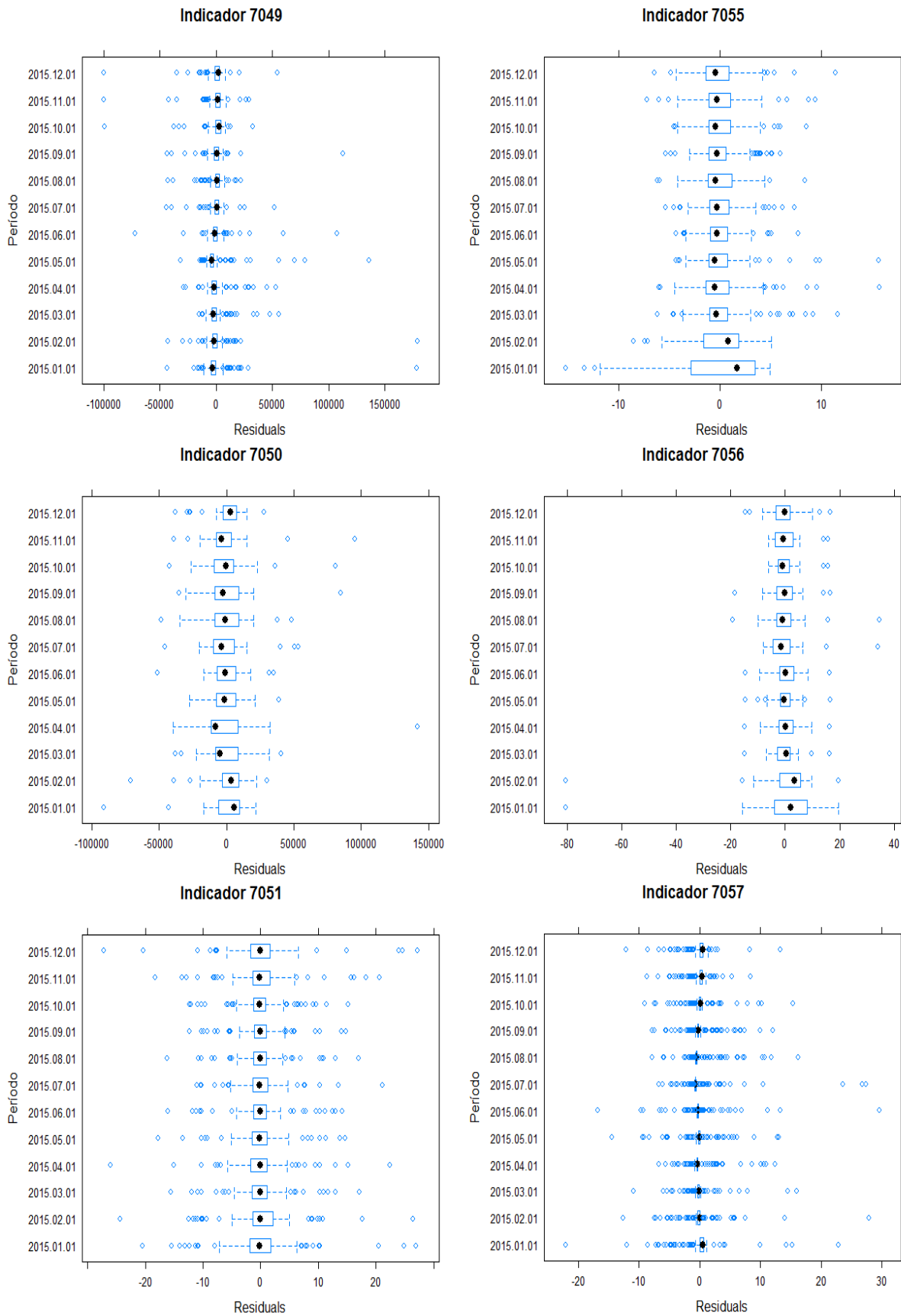
FONTE: O autor (2017).

GRÁFICO 8 – GRÁFICO DE RESÍDUOS VERSUS VALORES PREDITOS DO MODELO LINEAR MISTO – CONTINUAÇÃO.



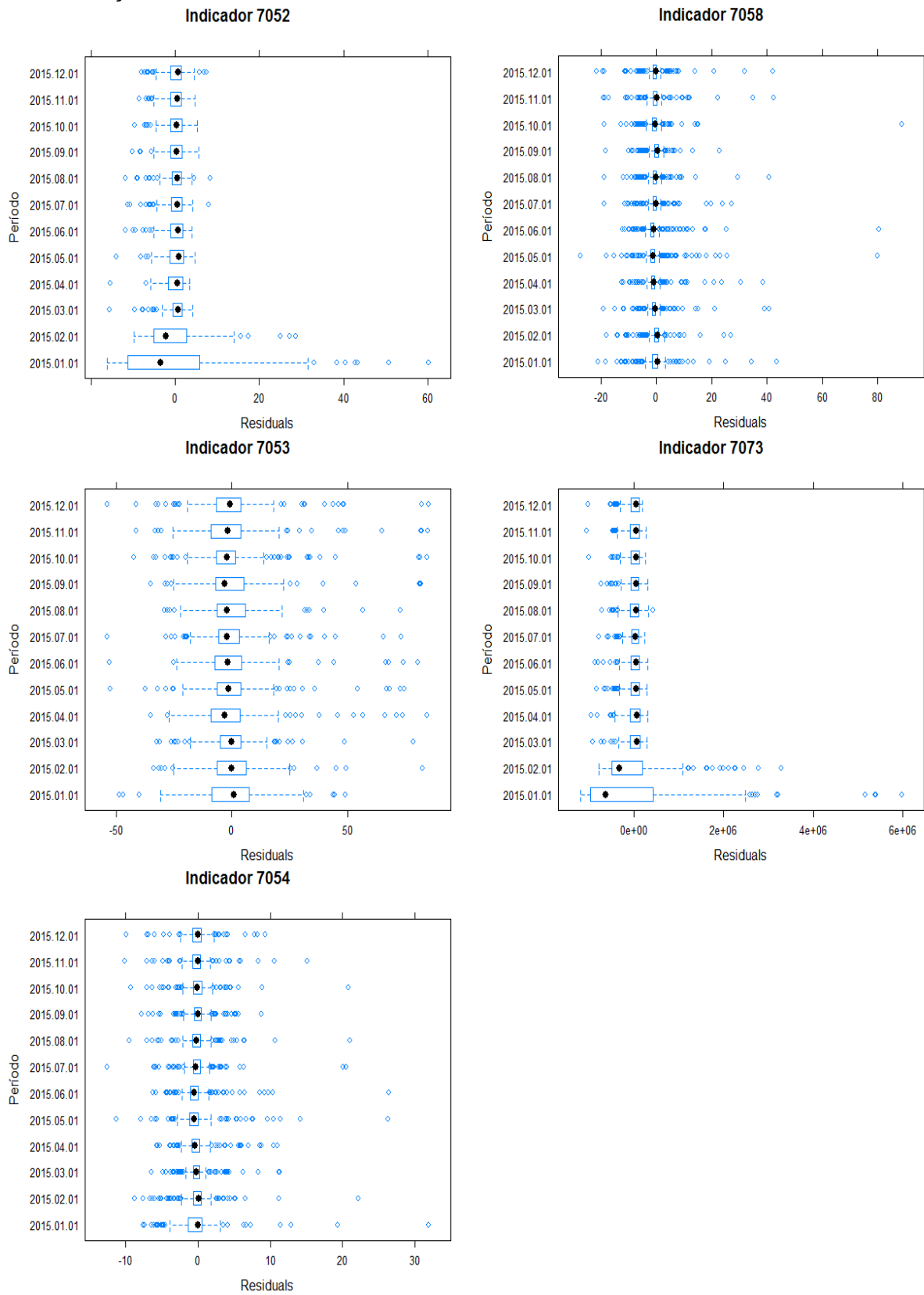
FONTE: O autor (2017).

GRÁFICO 9 – GRÁFICO DE RESÍDUOS VERSUS PERÍODO DO MODELO LINEAR MISTO.



FONTE: O autor (2017).

GRÁFICO 10 – GRÁFICO DE RESÍDUOS VERSUS PERÍODO DO MODELO LINEAR MISTO – CONTINUAÇÃO.



FONTE: O autor (2017).

**ANEXOS**

**ANEXO 1** – Tabela P de valores críticos de  $r_s$  do coeficiente de correlação de Spearman

TABELA 11 – TABELA P PARA O COEFICIENTE DE CORRELAÇÃO DE SPEARMAN.

N	Nível de significância (teste unilateral)	
	0,05	0,01
4	1,000	
5	0,900	1,000
6	0,829	0,943
7	0,714	0,893
8	0,643	0,833
9	0,600	0,783
10	0,564	0,746
12	0,506	0,712
14	0,456	0,645
16	0,425	0,601
18	0,399	0,564
20	0,377	0,534
22	0,359	0,508
24	0,343	0,485
26	0,329	0,465
28	0,317	0,448
30	0,306	0,432

FONTE: O autor (2017).