

**UNIVERSIDADE FEDERAL DE ALFENAS  
UNIFAL-MG**

**LAÍS BRAMBILLA STORTI FERREIRA**

**TESTES ASSINTÓTICOS PARA DETECTAR CONSENSO  
MULTIVARIADO EM PAINÉIS SENSORIAIS**

**ALFENAS/MG  
2017**

**LAÍS BRAMBILLA STORTI FERREIRA**

**TESTES ASSINTÓTICOS PARA DETECTAR CONSENSO MULTIVARIADO EM  
PAINÉIS SENSORIAIS**

Dissertação apresentada à Universidade Federal de Alfenas, como parte dos requisitos para obtenção do título de Mestre em Estatística Aplicada e Biometria.

Área de concentração: estatística aplicada e biometria.

Orientador: Prof<sup>o</sup>. Dr. Eric Batista Ferreira

Coorientadora: Prof<sup>a</sup>. Dra. Adriana Dias

**ALFENAS/MG**

**2017**

Dados Internacionais de Catalogação-na-Publicação (CIP)  
Biblioteca Central da Universidade Federal de Alfenas

Ferreira, Laís Brambilla Storti.

Testes assintóticos para detectar consenso multivariado em painéis sensoriais. / Laís Brambilla Storti Ferreira. -- Alfenas/MG, 2017.  
76 f.

Orientador: Eric Batista Ferreira.

Dissertação (Mestrado em Estatística Aplicada e Biometria) -  
Universidade Federal de Alfenas, 2017.

Bibliografia.

1. Análise multivariada. 2. Avaliação sensorial. 3. Análise de componentes principais. I. Ferreira, Eric Batista. II. Título.

CDD-519.535



MINISTÉRIO DA EDUCAÇÃO  
Universidade Federal de Alfenas / UNIFAL-MG  
Programa de Pós-graduação em Estatística Aplicada e Biometria

Rua Gabriel Monteiro da Silva, 700. Alfenas - MG CEP 37130-000  
Fone: (35) 3299-1392 (Secretaria) (35) 3299-1121 (Coordenação)  
<https://www.unifal-mg.edu.br/ppgeab/>



LAÍS BRAMBILLA STORTI FERREIRA

“TESTES ASSINTÓTICOS PARA DETECTAR CONSENSO MULTIVARIADO EM  
PAINÉIS SENSORIAIS”

A Banca Examinadora, abaixo assinada, aprova a  
Dissertação apresentada como parte dos requisitos para  
a obtenção do título de Mestre em Estatística Aplicada  
e Biometria pela Universidade Federal de Alfenas.  
Área de Concentração: Estatística Aplicada e  
Biometria

Aprovado em: 16 de dezembro de 2016.

Prof. Dr. Eric Batista Ferreira  
Instituição: UNIFAL-MG

Assinatura: \_\_\_\_\_

Prof. Dr. Davi Butturi Gomes  
Instituição: UNIFAL-MG

Assinatura: \_\_\_\_\_

Prof. Dr. Daniel Furtado Ferreira  
Instituição: UFLA

Assinatura: \_\_\_\_\_

Dedico a minha família pelo apoio incondicional ao longo desta caminhada.

## AGRADECIMENTOS

Agradeço, primeiramente, a Deus, pela força concedida para realização deste trabalho.

À toda minha família, por sempre acreditarem no meu potencial. Em especial aos meus pais, Claudinei e Cláudia, por acreditarem e apoiarem todos os meus sonhos, a minha irmã e afilhado, Larissa e Vitor, por serem fonte de inspiração no decorrer desta caminhada e ao meu marido, Rafael, por sempre estar ao meu lado e ser meu ponto de apoio nos momentos difíceis.

Aos amigos que me apoiaram com palavras de conforto e incentivo durante todo este processo.

Ao meu orientador, professor Eric Batista Ferreira, pela dedicação, competência e paciência durante todo este processo e, principalmente, pela troca de conhecimento que foi fundamental para meu crescimento acadêmico. Agradeço também à minha coorientadora, professora Adriana Dias, pela dedicação e grande colaboração no decorrer deste trabalho.

Aos membros da banca dos exames de qualificação e defesa de mestrado, Prof. Dr. Denismar Alves Nogueira e Prof<sup>a</sup>. Dra. Roberta Bessa Veloso Silva, Prof. Dr. Daniel Furtado Ferreira e Prof. Dr. Davi Butturi Gomes, pelas contribuições e orientações.

À Universidade Federal de Alfenas e ao Programa de Pós-Graduação em Estatística Aplicada e Biometria, juntamente com seus docentes e servidores, por todas as oportunidades de aprendizado oferecidas.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), pela concessão da bolsa de estudos.

Por fim, agradeço a todos que contribuíram, direta ou indiretamente, para realização deste trabalho.

*“It always seems impossible until  
it’s done”*

*Nelson Mandela*

## RESUMO

A unidimensionalidade de um painel sensorial está diretamente relacionada com a consonância do mesmo, ou seja, um painel é considerado unidimensional quando os provadores pontuam da mesma forma um determinado atributo. Devido a importância da concordância do painel para a confiabilidade da análise sensorial os provadores devem ser treinados de modo que concordem entre si em relação às características deste atributo. Na literatura é possível encontrar métodos que vem sendo propostos para avaliar esta concordância, porém os métodos existentes avaliam as notas dadas pelos provadores para um atributo de cada vez, tornando a análise mais lenta. Assim, o objetivo deste trabalho é generalizar o teste assintótico de autovalores proposto por Ferreira (2008a), a fim de inferir sobre o consenso multivariado de painéis sensoriais. A partir da generalização do teste assintótico de autovalores foi possível obter quatro novas estatísticas de teste. A avaliação dos testes foi realizada via simulação Monte Carlo, na qual foram avaliados diferentes cenários resultantes da combinação dos números de provadores (2, 5, 10 e 15), atributos (2, 5, 10 e 20), observações (10, 20, 30, 40, 50, 100 e 200), grau de treinamento do painel sensorial ( $0,1 \geq \rho^2 \geq 0,99$ ) e da restrição  $n \geq pq$ . De maneira geral, analisando a taxa de erro tipo I e o poder dos testes, o teste  $InvH_2$  mostrou-se ser o mais eficiente.

Palavras-chave: Consonância de painel. Sensometria. Componentes principais.

## ABSTRACT

The unidimensionality of a sensory panel is directly related to the panel consonance, i.e., a panel is considered unidimensional when assessors score in the same way a particular attribute. Due to the importance of the panel agreement to the reliability of sensory analysis panelists should be trained in order that they agree with each other regarding the characteristics of an attribute. In the literature several methods have been proposed to assess such agreement, although existing methods evaluate the marks for one attribute at a time, making the analysis slower. The objective of this study is to generalize the asymptotic test eigenvalues proposed by Ferreira (2008a), in order to infer about the multivariate consensus of sensory panels. From the generalization of the asymptotic test of eigenvalues it was possible to obtain four new test statistics. The evaluation of the tests was conducted via Monte Carlo simulation, in which were evaluated different scenarios resulting from the combination of the numbers of panelists (2, 5, 10 and 15), attributes (2, 5, 10 and 20), observations (10, 20, 30, 40, 50, 100 and 200), degree of training of the sensory panel ( $0,1 \geq \rho^2 \geq 0,99$ ) and the restriction  $n \geq pq$ . Overall, analyzing the type I error and the power function of the tests, the test  $InvH_2$  was more efficient.

Key words: Panel consonance. Sensometrics. Principal components.

## LISTA DE FIGURAS

Figura 1 –	Representação da hipermatriz $\mathbf{X}$ e sua respectiva fatia para o atributo $j$ . Fonte: adaptado de Dijksterhuis (1995). . . . .	14
Figura 2 –	Exemplo de um <i>scree plot</i> . Fonte: da autora. . . . .	23
Figura 3 –	Matriz desdobrada da qual são obtidas as observações via simulação. Fonte: da autora. . . . .	30
Figura 4 –	Esquema para proposição dos testes. Fonte: da autora. . . . .	35
Figura 5 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 2$ provadores, $q = 2$ atributos e $n = 30$ observações. Fonte: da autora. . . . .	40
Figura 6 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 2$ provadores, $q = 2$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	41
Figura 7 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 5$ provadores, $q = 2$ atributos e $n = 20$ observações. Fonte: da autora. . . . .	42
Figura 8 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 5$ provadores, $q = 2$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	42
Figura 9 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 5$ provadores, $q = 20$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	43
Figura 10 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 10$ provadores, $q = 2$ atributos e $n = 20$ observações. Fonte: da autora. . . . .	44
Figura 11 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 10$ provadores, $q = 2$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	45
Figura 12 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 15$ provadores, $q = 2$ atributos e $n = 30$ observações. Fonte: da autora. . . . .	46
Figura 13 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 8$ provadores, $q = 3$ atributos e $n = 30$ observações. Fonte: da autora. . . . .	47
Figura 14 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 2$ provadores, $q = 2$ atributos e $n = 20$ observações. Fonte: da autora. . . . .	56
Figura 15 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 2$ provadores, $q = 2$ atributos e $n = 40$ observações. Fonte: da autora. . . . .	56
Figura 16 –	Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 2$ provadores, $q = 2$ atributos e $n = 50$ observações. Fonte: da autora. . . . .	57

Figura 17 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 2$ prova- dores, $q = 2$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	57
Figura 18 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 2$ prova- dores, $q = 5$ atributos e $n = 50$ observações. Fonte: da autora. . . . .	58
Figura 19 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 2$ prova- dores, $q = 5$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	58
Figura 20 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 2$ prova- dores, $q = 5$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	59
Figura 21 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 2$ prova- dores, $q = 10$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	59
Figura 22 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 2$ prova- dores, $q = 10$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	60
Figura 23 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 2$ prova- dores, $q = 20$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	60
Figura 24 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ prova- dores, $q = 2$ atributos e $n = 10$ observações. Fonte: da autora. . . . .	61
Figura 25 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ prova- dores, $q = 2$ atributos e $n = 30$ observações. Fonte: da autora. . . . .	61
Figura 26 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ prova- dores, $q = 2$ atributos e $n = 40$ observações. Fonte: da autora. . . . .	62
Figura 27 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ prova- dores, $q = 2$ atributos e $n = 50$ observações. Fonte: da autora. . . . .	62
Figura 28 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ prova- dores, $q = 2$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	63
Figura 29 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ prova- dores, $q = 5$ atributos e $n = 30$ observações. Fonte: da autora. . . . .	63
Figura 30 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ prova- dores, $q = 4$ atributos e $n = 40$ observações. Fonte: da autora. . . . .	64
Figura 31 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ prova- dores, $q = 5$ atributos e $n = 50$ observações. Fonte: da autora. . . . .	64
Figura 32 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ prova- dores, $q = 5$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	65

Figura 33 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ provedores, $q = 5$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	65
Figura 34 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ provedores, $q = 10$ atributos e $n = 50$ observações. Fonte: da autora. . . . .	66
Figura 35 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ provedores, $q = 10$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	66
Figura 36 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ provedores, $q = 10$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	67
Figura 37 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 5$ provedores, $q = 20$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	67
Figura 38 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 2$ atributos e $n = 30$ observações. Fonte: da autora. . . . .	68
Figura 39 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 2$ atributos e $n = 40$ observações. Fonte: da autora. . . . .	68
Figura 40 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 2$ atributos e $n = 50$ observações. Fonte: da autora. . . . .	69
Figura 41 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 2$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	69
Figura 42 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 5$ atributos e $n = 50$ observações. Fonte: da autora. . . . .	70
Figura 43 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 5$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	70
Figura 44 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 5$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	71
Figura 45 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 10$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	71
Figura 46 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 10$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	72
Figura 47 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 10$ provedores, $q = 20$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	72
Figura 48 – Poder dos testes $IAH_1, IAH_2, InvH_1$ e $InvH_2$ para o caso $p = 15$ provedores, $q = 2$ atributos e $n = 40$ observações. Fonte: da autora. . . . .	73

Figura 49 – Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 15$ providores, $q = 2$ atributos e $n = 50$ observações. Fonte: da autora. . . . .	73
Figura 50 – Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 15$ providores, $q = 2$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	74
Figura 51 – Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 15$ providores, $q = 2$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	74
Figura 52 – Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 15$ providores, $q = 5$ atributos e $n = 100$ observações. Fonte: da autora. . . . .	75
Figura 53 – Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 15$ providores, $q = 5$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	75
Figura 54 – Poder dos testes $IAH_1$ , $IAH_2$ , $InvH_1$ e $InvH_2$ para o caso $p = 15$ providores, $q = 10$ atributos e $n = 200$ observações. Fonte: da autora. . . . .	76

## LISTA DE TABELAS

Tabela 1 – Valores de $p$ provadores, $q$ atributos e $n$ observações para os quais a porcentagem de rejeição dos testes é zero. . . . .	39
--	----

## SUMÁRIO

1	<b>INTRODUÇÃO</b>	10
2	<b>REFERENCIAL TEÓRICO</b>	12
2.1	ANÁLISE SENSORIAL E SENSOMETRIA	12
2.2	ANÁLISE DE COMPONENTES PRINCIPAIS	18
2.2.1	Retenção de componentes principais	22
2.3	TAXA DE ERRO TIPO I E PODER DE UM TESTE	24
2.4	SIMULAÇÃO MONTE CARLO	25
3	<b>MATERIAL E MÉTODOS</b>	26
3.1	BASE TEÓRICA PARA A PROPOSIÇÃO DOS TESTES	26
3.2	ESTUDO DE SIMULAÇÃO DE DADOS	29
3.3	EXPERIMENTOS COM DADOS REAIS	32
4	<b>RESULTADOS E DISCUSSÃO</b>	34
4.1	RESULTADOS METODOLÓGICOS	34
4.2	ESTUDO DE SIMULAÇÃO	38
4.3	EXPERIMENTO COM DADOS REAIS	47
5	<b>CONCLUSÃO</b>	49
	<b>REFERÊNCIAS</b>	50
	<b>APÊNDICES</b>	54

## 1 INTRODUÇÃO

Há registros de que a análise sensorial surgiu na Europa, onde um grupo de pessoas se reunia para degustação em cervejarias e destilarias. Nos Estados Unidos, a análise sensorial surgiu com a necessidade de produzir alimentos de qualidade para o exército americano durante a segunda guerra mundial. No Brasil, a prática de degustação com o objetivo de analisar/avaliar produtos adveio da necessidade de classificar o café enquanto bebida (MONTEIRO, 1984).

Desde então, a análise sensorial vem sendo muito utilizada na indústria de alimentos, uma vez que, por meio dela, é possível analisar a qualidade e a aceitação dos produtos avaliados. Para realizar a análise sensorial, um conjunto de pessoas (provadores) é treinado com o objetivo de avaliar as características dos produtos submetidos à análise.

Pelo fato dos produtos serem avaliados por seres humanos, a confiabilidade da análise sensorial está totalmente ligada ao desempenho do painel sensorial, isto é, quanto melhor for o desempenho do painel mais confiável será o resultado da análise sensorial.

A consonância do painel consiste na concordância entre os provadores, isto é, se os provadores pontuam um determinado atributo de maneira semelhante. Esta consonância está diretamente relacionada com o grau de treinamento do mesmo, ou seja, quanto mais bem treinado um painel está, mais os provadores pontuam os atributos de maneira semelhante.

Uma vez que a confiabilidade da análise sensorial está relacionada com a concordância dos provadores, é interessante para o pesquisador que haja a consonância do painel e métodos capazes de avaliar de forma prática esta concordância entre os provadores.

Na literatura é possível encontrar métodos capazes de avaliar a consonância de um painel sensorial, porém estes métodos avaliam a concordância entre provadores para um atributo de cada vez, o que acarreta em diversas análises para um mesmo conjunto de dados e torna a obtenção de resultados mais lenta. Logo, é interessante o desenvolvimento de testes que avaliem a consonância do painel de forma multivariada, ou seja, em que todos os atributos são avaliados simultaneamente.

Assim, o objetivo geral deste trabalho foi generalizar o teste assintótico de autovalores (AVa), que foi formalizado por Ferreira (2008a), advindo de propriedades assintóticas descritas por Girshick (1939), Lawley (1956) e Anderson (1963), e trazido para o contexto sensorial por Fernandes (2012), para inferir sobre o consenso multivariado de painéis sensoriais.

A generalização proposta utiliza premissas como a igualdade e invariância dos estima-

dores dos autovalores da matriz de covariâncias para a proposição de novos testes. A partir dessas premissas derivam-se quatro testes que foram analisados sob dois pares de hipóteses diferentes, ambos dispostos a acessar o consenso multivariado dos painéis. A avaliação do desempenho dos testes propostos se deu por meio de um estudo de simulação Monte Carlo, em que foram avaliados a taxa de erro tipo I e o poder dos testes propostos em diversos cenários, resultantes da combinação do número de provadores, atributos, observações e diferentes graus de treinamento.

O trabalho está dividido em cinco seções: na seção 2 foi realizada uma revisão de literatura com o objetivo de buscar trabalhos que propõem métodos para a avaliação da unidimensionalidade de um painel sensorial, bem como a revisão da teoria de componentes principais que fundamenta a generalização proposta, a teoria de taxa de erro tipo I e o poder de testes que foram utilizados para a avaliação dos testes e a teoria que fundamenta a simulação Monte Carlo.

A seção 3 apresenta a metodologia utilizada para a generalização do teste assintótico de autovalores, a teoria para o estudo de simulação e a descrição do experimento com dados reais, no qual oito provadores avaliam vinte e três atributos de oito marcas de *cream cheese*, utilizado para ilustrar os testes propostos.

A proposição dos testes, os resultados do estudo da simulação Monte Carlo e os resultados da ilustração dos testes propostos com um experimento com dados reais foram apresentados e discutidos na seção 4.

Por fim, as conclusões deste trabalho foram apresentadas na seção 5.

## 2 REFERENCIAL TEÓRICO

O objetivo desta revisão de literatura foi recuperar textos clássicos e atuais que tratam sobre a unidimensionalidade de painéis sensoriais. Para isto, serão abordados tópicos sobre análise sensorial e Sensometria, análise de componentes principais, taxas de erro tipo I e poder de um teste e simulação Monte Carlo.

### 2.1 ANÁLISE SENSORIAL E SENSOMETRIA

De acordo com Stone e Sidel (2004), a análise sensorial é definida pela Divisão de Avaliação Sensorial do Instituto de Tecnologia de Alimentos como a disciplina científica usada para evocar, medir, analisar e interpretar as reações devidas às características dos alimentos e materiais tais como são percebidas pelos sentidos da visão, olfato, paladar, tato e audição.

Os autores afirmam ainda que a análise sensorial ganhou força durante as décadas de 40 e 50 quando o exército dos Estados Unidos financiou uma pesquisa sobre a aceitação dos alimentos nas forças armadas.

Atualmente a análise sensorial vem sendo muito utilizada nas indústrias de alimentos, bebidas, cosméticos, automobilística, entre outras, com o objetivo de desenvolver novos produtos, determinar a preferência do consumidor e atuar no controle de qualidade (AUST et al., 1987; FERREIRA; OLIVEIRA, 2007; TEIXEIRA, 2009; ISAAC et al., 2012).

Minim (2006) destaca que a análise sensorial consiste, basicamente, em realizar a identificação dos atributos sensoriais, escolher e aplicar o método para quantificação de tais atributos e selecionar os métodos estatísticos para a análise e interpretação dos dados.

A área da ciência responsável pelos métodos estatísticos para a realização da análise sensorial é a Sensometria, cujo objetivo é medir, por meio de ferramentas estatísticas, os estímulos percebidos pelos sentidos humanos (FERREIRA; OLIVIERA, 2007). Pelo fato da realização da análise sensorial se dar por meio dos sentidos, o instrumento de medida utilizado é o ser humano, o qual é chamado de provador. O conjunto de provadores é denominado painel (DUT-COSKY, 2011).

De acordo com Amorim (2009), os provadores são como instrumentos utilizados para medir, classificar, diferenciar, avaliar e caracterizar um determinado produto. Assim, a vali-

dade dos resultados dependem da sensibilidade e da capacidade dos membros do painel em reproduzir seus julgamentos.

McEwan et al. (2002) afirmam que, em análise sensorial, testar a concordância do painel é um passo importante para demonstrar que os dados obtidos a partir das notas dos provadores é tão confiável quanto qualquer outro dado obtido a partir de outra ferramenta de medida. Dessa forma, os provadores devem ser preparados da melhor forma possível por meio de treinamentos que variam de acordo com o objetivo da análise sensorial.

Assim, o treinamento de um painel é importante para a credibilidade dos resultados da análise sensorial, uma vez que a concordância dos provadores influencia no resultado da análise. Por esta razão é interessante estudar métodos que são capazes de quantificar e avaliar a qualidade deste treinamento. Desta forma serão apresentados alguns estudos que propõe métodos para a avaliação da concordância de painéis sensoriais.

O método proposto por Dijksterhuis (1995) baseia-se na análise de componentes principais (ACP) para avaliar a consonância de um painel em relação a um determinado atributo. Os componentes principais são analisados e, caso os provadores pontuem da mesma forma o atributo avaliado, a maior parte da variância será explicada pelo primeiro componente principal e, assim, o painel será considerado unidimensional, ou seja, os provadores concordam em relação a este atributo.

Para entender o método proposto por Dijksterhuis (1995) considere um experimento com  $p$  provadores,  $n$  produtos e  $q$  atributos. Os dados podem ser dispostos em uma hipermatriz  $\mathbf{X}$  de dimensão  $n \times q \times p$ , com elementos  $x_{ijk}$ , em que  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, q$  e  $k = 1, 2, \dots, p$ .

A composição da hipermatriz  $\mathbf{X}$  se dá por meio de matrizes  $\mathbf{X}_j$  de ordem  $n \times p$  e um atributo fixo, ou seja, tomando uma matriz  $\mathbf{X}_j$ , com  $j = 1, 2, \dots, q$ , cada uma das  $p$  colunas refere-se às notas de cada provador para o atributo  $j$ , como ilustrado na Figura 1.

De acordo com o autor, ao aplicar a análise de componentes principais na matriz  $\mathbf{X}_j$  e obter um primeiro autovalor grande ou um alto percentual de explicação da variância, o grau de concordância entre os provadores para o atributo fixado é alto, ou seja, o painel pode ser considerado unidimensional em relação a esse atributo.

Um painel unidimensional dificilmente será obtido na prática. Assim uma forma de comparar a primeira maior dimensão e as demais dimensões de um mesmo atributo  $j$  é por meio da seguinte expressão:

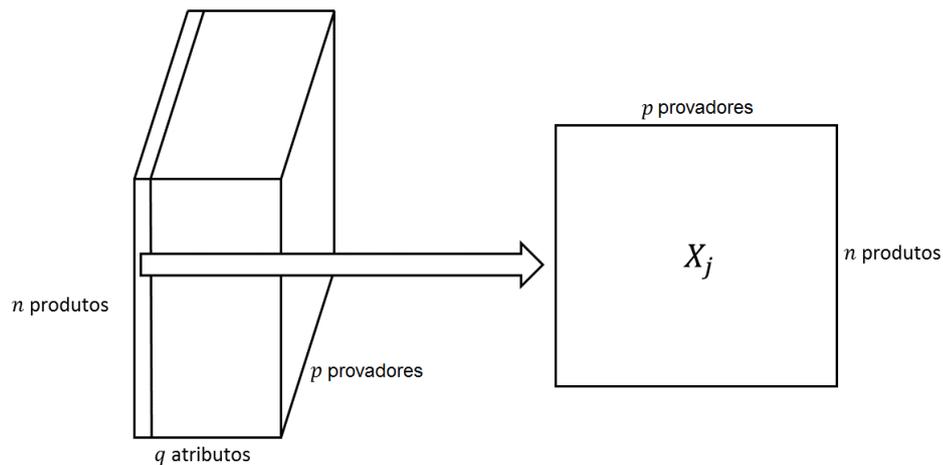


Figura 1 – Representação da hipermatriz  $X$  e sua respectiva fatia para o atributo  $j$ .  
 Fonte: adaptado de Dijksterhuis (1995).

$$C_j = \frac{\lambda_1^2}{\sum_2^p \lambda_i^2},$$

em que  $\lambda_1^2$  é o maior autovalor e  $\lambda_i^2$  é o  $i$ -ésimo autovalor da matriz  $X_j'X_j$ . Se  $C_j < 1$ , o denominador apresenta a maior das dimensões, caso contrário,  $C_j \geq 1$ , a dimensão do numerador é a maior e indica que os provadores estão concordando. Assim, pode-se determinar para qual atributo  $j$  o painel está concordando.

Derndorfer et al. (2005) apresentam um método para avaliar características de painéis como a discriminação entre produtos, a repetibilidade das avaliações, o uso de uma escala calculada pela diferença entre a maior e a menor avaliação e a concordância entre os provadores.

Para avaliar se os provadores concordam em relação a um atributo específico, o autor deste método inicialmente determina intervalos com 95% de confiança para as médias dos produtos e obtém os erros padrão destes intervalos a partir dos resíduos da análise de variância para cada interação do modelo, *atributo*  $\times$  *produto* e *provador*  $\times$  *produto*. Em seguida, são calculadas as médias das notas dadas pelos provadores, em todas as repetições, para um atributo específico e a porcentagem de médias dos produtos que se encontram dentro do intervalo de confiança representa a concordância de cada provador com os demais provadores do painel.

Já o método proposto por Latreille et al. (2006) para avaliar critérios de confiabilidade de um painel consiste em testar as seguintes características: discriminação, repetibilidade, severidade e concordância com o painel. Este método é baseado em modelos lineares mistos e

considera tanto o desempenho global do painel quanto o desempenho individual dos provadores. O modelo proposto é descrito a seguir:

$$Y_{ijk} = \delta + \alpha_j + s_i + b_k + a_{ji} + c_{jk} + d_{ik} + \varepsilon_{ijk},$$

em que  $Y_{ijk}$  é o valor atribuído pelo provador  $i$  ao produto  $j$  na sessão<sup>1</sup>  $k$ ,  $\delta$  é o intercepto,  $\alpha_j$  é o efeito fixo do produto,  $s_i$  é o efeito aleatório de provador ( $s_i \sim N(0, \sigma_{\text{prov}}^2)$ ),  $b_k$  é o efeito aleatório de sessão ( $b_k \sim N(0, \sigma_{\text{sessão}}^2)$ ),  $a_{ji}$ ,  $c_{jk}$  e  $d_{ik}$  são as interações *produto*  $\times$  *provador*, *produto*  $\times$  *sessão* e *provador*  $\times$  *sessão*, respectivamente ( $a_{ji} \sim N(0, \sigma_{\text{prod} \times \text{prov}}^2)$ ,  $c_{jk} \sim N(0, \sigma_{\text{prod} \times \text{sessão}}^2)$ ,  $d_{ik} \sim N(0, \sigma_{\text{prov} \times \text{sessão}}^2)$ ), e  $\varepsilon_{ijk}$  é o erro aleatório ( $\varepsilon_{ijk} \sim N(0, \sigma^2)$ ).

Os contrastes da interação *produto*  $\times$  *provador* são utilizados para testar a concordância de um provador com o painel. Entretanto, segundo os autores, o método mostrou-se adequado apenas para uma pequena quantidade de produtos e sessões, assim, indicam a utilização de modelos mais simples ou métodos de análise multivariada.

Fernandes (2012) propôs uma versão Monte Carlo do teste assintótico de autovalores formalizado por Ferreira (2008a), com o objetivo de obter um teste para unidimensionalidade de painéis sensoriais com poder e taxa de erro tipo I aceitáveis para amostras pequenas. Este teste tem como par de hipóteses e estatística de teste, respectivamente:

$$\begin{cases} H_0 : \lambda_2 = \lambda_3 = \dots = \lambda_p = 0. \\ H_1 : \text{ao menos um } \lambda_i, i > 1, \text{ é diferente de zero.} \end{cases}$$

e

$$Z_c = \frac{-\sum_{i=2}^p \hat{\lambda}_i}{\sum_{i=1}^p \hat{\lambda}_i \sqrt{\frac{2}{n-1}}} \sim N(0,1). \quad (2.1)$$

Para realizar os testes são geradas  $B$  amostras de uma variável aleatória  $X_{p \times 1} \sim N_p(\mathbf{0}, \Sigma)$ , em que  $\Sigma$  é a matriz de covariâncias equicorrelacionada. A variável  $X_{p \times 1}$  contém as notas dos  $p$  provadores para os  $n$  produtos.

---

<sup>1</sup>Em análise sensorial o momento em que os provadores realizam a degustação dos produtos é denominado sessão.

Em seguida, deve-se realizar a análise de componentes principais e determinar o primeiro autovalor de cada uma das  $B$  amostras  $\hat{\lambda}_{1l}$ , em que  $l \in 1, 2, \dots, B$ . O próximo passo é calcular a estatística de teste representada pela equação 2.1 para cada  $\hat{\lambda}_{1l}$  e ordenar a estatística de teste da  $l$ -ésima amostra de *bootstrap*  $Z_{cl}$  e é dada por

$$Z_{cl} = \frac{R_{1l}^2 - \rho_{10}^2}{\frac{\omega_{0l}}{\sqrt{n}}},$$

em que  $R_{1l}^2 = \frac{\lambda_1^*}{\sum_{i=1}^p \lambda_i^*}$ , no qual  $\lambda_i^*$  são os autovalores da matriz de covariâncias  $S_l^*$  que é calculada na  $l$ -ésima amostra *bootstrap*,  $\rho_1^2$  é a proporção acumulada da explicação da variação total do primeiro componente principal e o índice 0 representa o valor hipotético para  $\rho_1^2$  e  $\omega_{0l} = \sqrt{\frac{2tr(s_l^{*2})}{tr(s_l^*)^2} [(\rho_{10}^2)^2 - 2\hat{\beta}_l \rho_{10}^2 + \hat{\beta}_l]}$ , no qual

$$\hat{\beta}_l = \frac{\lambda_1^{*2}}{\sum_{i=1}^p \lambda_i^{*2}}.$$

Fixado um nível  $\alpha$  de significância, calcula-se o primeiro autovalor da matriz de dados observados e obtém-se a estatística de teste  $Z_c$  amostral. A hipótese nula  $H_0$  é rejeitada se valor-p =  $\frac{\sum_{l=1}^B I(Z_{cl} \leq Z_c)}{B}$  for menor que o nível de significância adotado.

Fernandes (2012) concluiu que o teste de autovalores Monte Carlo (AVmc) proposto apresentou melhor desempenho que o teste assintótico de autovalores (AVa), apresentando maior poder e controle das taxas de erro tipo I em todos os casos estudados.

Pinto et al. (2014) propuseram um método para análise da concordância de um painel baseado no coeficiente alfa de Cronbach (1951), que é utilizado para medir a consistência de respostas em questionários e pode ser utilizado para medir a similaridade entre as avaliações dos provadores. O método proposto pelos autores supra citados foi comparado com o método proposto por Dijksterhuis (1995) para avaliar a consonância de um painel.

Para entender o método proposto por Pinto et al. (2014) considere um experimento com  $p$  provadores,  $n$  produtos,  $q$  atributos e  $r$  sessões, assim, para cada atributo, tem-se  $n \times r$  notas

para cada provador. O coeficiente alfa de Cronbach é dado pela seguinte expressão:

$$CA_G = \frac{p\bar{c}}{\bar{v} + (p-1)\bar{c}}, \quad (2.2)$$

em que  $p$  é o número de provadores,  $\bar{c}$  é a covariância média entre todos os possíveis pares de provadores, e  $\bar{v}$  é a variância média de todos os provadores.

Os autores propõem duas abordagens para o uso do coeficiente, porém neste trabalho é de interesse apenas a abordagem que consiste na obtenção de uma avaliação global do desempenho do painel, considerando, simultaneamente, todos os produtos e todos os atributos. Esta abordagem utiliza a hipermatriz  $\mathbf{X}$  de  $n \times q \times r$  linhas e  $p$  colunas, a partir da qual são calculados as médias das sessões e o coeficiente alfa de Cronbach determinado na expressão 2.2 para cada atributo.

Para a avaliação do desempenho do  $p$ -ésimo provador ( $CA_p$ ) individualmente, deve-se calcular o coeficiente alfa de Cronbach da matriz  $\mathbf{X}$ , excluindo-se a  $p$ -ésima coluna e, em seguida, calcula-se o coeficiente alfa gradiente ( $CA_{\Delta p}$ ) para o provador com o objetivo de controlar a redução que a exclusão da  $p$ -ésima coluna causa no valor do  $CA_G$ . O coeficiente alfa gradiente é dado por:

$$CA_{\Delta p} = \frac{(CA_G - CA_p)}{CA_G} \times 100. \quad (2.3)$$

Ao se excluir o  $p$ -ésimo provador da hipermatriz  $\mathbf{X}$  o coeficiente alfa gradiente fornecerá o crescimento ou decréscimo do coeficiente  $CA_G$ . Se  $CA_{\Delta p} > 0$ , então um julgador é consonante com o restante do painel. Por outro lado, se  $CA_{\Delta p} < 0$ , então o painel é considerado menos consistente, ou seja, mais de um provador pontua um determinado atributo de forma diferente.

Ao aplicar o índice alfa de Cronbach em um conjunto de dados e compará-lo com o método proposto por Dijksterhuis (1995) os autores concluíram que o método de Dijksterhuis permite identificar para qual atributo o painel concorda, porém não é capaz de identificar qual provador discorda do painel. Logo, segundo os autores, o método proposto por eles foi considerado mais eficiente por permitir classificar os provadores em relação à sua concordância com o restante do painel.

Rocha (2015) propôs a generalização de cinco testes: o teste Monte Carlo para unidimensionalidade proposto por Amorim et al. (2010), dois testes *bootstrap* paramétricos pro-

postos por Gebert (2010), o teste de Fujikoshi proposto por Ferreira (2011) e o teste sobre autovalores Monte Carlo proposto por Fernandes (2012).

Que se tem conhecimento, os testes propostos por Rocha (2015) são os únicos que realizam a análise em um consenso multivariado. Assim, são os únicos testes que podem ser utilizados para comparação com os testes propostos neste trabalho.

A matriz de covariância amostral utilizada nos algoritmos destes cinco testes foi substituída pela matriz de correlação canônica com o objetivo de inferir sobre a concordância do painel considerando todos os atributos simultaneamente. Rocha (2015) recomenda que, para testar o consenso multivariado de painéis sensoriais, o teste sobre autovalores Monte Carlo seja utilizado, pois apresentou desempenho igual ou superior aos demais testes.

## 2.2 ANÁLISE DE COMPONENTES PRINCIPAIS

A análise de componentes principais (ACP) é uma metodologia estatística introduzida por Pearson (1901) e desenvolvida por Hotelling (1933) e tem por objetivos reduzir a dimensionalidade original dos dados e facilitar a interpretação das análises realizadas.

De acordo com Johnson e Wichern (2007), os componentes principais são, algebricamente, combinações lineares particulares de  $p$  variáveis aleatórias correlacionadas  $X_1, X_2, \dots, X_p$ . Geometricamente, estas combinações lineares representam um novo sistema de coordenadas obtido por meio da rotação do sistema original, no qual as variáveis  $X_1, X_2, \dots, X_p$  são posições dos eixos coordenados. O eixo de maior importância dentre os demais eixos representa a direção com variabilidade máxima e fornece uma descrição mais simples da covariância. Além disso, os componentes principais dependem unicamente da matriz de covariância  $\Sigma_{p \times p}$ , de  $X_1, X_2, \dots, X_p$  ou da matriz de correlação  $\rho$  determinada por meio da matriz  $\Sigma$ .

Para derivação dos componentes principais considere o vetor aleatório  $\mathbf{x} = [X_1, X_2, \dots, X_p]'$  cuja matriz de covariância  $\Sigma$  apresenta os autovalores  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_i \geq 0$  (JOHNSON; WHICHERN, 2007). Considere ainda as seguintes combinações lineares:

$$\begin{aligned}
 Y_1 &= \mathbf{a}'_1 \mathbf{x} = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\
 Y_2 &= \mathbf{a}'_2 \mathbf{x} = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p \\
 &\quad \vdots \\
 Y_p &= \mathbf{a}'_p \mathbf{x} = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p,
 \end{aligned}$$

no qual  $\mathbf{a}'_i$  é o vetor que determina a  $i$ -ésima combinação linear,  $i \in \{1, 2, \dots, p\}$ .

A combinação linear apresentada acima pode ser representada em notação matricial como segue:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \mathbf{A}\mathbf{x},$$

a qual tem as seguintes medidas estatísticas:

- Média:

$$\begin{aligned}
 \boldsymbol{\mu}_Y &= E[\mathbf{Y}] \\
 &= E[\mathbf{A}\mathbf{x}] \\
 &= \mathbf{A}E[\mathbf{x}] \\
 &= \mathbf{A}\boldsymbol{\mu}_X,
 \end{aligned}$$

onde  $\boldsymbol{\mu}$  é o vetor de médias.

- Covariância:

$$Cov(\mathbf{Y}_i, \mathbf{Y}_j) = \mathbf{a}'_i \boldsymbol{\Sigma} \mathbf{a}_j, \text{ com } i \neq j, \text{ e}$$

- Variância:

$$Var(\mathbf{Y}_i) = \mathbf{a}'_i \boldsymbol{\Sigma} \mathbf{a}_i.$$

Os componentes principais são as combinações lineares não correlacionadas que apresentam a maior variância possível. O primeiro componente principal é a combinação linear com máxima variância, porém esse máximo não existe uma vez que a variância do  $i$ -ésimo componente principal cresce para infinito à medida que os componentes do vetor  $\mathbf{a}_i$  crescem. Para eliminar esta indeterminação deve-se impor a restrição de que a soma ao quadrado dos coeficientes do vetor, que estabelece a combinação linear dos componentes principais, seja igual à unidade (FERREIRA, 2008a).

$$\mathbf{a}'_i \mathbf{a}_i = 1. \quad (2.4)$$

Os componentes principais também podem ser determinados em função dos autovalores e autovetores da matriz de covariâncias  $\Sigma$ . O  $i$ -ésimo componente principal cuja matriz de covariâncias  $\Sigma$  é associada ao vetor aleatório  $\mathbf{x} = [X_1, X_2, \dots, X_p]'$  e tem como autovalores e autovetores os seguintes pares  $(\lambda_1, \mathbf{a}_1), (\lambda_2, \mathbf{a}_2), \dots, (\lambda_p, \mathbf{a}_p)$ , com  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ , é definido como:

$$Y_i = \mathbf{a}'_i \mathbf{x} = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p.$$

Por definição, se  $\mathbf{A}$  é uma matriz quadrada de ordem  $n$ , então o vetor não-nulo  $\mathbf{a}$  é chamado de autovetor de  $\mathbf{A}$  se  $\mathbf{A}\mathbf{a} = \lambda\mathbf{a}$ , para algum escalar  $\lambda$ , denominado autovalor e que está associado ao autovetor  $\mathbf{a}$  (HEFFERON, 2014).

Para determinar os autovalores basta resolver a equação característica de uma matriz quadrada que é dada por

$$|\mathbf{A} - \lambda\mathbf{I}| = 0,$$

em que  $\mathbf{I}$  é a matriz identidade de dimensão  $n \times n$ . Já os autovetores são obtidos por meio da resolução do sistema linear

$$(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{a}_i = 0.$$

Para maximizar a variância sujeita a restrição representada na equação 2.4, Ferreira (2008a) sugere a utilização da técnica de multiplicadores de Lagrange, a qual consiste em ma-

maximizar a seguinte expressão em relação a  $\mathbf{a}_i$ :

$$\max[\mathbf{a}'_i \Sigma \mathbf{a}_i - \lambda_i (\mathbf{a}'_i \mathbf{a}_i - 1)], \quad (2.5)$$

em que  $\lambda_i$  é o multiplicador de Lagrange.

Ao derivar a expressão 2.5 em relação a  $\mathbf{a}_i$  e igualar a derivada a zero tem-se a seguinte expressão:

$$(\Sigma - \lambda_i \mathbf{I}) \mathbf{a}_i = \mathbf{0}, \quad (2.6)$$

em que  $\mathbf{I}$  é a matriz identidade de dimensão  $p \times p$ . De 2.6 segue que

$$\Sigma \mathbf{a}_i = \lambda_i \mathbf{a}_i, \quad (2.7)$$

em que os vetores  $\mathbf{a}_i$  e os multiplicadores de Lagrange  $\lambda_i$  ( $i \in \{1, 2, \dots, p\}$ ) são, respectivamente, os autovetores e autovalores da matriz de covariância  $\Sigma$ .

Logo, é possível verificar, utilizando 2.7, que:

$$\text{Var}(Y_i) = \mathbf{a}'_i \Sigma \mathbf{a}_i = \mathbf{a}'_i \lambda_i \mathbf{a}_i = \lambda_i \mathbf{a}'_i \mathbf{a}_i = \lambda_i, \quad (2.8)$$

e

$$\text{Cov}(Y_i, Y_j) = \mathbf{a}'_i \Sigma \mathbf{a}_j = \mathbf{a}'_i \lambda_j \mathbf{a}_j = \lambda_j \mathbf{a}'_i \mathbf{a}_j = 0,$$

se  $i \neq j$ , uma vez que  $\mathbf{a}_i$  e  $\mathbf{a}_j$  são ortogonais, ou seja, o produto interno <sup>2</sup> de  $\mathbf{a}_i$  e  $\mathbf{a}_j$  é igual a zero.

Além disso, os componentes principais são determinados de acordo com as variâncias, ou seja, o componente principal de maior importância é aquele que apresenta maior variância, o segundo mais importante apresenta a segunda maior variância, e assim por diante.

Outro resultado apresentado por Johnson e Wichern (2007) com relação aos componentes principais relaciona o traço da matriz de correlação  $\Sigma$ , ou seja, a soma da diagonal principal de  $\Sigma$ , com os seus autovalores e será apresentado no Teorema 1.

---

<sup>2</sup>O produto interno de dois vetores  $\mathbf{u}$  e  $\mathbf{v}$  é definido por:  $\mathbf{u} \cdot \mathbf{v} = \mathbf{v}' \mathbf{u} = \sum_{i=1}^n v_i u_i$ , em que  $u_i$  e  $v_i$  é a  $i$ -ésima coordenada dos vetores  $\mathbf{u}$  e  $\mathbf{v}$ , respectivamente.

**TEOREMA 1:** Seja o vetor  $\mathbf{x} = [X_1, X_2, \dots, X_p]'$ , cuja matriz de correlação  $\Sigma$  tem como autovalores e autovetores os seguintes pares  $(\lambda_1, \mathbf{a}_1), (\lambda_2, \mathbf{a}_2), \dots, (\lambda_p, \mathbf{a}_p)$ , com  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ . Considere que os componentes principais são dados por  $Y_i = \mathbf{a}_i' \mathbf{x}$ . Então

$$tr(\Sigma) = \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \sum_{i=1}^p Var(X_i) = \lambda_1 + \lambda_2 + \dots + \lambda_p = \sum_{i=1}^p Var(Y_i),$$

sendo  $\sigma_{ii}, i \in \{1, 2, \dots, p\}$ , as variâncias das variáveis  $X_1, X_2, \dots, X_p$ .

A partir deste teorema e da equação 2.8 pode-se concluir que a variabilidade total das variáveis originais é igual a variabilidade dos componentes principais.

Após a determinação dos componentes principais é necessário determinar quantos deles sintetizam as informações das variáveis originais em um modelo mais simples. Para definir o número ideal de componentes principais diversos testes são encontrados na literatura. Alguns destes teste serão apresentados na subseção seguinte.

### 2.2.1 Retenção de componentes principais

A utilização de componentes principais pode ser aplicada quando deseja-se avaliar a unidimensionalidade de um painel sensorial, uma vez que é possível resumir a informação das  $p$ -variáveis originais em  $q$  componentes principais tais que  $q < p$ . Para isto é necessário saber quantos componentes serão retidos (MINGOTI, 2005). Uma vez que não existe um consenso com relação a qual é a porcentagem ideal de informação que deseja-se que seja explicada pelos primeiros componentes principais, neste trabalho optou-se por trabalhar com os  $q$  primeiros componentes principais que retenham 70% da explicação da variância total dos dados originais por ser uma porcentagem bastante utilizada nas análises sensoriais.

Johnson e Wichern (2007) afirmam que não existe uma resposta definitiva para o número de componentes a serem retidos e que devem ser consideradas a quantidade total da variância amostral que foi explicada, as variâncias dos componentes amostrais e as interpretações dos componentes. Os autores ainda sugerem um método visual para a determinação da quantidade de componentes principais a serem retidos denominado *scree plot*.

Nesta representação os valores numéricos dos autovalores  $\lambda_i$  são ordenados, em ordem decrescente, e representados no eixo das ordenadas. No eixo das abcissas é plotada a ordem  $i$  dos autovalores (ou número de componentes principais), como pode ser observado na Figura 2. O número de componentes principais a serem retidos é o ponto no qual os autovalores que o seguem são relativamente pequenos e, praticamente, com mesmo valor, ou seja, o ponto a partir do qual quase não exista variação.

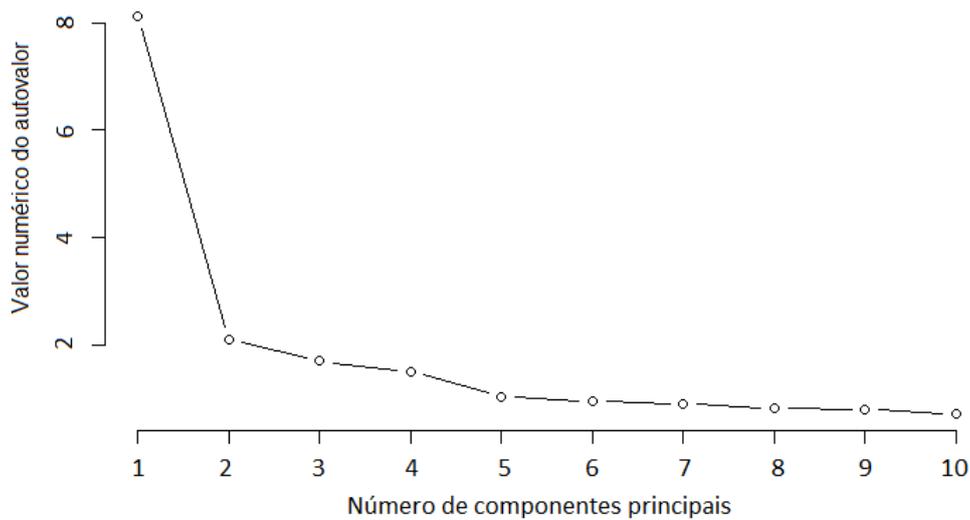


Figura 2 – Exemplo de um *scree plot*.  
Fonte: da autora.

Na Figura 2 pode-se observar que a partir do quinto componente principal os autovalores possuem aproximadamente o mesmo valor e se aproximam de zero. Logo, neste caso, quatro componentes principais resumiriam a variação total da amostra.

Outra forma de se obter o número de componentes a ser retido é proposta por Ferreira (2008a) e consiste na avaliação da proporção acumulada da explicação da variação total que é dada por:

$$R_k^2 = \frac{\sum_{i=1}^k \hat{\lambda}_i}{\sum_{i=1}^p S_{ii}} \times 100,$$

no qual  $S_{ii}$  é o elemento da diagonal da matriz de covariância amostral  $S$ ,  $k$  é o número de componentes principais retidos e  $p$  o número de variáveis originais.

Ferreira (2008a) ainda propõe um método para determinar o número de componentes principais a serem retidos com base no teste de hipótese de igualdade dos últimos autovalores,

que é representado pelo seguinte par de hipóteses:

$$\begin{cases} H_0 : \lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_p \\ H_1 : \text{pelo menos um dos autovalores difere dos demais.} \end{cases} \quad (2.9)$$

Quando a hipótese nula ( $H_0$ ) for rejeitada deve-se adicionar um componente principal ao modelo reduzido que foi escolhido e em seguida deve-se testar novamente a hipótese nula. Este é um teste recursivo, o qual deve ser repetido até que o menor valor de  $k$  seja obtido mediante a não rejeição da hipótese nula. Neste caso, o teste exige que os dados tenham distribuição normal multivariada, uma vez que é baseado em um teste de razão de verossimilhança, que assume como verdadeiro o modelo normal.

### 2.3 TAXA DE ERRO TIPO I E PODER DE UM TESTE

Em pesquisas de diversas áreas do conhecimento a Estatística é utilizada para a avaliação de suas hipóteses. A área da Estatística responsável por tal avaliação é a Inferência Estatística, em que as conclusões referentes ao par de hipóteses estão sujeitas a dois erros: erro tipo I e erro tipo II.

Ao se rejeitar a hipótese de nulidade ( $H_0$ ) quando esta é verdadeira comete-se o erro tipo I e a probabilidade de ocorrência deste erro é o nível de significância  $\alpha$ , que é estipulado pelo pesquisador. Já o erro tipo II é cometido quando não se rejeita a hipótese de nulidade ( $H_0$ ) quando esta é falsa e a probabilidade de cometê-lo é  $\beta$  (MOOD; GRAYBILL; BOES, 1974). Neste trabalho o erro tipo I significa concluir que há pelo menos um julgador mal treinado, quando o painel está bem treinado.

Além das taxas de erro, por meio dos testes de hipóteses pode-se avaliar o poder de um teste que é definido como a probabilidade de se rejeitar a hipótese de nulidade quando esta é realmente falsa e é dada por  $1 - \beta$ .

Em um teste de hipóteses deseja-se que as probabilidades de se cometer os erros tipo I e tipo II sejam pequenas, porém estes erros são inversamente proporcionais, ou seja, quanto menor a significância ( $\alpha$ ) do teste maior será o erro tipo II ( $\beta$ ) e, conseqüentemente, menor será o poder do teste. A taxa do erro tipo II pode ser menor quanto maior for o tamanho da amostra

(FERREIRA, 2008a).

## 2.4 SIMULAÇÃO MONTE CARLO

De acordo com Harrell, Gohsh e Bowden (2000) a simulação é a importação da realidade para um ambiente controlado, no qual pode-se avaliar e melhorar seu desempenho, por meio do estudo de seu comportamento. Além disso, com a simulação é possível identificar e eliminar problemas e ineficiências antes de implementar o projeto. Logo, a simulação é uma técnica que permite reproduzir uma situação real com o auxílio de um modelo matemático e gerar expectativas de resultados.

O método de Monte Carlo é uma técnica estatística utilizada em simulações para problemas que apresentam base probabilística ou determinística de acordo, ou não, com o tempo (HAMMERSLEY; HANDSCOMB, 1975).

Este método utiliza um processo aleatório para a geração de observações de uma determinada distribuição de probabilidades para aproximar a função de interesse (FERNANDES, 2012). De acordo com Barreto e Howland (2006), é necessário definir um ambiente artificial em que os parâmetros e o número de vezes que o processo irá executar são especificados para que após a execução do processo o computador mostre o resultado do experimento.

A simulação de Monte Carlo é utilizada em diversas áreas, desde simulação de fenômenos físicos, como o transporte de radiação da atmosfera da terra, até situações cotidianas, como simulação do jogo do bingo (BARRETO; HOWLAND, 2006).

### 3 MATERIAL E MÉTODOS

Nesta seção foi apresentada a metodologia utilizada para a proposição de quatro testes, além da metodologia para o estudo de simulação realizado para a avaliação e comparação destes testes e o procedimento para a realização do experimento com dados reais.

#### 3.1 BASE TEÓRICA PARA A PROPOSIÇÃO DOS TESTES

Para realizar a generalização do teste assintótico de autovalores é necessário compreender como este teste foi formalizado por Ferreira (2008a). Para isto, seja a amostra aleatória  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  de uma população normal  $k$ -variada, parametrizada por  $\boldsymbol{\mu}$  e  $\boldsymbol{\Sigma}$ . Ainda, admita que a matriz de covariâncias  $\boldsymbol{\Sigma}$  possua autovalores  $\lambda_1 > \lambda_2 > \dots > \lambda_k$  e autovetores  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ . Seja ainda  $\mathbf{S}$  a matriz de covariâncias amostrais, com autovalores  $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots > \hat{\lambda}_k$  e autovetores  $\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_k$ . Então, têm-se as seguintes decomposições espectrais:  $\mathbf{S} = \hat{\mathbf{P}}\hat{\boldsymbol{\Lambda}}\hat{\mathbf{P}}'$  e  $\boldsymbol{\Sigma} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'$ , nas quais  $\hat{\mathbf{P}}$  é uma matriz cujas colunas contêm os autovetores de  $\mathbf{S}$  e  $\hat{\boldsymbol{\Lambda}}$  é uma matriz diagonal composta pelos autovalores de  $\mathbf{S}$ ,  $\mathbf{P}$  e  $\boldsymbol{\Lambda}$  (diagonal) são matrizes definidas como as anteriores, porém seus autovetores e autovalores são da matriz  $\boldsymbol{\Sigma}$ . Note que o sinal  $'$  denota a matriz transposta.

Agora, considere que os componentes principais, segundo Ferreira (2008a), têm as seguintes propriedades assintóticas que foram demonstradas por Girshick (1939), Lawley (1956) e Anderson (1963):

- a)  $\sqrt{n-1}(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda})$  segue assintoticamente uma distribuição normal  $k$ -variada com vetor de médias  $\mathbf{0}$  e matriz de covariâncias  $2\boldsymbol{\Lambda}^2$ ,  $N_k(\mathbf{0}, 2\boldsymbol{\Lambda}^2)$ , em que  $\hat{\boldsymbol{\lambda}} = [\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_k]'$  e  $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_k]'$  são as diagonais das matrizes diagonais  $\hat{\boldsymbol{\Lambda}}$  e  $\boldsymbol{\Lambda}$ , respectivamente;
- b)  $\sqrt{n-1}(\hat{\mathbf{a}}_i - \mathbf{a}_i) \sim N_{k-1}(\mathbf{0}, \boldsymbol{\Upsilon}_i)$ , em que

$$\begin{aligned} \boldsymbol{\Upsilon}_i &= \lambda_i \sum_{m=1}^k \frac{\lambda_k}{(\lambda_m - \lambda_i)^2} \mathbf{a}_m \mathbf{a}_m' \\ &= \lambda_i \mathbf{P}_i \boldsymbol{\Gamma}_i \mathbf{P}_i', \end{aligned}$$

em que  $\mathbf{P}_i = [\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_k]$  uma matriz de dimensões  $k \times (k-1)$  na qual em suas colunas estão os autovetores  $\mathbf{a}_i$ , com  $i = 1, \dots, i-1, i+1, \dots, k$ , e a matriz diagonal  $\mathbf{\Gamma}_i$ , de dimensões  $(k-1) \times (k-1)$  é dada por

$$\mathbf{\Gamma}_i = \begin{bmatrix} \frac{\lambda_1}{(\lambda_1 - \lambda_i)^2} & 0 & \cdots & 0 \\ 0 & \frac{\lambda_2}{(\lambda_2 - \lambda_i)^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\lambda_k}{(\lambda_k - \lambda_i)^2} \end{bmatrix},$$

na qual elimina-se a  $i$ -ésima linha e a  $i$ -ésima coluna correspondente ao autovalor  $\lambda_i$  da matriz de covariâncias  $\mathbf{\Sigma}$ .

c) A covariância entre dois autovetores amostrais distintos é dada pela matriz

$$\text{Cov}(\hat{a}_i, \hat{a}_j) = \mathbf{\Upsilon}_{ij} = \begin{bmatrix} -\frac{\lambda_i \lambda_j \mathbf{a}_{i1} \mathbf{a}_{j1}}{(n-1)(\lambda_i - \lambda_j)^2} & \cdots & -\frac{\lambda_i \lambda_j \mathbf{a}_{i1} \mathbf{a}_{jp}}{(n-1)(\lambda_i - \lambda_j)^2} \\ \vdots & \ddots & \vdots \\ -\frac{\lambda_i \lambda_j \mathbf{a}_{ik} \mathbf{a}_{j1}}{(n-1)(\lambda_i - \lambda_j)^2} & \cdots & -\frac{\lambda_i \lambda_j \mathbf{a}_{ik} \mathbf{a}_{jk}}{(n-1)(\lambda_i - \lambda_j)^2} \end{bmatrix}.$$

d) Os elementos de  $\hat{\boldsymbol{\lambda}}$  são assintoticamente independentes dos elementos de  $\hat{\mathbf{P}}$ .

A partir da propriedade (a) pode-se concluir que  $\hat{\lambda}_i \sim N\left(\lambda_i, \lambda_i^2 \left(\frac{2}{n-1}\right)\right)$ .

De fato, pela propriedade (a) descrita anteriormente segue que

$$\sqrt{n-1}(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}) \sim N_k(\mathbf{0}, 2\mathbf{\Lambda}^2).$$

Daí,

$$(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}) \sim N_k\left(\frac{\mathbf{0}}{\sqrt{n-1}}, \frac{2\mathbf{\Lambda}^2}{n-1}\right).$$

Isolando  $\hat{\boldsymbol{\lambda}}$  tem-se

$$\hat{\boldsymbol{\lambda}} \sim N_k\left(\boldsymbol{\lambda}, \frac{2\mathbf{\Lambda}^2}{n-1}\right). \quad (3.1)$$

Agora, observe que

$$\begin{aligned} \mathbf{\Lambda}^2 &= \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{pk} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k^2 \end{bmatrix}. \end{aligned} \quad (3.2)$$

Substituindo 3.2 na equação 3.1 obtém-se

$$\hat{\boldsymbol{\lambda}} \sim N_k \left( \boldsymbol{\lambda}, \frac{2}{n-1} \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k^2 \end{bmatrix} \right). \quad (3.3)$$

A partir do Teorema 2, que será descrito a seguir e foi enunciado e demonstrado na página 119 do livro Estatística Multivariada de Ferreira (2008a), é possível concluir sobre a independência dos autovalores.

**TEOREMA 2:** Se o vetor aleatório  $\hat{\boldsymbol{\lambda}} = [\lambda_1, \lambda_2, \dots, \lambda_k]'$  que segue uma distribuição normal multivariada com parâmetros  $\boldsymbol{\mu}$  e  $\mathbf{\Lambda}^2 \left( \frac{2}{n-1} \right)$  em que

$$\mathbf{\Lambda}^2 = \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k^2 \end{bmatrix}$$

e  $\lambda_i$  são os autovalores da matriz de covariâncias  $\boldsymbol{\Sigma}$ , então os componentes do vetor  $\hat{\boldsymbol{\lambda}}$  são variáveis normais independentemente distribuídas.

Logo, por meio deste teorema, é possível concluir que os autovalores são independentes e assintoticamente normais. Portanto,

$$\hat{\lambda}_i \sim N \left( \lambda_i, \left( \frac{2}{n-1} \right) \lambda_i^2 \right).$$

A partir da distribuição dos autovalores amostrais, Ferreira (2008a) formaliza um teste de hipóteses para inferir sobre os autovalores populacionais. O par de hipóteses e a estatística para este teste são dados, respectivamente, por:

$$\begin{cases} H_0 : \lambda_i = \lambda_{i0} \\ H_1 : \lambda_i \neq \lambda_{i0}. \end{cases} \quad (3.4)$$

em que  $\lambda_{i0}$  é o valor hipotético para o parâmetro  $\lambda_i$ , e

$$Z = \frac{\hat{\lambda}_i - \lambda_{i0}}{\lambda_{i0} \sqrt{\frac{2}{n-1}}} \sim N(0,1). \quad (3.5)$$

Este teste é chamado de teste AVa uma vez que é um teste assintótico sobre autovalores e foi generalizado neste trabalho a fim de inferir sobre o consenso multivariado de painéis sensoriais.

### 3.2 ESTUDO DE SIMULAÇÃO DE DADOS

A comparação e a avaliação dos quatro testes propostos foram realizadas por meio de um estudo de simulação. A comparação dos testes é realizada por meio da avaliação do desempenho de cada um deles em termos de taxa de erro tipo I e poder. Para isso, foram implementadas rotinas no software R (R CORE TEAM, 2016), tanto para a programação dos testes quanto para a avaliação e comparação dos mesmos.

O estudo de simulação foi realizado via simulação Monte Carlo e foram considerados 74 cenários, determinados por meio da combinação do número de atributos  $q \in \{2, 5, 10, 20\}$ , número de provedores  $p \in \{2, 5, 10, 15\}$ , o número de observações (ou produtos)  $n \in \{10, 20, 30, 40, 50, 100, 200\}$ , a proporção da variação total explicada pelos  $pq$  componentes principais, ou grau de treinamento  $\rho^2 \in \{0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9; 0,99\}$  e da restrição de que  $n \geq pq$ . Esta restrição é imposta pela necessidade de existir graus de liberdade suficientes para a estimação dos  $pq$   $\lambda$  na estatística de teste (3.5). Além disso, para cada cenário foram simuladas

$N = 1000$  amostras de Monte Carlo.

Cada observação foi obtida a partir do desdobramento de uma hipermatriz de dimensão  $n \times q \times p$ , ou seja, a partir do concatenamento das  $p$  fatias de dimensão  $n \times q$  da hipermatriz  $M$ , como ilustrado na Figura 3.

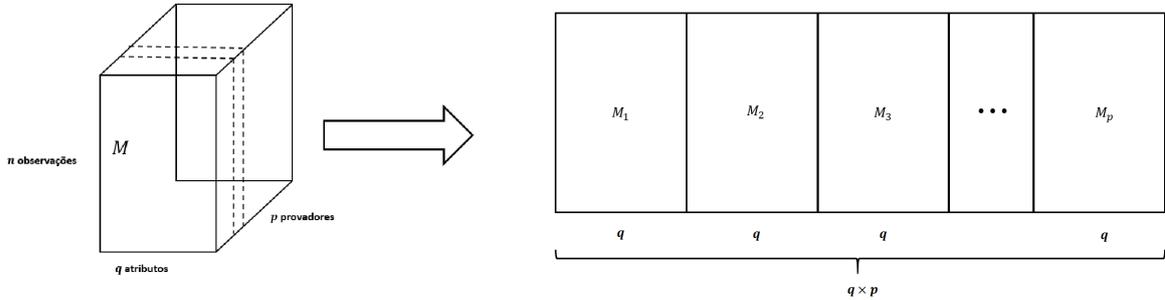


Figura 3 – Matriz desdobrada da qual são obtidas as observações via simulação.

Fonte: da autora.

Um ponto importante que deve ser destacado é que existe correlação entre os atributos sensoriais. Porém, neste trabalho, esta correlação foi fixada em zero ( $\rho_v = 0$ ), pois é de interesse estudar a correlação entre provedores. Além disso, quando a correlação entre atributos é levada em consideração pode existir o confundimento desta correlação com a correlação entre os provedores.

Para a simulação dos dados deste estudo foi utilizado o seguinte modelo linear misto:

$$Y_{ijm} = \mu + \tau_j + \beta_m + e_{ijm},$$

no qual  $i \in \{1, 2, \dots, n\}$ ,  $j \in \{1, 2, \dots, q\}$  e  $m \in \{1, 2, \dots, p\}$ ,  $Y_{ijm}$  é a  $j$ -ésima observação (nota) do produto  $i$  atribuída pelo o provedor  $m$  para o atributo  $j$ ,  $\mu$  é a constante comum a todas as observações, que pertence ao vetor de média  $\boldsymbol{\mu}$ ,  $\tau_j$  é o efeito fixo do atributo  $j$ , pertencente a um vetor de efeitos fixos  $\boldsymbol{\tau}$  de comprimento  $q$ ,  $\beta_m$  é o efeito aleatório do provedor  $m$ , que pertence a um vetor de efeitos aleatórios  $\boldsymbol{\beta}$  de comprimento  $p$ , e  $e_{ijm}$  é o erro aleatório associado a cada observação  $Y_{ijm}$ , vindo de uma normal padrão univariada. Neste trabalho foi assumido, sem perda de generalidade, que  $\mu = 0$ .

Os efeitos de provedor  $\boldsymbol{\beta}$  foram sorteados de uma população normal multivariada com parâmetros  $\boldsymbol{\mu} = \mathbf{0}$  e  $\boldsymbol{\Sigma}_p$  a fim de que o grau de treinamento ( $\rho^2$ ) dos provedores variasse no intervalo entre 0 e 1. Para isso, foi fixada a estrutura da matriz de covariâncias  $\boldsymbol{\Sigma}_p$ , de dimensão

$p \times p$ , equicorrelacionada:

$$\Sigma_p = \sigma^2 \mathbf{R} = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}, \quad (3.6)$$

em que é fixado  $\sigma^2 = 1$  e  $\rho$  é determinado de acordo com o  $\rho^2$  previamente escolhido.

Para determinação de  $\rho$ , observe que, segundo Ferreira (2008a), para a matriz de covariâncias definida acima, os componentes principais, que coincidem com os autovalores, são definidos em dois grupos. No primeiro, encontra-se o componente principal de maior importância e é definido como

$$\lambda_1 = \sigma^2[1 + (p - 1)\rho], \quad (3.7)$$

em que  $\rho$  é a correlação entre as variáveis. Por outro lado, no segundo grupo encontram-se os demais componentes principais que são dados por

$$\lambda_i = \sigma^2(1 - \rho), \quad (3.8)$$

com  $i \in \{2, \dots, p\}$ .

Por  $\Sigma_p$  ser uma matriz de covariâncias ela deve ser positiva definida. Logo, para que isto aconteça, os autovalores e a proporção da variação total devem ser todos positivos. Assim, fazendo  $\lambda_1 > 0$  tem-se que

$$\rho > \frac{-1}{p-1}. \quad (3.9)$$

Agora, sabe-se que a proporção da variação total pode ser estimada por:

$$\rho^2 = \frac{\lambda_1}{\sum_{i=1}^p \lambda_i} = \frac{1 + (p-1)\rho}{p}. \quad (3.10)$$

Logo,  $\rho^2 > 0$  obtém-se que

$$\rho > \frac{-1}{p-1}, \quad (3.11)$$

que coincide com a restrição 3.9. Assim, a matriz de covariância  $\Sigma_p$  será positiva definida se a restrição 3.9 for satisfeita. O desenvolvimento algébrico para determinação destas restrições é apresentado no Apêndice A.

A fim de atender as restrições impostas anteriormente, foi possível observar que para os casos em que  $p = 2$  provadores a restrição 3.9 é satisfeita apenas quando  $\rho^2 > 0,5$ . Já para  $p = 5$  provadores é necessário que  $\rho^2 > 0,2$ . Para os demais valores de número de provadores ( $p = 10$  e  $p = 15$ ), a restrição 3.9 é satisfeita quando  $\rho^2 > 0,1$ .

Com base nas considerações realizadas anteriormente, foram aplicados em cada amostra simulada os quatro testes propostos na seção 4.1, fixando-se a correlação ideal em  $\rho_0 = 0,70$ , como justificado na subseção 2.2.1. As porcentagens de rejeição da hipótese nula no decorrer das  $N$  amostras foram avaliadas ao nível de 5% de significância ( $\alpha = 5\%$ ) e a proporção de rejeições da hipótese nula, ou seja, o estimador da função poder é definido por:

$$FP = \frac{\sum_{l=1}^N I(\text{valor-}p_l \leq \alpha)}{N}, \quad (3.12)$$

sendo  $\text{valor-}p_l$  o valor-p da  $l$ -ésima amostra Monte Carlo simulada e  $I$  a função indicadora que retorna 1 se  $\text{valor-}p_l \leq \alpha$  e zero, caso  $\text{valor-}p_l > \alpha$ . Caso a amostra gerada seja proveniente de uma população na qual  $\rho < \rho_0$ , então a  $FP$  é considerada uma estimativa do poder, caso contrário  $FP$  será considerada uma estimativa do erro tipo I.

Além disso, foi determinado no software Sisvar (FERREIRA, 2008b), utilizando um teste binomial exato para uma proporção, um intervalo com 99% de confiança para o nível de significância  $\alpha$ ,  $[0,033927; 0,070504]$ .

Os testes foram comparados a partir da proporção de rejeição da hipótese nula nas  $N$  simulações Monte Carlo.

### 3.3 EXPERIMENTOS COM DADOS REAIS

Os dados utilizados para ilustrar a generalização do teste assintótico de autovalores realizado por Bro et al. (2008) são provenientes de um experimento com *cream cheese*<sup>3</sup>.

A análise sensorial foi realizada no Departamento de Ciências dos Alimentos da Uni-

<sup>3</sup>Os dados da análise sensorial estão disponíveis no site <http://www.models.life.ku.dk/Cream>.

versidade de Copenhague na Dinamarca. Os provadores fazem parte do painel permanente da universidade e foram treinados durante cinco sessões para escolherem os atributos a serem avaliados posteriormente. O conjunto de dados é composto por notas atribuídas por oito provadores para cada um dos vinte e três atributos sensoriais avaliados de dez produtos (oito deles diferentes e dois iguais) em três repetições amostrais, resultando em  $n = 30$  observações.

Os atributos pontuados pelos provadores foram divididos em categorias relacionadas ao olfato (4 atributos), visão (6 atributos), tato (1 atributo) e paladar (12 atributos).

Neste trabalho, seguindo a restrição  $n \geq pq$ , foram analisadas as notas dos oito provadores para três atributos relacionados ao olfato, a saber: aroma de creme, aroma ácido e aroma de manteiga. Foi fixado o grau de treinamento  $\rho^2 = 0,70$  e o par de hipótese que deseja-se testar é:

$$\begin{cases} H_0 : \text{O painel apresenta consenso multivariado} \\ H_1 : \text{O painel não apresenta consenso multivariado.} \end{cases}$$

A análise do comportamento dos testes para o cenário dos dados reais ( $p = 8$ ,  $q = 3$  e  $n = 30$ ) e a curva de poder dos testes foram realizadas por meio de um estudo de simulação de acordo com a metodologia descrita na subseção 3.2. Em seguida, foram realizados os quatro testes propostos neste trabalho pra verificar o consenso multivariado do painel.

## 4 RESULTADOS E DISCUSSÃO

Nesta seção são discutidos os resultados metodológicos decorrentes da proposição dos testes, bem como os resultados obtidos a partir do estudo de simulação e uma aplicação dos testes em um experimento com dados reais.

### 4.1 RESULTADOS METODOLÓGICOS

Os testes propostos nesta subseção são generalizações do teste assintóticos de autovalores proposto por Ferreira (2008a). O objetivo da generalização deste teste é dar a ela a capacidade testar o consenso multivariado de painéis sensoriais, ou seja, verificar a concordância dos provedores para vários atributos avaliados simultaneamente. A seguir é descrito o processo de generalização deste teste.

Seja  $\Omega$  um subconjunto do conjunto de índices  $I = \{1, 2, \dots, pq\}$  ( $\Omega \subseteq I$ ). Considere a variável aleatória definida por  $Y_\Omega = \sum_{\Omega} \hat{\lambda}_i$ . Pode-se observar, de acordo com o que foi abordado na seção 3.1, que  $Y_\Omega \sim N\left(\sum_{\Omega} \lambda_i, \left(\frac{2}{n-1}\right) \sum_{\Omega} \lambda_i^2\right)$ .

Primeiramente foi necessário realizar a padronização de  $Y$  e definir o par de hipóteses geral, que são dados, respectivamente, por:

$$Z = \frac{Y - E[Y]}{\sigma_Y} = \frac{Y - \sum_{\Omega} \lambda_i}{\sqrt{\frac{2}{n-1} \sum_{\Omega} \lambda_i^2}} \sim N(0,1) \quad (4.1)$$

e

$$\begin{cases} H_0 : & \text{O painel apresenta consenso multivariado} \\ H_1 : & \text{O painel não apresenta consenso multivariado.} \end{cases} \quad (4.2)$$

É a partir da estatística de teste  $Z$  e do par de hipóteses geral que foi realizada a generalização do teste assintótico de autovalores a fim de alcançar o objetivo geral deste trabalho.

Assim, diante das colocações acima, foi possível obter quatro testes que são ilustrados em forma de esquema na Figura 4 e são descritos a seguir.

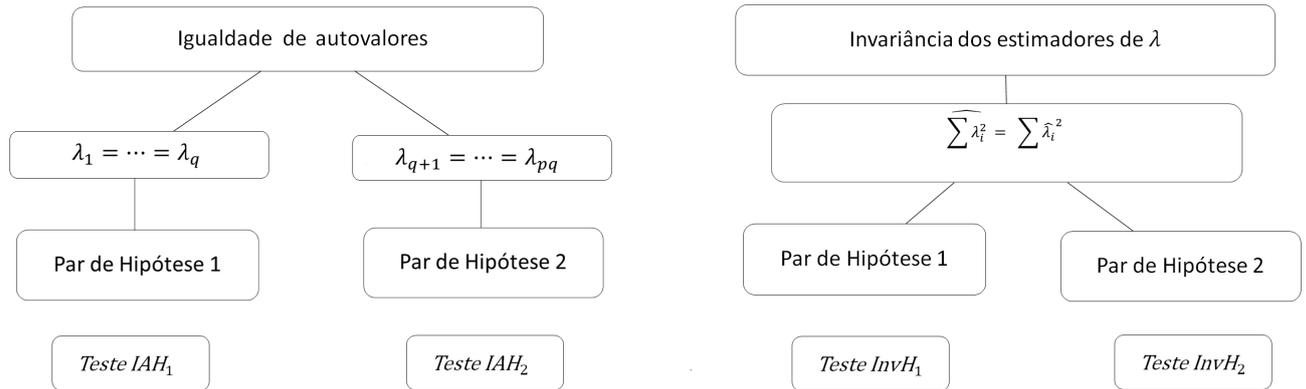


Figura 4 – Esquema para proposição dos testes.  
Fonte: da autora.

A generalização do teste AVa segue a partir de duas premissas iniciais: a igualdade dos autovalores  $\lambda$  e a propriedade de invariância dos estimadores de  $\lambda$ <sup>4</sup>.

Como pode ser observado na Figura 4, cada teste foi avaliado sob dois pares de hipóteses, a saber:

$$\left\{ \begin{array}{l} H_0 : \sum_{i=1}^q \lambda_i \geq \rho_0 \sum_{i=1}^{pq} \lambda_i = \rho_0 tr(\Sigma) \\ H_1 : \sum_{i=1}^q \lambda_i < \rho_0 \sum_{i=1}^{pq} \lambda_i = \rho_0 tr(\Sigma), \end{array} \right.$$

em que  $\rho_0$  é uma proporção aceitável fixada pela pesquisador,  $p$  o número de provedores,  $q$  o número de atributos e  $\Sigma$  é a matriz de covariância da combinação entre os provedores e atributos, com dimensão  $pq \times pq$ . O segundo par de hipóteses é dado por:

$$\left\{ \begin{array}{l} H_0 : \sum_{i=q+1}^{pq} \lambda_i \leq (1 - \rho_0) \sum_{i=1}^{pq} \lambda_i = (1 - \rho_0) \rho_0 tr(\Sigma) \\ H_1 : \sum_{i=q+1}^{pq} \lambda_i > (1 - \rho_0) \sum_{i=1}^{pq} \lambda_i = (1 - \rho_0) \rho_0 tr(\Sigma). \end{array} \right.$$

Porém, nas situações simuladas neste trabalho, assume-se que  $\Sigma = \mathbf{R}$ , como ilustrado na expressão 3.6 da sessão 3.2, então  $tr(\Sigma) = tr(\mathbf{R}) = pq$ . Logo, o par de hipóteses utilizados nos testes propostos neste trabalho são dados, respectivamente, por:

<sup>4</sup>A propriedade de invariância para estimadores de máxima verossimilhança é enunciada, por Mood, Graybill e Boes (1974), no teorema a seguir:

**TEOREMA:** Seja  $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$  o estimador de máxima verossimilhança de  $\theta$  com função densidade  $f(x; \theta)$ , no qual  $\theta$  é assumido unidimensional. Se  $\tau(\cdot)$  é uma função biunívoca, então o estimador de máxima verossimilhança de  $\tau(\theta)$  é  $\tau(\hat{\theta})$ .

Assim, como  $S$  é um estimador de máxima verossimilhança de  $\Sigma$ , então, pela propriedade de invariância, os  $\hat{\lambda}$  são estimadores de máxima verossimilhança dos autovalores de  $\Sigma$ .

$$\left\{ \begin{array}{l} H_0 : \sum_{i=1}^q \lambda_i \geq \rho_0 \sum_{i=1}^{pq} \lambda_i = \rho_0 pq \\ H_1 : \sum_{i=1}^q \lambda_i < \rho_0 \sum_{i=1}^{pq} \lambda_i = \rho_0 pq, \end{array} \right. \quad (4.3)$$

e

$$\left\{ \begin{array}{l} H_0 : \sum_{i=q+1}^{pq} \lambda_i \leq (1 - \rho_0) \sum_{i=1}^{pq} \lambda_i = (1 - \rho_0) pq \\ H_1 : \sum_{i=q+1}^{pq} \lambda_i > (1 - \rho_0) \sum_{i=1}^{pq} \lambda_i = (1 - \rho_0) pq. \end{array} \right. \quad (4.4)$$

Além disso, as estatísticas para os quatro testes propostos neste trabalho derivam da padronização 4.1 e utilizam o seguinte resultado:

**AFIRMAÇÃO:** Se  $\lambda_r = \dots = \lambda_s$ , em que os índices  $r, s \in \Omega$ , então  $\sum_{\Omega} \lambda_i^2 = \frac{\left(\sum_{\Omega} \lambda_i\right)^2}{l}$ , em que  $l$  é a cardinalidade do conjunto  $\Omega$  ( $|\Omega| = l$ ), ou seja,  $l$  é o número de elementos do conjunto  $\Omega$ .

De fato,

$$\begin{aligned} \frac{\left(\sum_{\Omega} \lambda_i\right)^2}{l} &= \frac{\left(\sum_{\Omega} \lambda\right)^2}{l} \\ &= \frac{(l\lambda)^2}{l} \\ &= \frac{l^2 \lambda^2}{l} \\ &= l\lambda^2 \\ &= \sum_{\Omega} \lambda^2 \\ &= \sum_{\Omega} \lambda_i^2. \end{aligned} \quad (4.5)$$

Agora, utilizando a padronização 4.1 e o resultado da afirmação anterior foi possível obter a estatística de teste geral, que é a base para as estatísticas dos quatro testes que foram descritas posteriormente sob a premissa da igualdade dos autovalores e a propriedade de invariância dos estimadores de  $\lambda$  (FIGURA 4), e é dada:

$$\begin{aligned}
Z &= \frac{Y - \sum_{\Omega} \lambda_i}{\sqrt{\frac{2}{n-1} \sum_{\Omega} \lambda_i^2}} \\
&= \frac{\sum_{\Omega} \hat{\lambda}_i - \sum_{\Omega} \lambda_i}{\sqrt{\frac{2}{n-1} \sum_{\Omega} \lambda_i^2}} \\
&= \frac{\sum_{\Omega} \hat{\lambda}_i - \sum_{\Omega} \lambda_i}{\sqrt{\frac{2}{n-1} \frac{\left(\sum_{\Omega} \lambda_i\right)^2}{l}}} \\
&= \frac{\sum_{\Omega} \hat{\lambda}_i - \sum_{\Omega} \lambda_i}{\sum_{\Omega} \lambda_i \sqrt{\frac{2}{l(n-1)}}}, \tag{4.6}
\end{aligned}$$

na qual  $Z \sim N(0,1)$  e  $l$  é o número de elementos do conjunto  $\Omega$ .

Para melhor compreender a utilização do conjunto  $\Omega$  no decorrer do trabalho observe que, por exemplo, se  $\Omega = \{1, 2, \dots, q\}$ , então é obtida a estatística de teste para  $IAH_1$ , pois este teste tem como premissa a igualdade dos  $q$  primeiros autovalores. Agora, se  $\Omega = \{q+1, q+2, \dots, pq\}$ , obtém-se a estatística para  $IAH_2$ , por assumir como premissa a igualdade dos últimos autovalores.

Como pode ser observado na Figura 4 o teste  $IAH_1$  tem como premissa a igualdade dos  $q$  primeiros autovalores, e embora pareça ser um premissa muito forte, sua comprovação será, ou não, confirmada após o estudo de simulação. O par de hipóteses para este teste é dado por 4.3.

Neste teste deseja-se testar se a variabilidade dos dados é explicada pelos  $q$  primeiros componentes principais. Assim, escrevendo-se a estatística de teste 4.6 sob  $H_0$  tem-se:

$$Z_1 = \frac{\sum_{i=1}^q \hat{\lambda}_i - \rho_0 \sum_{i=1}^{pq} \lambda_i}{\rho_0 \sum_{i=1}^{pq} \lambda_i \sqrt{\frac{2}{q(n-1)}}} = \frac{\sum_{i=1}^q \hat{\lambda}_i - \rho_0 pq}{\rho_0 pq \sqrt{\frac{2}{q(n-1)}}} \sim N(0,1).$$

Já no teste  $IAH_2$  assume-se a premissa de igualdade dos  $pq$  últimos autovalores, ou seja,  $\lambda_{q+1} = \dots = \lambda_{pq}$ . Pelo fato de desejar-se avaliar o consenso multivariado do painel sensorial, espera-se que a variabilidade dos dados seja explicada pelos  $q$  primeiros componentes

principais. Logo, supor a igualdade dos últimos autovalores pode ser considerada uma premissa menos forte que a premissa assumida no teste  $IAH_1$ , uma vez que a variabilidade explicada pelos últimos autovalores é, teoricamente, pequena e com valores próximos de zero.

O par de hipóteses para o teste  $IAH_2$  é dado por 4.4 e sua estatística de teste é definida por:

$$Z_2 = \frac{\sum_{i=q+1}^{pq} \hat{\lambda}_i - (1 - \rho_0)pq}{(1 - \rho_0)pq \sqrt{\frac{2}{q(p-1)(n-1)}}} \sim N(0,1).$$

Para os demais testes,  $InvH_1$  e  $InvH_2$ , utiliza-se a propriedade de invariância dos estimadores dos autovalores, ou seja, baseia-se na seguinte relação:  $\widehat{\sum \lambda_i^2} = \sum \hat{\lambda}_i^2$ . Além disso, estes testes têm, respectivamente como pares de hipóteses, os mesmos dos testes  $IAH_1$  e  $IAH_2$ .

Assim, o teste  $InvH_1$  é baseado no par de hipóteses 4.3 e tem a seguinte estatística de teste:

$$Z_3 = \frac{\sum_{i=1}^q \hat{\lambda}_i - \rho_0 pq}{\sqrt{\frac{2}{n-1} \sum_{i=1}^q \hat{\lambda}_i^2}} \sim N(0,1).$$

E, por fim, o teste  $InvH_2$  apresenta o par de hipóteses definido por 4.4 e sua estatística de teste é dada por:

$$Z_4 = \frac{\sum_{i=q+1}^{pq} \hat{\lambda}_i - (1 - \rho_0)pq}{\sqrt{\frac{2}{n-1} \sum_{i=q+1}^{pq} \hat{\lambda}_i^2}} \sim N(0,1).$$

O desempenho dos testes foi avaliado por meio de um estudo de simulação e seus resultados estão apresentados na seção 4.2.

## 4.2 ESTUDO DE SIMULAÇÃO

No decorrer desta seção são apresentados e discutidos os resultados obtidos por meio de um estudo de simulação. Para isto são apresentados gráficos referentes aos quatro testes propostos na seção 4.1, nos quais são ilustradas as curvas de poder de cada teste. Além disso,

os gráficos apresentam duas linhas pontilhadas pretas e duas vermelhas. A linha preta paralela ao eixo ordenado, que cruza o eixo das abcissas no ponto em que o grau de treinamento do painel sensorial corresponde à correlação ideal definida anteriormente de  $\rho_0 = 0,70$ , representa a separação das regiões de  $H_1$ , à esquerda da linha pontilhada, e  $H_0$ , à direita. Já a linha preta pontilhada paralela ao eixo das abcissas representa o nível de significância adotado, a saber  $\alpha = 5\%$ . As duas linhas vermelhas determinam o intervalo de confiança para o nível de significância determinado na seção 3.2.

Para os cenários definidos na seção 3.2 foi possível observar alguns casos nos quais não houve rejeição da hipótese nula ( $H_0$ ), mesmo quando  $H_0$  era nitidamente falsa, e que o poder dos testes é zero e (TABELA 1). Isso se deveu ao tamanho amostral insuficiente para o teste detectar  $H_0$  falsa, devido a seu caráter assintótico. Assim, sugere-se a busca por diferentes testes que atendam a esses casos.

Tabela 1 – Valores de  $p$  provadores,  $q$  atributos e  $n$  observações para os quais a porcentagem de rejeição dos testes é zero.

$p$ provadores	$q$ atributos	$n$ observações
2	2	10
2	5	10
2	5	20
2	5	30
2	5	40
2	10	20
2	10	30
2	10	40
2	10	50
2	20	40
2	20	50
2	20	100

Assim, para os casos em que o número de provadores é pequeno ( $p = 2$ ), Rocha (2015) sugere a utilização do teste de autovalores Monte Carlo (AVmc), visto que este foi o único teste generalizado neste trabalho que apresentou diferenças entre as simulações realizadas sob  $H_0$  e  $H_1$ . Já para amostras grandes, quando há o aumento no número de atributos, sugere-se a utilização dos testes sobre autovalores Monte Carlo (AVmc) e Fujikoshi.

Além dos casos mencionados anteriormente foi possível observar um comportamento comum em todos os cenários analisados. Os quatro testes não rejeitam a hipótese nula a partir do grau de treinamento  $\rho^2 = 0,80$ , o que indica que os testes propostos podem ser utilizados a partir deste grau de treinamento, com a ressalva de que alguns deles são mais poderosos que

outros.

A seguir, serão apresentados os comportamentos dos testes para cenários específicos, bem como o comportamento dos mesmos quando ocorre o aumento no número de um dos fatores (provedor, atributo e produto). Vale observar que alguns dos casos analisados não condizem com a realidade uma vez que as análises sensoriais realizadas com provedores treinados utilizam de 7 a 15 provedores, avaliam de 10 a 15 atributos de 5 marcas diferentes, porém neste trabalho, por serem testes assintóticos, tem-se o interesse em analisar o comportamento dos testes também em casos extremos (TEIXEIRA, 2009).

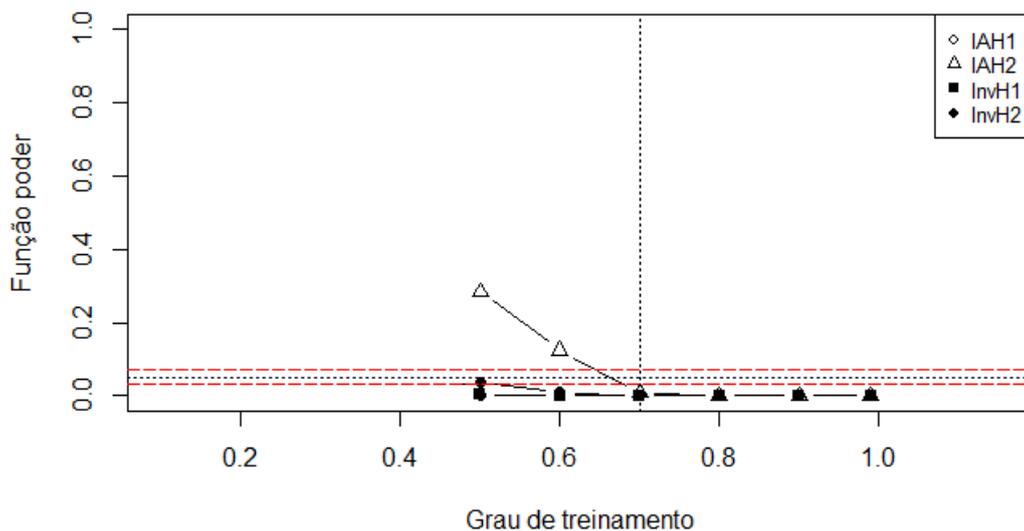


Figura 5 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provedores,  $q = 2$  atributos e  $n = 30$  observações.

Fonte: da autora.

Ao analisar o caso em que  $p = 2$  provedores,  $q = 2$  atributos e  $n = 30$  observações foi possível observar que nenhum dos testes apresentam comportamento satisfatório, uma vez que os testes  $IAH_1$ ,  $InvH_1$  e  $InvH_2$  não rejeitam a hipótese nula porém apresentam poder nulo ou próximo de zero, como pode ser observado na Figura 5. Além disso, o teste  $IAH_2$  é o teste que apresenta o maior poder (29,3%), porém é considerado conservador, assim como os demais, pois a probabilidade real de se cometer o erro tipo I é menor do que o nível de significância estipulado.

Apesar de os testes não apresentarem um comportamento satisfatório para o caso em que o número de provedores e atributos são iguais a dois ( $p = 2$  e  $q = 2$ , respectivamente), este

resultado não compromete os testes propostos neste trabalho, uma vez que, em casos de análises reais, raramente o número de provedores treinados e atributos avaliados são tão pequeno quanto neste caso de simulação.

Para os casos em que  $p = 2$  provedores,  $q = 2$  atributos e  $n = 30$  observações, o teste de autovalores Monte Carlo generalizado por Rocha (2015) apresenta maior poder que os testes propostos neste trabalho.

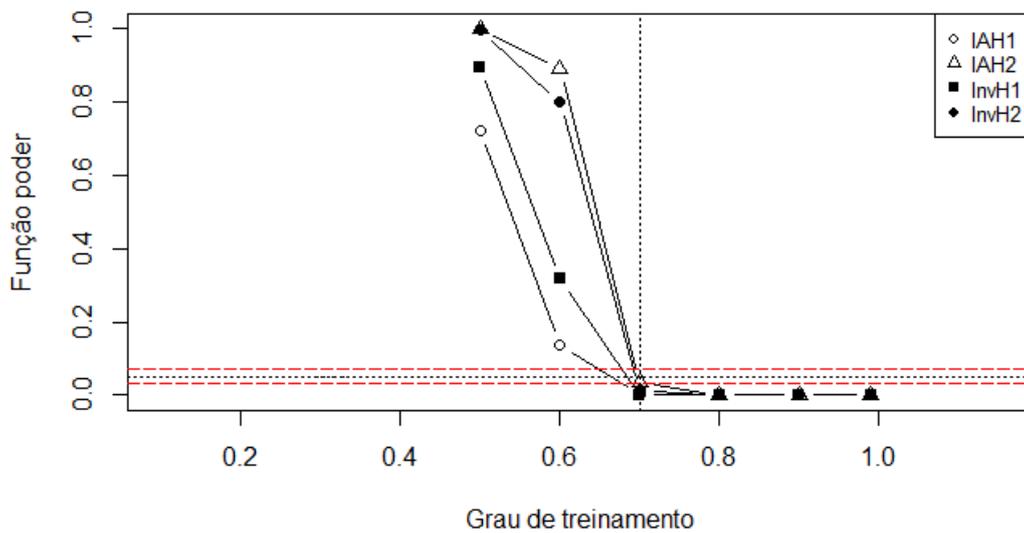


Figura 6 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provedores,  $q = 2$  atributos e  $n = 100$  observações.  
Fonte: da autora.

Nas situações em que o número de observações  $n$  aumenta para 100 ( $n = 100$ ), considerando ainda  $p = 2$  provedores e  $q = 2$  atributos, o poder dos quatro testes aumenta consideravelmente atingindo 99,9% nos testes  $IAH_2$  e  $InvH_2$ , 71,6% no teste  $IAH_1$ , 87,8% no teste  $InvH_1$ . Entretanto o teste  $IAH_2$  passa a ser considerado exato, como ilustrado na Figura 6 e nas Figuras do Apêndice B. Assim, neste caso o teste  $IAH_2$  é aconselhável.

Quando aumenta-se o número de provedores para cinco ( $p = 5$ ) e mantém-se constante os valores de  $q = 2$  atributos e  $n = 20$  observações, pode-se observar que na vizinhança do grau de treinamento definido os testes  $IAH_2$  e  $InvH_2$  possuem poder considerável, porém o teste com maior poder é o  $IAH_2$  (58,7%) e o único teste considerado liberal, como pode ser observado na Figura 7. Como os demais testes são considerados conservadores, nesta situação aconselha-se a utilização do teste  $IAH_2$ , pois apesar de ser liberal é recomendada a utilização

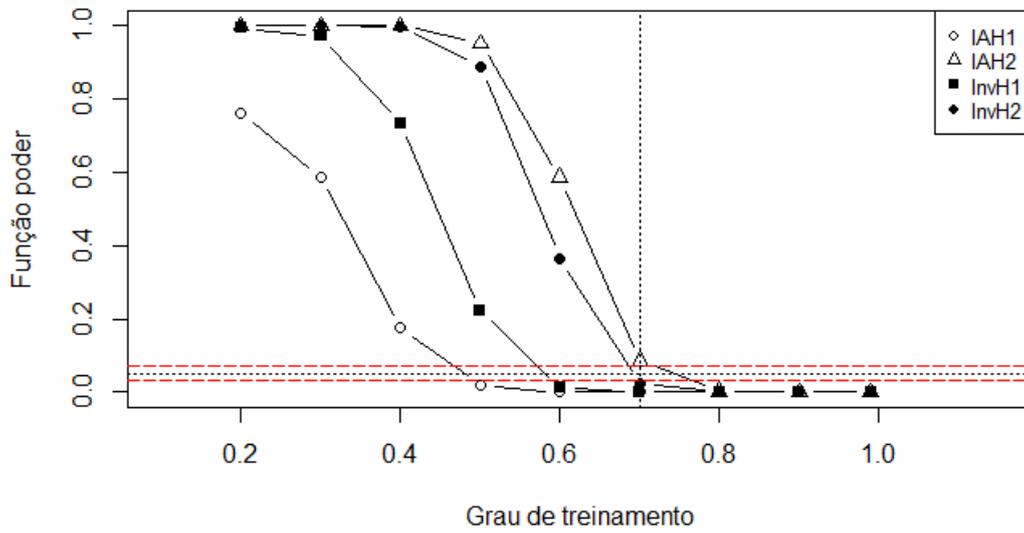


Figura 7 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 2$  atributos e  $n = 20$  observações.

Fonte: da autora.

de um teste liberal do que um teste conservador, uma vez que afirmar que um provedor não está treinado quando na realidade ele está é menos prejudicial para o resultado da análise do que afirmar o contrário.

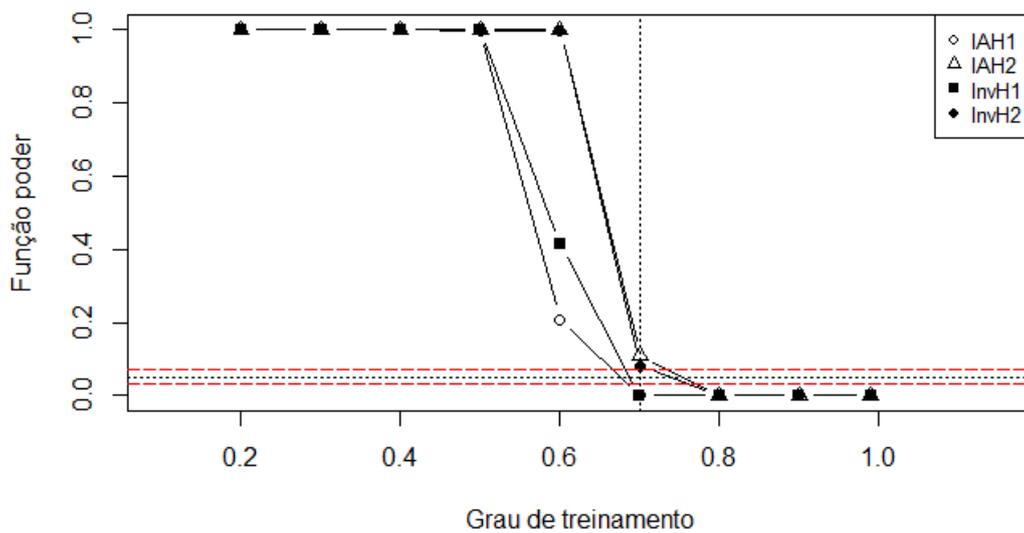


Figura 8 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 2$  atributos e  $n = 100$  observações.

Fonte: da autora.

Com o aumento do número de observações para 100 ( $n = 100$ ) pode-se observar, na Figura 8 e nas Figuras do Apêndice C, que o poder dos quatro testes aumentam consideravelmente, porém os testes  $IAH_2$  e  $InvH_2$  tornam-se mais liberais. Por outro lado, o teste  $InvH_2$ , por ser o menos liberal, é o aconselhável para este cenário.

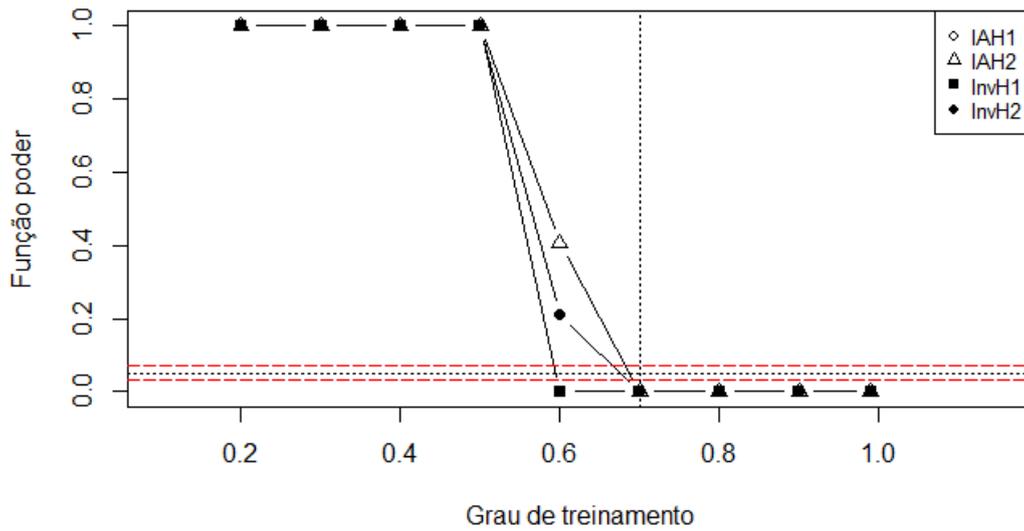


Figura 9 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provadores,  $q = 20$  atributos e  $n = 100$  observações.  
Fonte: da autora.

Ao manter-se os valores de provadores e observações fixos em  $p = 5$  e  $n = 100$ , respectivamente, e aumentando-se o número de atributos para  $q = 20$  é possível observar que há uma queda na curva de poder dos quatro testes no grau de treinamento  $\rho^2 = 0,60$ , como pode ser observado na Figura 9. Quando adotado o grau de treinamento  $\rho^2 = 0,60$ , recomenda-se a utilização do teste  $InvH_2$  por ser o teste menos liberal que os demais. Além disso, os testes passam a não rejeitar a hipótese nula a partir de  $\rho^2 = 0,70$ .

Para os casos em que o número de provadores e observações são mantidos fixos e aumenta-se o número de atributos, Rocha (2015) propõe a utilização do teste de autovalores Monte Carlo generalizado, porém este teste neste cenário não controla as taxas de erro tipo I. Assim, para estes casos, é recomendado a utilização de um dos testes propostos neste trabalho para o grau de treinamento  $\rho^2 = 0,70$ . Caso o pesquisador estipule um grau de treinamento menor, é recomendado a utilização do teste  $InvH_2$ .

Ao aumentar-se o número de provadores para  $p = 10$  e manter-se constante os valores

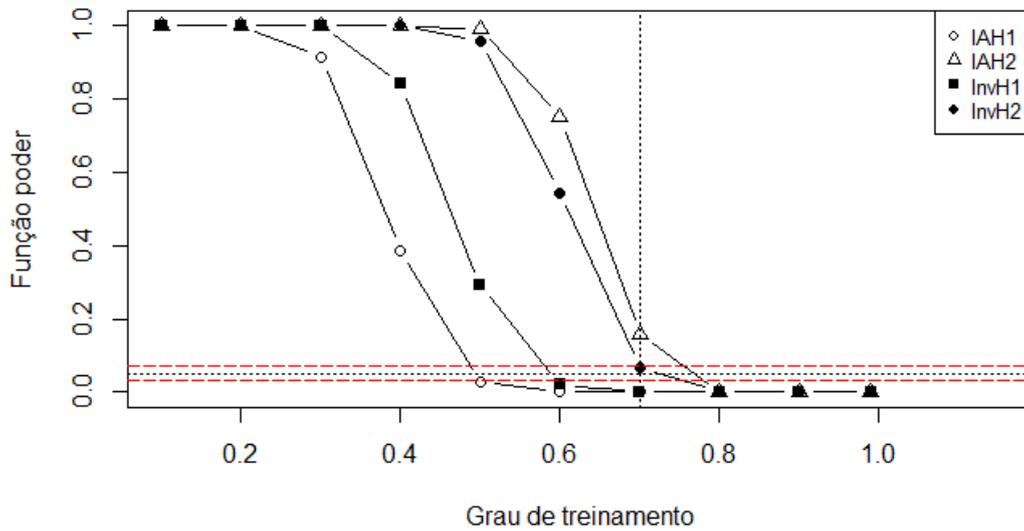


Figura 10 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 2$  atributos e  $n = 20$  observações.  
Fonte: da autora.

de  $q = 2$  atributos e  $n = 20$  observações (Figura 10), o teste  $InvH_2$  apresenta o segundo maior poder (54,5%) e pode ser considerado exato por estar dentro do intervalo de confiança para o nível de significância  $\alpha$ . Já o teste  $IAH_2$ , apesar de ser o mais poderoso (75%), é liberal deixando de ser aconselhável. Além disso, o teste  $IAH_1$  apresenta o menor poder (1%) e é considerado conservador. Assim, para este caso aconselha-se a utilização do teste  $InvH_2$ .

Nas situações em que aumenta-se o número de observações para  $n = 100$  e conserva-se o número de provadores ( $p = 10$ ) e atributos ( $q = 2$ ) pode-se observar que os quatro testes controlam melhor as taxas de erro tipo I, porém os teste  $IAH_2$  e  $InvH_2$  se tornam mais liberais, como apresentado na Figura 11. Embora os testes  $IAH_1$  e  $InvH_1$  apresentarem um comportamento adequado neste cenário é aconselhável a utilização de um teste mais liberal do que conservador. Assim, o teste  $InvH_2$ , por ser menos liberal que o teste  $IAH_2$ , é o teste aconselhável.

Para as situações em que  $p = 10$  provadores e  $q = 2$  atributos, os testes generalizados por Rocha (2015) apesar de apresentarem poder alto não controlam as taxas de erro tipo I na região próxima ao ponto crítico, que no caso da autora é  $\rho^2 = 0,75$ . Este comportamento não é observado nos testes propostos neste trabalho, uma vez que neste caso os quatro testes apresentam alto poder e controlam as taxas de erro tipo I na região do ponto crítico determinado

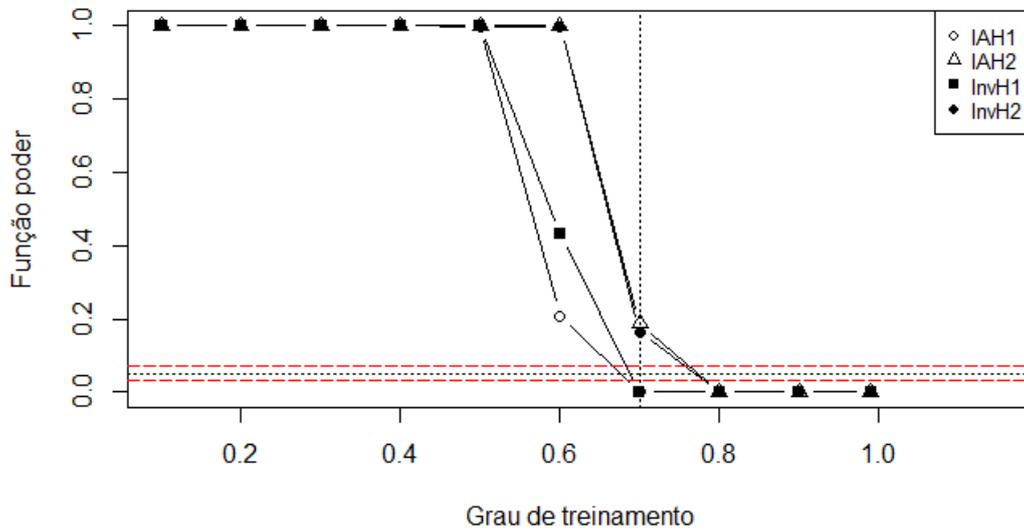


Figura 11 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 2$  atributos e  $n = 100$  observações.

Fonte: da autora.

pela autora supracitada, como pode ser observado nas Figuras 10, 11 e nas Figuras do Apêndice D.

Ao se comparar as Figuras 5, 7 e 10 é possível analisar o comportamento dos testes com relação ao aumento da quantidade de provadores, em que observa-se um maior controle da taxa de erro tipo I com o aumento do número de provadores. Além disso, o comportamento dos testes é semelhante a partir de  $p = 5$  (FIGURA 7 e FIGURA 10).

Quando os valores das variáveis são fixados em  $p = 15$  provadores,  $q = 2$  atributos é possível observar que este é o caso que apresenta a menor quantidade de cenários a serem analisados, uma vez que ao aumentar o número de atributos para  $q = 5$  e  $q = 10$  atributos os casos que satisfazem as condições impostas na seção 3.2 são apenas para  $n > 100$  observações.

Ao considerar os casos em que  $p = 15$  produtos,  $q = 2$  atributos e  $n = 30$  observações, pode-se observar que os testes  $IAH_2$  e  $InvH_2$  são considerados liberais e poderoso, com 72,6% e 86,8%, respectivamente, de poder (FIGURA 12). Além disso, os testes  $IAH_1$  e  $InvH_1$  continuam conservadores e com baixo poder (3% e 4,7%, respectivamente). Logo, para este caso é recomendado a utilização do teste  $InvH_2$ , por ser o menos liberal que o teste  $IAH_2$ .

Os demais cenários para o caso em que o número de provadores  $p = 15$  que satisfazem as restrições impostas na seção 3.2 são apenas os casos em que  $n \geq 100$  e  $5 \leq q \leq 10$  e seus

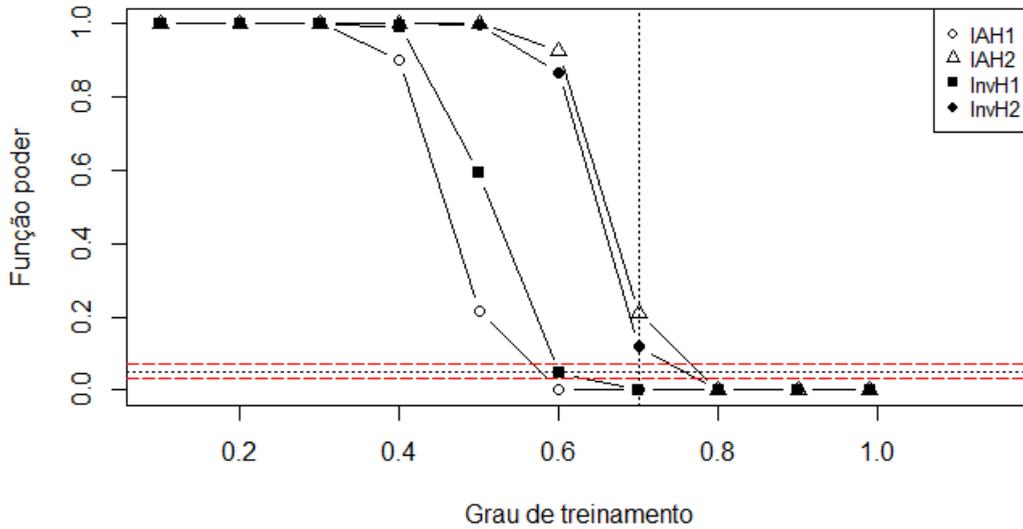


Figura 12 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 15$  provadores,  $q = 2$  atributos e  $n = 30$  observações.

Fonte: da autora.

gráficos podem ser observados no Apêndice E.

Para os cenários em que o número de provadores é fixado em  $p = 15$  os testes *bootstrap* propostos por Rocha (2015) não controlam as taxas de erro tipo I com o aumento do número de atributos  $q$  e o teste de autovalores Monte Carlo é liberal sempre que  $p = n$ . Assim como para os testes propostos neste trabalho, os testes não apresentaram um comportamento adequado, como pode ser observado nas Figuras do Apêndice E.

Em geral, os testes propostos neste trabalho controlam melhor a taxa de erro tipo I e apresentam maior poder conforme aumenta-se o número de observações ( $n$ ), porém quando isto ocorre os testes passam a ser mais liberais. Além disso, como pode ser observado nas figuras em que o número de observações é 200 ( $n = 200$ ) que encontram-se nos Apêndices, os quatro testes podem ser considerados assintoticamente iguais, uma vez que quando o número de observações tende ao infinito o comportamento dos testes tendem a ser iguais.

Com o aumento no número de provadores ( $p$ ) também foi possível observar que os testes ficam mais poderosos, porém para o caso em que  $p = 15$  e  $q \geq 5$  os cenários que satisfazem as restrições impostas na seção 3.2 são apenas para os casos em que  $n \geq 100$ , o que não condiz com a realidade das análises sensoriais realizadas com provadores treinados.

Na seção 4.1 foi mencionado que a premissa da igualdade dos  $q$  primeiros autovalores

para o teste  $IAH_1$  poderia ser uma premissa forte, o que se confirmou durante o estudo de simulação, uma vez que o teste  $IAH_1$  não apresenta comportamento satisfatório nos cenários estudados.

### 4.3 EXPERIMENTO COM DADOS REAIS

Nesta seção são apresentados os resultados obtidos da análise do experimento com dados reais descrito na seção 3.3.

Inicialmente foi realizada a simulação considerando a mesma situação dos dados do experimento com *cream cheese*, ou seja,  $p = 8$  provadores,  $q = 3$  atributos e  $n = 30$  observações, para avaliar o desempenho dos quatro testes propostos neste trabalho. Além disso, o grau de treinamento adotado foi  $\rho^2 = 0,70$  e o par de hipóteses testado foi:

$$\begin{cases} H_0 : \text{O painel apresenta consenso multivariado} \\ H_1 : \text{O painel não apresenta consenso multivariado.} \end{cases}$$

Como pode ser observado na Figura 13, os testes  $IAH_2$  e  $InvH_2$  são os testes mais poderoso e podem ser considerados exato. Assim, para este caso simulado, o teste o pesquisador pode

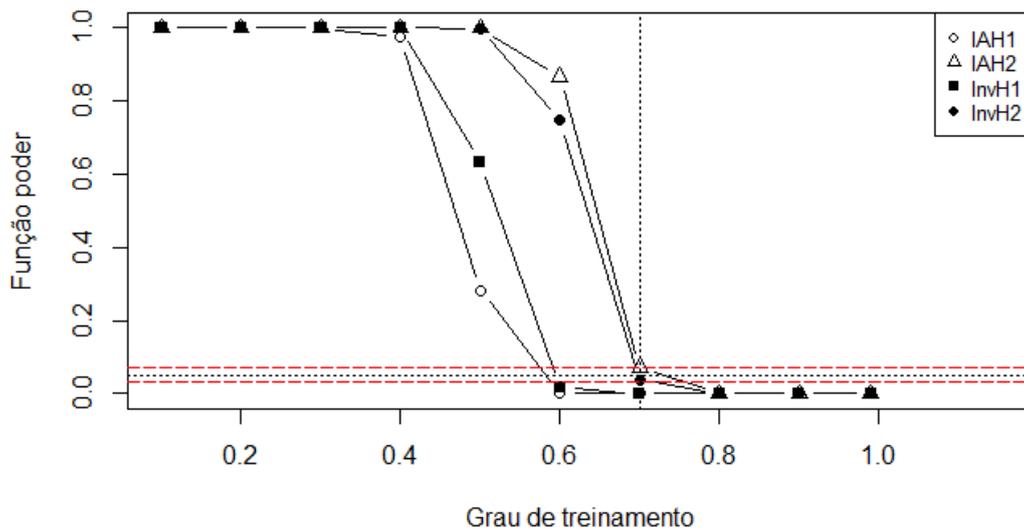


Figura 13 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 8$  provadores,  $q = 3$  atributos e  $n = 30$  observações.  
Fonte: da autora.

Embora os testes  $IAH_2$  e  $InvH_2$  tenham apresentado um comportamento satisfatório no estudo de simulação, os quatro testes apresentados na seção 4.1 rejeitam a hipótese de consenso multivariado ao nível de 5% de significância, ou seja, ao considerar simultaneamente os atributos, os provadores não pontuam os atributos da mesma forma. Os valores-p dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  são, respectivamente, 0,002,  $5,59 \times 10^{-70}$ ,  $2,72 \times 10^{-7}$  e  $3,23 \times 10^{-10}$ .

Como a hipótese de consenso multivariado foi rejeitada, o teste  $InvH_2$  seria o teste mais crível, pois, no estudo de simulação para os casos com valores próximos ao do experimento com *cream cheese*, estes foram os testes que apresentaram melhor desempenho, se mostrando poderosos e controlando as taxas de erro tipo I.

## 5 CONCLUSÃO

A fim de cumprir o objetivo deste trabalho pode-se generalizar o teste assintótico de autovalores obtendo-se outros quatro testes para testar o consenso multivariado de painéis sensoriais, além de avaliar o comportamento dos mesmos. Em geral, ao avaliar o poder e a taxa de erro tipo I, foi possível concluir que os quatro testes proposto são assintoticamente iguais, os testes  $IAH_2$  e  $InvH_2$  são os melhores testes, porém o teste  $InvH_2$  foi exato em mais cenários.

Os quatro testes apresentaram comportamento semelhante em algumas situações. A partir do grau de treinamento  $\rho^2 = 0,8$  todos os testes podem ser utilizados. Além disso, entre os quatros testes propostos, o desempenho do teste  $IAH_1$  não foi tão satisfatório quanto o dos demais testes, pois em todos os cenários é o teste que é sempre conservador, o que acarreta na falta de poder observada durante o estudo de simulação. Por outro lado, o teste  $IAH_2$  é o teste mais liberal em todos os casos e, conseqüentemente, é o teste mais poderoso.

Os melhores desempenhos foram observados nos testes  $IAH_2$  e  $InvH_2$ , o que sugere que é mais adequado testar os últimos autovalores da matriz de covariâncias  $\Sigma$ . Além disso, aumentando-se o número de observações, os testes se tornam mais poderosos, porém os testes  $IAH_2$  e  $InvH_2$  tendem a ficar mais liberais. Apesar disso, o teste  $InvH_2$  é exato em mais cenários do que o teste  $IAH_2$ .

O teste  $InvH_2$  é indicado para os casos em que conserva-se o número de produto e atributo e aumenta-se o número de observações. O teste  $InvH_2$  é novamente indicado para os casos em que o número de atributos e observações são mantidos constantes. Por outro lado, aumentando o número de atributos e conservando o número de provadores e observações os testes  $IAH_2$  e  $InvH_2$  são aconselháveis.

Vale ressaltar que, pelo fato de serem testes assintóticos, os testes propostos neste trabalho necessitam de um número grande de observações para apresentarem um comportamento satisfatório, o que pode ser uma limitação ao se trabalhar com análises sensoriais reais, uma vez que análises com grande quantidade de produtos a serem avaliados seria muito cansativo para o provador, podendo causar fadiga gustatória e olfativa, por exemplo.

Assim, para trabalhos futuros, sugere-se investigar variações dos testes propostos neste trabalho, a fim de obter-se desempenho satisfatório para pequenas amostras.

## REFERÊNCIAS

AMORIM, I. S. **Teste Monte Carlo na avaliação da unidimensionalidade de painéis sensoriais ara uma variável**. 62 f. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2009.

AMORIM et al. Monte Carlo based test for inferring about the unidimensionality of a Brazilian coffee sensory panel. **Food Quality and Preference**, v. 21, p. 319-323, 2010.

ANDERSON, T. W. Asymptotic theory for principal components. **Annals of Mathematical Statistics**, Ann Arbor, v. 34, p. 122 - 148, 1963.

AUST, L. B. et al. The descriptive analysis of skin care products by a trained panel of judges. **Journal of the Society of Cosmetic Chemists**, v. 38, p. 443 - 449, 1987.

BARRETO, H.; HOWLAND, F. M. **Introductory econometrics**: using Monte Carlo simulation with Microsoft Excel. Cambridge: Cambridge university press, 2006. 774 p.

BRO, R. et al. Multiway models for sensory profiling data. **Journal of Chemometrics**, v. 22, p. 36-45, 2008.

CRONBACH, L. J. Coefficient alpha and the internal structure of tests. **Psychometrika**, n. 16, p. 297-334, 1951.

DERNDORFER, E. et al. A panel performance procedure implemented in R. **Journal of Sensory Studies**, n. 20, p. 217 - 227, 2005.

DUTCOSKY, S. D. **Análise sensorial de alimentos**. Curitiba: Champagnat, 2011. 426 p.

DIJKSTERHUIS, G. Assessing panel consonance. **Food Quality and Preference**, v. 6, p. 7 - 14, 1995.

FERNANDES, F. M. O. **Proposta de um Teste Monte Carlo para unidimensionalidade de painéis sensoriais**. 80 f. Tese (Doutorado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2012.

FERREIRA, D. F. **Estatística Multivariada**. 1<sup>a</sup> ed. Lavras: Editora UFLA, 2008a. 662 p.

FERREIRA, D. F. **Estatística Multivariada**. 2<sup>a</sup> ed. Lavras: Editora UFLA, 2011. 676 p.

FERREIRA, D. F. SISVAR: um programa para análises e ensino de estatística. **Revista Symposium**, (Lavras), v. 6, p. 36-41, 2008b.

FERREIRA, E. B.; OLIVEIRA, M; S. **Sensometria**: uma abordagem com ênfase em Procrustes. Santa Maria: UFSM, 2007. 71 p. Apostila.

GEBERT, D.M.P. **Proposta de testes *bootstrap* para inferir sobre o número de componentes principais retidos**. 2010. 75 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária)- Universidade Federal de Lavras, Lavras, 2010.

GIRSHCK, M. A. On the sampling theory of roots of determinantal equations. **The Annals of Mathematical Statistics**, v. 10, p. 203 - 224, 1939.

HAMMERSLEY, J. M.; HANDSCOMB, D. C. **Monte Carlo methods**. London: Methuen & Co ltd, 1975. 178 p.

HARRELL, C. R.; GOHSH, B. K.; BOWDEN, R. **Simulation using ProModel**. Boston: McGraw-Hill, 2000. 603 p.

HEFFERON, J. **Linear algebra**. 2014. 498 p. Disponível em: <<http://joshua.smcvt.edu/linearalgebra>>. Acesso em: 14 mar. 2016

HOTELLING, H. Analysis of a complex of statistical variables into principal components. **Journal of Educational Psychology**, Arlington, v. 24, p. 417 - 441, 1933.

ISAAC, V. et al. Análise sensorial como ferramenta útil no desenvolvimento de cosméticos. **Revista de Ciências Farmacêuticas Básica e Aplicada**, v. 33, n. 4, p. 479 - 488, 2012.

JOHNSON, R. A.; WHICHREN, D. W. **Applied multivariate statistical analysis**. 6th ed. New Jersey: Pearson Prentice Hall, 2007. 773 p.

LATREILLE, J. et al. Measurement of the reliability of sensory panel performances. **Food**

**Quality and Preference**, Barking, n. 17, p. 369 - 365, 2006.

LAWLEY, D. N. Tests of significance of the latent roots and correlation matrices. **Biometrika**, London, v. 43, p. 128 - 136, 1956.

McEWAN, J. A. et al. Proficiency testing for sensory profile panels: measuring panel performance. **Food Quality and Preference**, Barking, n. 13, p. 181 - 190, 2002.

MINIM, V. P. R. **Análise sensorial: estudos com consumidores**. Viçosa: Editora UFV, 2006. 225p.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: UFMG, 2005. 297p.

MONTEIRO, C. L. B. Análise sensorial - seleção e treinamento de equipes de degustadores. **Boletim do Centro de Pesquisa de Processamento de Alimentos**, Curitiba, v. 2, n. 1, 1984.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. 3rd ed. Singapore: McGraw-Hill, 1974. 564 p.

PEARSON, K. On lines and planes of closest fit to systems of points in space **Philosophical Magazine**, London, v. 6, n. 2, p. 559 - 572, 1901.

PINTO, F. S. T., et al. A method for panelists consistency assessment in sensory evaluations based on the Cronbach's alpha coefficient. **Food Quality and Preference**, Barking, v. 32, p. 41 - 47, 2014.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2016. Software.

ROCHA, M. C. **Testes baseados em correlação canônica para avaliar consonância de painéis sensoriais**. 71 f. Dissertação (Mestrado em Estatística Aplicada e Biometria) - Universidade Federal de Alfenas, Alfenas, 2015.

STONE, H.; SIDEL, J.L. **Sensory Evaluation Practices**. 3<sup>a</sup> ed. San Diego, CA: Elsevier Academic Press, 2004. 377 p.

TEIXEIRA, L. V. Análise sensorial na indústria de alimentos. **Revista do Instituto de Laticínios “Cândido Tostes”**, v. 64, n. 366, p. 12 - 21, 2009.

## APÊNDICES

APÊNDICE A - Desenvolvimento algébrico das restrições 3.9 e 3.11 descritas na seção 3.2.

Na seção 3.2 determina-se duas restrições para correlação entre os provedores ( $\rho$ ). Para encontrar a restrição 3.9 parte-se da pressuposição de que a matriz  $\Sigma_p$  deve ser positiva definida e para isto os os autovalores devem ser positivos. Assim, partindo da equação 3.7, tem-se que

$$\lambda_1 = \sigma^2[1 + (p - 1)\rho] > 0.$$

Como é fixado  $\sigma^2 = 1$ , então

$$\begin{aligned} 1 + (p - 1)\rho &> 0 \\ \Rightarrow (p - 1)\rho &> -1 \\ \Rightarrow \rho &> \frac{-1}{p - 1}. \end{aligned}$$

Agora, garantir que a matriz  $\Sigma$  seja positiva definida a proporção da variação total também deve ser positiva. Logo, para determinar a restrição 3.11 utiliza-se a proporção acumulada da explicação da variação total que pode ser estimada por:

$$\rho^2 = \frac{\lambda_1}{\sum_{i=1}^p \lambda_i}$$

Daí,

$$\begin{aligned} \rho^2 &= \frac{1 + (p - 1)\rho}{\lambda_1 + \sum_{i=2}^p \lambda_i} \\ &= \frac{1 + (p - 1)\rho}{1 + (p - 1)\rho + (p - 1)(1 - \rho)} \\ &= \frac{1 + (p - 1)\rho}{1 + p\rho - \rho + p - p\rho - 1 + \rho} \\ &= \frac{1 + (p - 1)\rho}{p} \end{aligned}$$

Logo,

$$\rho^2 = \frac{1 + (p - 1)\rho}{p}.$$

Fazendo  $\rho^2 > 0$  tem-se que

$$\begin{aligned}\frac{1 + (p - 1)\rho}{p} &> 0 \\ \Rightarrow 1 + (p - 1)\rho &> 0 \\ \Rightarrow (p - 1)\rho &> -1 \\ \Rightarrow \rho &> \frac{-1}{(p - 1)}\end{aligned}$$

Logo, as restrições para  $\rho$  coincidem e estão definidas.

APÊNDICE B - Gráficos das curvas de poderes dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para os casos com  $p = 2$  provadores

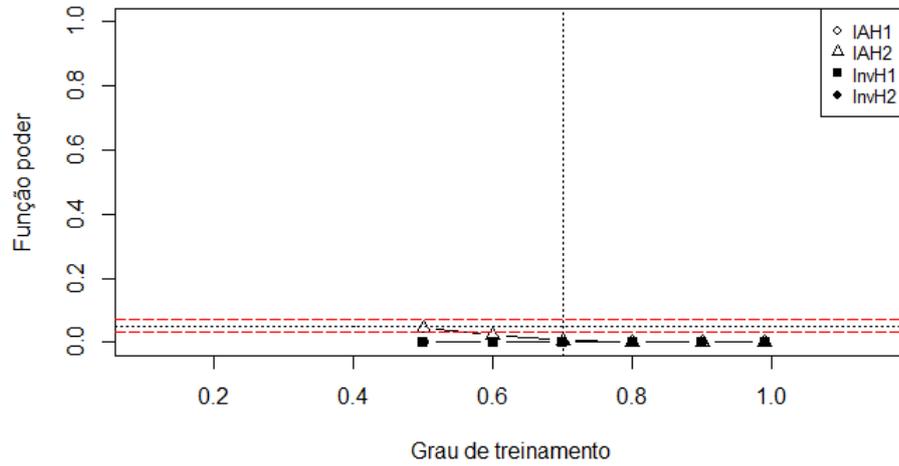


Figura 14 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 2$  atributos e  $n = 20$  observações.

Fonte: da autora.

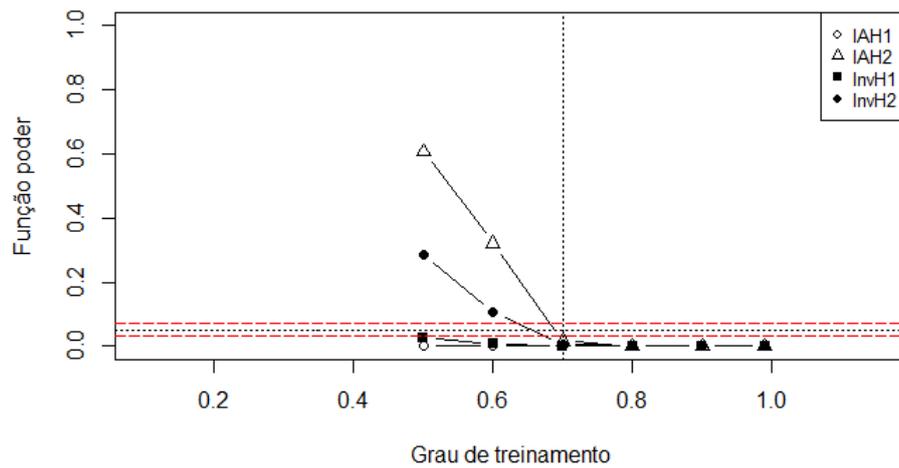


Figura 15 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 2$  atributos e  $n = 40$  observações.

Fonte: da autora.

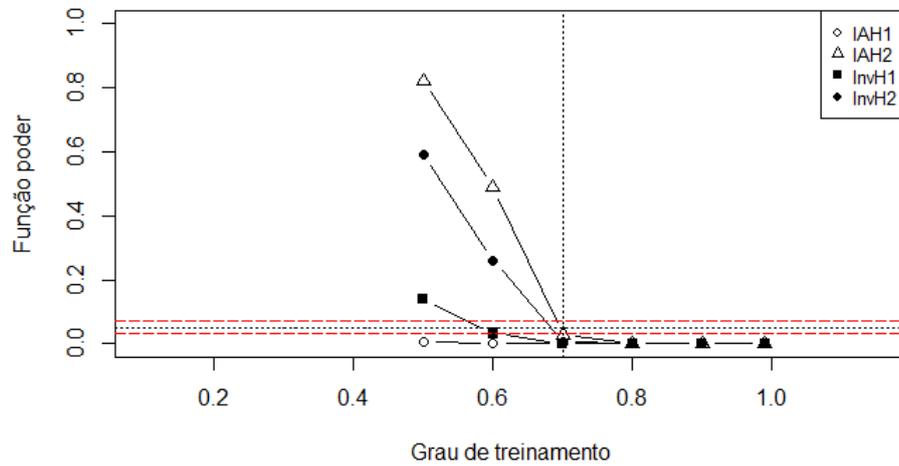


Figura 16 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 2$  atributos e  $n = 50$  observações.

Fonte: da autora.

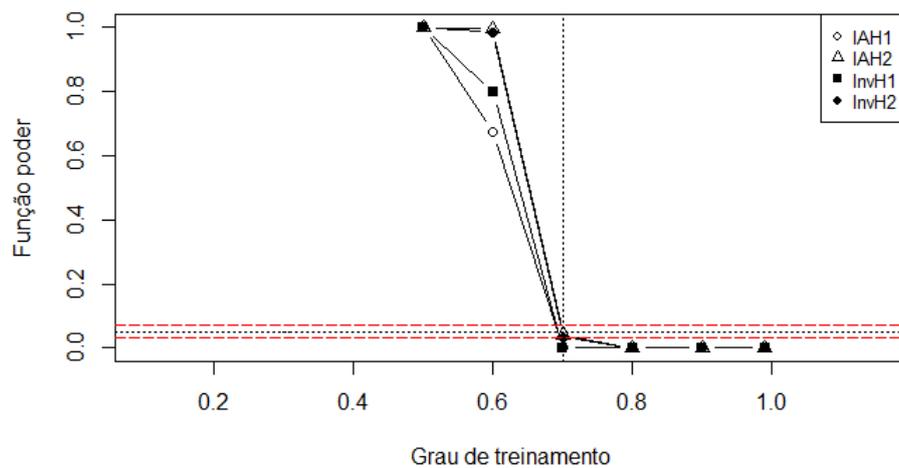


Figura 17 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 2$  atributos e  $n = 200$  observações.

Fonte: da autora.

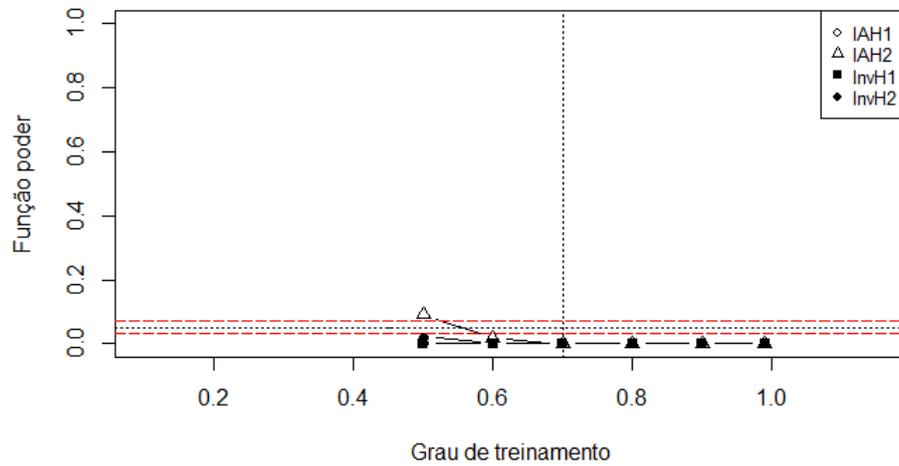


Figura 18 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 5$  atributos e  $n = 50$  observações.

Fonte: da autora.

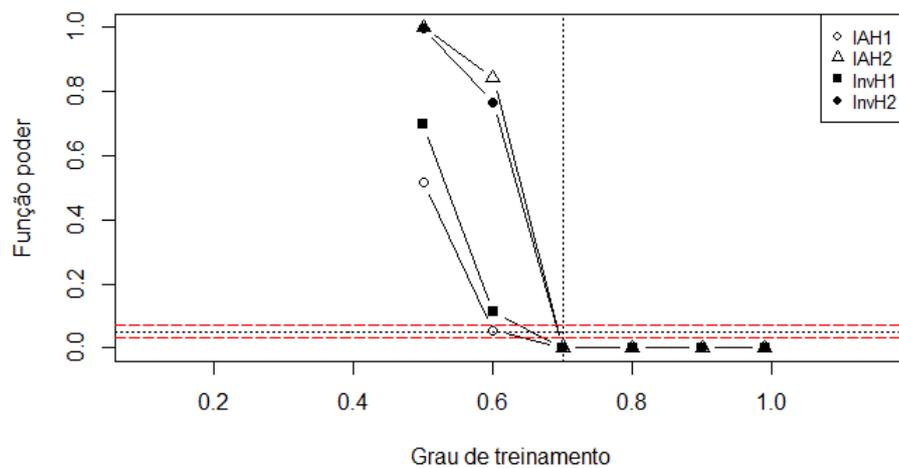


Figura 19 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 5$  atributos e  $n = 100$  observações.

Fonte: da autora.

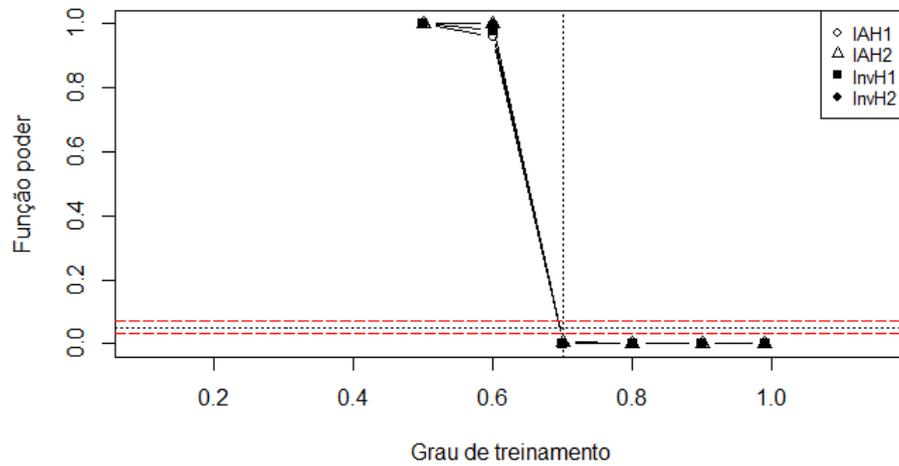


Figura 20 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 5$  atributos e  $n = 200$  observações.

Fonte: da autora.

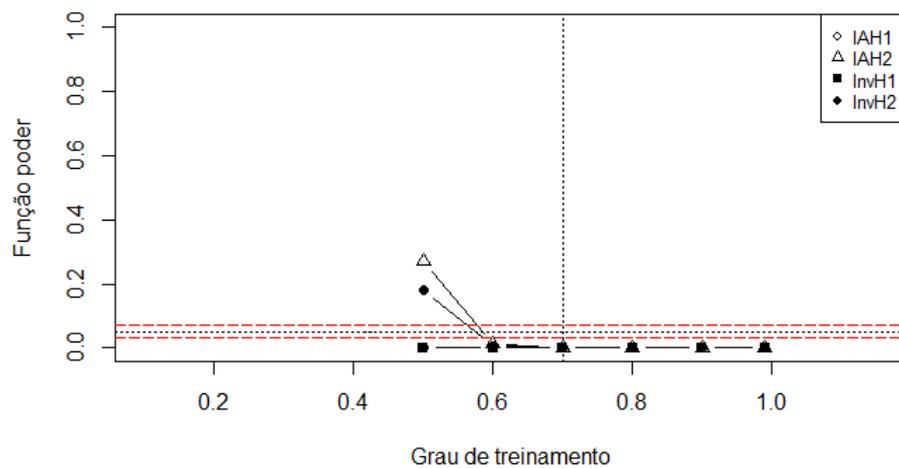


Figura 21 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 10$  atributos e  $n = 100$  observações.

Fonte: da autora.

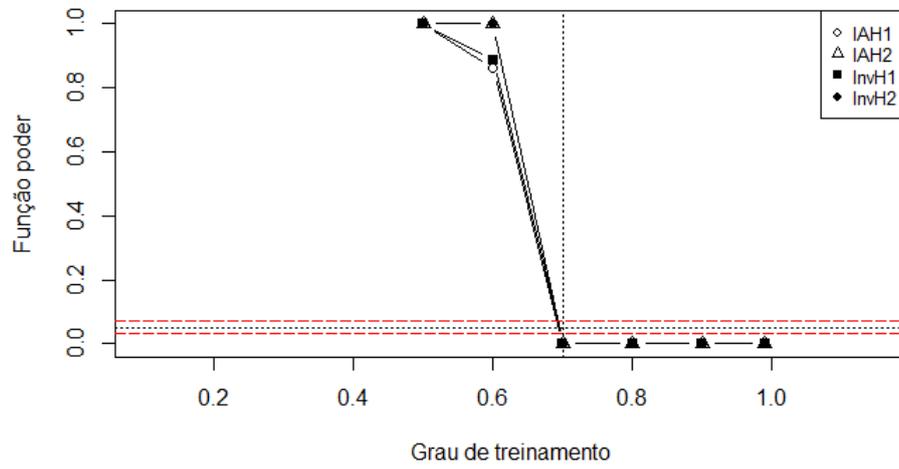


Figura 22 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 10$  atributos e  $n = 200$  observações.

Fonte: da autora.

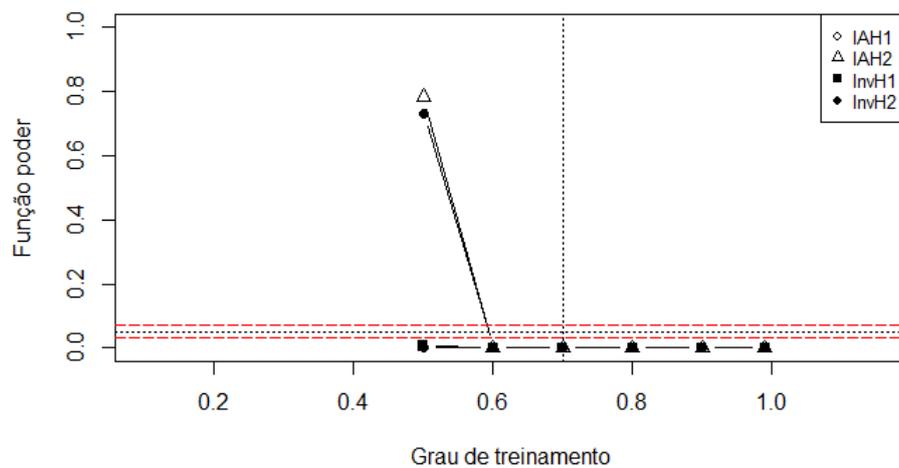


Figura 23 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 2$  provadores,  $q = 20$  atributos e  $n = 200$  observações.

Fonte: da autora.

APÊNDICE C - Gráficos das curvas de poderes dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para os casos com  $p = 5$  provadores

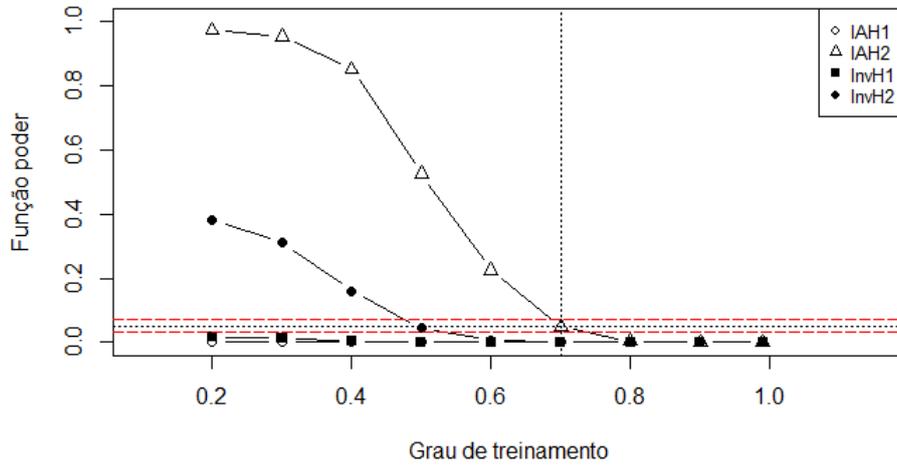


Figura 24 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provadores,  $q = 2$  atributos e  $n = 10$  observações.

Fonte: da autora.

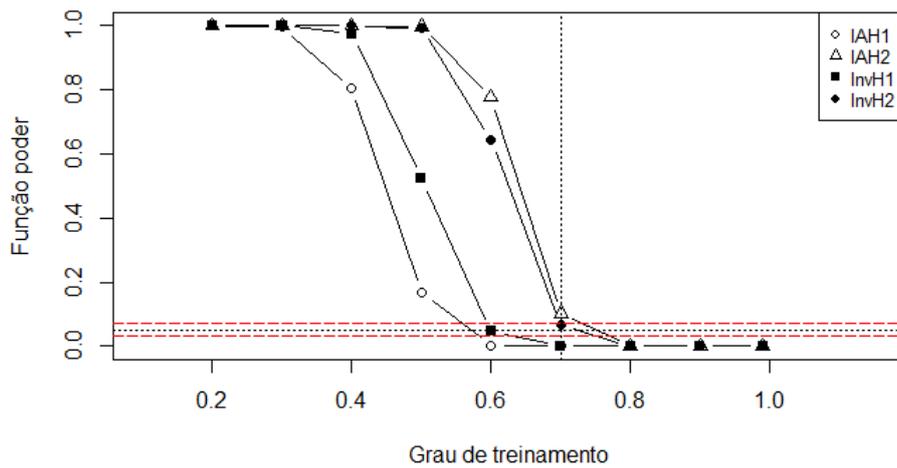


Figura 25 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provadores,  $q = 2$  atributos e  $n = 30$  observações.

Fonte: da autora.

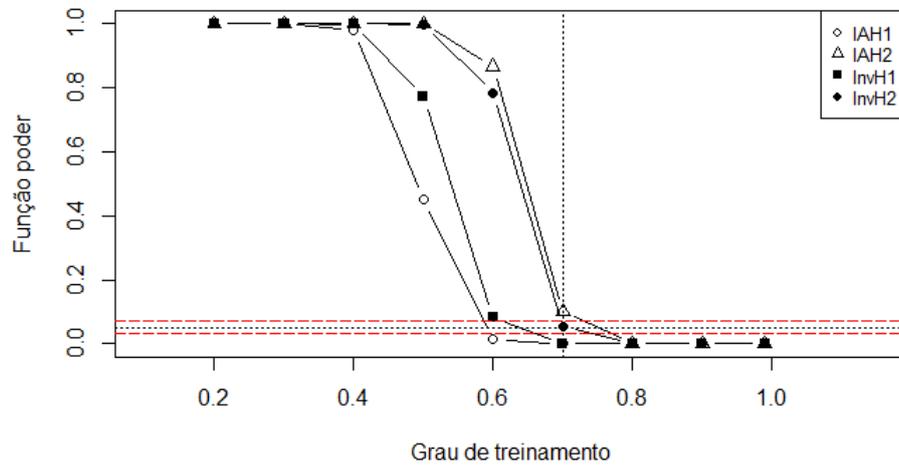


Figura 26 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 2$  atributos e  $n = 40$  observações.  
Fonte: da autora.

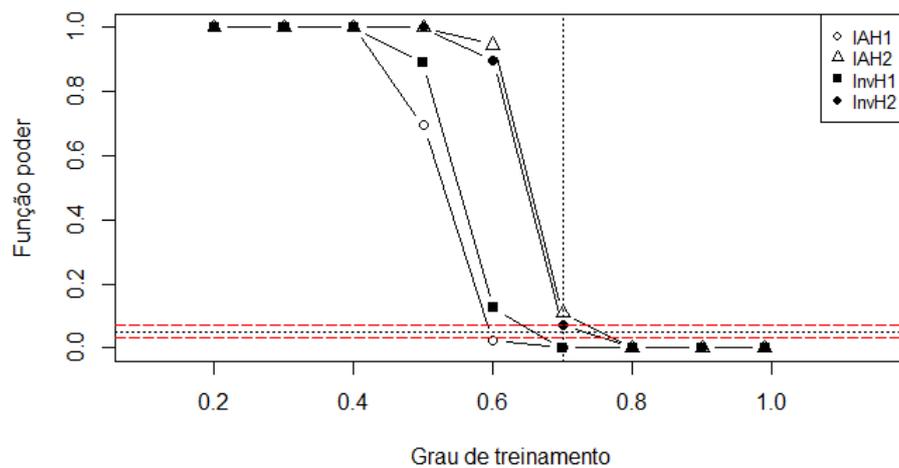


Figura 27 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 2$  atributos e  $n = 50$  observações.  
Fonte: da autora.

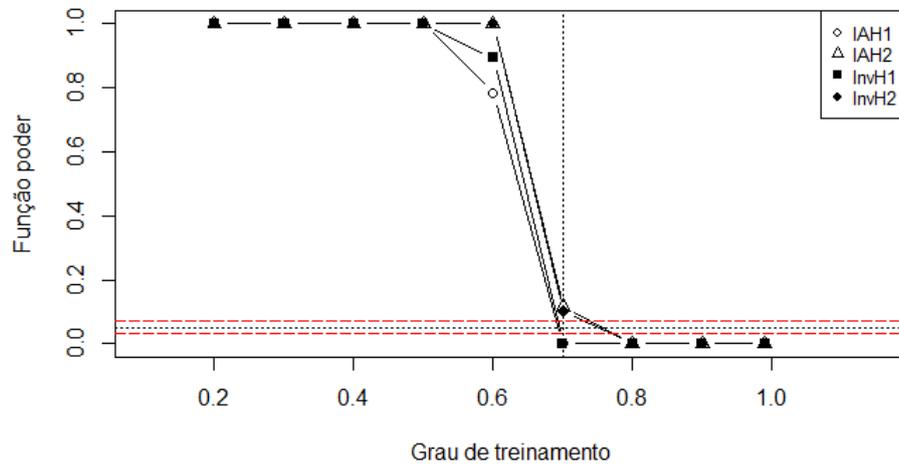


Figura 28 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 2$  atributos e  $n = 200$  observações.

Fonte: da autora.

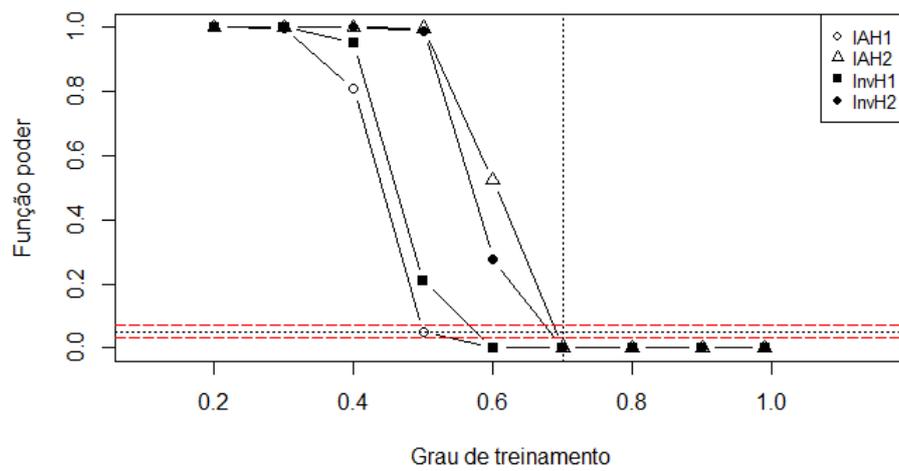


Figura 29 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 5$  atributos e  $n = 30$  observações.

Fonte: da autora.

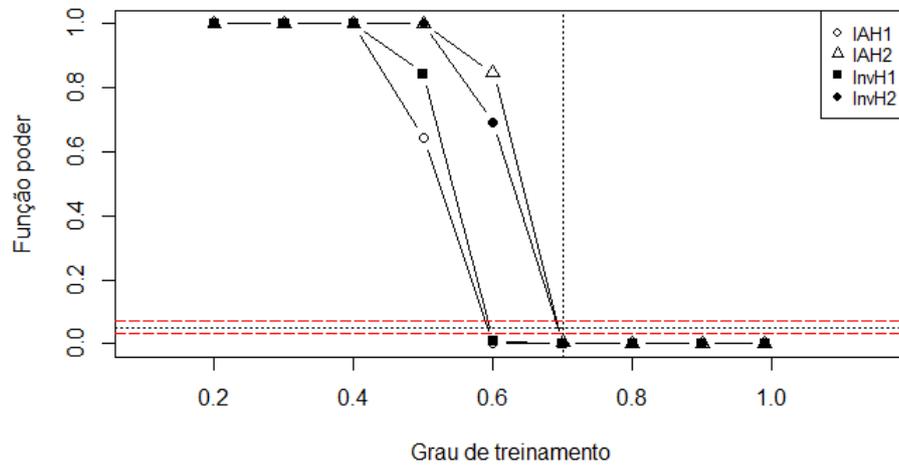


Figura 30 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 4$  atributos e  $n = 40$  observações.  
Fonte: da autora.

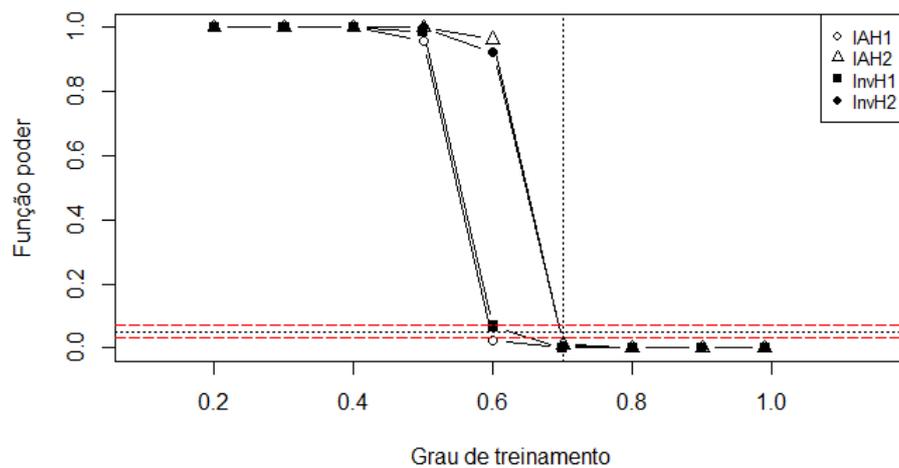


Figura 31 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 5$  atributos e  $n = 50$  observações.  
Fonte: da autora.

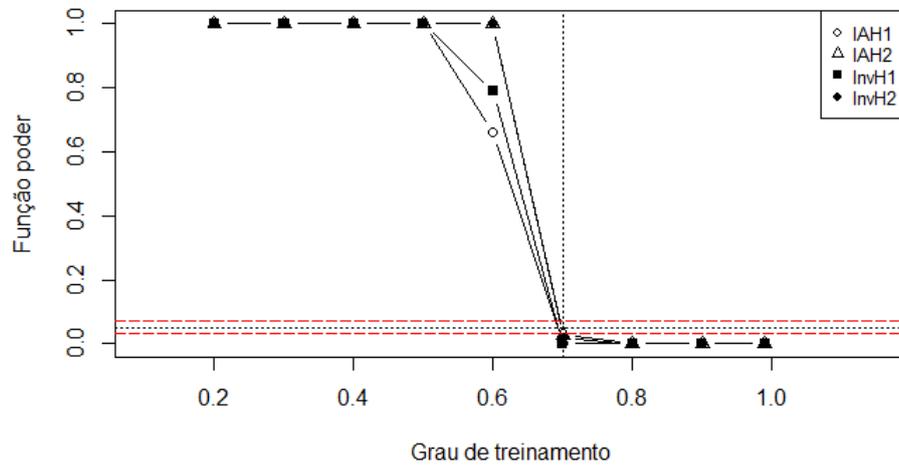


Figura 32 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provadores,  $q = 5$  atributos e  $n = 100$  observações.

Fonte: da autora.

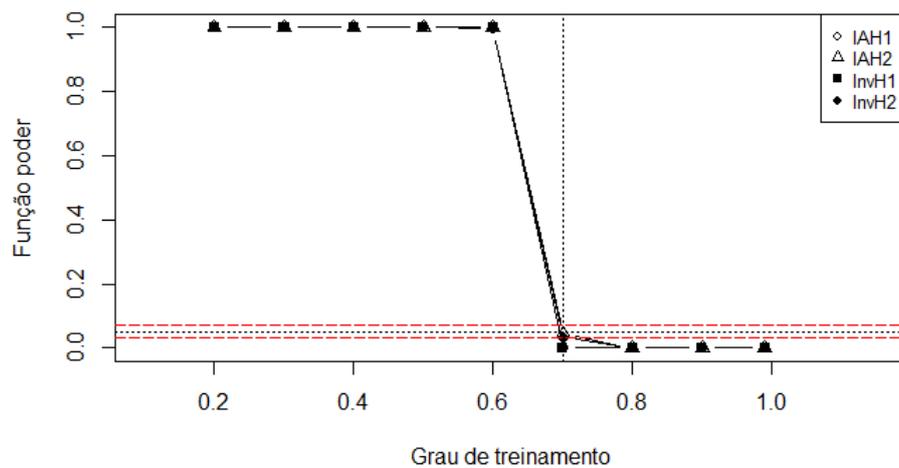


Figura 33 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provadores,  $q = 5$  atributos e  $n = 200$  observações.

Fonte: da autora.

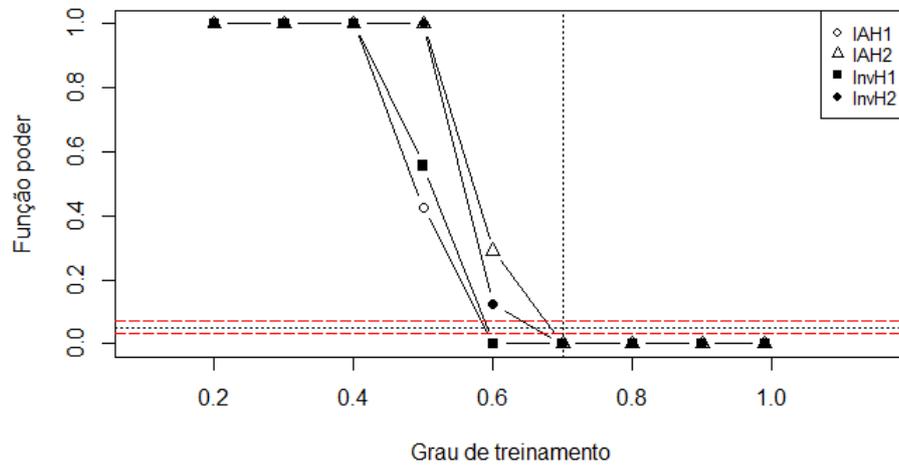


Figura 34 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 10$  atributos e  $n = 50$  observações.

Fonte: da autora.

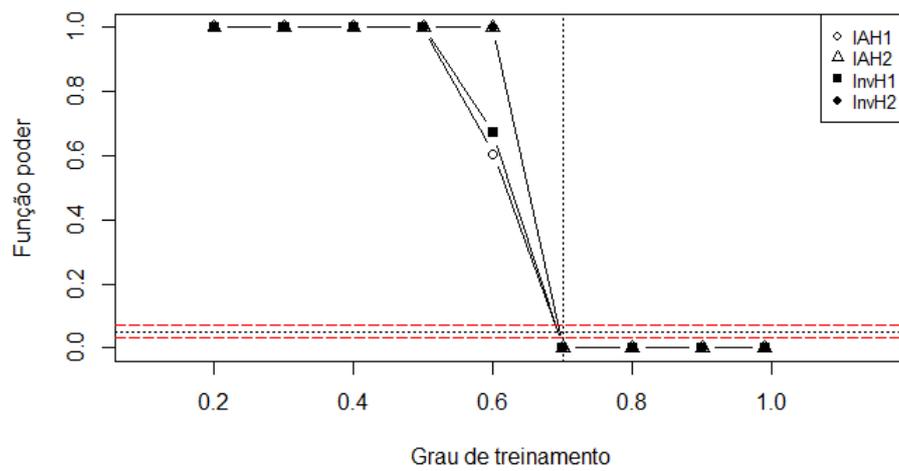


Figura 35 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provedores,  $q = 10$  atributos e  $n = 100$  observações.

Fonte: da autora.

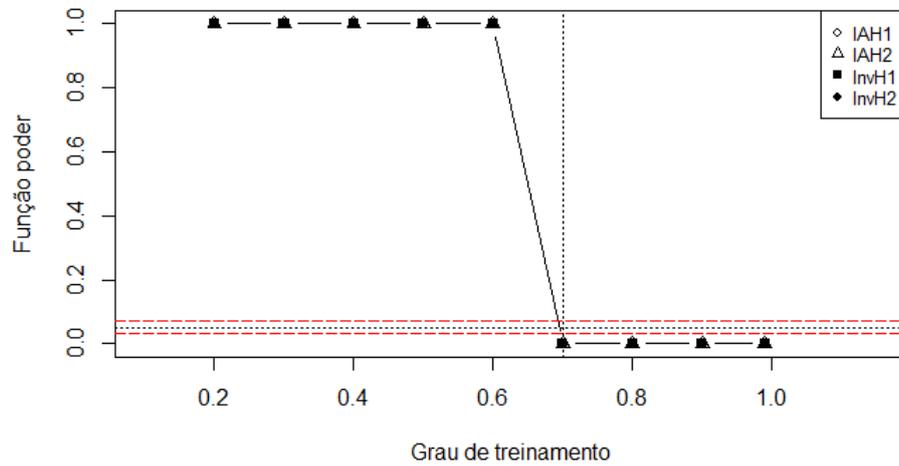


Figura 36 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provadores,  $q = 10$  atributos e  $n = 200$  observações.

Fonte: da autora.

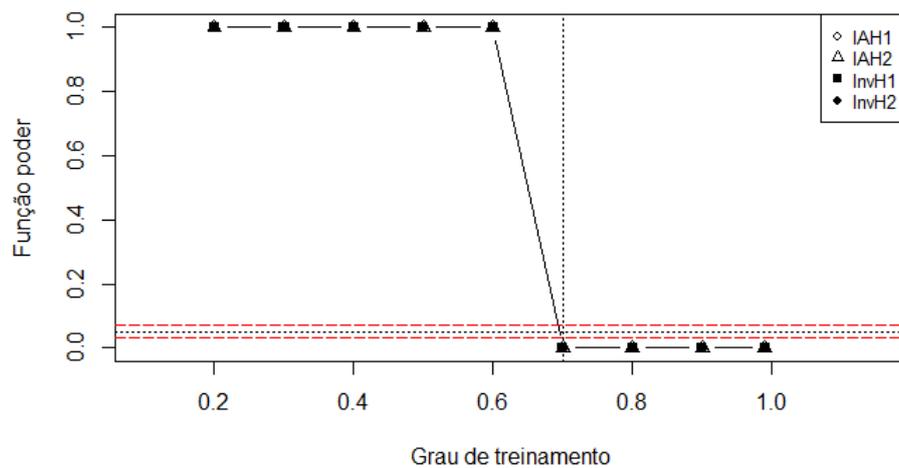


Figura 37 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 5$  provadores,  $q = 20$  atributos e  $n = 200$  observações.

Fonte: da autora.

APÊNDICE D - Gráficos das curvas de poderes dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para os casos com  $p = 10$  provadores

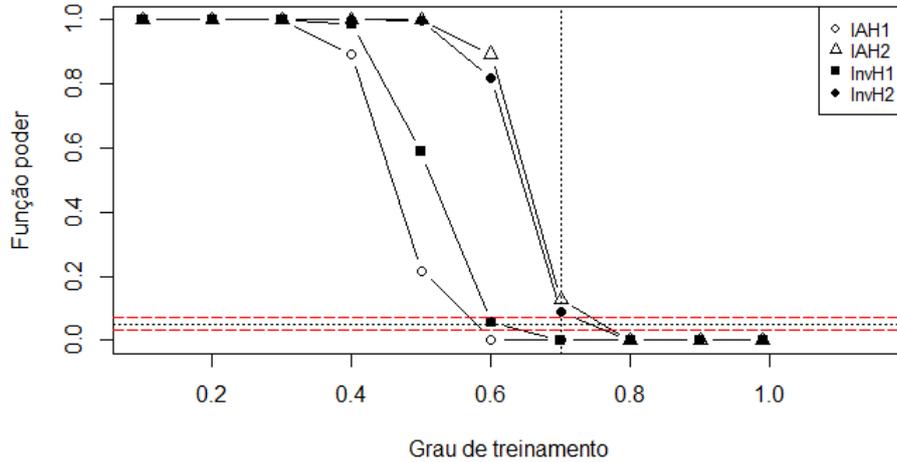


Figura 38 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 2$  atributos e  $n = 30$  observações.

Fonte: da autora.

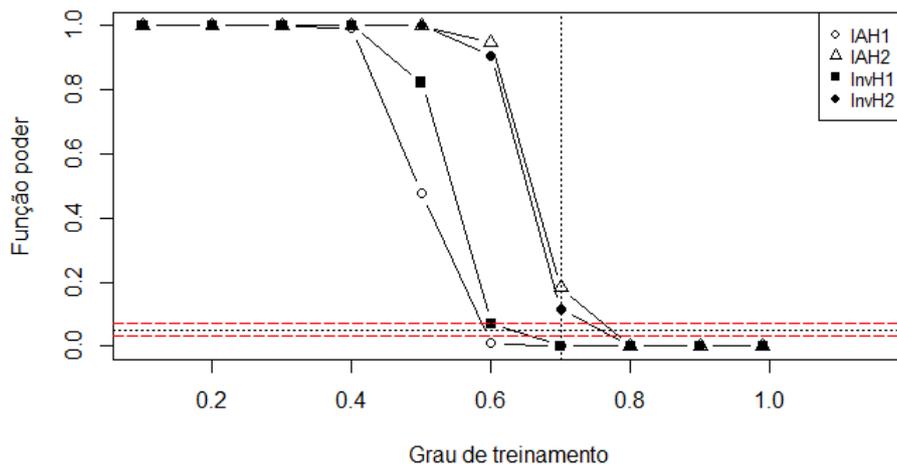


Figura 39 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 2$  atributos e  $n = 40$  observações.

Fonte: da autora.

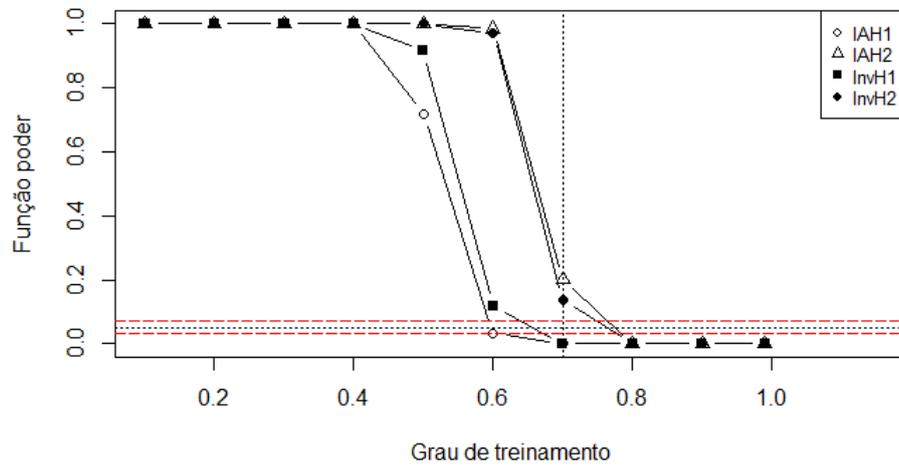


Figura 40 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 2$  atributos e  $n = 50$  observações.

Fonte: da autora.

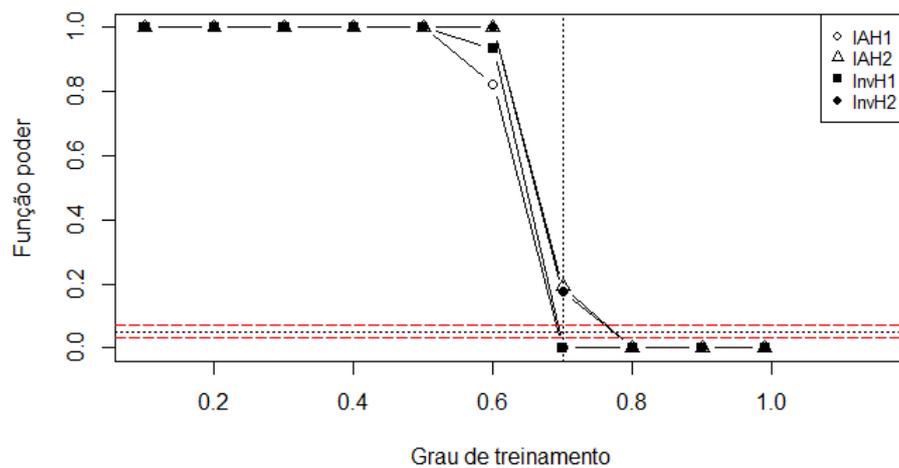


Figura 41 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 2$  atributos e  $n = 200$  observações.

Fonte: da autora.

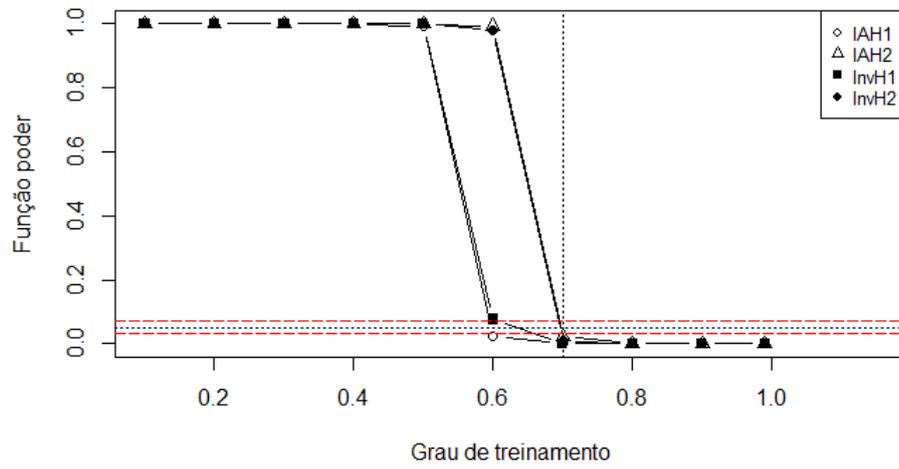


Figura 42 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 5$  atributos e  $n = 50$  observações.

Fonte: da autora.

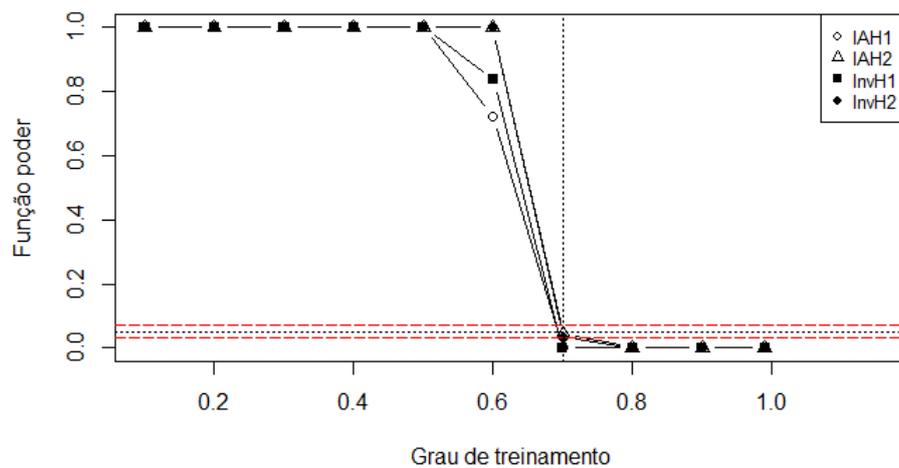


Figura 43 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 5$  atributos e  $n = 100$  observações.

Fonte: da autora.

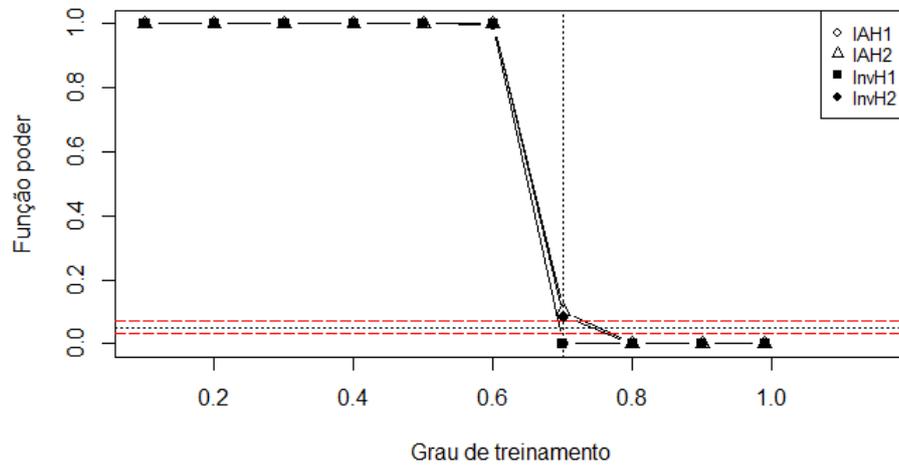


Figura 44 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 5$  atributos e  $n = 200$  observações.

Fonte: da autora.

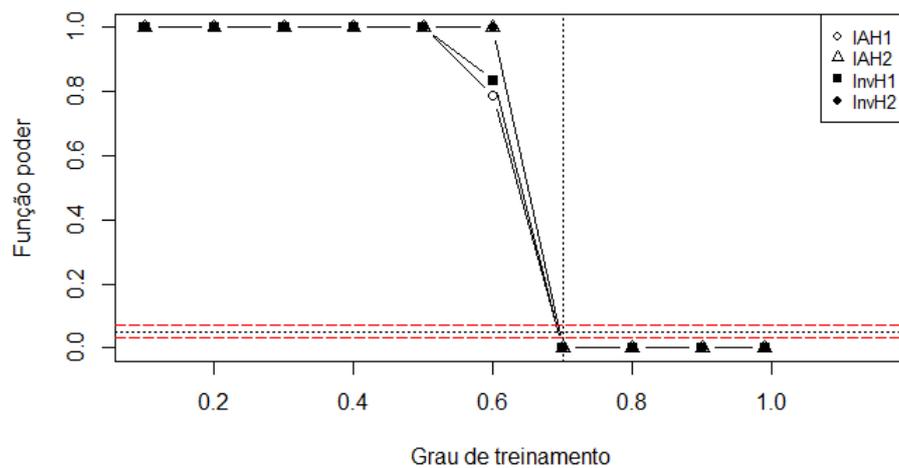


Figura 45 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provadores,  $q = 10$  atributos e  $n = 100$  observações.

Fonte: da autora.

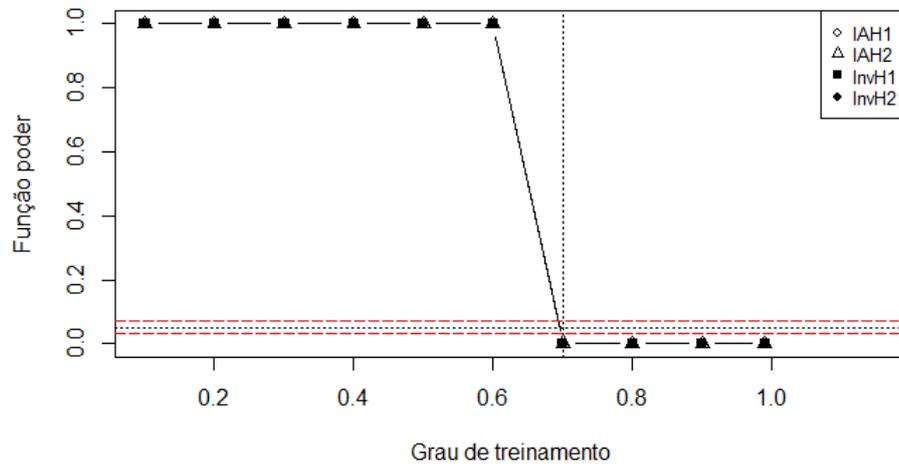


Figura 46 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provedores,  $q = 10$  atributos e  $n = 200$  observações.  
Fonte: da autora.

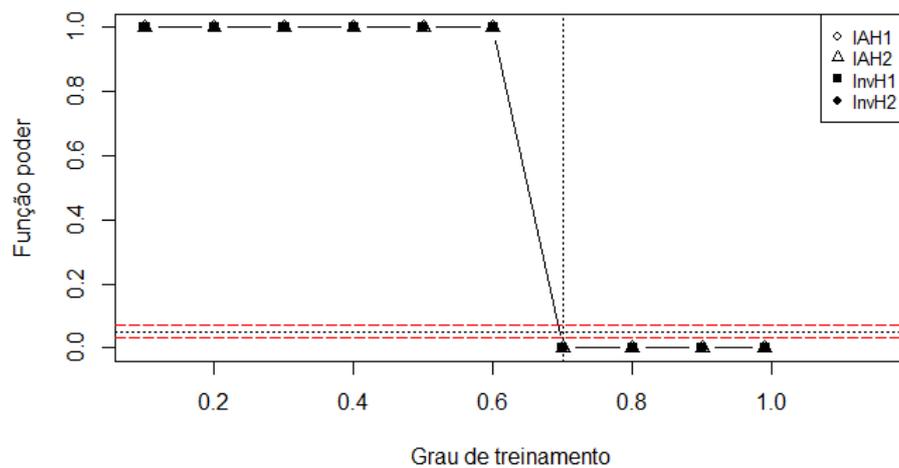


Figura 47 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 10$  provedores,  $q = 20$  atributos e  $n = 200$  observações.  
Fonte: da autora.

APÊNDICE E - Gráficos das curvas de poderes dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para os casos com  $p = 15$  provadores

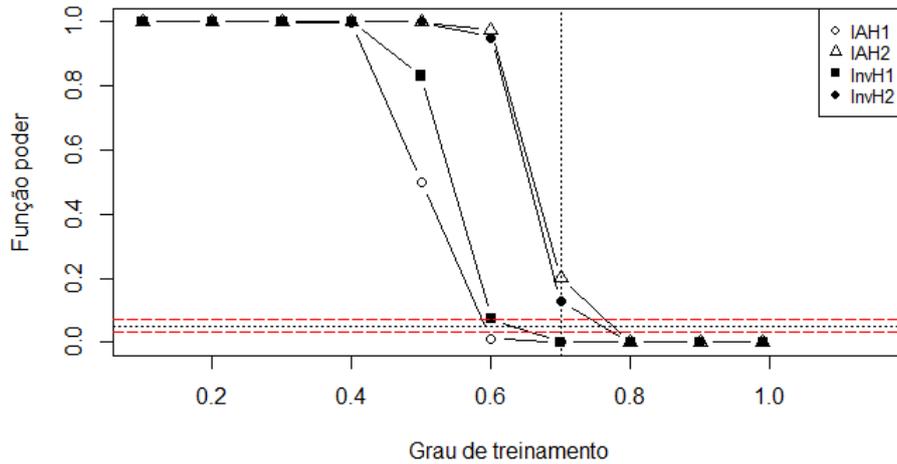


Figura 48 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 15$  provadores,  $q = 2$  atributos e  $n = 40$  observações.  
Fonte: da autora.

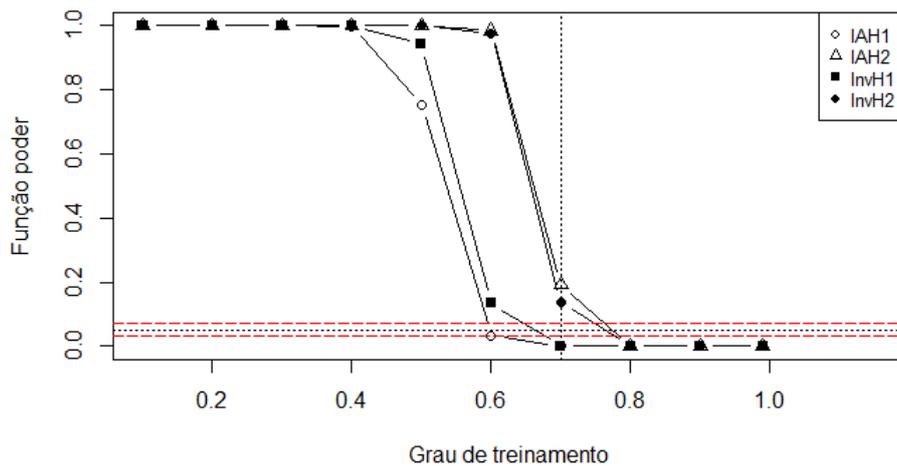


Figura 49 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 15$  provadores,  $q = 2$  atributos e  $n = 50$  observações.  
Fonte: da autora.

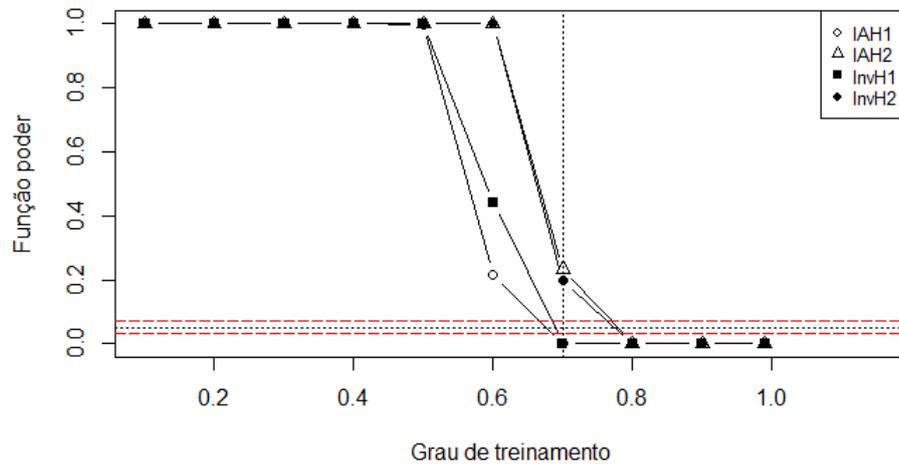


Figura 50 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 15$  provadores,  $q = 2$  atributos e  $n = 100$  observações.

Fonte: da autora.

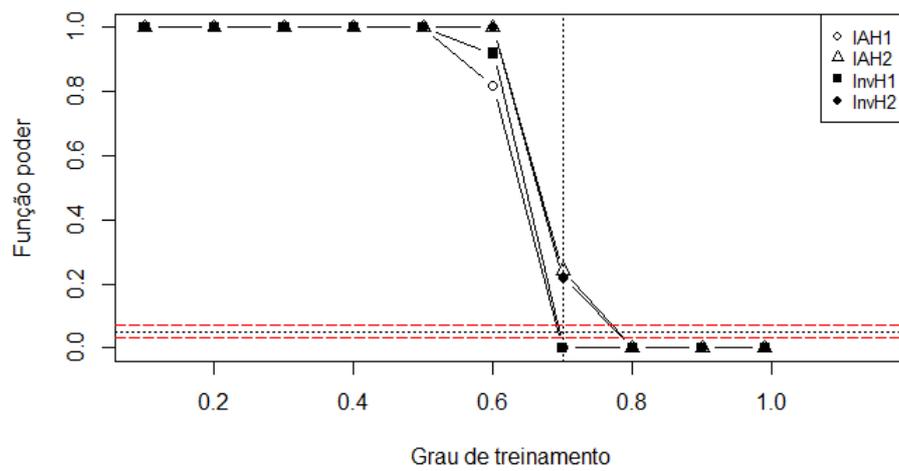


Figura 51 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 15$  provadores,  $q = 2$  atributos e  $n = 200$  observações.

Fonte: da autora.

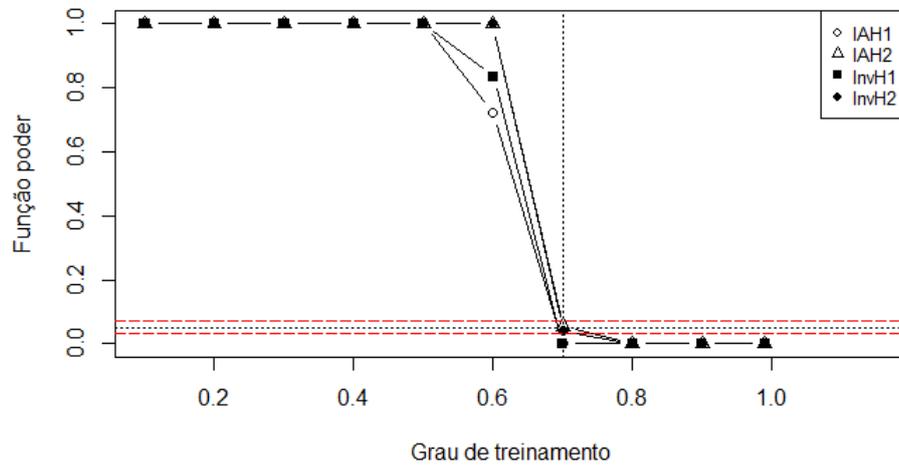


Figura 52 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 15$  provadores,  $q = 5$  atributos e  $n = 100$  observações.

Fonte: da autora.

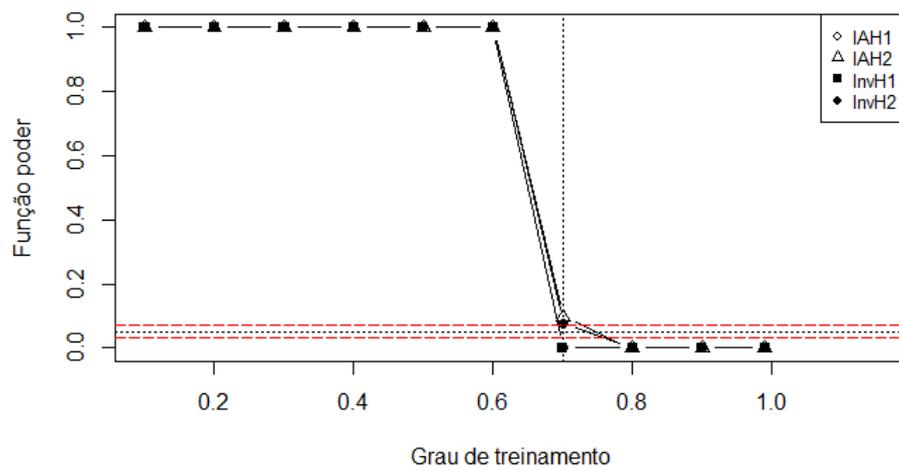


Figura 53 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 15$  provadores,  $q = 5$  atributos e  $n = 200$  observações.

Fonte: da autora.

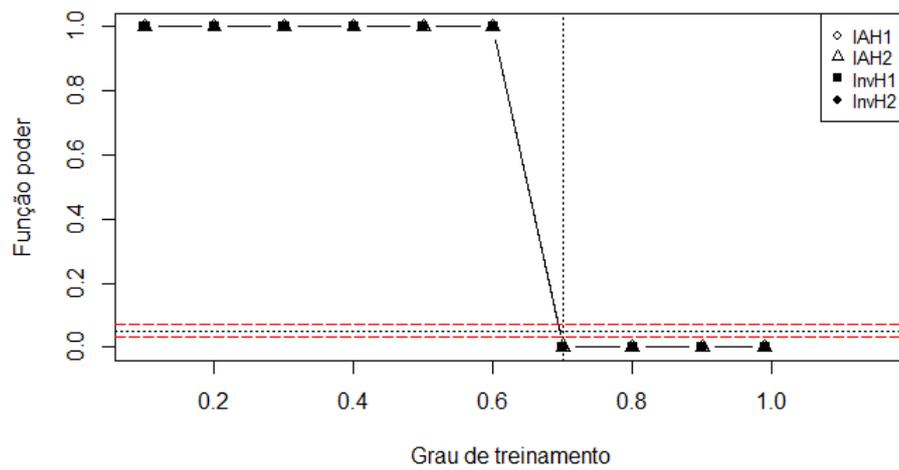


Figura 54 – Poder dos testes  $IAH_1$ ,  $IAH_2$ ,  $InvH_1$  e  $InvH_2$  para o caso  $p = 15$  provedores,  $q = 10$  atributos e  $n = 200$  observações.  
Fonte: da autora.