

UNIVERSIDADE FEDERAL DE ALFENAS

Iasmine Queiroga de Paula

**Testes para a seleção de provedores com base na proporção
de acertos ao longo de ensaios triangulares**

ALFENAS, MG

2020

IASMINE QUEIROGA DE PAULA

Testes para a seleção de provadores com base na proporção de acertos ao longo de ensaios triangulares

Dissertação apresentada ao Programa de Pós-Graduação em Estatística Aplicada e Biometria, Área de concentração em Estatística Aplicada e Biometria da Universidade Federal de Alfenas, MG, como parte dos requisitos para a obtenção do título de Mestre.

Linha de pesquisa: Estatística Aplicada e Biometria.
Orientador: Prof. Dr. Eric Batista Ferreira

ALFENAS, MG

2020

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal de Alfenas

Paula, Iasmirine Queiroga de .
P324t Testes para a seleção de provedores com base na proporção de
acertos ao longo de ensaios triangulares / Iasmirine Queiroga de Paula
– Alfenas/MG, 2020.
115 f.: il. -

Orientador: Eric Batista Ferreira.
Dissertação (Mestrado em Estatística Aplicada e Biometria)
Universidade Federal de Alfenas, 2020.
Bibliografia.

1. Análise sensorial. 2. Binomial. 3. Estatística. I. Ferreira, Eric
Batista. II. Título.

CDD-519.5



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 37130-001
Fone: (35) 3701-9600
<https://www.unifal-mg.edu.br/ppgeab/>



IASMINE QUEIROGA DE PAULA

“TESTES PARA A SELEÇÃO DE PROVADORES COM BASE NA PROPORÇÃO DE ACERTOS AO LONGO DE ENSAIOS TRIANGULARES”

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: 30 de janeiro de 2020.

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

Assinatura: Eric Batista Ferreira

Profa. Dra. Flávia Della Lucia
Instituição: UNIFAL-MG

Assinatura: Flávia Della Lucia

Profa. Dra. Juliana Petrini
Instituição: UNIFAL-MG

Assinatura: Juliana Petrini

AGRADECIMENTOS

Agradeço a Deus por toda benção recebida, por mais uma etapa que se concretiza e me transforma. Por sempre trilhar meus caminhos na direção certa!

Ao Prof. Eric, por me receber como orientador. Pela oportunidade em trabalhar com Sensometria e Estatística Computacional, o que me permitiu imergir nesse Universo e aprofundar mais sobre o assunto.

Às Prof^{as}. Juliana e Flávia, por aceitarem participar dessa etapa importante para mim. Por se dedicarem a analisar meu trabalho e contribuir ainda mais com o meu desenvolvimento junto a ele. Agradeço também a Prof^a. Rosemary e ao Dr. Éderson por acrescentarem e colaborarem com a melhoria do trabalho.

Ao Prof. Fabrício, por todas as palavras de incentivo nos momentos mais necessários, que me ajudaram a seguir adiante com mais leveza. Ao Prof. Luiz Beijo, por me ensinar a ver o conhecimento de outra forma, o que me motivou a buscá-lo ainda mais e trabalhá-lo em mim mesma.

Aos Profs. Denismar e Natália, por participarem da minha formação e trajetória. Agradeço por estarem à disposição sempre que foi preciso.

À Martha, por todo carisma! Agradeço por ter tornado nossos dias ainda melhores com sua presença e energia positiva.

Aos meus pais e meu irmão, que mesmo com a distância física, sempre estiveram presentes em todos os momentos. Obrigada por vivenciarem comigo cada dia, dando todo apoio em mais uma etapa da minha vida!

RESUMO

Os consumidores têm se tornado cada vez mais críticos e exigentes quando se trata da escolha dos alimentos. Buscando aprimorar a qualidade desses produtos, muitos estudos que envolvem a pesquisa sensorial se tornam relevantes. A análise sensorial é um procedimento experimental que permite medir, analisar e interpretar através dos sentidos reações estimuladas pelos alimentos. Essa análise é realizada através de painéis sensoriais, que constituem pessoas responsáveis por analisar sensorialmente o produto. Para selecionar degustadores e submetê-los ao treinamento, testes triangulares são comumente utilizados a fim de avaliar a capacidade discriminativa dos juízes. Deseja-se cometer o menor erro possível durante a seleção, levando em consideração que esses avaliadores podem ter uma proporção de acerto constante ou não. Nesta última situação, os juízes podem adquirir queda da proporção de acertos (devido à fadiga), ou até mesmo apresentarem elevação da proporção de acertos (aprendizagem). Dessa forma, o objetivo deste trabalho é fazer um estudo de estimadores intervalares, reescrevendo-os como testes a serem aplicados para seleção de provadores, a fim de recomendar o melhor em termos de menor taxa de erro tipo I e maior poder. Para isso, foram realizadas simulações de Monte Carlo considerando provadores que acertam a uma proporção (p) constante ao longo dos ensaios (n) e também aqueles que desenvolvem fadiga ou aprendem de forma linear e não linear, avaliando o comportamento dos testes com a quebra de pressuposição (p constante). Adicionalmente, foram modelados dados reais para candidatos submetidos à análise olfativa de café, com o uso da regressão segmentada, que permite verificar pontos de mudança, auxiliar a decisão do teste e recomendar número de ensaios a serem aplicados. Dos seis testes analisados, obtidos de estimadores da Normal (TN1, TN2 e TN3), da Distribuição F (TF), do Teste Sequencial (TS) e da Poisson (TP), o TS apresentou menor erro tipo I e maior poder para proporção constante e $n < 20$, enquanto TN3 apresentou essas características para $n \geq 20$. Com a quebra da pressuposição de p constante, observou-se que as taxas de erro tipo I e poder são bastante afetadas. Para o cansaço, há uma queda no poder, devido ao pequeno valor de n quando p aumenta, enquanto na aprendizagem há o aumento do poder simultâneo a n e p . Nota-se que a não linearidade das funções afeta a velocidade de alcance do poder final. Já a regressão segmentada é uma grande ferramenta a ser utilizada para se tomar uma decisão quando a proporção de acertos converge ou reduz a variabilidade. A recomendação foi de 15 ensaios para análise olfativa de café.

Palavras-chave: Análise sensorial. Binomial. Estatística.

ABSTRACT

Consumers have become increasingly critical and demanding when it comes to food choice. Seeking to improve the quality of these products, many studies involving sensory research become relevant. Sensory analysis is an experimental procedure that allows to measure, analyze and interpret reactions stimulated by food. This analysis is performed through sensory panels, formed by people responsible for sensorially analyzing the product. To select tasters and subject them to training, triangular tests are commonly used in order to assess the discriminative ability of judges. It's desired to make the smallest possible error during selection, considering that these evaluators may have a constant hit proportion or not. In the latter situation, judges hit ratio decrease (due to fatigue), or increase (due to learning). Thus, the aim of this paper was to study interval estimators, rewriting them as tests to be applied for tasters selection, in order to recommend the best in terms of the lowest type I error rate and the highest power. For this, Monte Carlo simulations were performed considering tasters who hit a constant proportion (p) throughout the trials (n) and also those tasters who developed fatigue or learn in a linear and nonlinear way, evaluating test behavior without assumption (p constant). Additionally, real data were modeled for candidates submitted to olfactory coffee analysis, using segmented regression, which allows to verify changing points, helping in test decision and recommending the number of trials to be applied. Of the six tests analyzed, obtained from Normal estimators (TN1, TN2 and TN3), F-Distribution (TF), Sequential Test (TS) and Poisson (TP), TS presented the lowest type I error and the highest power for constant ratio and $n < 20$, while TN3 showed these characteristics for $n \geq 20$. Without p constant assumption, type I error and power rates were very affected. For fatigue, there was a decrease in power, due to the small value of n when p increases, while with learning there was power increase simultaneous to n and p . Note that nonlinear functions affects the speed range of final power. Also, segmented regression is a great tool to be used to make a decision when the ratio of hits converges or reduces its variability. The recommendation was 15 trials for coffee olfactory analysis.

Keywords: Sensory evaluation. Binomial. Statistics.

LISTA DE FIGURAS

Figura 1 –	Figura representativa do teste triangular.	20
Figura 2 –	Gráfico representativo do teste sequencial triangular para seleção de provadores.	29
Figura 3 –	Fluxograma representativo da simulação MC.	32
Figura 4 –	Gráfico dos testes aplicados a um provador virtual com 0,7 de acerto.	46
Figura 5 –	Gráfico do poder para 5, 20, 50 e 100 ensaios.	47
Figura 6 –	Gráfico da taxa de decisão em relação a p , para diferente números de ensaios (5, 20, 50 e 100), no TS.	49
Figura 7 –	Gráfico da variação de pq em relação a p	50
Figura 8 –	Figura representativa do cansaço para o modelo linear.	61
Figura 9 –	Figura representativa do cansaço para o modelo não linear.	62
Figura 10 –	Figura representativa do aprendizado para o modelo linear.	63
Figura 11 –	Figura representativa do aprendizado para o modelo não linear.	64
Figura 12 –	Gráfico do poder para o modelo linear do cansaço.	65
Figura 13 –	Gráfico do poder para modelo não linear do cansaço.	66
Figura 14 –	Gráfico do poder para o modelo linear do aprendizado.	67
Figura 15 –	Gráfico do poder para o modelo não linear do aprendizado.	68
Figura 16 –	Exemplo de simulação para a lei fraca dos grandes números de uma Binomial.	75
Figura 17 –	Ficha de avaliação para teste triangular.	77
Figura 18 –	Aplicação do teste triangular.	78
Figura 19 –	Teste aplicado em dados obtidos pela análise sensorial de três candidatos, com o auxílio da regressão segmentada.	81
Figura 20 –	Gráficos obtidos para um candidato sem <i>breakpoint</i> e um candidato cuja proporção de acertos ainda não convergiu.	83
Figura 21 –	Simulação de candidato virtual com proporção de acertos convergente para $1/3$	84
Figura 22 –	Gráfico das taxas de rejeição de H_0 para 5 e 10 ensaios.	90
Figura 23 –	Gráfico das taxas de rejeição de H_0 para 15 e 20 ensaios.	90
Figura 24 –	Gráfico das taxas de rejeição de H_0 para 25 e 30 ensaios.	91

Figura 25 – Gráfico das taxas de rejeição de H_0 para 35 e 40 ensaios.	91
Figura 26 – Gráfico das taxas de rejeição de H_0 para 45 e 50 ensaios.	92
Figura 27 – Gráfico das taxas de rejeição de H_0 para 55 e 60 ensaios.	92
Figura 28 – Gráfico das taxas de rejeição de H_0 para 65 e 70 ensaios.	93
Figura 29 – Gráfico das taxas de rejeição de H_0 para 75 e 80 ensaios.	93
Figura 30 – Gráfico das taxas de rejeição de H_0 para 85 e 90 ensaios.	94
Figura 31 – Gráfico das taxas de rejeição de H_0 para 95 e 100 ensaios.	94
Figura 32 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 5 e 10 ensaios.	95
Figura 33 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 15 e 20 ensaios.	95
Figura 34 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 25 e 30 ensaios.	96
Figura 35 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 35 e 40 ensaios.	96
Figura 36 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 45 e 50 ensaios.	97
Figura 37 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 55 e 60 ensaios.	97
Figura 38 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 65 e 70 ensaios.	98
Figura 39 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 75 e 80 ensaios.	98
Figura 40 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 85 e 90 ensaios.	99
Figura 41 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 95 e 100 ensaios.	99
Figura 42 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 5 e 10 ensaios.	100
Figura 43 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 15 e 20 ensaios.	100

Figura 44 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 25 e 30 ensaios.	101
Figura 45 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 35 e 40 ensaios.	101
Figura 46 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 45 e 50 ensaios.	102
Figura 47 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 55 e 60 ensaios.	102
Figura 48 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 65 e 70 ensaios.	103
Figura 49 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 75 e 80 ensaios.	103
Figura 50 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 85 e 90 ensaios.	104
Figura 51 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 95 e 100 ensaios.	104
Figura 52 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 5 e 10 ensaios.	105
Figura 53 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 15 e 20 ensaios.	105
Figura 54 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 25 e 30 ensaios.	106
Figura 55 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 35 e 40 ensaios.	106
Figura 56 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 45 e 50 ensaios.	107
Figura 57 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 55 e 60 ensaios.	107
Figura 58 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 65 e 70 ensaios.	108
Figura 59 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 75 e 80 ensaios.	108

Figura 60 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 85 e 90 ensaios.	109
Figura 61 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 95 e 100 ensaios.	109
Figura 62 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 5 e 10 ensaios.	110
Figura 63 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 15 e 20 ensaios.	110
Figura 64 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 25 e 30 ensaios.	111
Figura 65 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 35 e 40 ensaios.	111
Figura 66 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 45 e 50 ensaios.	112
Figura 67 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 55 e 60 ensaios.	112
Figura 68 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 65 e 70 ensaios.	113
Figura 69 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 75 e 80 ensaios.	113
Figura 70 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 85 e 90 ensaios.	114
Figura 71 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 95 e 100 ensaios.	114
Figura 72 – Gráficos obtidos, respectivamente, para os candidatos 1 e 2.	115
Figura 73 – Gráficos obtidos, respectivamente, para os candidatos 3 e 4.	115
Figura 74 – Gráficos obtidos, respectivamente, para os candidatos 5 e 6.	115
Figura 75 – Gráficos obtidos, respectivamente, para os candidatos 7 e 8.	116
Figura 76 – Gráficos obtidos, respectivamente, para os candidatos 9 e 10.	116
Figura 77 – Gráficos obtidos, respectivamente, para os candidatos 11 e 12.	116
Figura 78 – Gráficos obtidos, respectivamente, para os candidatos 13 e 14.	117
Figura 79 – Gráficos obtidos, respectivamente, para os candidatos 15 e 16.	117

Figura 80 – Gráficos obtidos, respectivamente, para os candidatos 17 e 18.	117
Figura 81 – Gráficos obtidos, respectivamente, para os candidatos 19 e 20.	118
Figura 82 – Gráficos obtidos, respectivamente, para os candidatos 21 e 22.	118
Figura 83 – Gráficos obtidos, respectivamente, para os candidatos 23 e 24.	118
Figura 84 – Gráficos obtidos, respectivamente, para os candidatos 25 e 26.	119
Figura 85 – Gráficos obtidos, respectivamente, para os candidatos 27 e 28.	119
Figura 86 – Gráficos obtidos, respectivamente, para os candidatos 29 e 30.	119
Figura 87 – Gráfico obtido para o candidato 31.	120

SUMÁRIO

	CAPÍTULO 1 - INTRODUÇÃO GERAL	13
1	REFERENCIAL TEÓRICO	15
1.1	ANÁLISE SENSORIAL	15
1.1.1	Análise sensorial de olfato de café	17
1.1.2	Qualidade dos alimentos	18
1.2	OS TESTES DISCRIMINATIVOS	19
1.2.1	Os usos do teste triangular	21
1.3	RECRUTAMENTO / PRÉ-SELEÇÃO, SELEÇÃO E TREINAMENTO	22
1.4	SENSOMETRIA	24
1.5	INFERÊNCIA SOBRE PROPORÇÃO DE SUCESSOS EM TESTES TRIANGULARES	25
1.6	OS ERROS TIPO I, TIPO II E PODER	31
1.7	A SIMULAÇÃO MONTE CARLO	31
1.8	REGRESSÃO SEGMENTADA	32
	REFERÊNCIAS	34
	CAPÍTULO 2 - COMPARAÇÃO DE TESTES PARA A SELEÇÃO DE POTENCIAIS PROVADORES	38
	RESUMO	38
2	INTRODUÇÃO	40
3	METODOLOGIA	42
3.1	ESTABELECIMENTO DOS TESTES	42
3.2	SIMULAÇÃO MONTE CARLO	45
4	RESULTADOS E DISCUSSÃO	46
5	CONCLUSÕES	51
	REFERÊNCIAS	52
	CAPÍTULO 3 - TESTES PARA PROPORÇÃO SOB QUEBRA DE PRESUPOSIÇÃO: OS CASOS DA FADIGA E MEMÓRIA SENSORIAIS	54
	RESUMO	54
6	INTRODUÇÃO	56
7	METODOLOGIA	58
7.1	TESTES ESTABELECIDOS	58
7.2	OS MODELOS LINEAR E NÃO LINEAR PARA O CANSAÇO	60
7.3	OS MODELOS LINEAR E NÃO LINEAR PARA O APRENDIZADO	62
7.4	DESEMPENHO DOS TESTES E SIMULAÇÃO	64
8	RESULTADOS E DISCUSSÃO	65
9	CONCLUSÕES	70
	REFERÊNCIAS	71
	CAPÍTULO 4 - ANÁLISE SENSORIAL DE OLFATO DE CAFÉ: A REGRESSÃO SEGMENTADA AUXILIANDO OS TESTES DE SELEÇÃO DE JUÍZES	72
	RESUMO	72
10	INTRODUÇÃO	74
11	METODOLOGIA	77
11.1	ANÁLISE SENSORIAL	77

11.2	MODELAGEM COM REGRESSÃO SEGMENTADA	79
11.3	ANÁLISE DO TESTE SEQUENCIAL	80
12	RESULTADOS E DISCUSSÃO	81
13	CONCLUSÕES	86
	REFERÊNCIAS	87
	CAPÍTULO 5 - CONSIDERAÇÕES FINAIS	89
	APÊNDICES	90

CAPÍTULO 1 - INTRODUÇÃO GERAL

Nos últimos anos, os consumidores têm se tornado cada vez mais críticos e exigentes quanto à escolha dos alimentos, sendo necessário produzi-los com qualidade (GRUNERT, 2005). Buscando esse aprimoramento, estudos envolvendo a pesquisa sensorial, que focam em como os sentidos reagem a determinados produtos, adquirem relevância quando aplicados ao controle de qualidade e desenvolvimento de novos produtos (BROCKHOFF, 2011).

A análise sensorial é um procedimento utilizado para medir, analisar e interpretar reações estimuladas por alimentos ou materiais, que são percebidas através dos sentidos (POSTE *et al.*, 1991). Esse tipo de análise, essencial na avaliação da aceitabilidade no mercado e qualidade do produto, é feita através de um grupo de pessoas que são responsáveis por analisar sensorialmente o produto, constituindo o painel sensorial (TEIXEIRA, 2009).

Ao selecionar juízes ou degustadores e submetê-los ao treinamento, pode-se trabalhar o desenvolvimento da memória e sensibilidade dos indivíduos, bem como estabelecer um vocabulário descritivo, a fim de obter pequenos painéis com medidas sensoriais consistentes e padronizadas. Na etapa de seleção, testes triangulares são comumente utilizados para avaliar habilidades discriminativas de um indivíduo, devido a sua simplicidade e facilidade de execução. Esse teste permite determinar se as amostras possuem diferenças sensoriais perceptíveis ou não (KUNERT; MEYNER, 1999; POSTE *et al.*, 1991). Além do mais, possuem diversas aplicações, apesar de seu uso para a seleção de provadores ainda apresentar critérios variados quanto à porcentagem requerida de acertos e número de ensaios efetuados.

O teste triangular deve ser criterioso a fim de minimizar possíveis erros na seleção de provadores, tendo em vista o impacto da má seleção no controle de qualidade do produto e a forma com que isso afetaria o mercado consumidor, o processamento industrial e investimentos no setor. Também, deve-se considerar a análise de diferentes perfis dos candidatos durante os ensaios dos testes, caso tenham comportamento constante ou variável.

Buscando reduzir esses erros no processo de seleção e seus consequentes impactos, deve-se obter um teste capaz de selecionar candidatos com acuidade sensorial e não selecionar candidatos sem essa acuidade específica ao produto. Através da estimação e da decisão é possível obter afirmações sobre os parâmetros, o que torna a transformação de um estimador em teste comumente praticável. Sendo assim, realizou-se um estudo de estimadores intervalares que foram reescritos como testes, adaptando-os para provadores submetidos a testes triangulares.

Nesse sentido, deseja-se recomendar o melhor teste (que apresente menor erro tipo I e maior poder) para a seleção de provadores com uma proporção constante de acertos; avaliar o comportamento desses testes para a quebra de pressuposição da proporção constante, considerando que o candidato possa atingir a fadiga ao longo dos ensaios ou desenvolver um aprendizado; modelar o comportamento da proporção de acertos para candidatos reais submetidos à análise sensorial olfativa de café, utilizando a regressão segmentada como ferramenta auxiliar na definição do número de ensaios em que o teste indicado deve ser aplicado. Para avaliar o desempenho dos testes, foram realizadas simulações de Monte Carlo.

1 REFERENCIAL TEÓRICO

Esta seção apresenta conceitos relevantes ao tema de pesquisa, envolvendo a análise sensorial, cuidados em laboratório, análise de café, testes aplicados, bem como etapas de recrutamento/pré-seleção, seleção e treinamento. Também, são abordados conteúdos sobre sensometria, inferência, simulação Monte Carlo e regressão segmentada, que fornecem fundamentação teórica para avaliação dos testes.

1.1 ANÁLISE SENSORIAL

Os métodos de degustação começaram a ser aplicados na Europa, a fim de controlar a qualidade de cervejarias e destilarias. Durante a Segunda Guerra Mundial, a necessidade de se produzir alimentos de qualidade para os soldados do exército, sem rejeição, fez com que a análise sensorial se tornasse base científica (TEIXEIRA, 2009). A análise sensorial teve crescimento rápido, em meados do século XX, devido à expansão de alimentos processados e o aumento do consumo de produtos industrializados (LAWLESS; HEYMANN, 2010). Essa prática chegou ao Brasil em 1954, com a avaliação do café no Instituto Agrônomo de Campinas (TEIXEIRA, 2009).

No entanto, quando pessoas são utilizadas para medir sensorialmente um produto, deve-se ter um controle rígido dos métodos do teste para reduzir os erros relacionados a fatores psicológicos ou influências externas (POSTE *et al.*, 1991). Há algumas variações nas respostas humanas que não podem ser totalmente controladas como humor, motivação dos participantes, sensibilidade fisiológica ao estímulo de cada um deles, sua história e familiaridade com os produtos (LAWLESS; HEYMANN, 2010).

O provador pode ser influenciado por alguns fatores tais como: tópicos relacionados ao erro de expectativas, quando ele ouve alguma informação a respeito do produto e associa à sensação produzida; o erro de estímulo, quando o provador associa uma característica irrelevante à sensação em estudo (por exemplo, associar a cor mais intensa ao sabor mais intenso); o erro lógico, quando o provador altera sua nota de acordo com critérios de sua mente; o erro de tendência central, quando o provador tende a dar notas intermediárias ao produto, evitando as extremidades da escala, por receio de encontrar um produto melhor ou pior depois (CHAM-

BERS IV; WOLF, 2005).

Dados gerados a partir de observações humanas são altamente variáveis. Por isso, a análise sensorial fornece diretrizes de preparação e entrega de amostras sob condições controladas, a fim de minimizar esses erros. Como qualquer outro teste analítico, a análise sensorial está preocupada com a precisão, acurácia, sensibilidade e evitar falsos resultados (LAWLESS; HEYMANN, 2010).

O ambiente de análise sensorial deve ser confortável, silencioso, bem ventilado, com temperatura em torno de 22°C e umidade controlada entre 44 - 55 %. Também deve conter cabines individuais com pias ou cuspideiras. A iluminação deve ser uniforme e não deve influenciar a aparência do produto, a não ser em casos em que se deseja eliminar diferenças de coloração entre as amostras, em que lâmpadas de coloração vermelha, verde e azul podem ser utilizadas. As amostras devem ser preparadas por métodos que mantenham as características do produto constantes e que não transmitam gostos ou odores estranhos (POSTE *et al.*, 1991).

Os sistemas sensoriais gustativo, olfativo, tátil, auditivo e visual permitem a avaliação de atributos dos alimentos tais como gosto, odor, textura, som e cor, respectivamente. O gosto é a propriedade relacionada ao paladar, através da identificação dos gostos primários dos alimentos pelas papilas gustativas. O odor é a propriedade perceptível à aspiração de certas substâncias voláteis pelo órgão olfativo, já o aroma é a percepção aromática pela via retronasal (TEIXEIRA, 2009). A temperatura da cavidade bucal é um dos fatores que contribuem para que os compostos voláteis cheguem ao bulbo olfativo pela cavidade retronasal, responsável pela percepção do aroma. Além disso, outros fatores como a mastigação, solubilização, saliva e mudanças no pH pode auxiliar nesse processo colaborando, conseqüentemente, com a percepção do sabor (SIMÕES; WASZCZYNSKYJ; WOSIACKI, 2009).

O sabor é um atributo complexo associado a experiências mistas gustativas, olfativas e táteis percebidas durante degustação. A textura está associada às propriedades reológicas e estruturais, quando o alimento sofre deformação de mordida, prensagem ou corte. O som é a característica produzida pelo alimento durante seu consumo. A cor está associada com a apresentação visual como aparência, tonalidade, intensidade e brilho (TEIXEIRA, 2009).

Por se tratar de um método científico utilizado para medir, analisar e interpretar respostas de produtos perceptíveis através dos sentidos, a análise sensorial é utilizada como uma ciência quantitativa cujos valores numéricos estão associados à relação entre as características do produto e a percepção humana (LAWLESS; HEYMANN, 2010).

1.1.1 Análise sensorial de olfato de café

A análise sensorial busca medir respostas humanas sobre os alimentos de forma exata, a fim de minimizar efeitos que possam enviesar a análise de dados, tais como marca e quaisquer influências sobre a percepção do consumidor. A ideia é isolar as propriedades sensoriais, que são informações relevantes e bastante úteis para desenvolvedores, cientistas dos alimentos e gerentes (LAWLESS; HEYMANN, 2010).

Na análise sensorial do café, alguns atributos principais são utilizados na avaliação como o aroma, doçura, amargor, corpo, gosto residual e acidez. O aroma é percebido através do olfato e depende da experiência do degustador. Ele pode ser classificado de suave a intenso, com aromas frutados, achocolatados, florais e cítricos. Um bom café é aquele cujo aroma é bem pronunciado (PAIVA, 2005).

O café é uma bebida mundialmente apreciada e consumida, devido ao seu sabor e aroma característicos. Essa alta aceitabilidade se deve, principalmente, ao aroma. O sabor e aroma do café são frequentemente estimados pela análise sensorial. Apesar de ser um método menos objetivo em relação aos instrumentais, ele permite determinar simultaneamente várias características como, por exemplo, a identificação e descrição dos diversos aromas no café que dependem da correlação entre atributos sensoriais e centenas de compostos voláteis presentes no produto (RODARTE, 2008).

A sensibilidade olfativa humana e de outros animais é marcante. Nossa habilidade para detectar odores em pequenas concentrações ainda supera qualquer instrumento de medida em análise química. Algumas moléculas orgânicas são muito efetivas na estimulação do sistema olfativo, como as que contém enxofre (ex. etil mercaptano) presente em repolho, outras contendo metoxi pirazina presente nos pimentões e os vastos compostos terpeno responsáveis por aromas cítricos ou com características aromáticas de ervas, hortelã e pinho que são perceptíveis na faixa de partes por milhão. No entanto, os limites são altamente variáveis nos indivíduos (LAWLESS; HEYMANN, 2010).

As etapas de torra, tempo do processo e temperatura influenciam a formação de compostos voláteis no café. A análise sensorial desses compostos é de suma importância e depende da concentração do produto, limiar de detecção, interações sinérgicas ou antagônicas com outros componentes voláteis ou não voláteis (RODARTE, 2008).

A qualidade do café tem se tornado indispensável ao consumidor, que se preocupa mais

com o sabor, aroma e higiene do produto (PAIVA, 2005). No Brasil, a qualidade dessa bebida é avaliada pelas características físicas dos grãos e análise sensorial. Os grãos podem ser avaliados quanto à cor, formato, tamanho e número de defeitos, enquanto a bebida pode ser classificada sensorialmente quanto ao sabor e aroma, apresentando variações entre estritamente mole, mole, apenas mole, dura, riado, rio e rio zona. Esse procedimento, tem se tornado bem específico, cujas etapas devem ser cuidadosamente padronizadas para identificação da potencialidade da bebida. Dentre os atributos avaliados, o aroma do café é o que mais se destaca, sendo determinante da qualidade. O aroma varia de acordo com a espécie do grão, região de produção e procedimentos pós-colheita (RODARTE, 2008)

1.1.2 Qualidade dos alimentos

Durante centenas de anos, os sentidos humanos têm sido utilizados na avaliação da qualidade dos alimentos. A produção de alimentos com boa qualidade dependia da análise sensorial exata de um *expert* que tomava decisões sobre as mudanças no processamento para que o produto apresentasse características desejáveis (LAWLESS; HEYMANN, 2010). Vale ressaltar que a qualidade dos alimentos tem uma vasta definição. Ela pode se referir tanto às características físicas dos alimentos, analisadas por engenheiros e tecnólogos, como características subjetivas percebidas pelos consumidores (GRUNERT, 2005).

A percepção de qualidade do consumidor pode estar relacionada ao gosto, aparência tradicional, se o produto é seguro ou até mesmo saudável. Observe que um alimento saudável depende da definição de cada indivíduo. Um alimento saudável pode ser aquele que contém produtos naturais, sem alterações químicas ou aditivos, como também um alimento pode ser caracterizado como não saudável se houver alto teor de gordura ou contaminantes microbianos, por exemplo (CAYOT, 2007).

De qualquer forma, a quantificação exata da diferença ou similaridade entre os alimentos, assim como a aceitação e preferência é necessária para otimizar e manter a qualidade dos alimentos (LEE; HOUT, 2009). Nesse sentido, os testes discriminativos têm sido bastante utilizados para o controle de qualidade (POSTE *et al.*, 1991).

1.2 OS TESTES DISCRIMINATIVOS

Na análise sensorial, existem os testes afetivos, descritivos e discriminativos. Os testes afetivos são utilizados para medir características subjetivas, como a aceitação, o quanto um consumidor gosta ou desgosta de um produto e a preferência em que o consumidor seleciona o produto preferido. Os testes descritivos permitem descrever características sensoriais do produto através de escalas que as quantificam e permitem identificar quais têm maior intensidade (POSTE *et al.*, 1991). Para detectar se os avaliadores estão identificando ou não diferenças nas amostras, como na seleção de potenciais provadores, utilizam-se os testes discriminativos.

Os testes discriminativos são utilizados para averiguar pequenas diferenças em determinadas características do produto ou para selecionar provadores. Algumas vezes a discriminação é desejável, como a melhoria de um produto comparado ao original. Outras, é indesejável, como a substituição de um ingrediente por um mais barato. Nesse tipo de teste, o participante deve escolher uma amostra que pode ser designada como correta ou incorreta. Se a frequência de respostas corretas for maior do que a esperada, então a diferença é declarada. Caso contrário, as amostras são ditas como as mesmas (CHAMBERS IV; WOLF, 2005).

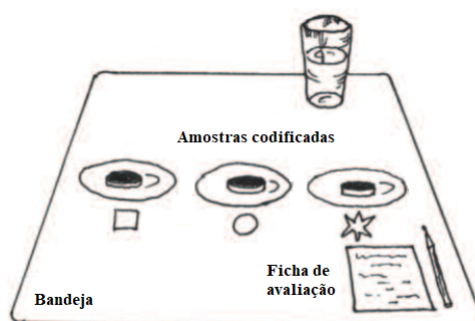
Existem vários testes discriminativos: Duo-Trio, A não A, Comparação Múltipla, Comparação Pareada, Dois em Cinco, Ordenação e Triangular. No teste Duo-Trio, uma amostra é apresentada como referência e depois, separadamente, são apresentadas outras duas amostras para identificação de qual delas equivale à amostra inicial (TEIXEIRA, 2009). No teste A não A, de forma similar, a amostra apresentada inicialmente é dita como referência, para posterior avaliação de uma sequência de amostras a serem identificadas, uma de cada vez, como equivalente à amostra de referência (A) ou não (não A) (PAES, 2005). Esses dois testes requerem uma certa memória sensorial.

No teste de comparação múltipla, uma amostra de referência é apresentada simultaneamente às demais amostras cujo atributo deve ser avaliado como mais intenso, menos intenso ou de intensidade igual. Esse teste fornece a magnitude e direção da diferença. Para realizá-lo, os provadores devem ser previamente treinados quanto às características sensoriais do produto e à utilização da técnica (PAES, 2005). No teste de comparação pareada são apresentadas duas amostras em cada sessão, em que o provador deve avaliar qual delas possui maior intensidade de um certo atributo. Esse teste pode ser utilizado na seleção de painéis, no entanto, deve-se observar que há uma chance de 50% de acerto ao acaso (POSTE *et al.*, 1991).

No teste Dois em Cinco, cinco amostras são apresentadas sendo duas provenientes de um tratamento A e, as três restantes, provenientes de outro tratamento B. Dessa forma, o provador deve identificar duas das cinco amostras que ele considere iguais. Esse teste requer um certo treinamento desses provadores. Já no teste de ordenação, o julgador deve classificar as amostras em ordem crescente ou decrescente de intensidade de acordo com alguma característica específica. Os julgadores devem estar familiarizados com os atributos envolvidos (PAES, 2005).

Por fim, no teste triangular, três amostras são apresentadas simultaneamente, sendo que duas amostras são a mesma e uma é diferente. O provador deve, então, selecionar a amostra diferente (CHAMBERS IV; WOLF, 2005). Esse teste é comumente empregado na seleção e treinamento de provadores, sendo que as amostras não devem induzir facilmente à fadiga ou adaptação sensorial (TEIXEIRA, 2009). No processo de aleatorização das amostras, há seis possíveis maneiras de apresentação: ABB, BBA, BAB, AAB, ABA, BAA; em que o resultado das avaliações se baseia no número absoluto de julgamentos corretos (POSTE *et al.*, 1991). Para ilustrar o teste triangular, apresenta-se a Figura 1.

Figura 1 – Figura representativa do teste triangular.



Fonte: Adaptado de PDST (2017, p. 47).

Através do teste triangular, determina-se se duas amostras são perceptivelmente diferentes ou não, porém não há indicação da direção da diferença, já que não se define qual(is) atributo(s) diferencia(m) as amostras. Um método similar ao teste triangular é o 3-AFC (*three-Alternative Forced Choise*) que indica se alguma característica sensorial específica é maior ou menor dentre as amostras apresentadas, ou seja, há uma comparação quanto à intensidade do atributo (LAWLESS; HEYMANN, 2010).

Pode-se observar que, pelo fato do teste triangular ser comumente empregado para seleção e treinamento de provadores, não necessita treinamento prévio ou familiarização com

algum atributo para aplicação do mesmo. Adicionalmente, por ser um teste que apresenta probabilidade de 1/3 (cerca de 33%) de acerto ao acaso, que é inferior a 50% (considerada uma taxa relativamente alta de acerto ao acaso), ele pode ser considerado uma boa opção para este trabalho.

1.2.1 Os usos do teste triangular

O teste triangular pode ser utilizado tanto na seleção de provadores como em análises para averiguar se houveram alterações perceptíveis entre amostras de um determinado produto que foi submetido a diferentes processamentos ou substituição de ingrediente (POSTE *et al.*, 1991).

Devido a sua simplicidade e facilidade de execução, esses testes triangulares são amplamente utilizados na análise sensorial. No entanto, para obter alto poder em pequenas diferenças entre produtos, o número de avaliadores pode não ser suficiente, sendo conveniente realizar repetições (KUNERT; MEYNER, 1999).

Existem muitos trabalhos que aplicaram esses testes de diferentes formas. Como por exemplo, em um estudo da análise de suplemento de fibra em pó a base de inulina comparado ao adoçante cristalizado de xilitol que foi utilizado como "placebo" quando misturado em um agente mascarante de suco de laranja. O teste triangular foi utilizado para averiguar se haveria diferença perceptível entre esses produtos (JONHSON; BONSU; MCSWEENEY, 2016).

Também, através da seleção de provadores que foram submetidos a testes triangulares de suco de maracujá com diferentes concentrações de sacarose (35 g/L e 50 g/L). Nesse estudo, foi avaliado o efeito de diferentes adoçantes utilizados no suco de maracujá e análise de possíveis substitutos à sacarose. A análise buscava avaliar o perfil descritivo, aceitação e preferência quanto aos diferentes tipos de adoçantes utilizados comparado à sacarose. Dessa forma, era desejável selecionar provadores com boa capacidade discriminativa em relação a essas concentrações (ROCHA; BOLINI, 2015).

Outro estudo utilizando o teste triangular foi feito para avaliar diferentes cultivares de trigo e suas influências na qualidade sensorial de pão. O teste triangular foi utilizado para avaliar diferentes pares de tratamentos em 5 pães a fim de verificar se os juízes percebiam diferença entre eles. Nessa situação, ainda foi solicitado aos avaliadores que relatassem os diferentes

atributos que eles fossem capazes de identificar a fim de estabelecer uma lista de atributos a ser utilizada posteriormente na análise sensorial (CALLEJO; VARGAS-KOSTIUK; QUIJANO, 2015).

Nota-se que houveram diversas situações em que os testes triangulares foram aplicados: para análise de produtos diferentes, com a finalidade de averiguar se as características sensoriais foram similares; para produtos com diferentes concentrações de um determinado ingrediente a fim de selecionar um provador; ou até mesmo para avaliar mais de dois produtos, cujas matérias primas foram oriunda de cultivares diferentes, de forma a compará-los aos pares no teste triangular e averiguar a interferência do cultivar na sensorial. Vale ressaltar que, mesmo quando são trabalhadas diferentes concentrações de um ingrediente ou processamentos, no teste triangular pede-se a amostra considerada diferente dentre as apresentadas e não aquela cuja intensidade é maior ou menor, como no 3-AFC.

A avaliação do teste triangular pode ser feita tanto quanto ao número de pessoas ou repetições dos testes. Dentre essas duas formas de avaliação, a interpretação pode ser feita da seguinte forma: se apenas uma pequena parcela de avaliadores identificam corretamente a amostra (ex. 35 de 100 avaliadores), então existem razões para que acreditar que os produtos não são significativamente diferentes. No entanto, se um avaliador identifica diferença em poucas repetições (ex. 35 de 100 repetições), então a diferença entre produtos é pequena para aquele avaliador (KUNERT; MEYNER, 1999).

1.3 RECRUTAMENTO / PRÉ-SELEÇÃO, SELEÇÃO E TREINAMENTO

Um painel treinado, utilizado para avaliar a alteração dos atributos de um produto, pode conter de 10 a 20 pessoas. Para análises sensoriais de aceitação, recomenda-se de 25 a 50 pessoas para prever as reações dos consumidores aos produtos. Grandes painéis de consumo podem conter mais de 100 pessoas (POSTE *et al.*, 1991).

Para o desenvolvimento de um painel, devem-se realizar etapas de recrutamento, seleção e treinamento. O recrutamento consiste em utilizar de meios de divulgação como telefone, anúncios publicitários e publicações, para recrutar pessoas internas e/ou externas à empresa. Nessa etapa também recolhem-se informações a respeito do interesse e motivação, disponibilidade e pontualidade, estado de saúde, consumo de tabaco, idade e sexo, dentre outras informações

necessárias ao estudo (SILVA, 2018). Adicionalmente, verifica-se a habilidade para quantificar escalas de intensidade e a familiaridade com os termos descritivos (ARAÚJO; SILVA; MINIM, 2003). Portanto, realiza-se uma seleção preliminar dos candidatos (SILVA, 2018).

Após coletar essas informações, os candidatos são submetidos a uma nova fase, a fim de garantir qualificações básicas como: capacidade de discriminação, acuidade sensorial, reprodução de resultados, comportamento apropriado de cooperação e rapidez nas respostas (ALVELOS, 2002).

É sugerido uma seleção de provadores com alta sensibilidade de gosto e olfato. Testes específicos podem ser realizados para identificar aqueles indivíduos com maior capacidade em reconhecer diferenças de gosto (HARRIS *et al.*, 1951). Os candidatos são ranqueados com base na porcentagem de respostas corretas. Aqueles que possuem uma alta porcentagem são selecionados, sendo que nenhum pode ter menos de 60% (CHAMBERS IV; WOLF, 2005).

Entretanto, na literatura existem variações nas proporções mínimas de acerto exigidos para declarar um candidato como apto ao treinamento bem como no número de ensaios utilizados para calcular tal proporção. A Tabela 1 apresenta alguns desses valores.

Tabela 1 – Tabela de percentuais de acertos e número de ensaios para seleção de provadores, em diferentes estudos sensoriais de alimentos e bebidas.

Autores	Percentuais de acerto	Ensaio
Araújo; Silva; Minim (2003).	50%	6
Carneiro <i>et al.</i> (2005).	50%	3
Scheid (2001).	50%	4
Chambers IV; Wolf (2005).	60%	-
Verruma-Bernardi; Damásio (2004).	60%	3
Alvelos (2002).	65%	-
Shirose; Mori (1996).	67%	11*
Minim <i>et al.</i> (2010).	75%	4
Oliveira; Benassi (2010).	75%	4
Rocha; Bolini (2015).	75%	10*
Spadoti; Dornellas; Roig (2005).	100%	3

Fonte: Da autora.

*média do número de ensaios para se tomar uma decisão no teste sequencial baseado em dados das referências.

Nota-se que os critérios são diversificados. Inclusive, segundo Alvelos (2002), candidatos com mais de 65% de repostas corretas podem ser selecionados e, em caso de número insuficiente de aprovados, podem ser realizadas repetições dos testes para candidatos que tenham entre 50% e 65% de acerto, aumentando a exigência para 70% de repostas corretas.

Quanto à capacidade de discriminar o alimento em análise, deve-se certificar que eles realmente compreenderam o método do teste, a ficha de avaliação e a técnica de análise. Os candidatos são ranqueados de acordo com suas habilidades em diferenciar os produtos ou descrevê-los. Então, são feitas repetições dos testes para medir a reprodutibilidade do avaliador. Vale ressaltar que o painel é selecionado de acordo com o produto, pois pessoas que discriminam bem um produto não necessariamente discriminam bem outros (POSTE *et al.*, 1991).

No teste triangular, o tamanho e a direção da diferença entre amostras não são determinadas. Também, não há indicação da característica responsável pela diferença (POSTE *et al.*, 1991). Quanto ao tamanho das amostras, a quantidade de alimento para o teste deve ser grande o suficiente para ativar os sentidos, mas não grande o suficiente para causar fadiga. Uma quantidade suficiente para duas ou três degustações é normalmente a quantidade ideal (HARRIS *et al.*, 1951).

Na fase do treinamento, os candidatos colocam em prática os conhecimentos adquiridos na seleção, desenvolvendo sua capacidade de detecção, reconhecimento e descrição dos estímulos sensoriais, para um produto específico. Além disso, devem ser objetivos e ignorar qualquer tipo de afinidade ou rejeição ao produto (SILVA, 2018).

O treinamento desenvolve e estabelece um vocabulário descritivo. A detecção e reconhecimento dos painéis devem estar em concordância com a conotação previamente estabelecida. Para isso, padrões de referência são utilizados para demonstrar as variações dos termos descritivos e auxiliar os membros dos painéis a serem mais consistentes nas avaliações. As técnicas de avaliação podem ser discutidas para qualquer atributo. O importante é que os provadores se familiarizem com o método do teste. O treinamento pode durar semanas ou até meses, dependendo do produto, do procedimento e da capacidade dos membros (POSTE *et al.*, 1991).

1.4 SENSOMETRIA

A Sensometria é o campo da Estatística preocupado com a aplicação, desenvolvimento e análise de dados da ciência sensorial. Os dados sensoriais são relevantes para pesquisas em ambientes acadêmicos, especializadas em estudos sensoriais e do consumidor existentes no mundo. Frequentemente, esses dados também são utilizados como base para a tomada de decisão nas indústrias de alimentos (BROCKHOFF, 2011).

Como todo campo da ciência, a Sensometria busca desenvolver sua tecnologia e aplicações. Tem sido perceptível a importância dessa ferramenta, não só por sua análise de caráter quantitativo, mas para responder problemas cognitivos em pesquisa de produtos, cujas decisões de negócios são baseadas nessas informações. Conceitos variados podem ser testados como a análise de fragrâncias, marca e emoções envolvidas (MOSKOWITZ; SILCHER, 2005).

A análise sensorial deve fornecer medidas confiáveis que conduzam a testes válidos. As medidas sensoriais podem assumir diversas interpretações de acordo com suas variedades de medidas: limiar sensorial, risco sensorial, intensidade do tempo, treinamento e desempenho de painéis. O teste de hipóteses é a base dos testes sensoriais (BI, 2015).

Segundo Lovric (2011, p. 1302) "a variabilidade hedônica e sensorial são características psicológicas fundamentais que devem ser explicitamente modeladas para que se possa desenvolver modelos estatísticos significativos de fenômenos sensoriais". Por isso, os modelos probabilísticos sobre esses fenômenos são um desafio, já que se baseiam em um sistema parcialmente não observado, o cérebro humano e o sistema perceptivo (LOVRIC, 2011).

1.5 INFERÊNCIA SOBRE PROPORÇÃO DE SUCESSOS EM TESTES TRIANGULARES

É sabido, e intuitivo que, em um único julgamento do teste triangular, a probabilidade do provador selecionar a amostra diferente, ao acaso, é $1/3$. Dessa forma, é possível calcular a probabilidade de determinado número de respostas ocorrerem ao acaso, em uma série de julgamentos independentes. Para n julgamentos, os números esperados de respostas corretas e incorretas são, respectivamente, $n/3$ e $2n/3$, se a probabilidade de sucesso for considerada constante (PAES, 2005).

Como cada ensaio representa um ensaio de Bernoulli, a soma dos x acertos em n ensaios tem distribuição Binomial com parâmetros n e p . Logo, o par de hipóteses para n ensaios triangulares equivale a (FERREIRA, 2005; NÆS; BROCKHOFF; TOMIC, 2010):

$$\begin{cases} H_0 : p = \frac{1}{3} \\ H_a : p > \frac{1}{3} \end{cases}$$

Se o avaliador acerta a uma proporção de $1/3$, então o acerto é ao acaso e ele não está identificando a amostra diferente, portanto não deve ser selecionado. Caso ele acerte a uma

proporção acima de 1/3, então está identificando a amostra diferente e deve ser selecionado.

A literatura está repleta de formas de se fazer inferência sobre a proporção de sucessos (p) dessa Binomial. Uma abordagem muito comum é aproximar a distribuição de \hat{p} pela distribuição Normal. Para essa abordagem tem-se que:

$$\hat{p} \sim N\left(p, \frac{pq}{n}\right)$$

Deve-se ressaltar que a qualidade dessa aproximação Normal requer grandes valores de n e uma proximidade de p ao valor 1/2 (FERREIRA, 2005). Então, algumas condições são necessárias para utilizar essa aproximação como: $np \geq 5$, $nq \geq 5$ e $npq > 9$ (LEEMIS; TRIVEDI, 1996). Também, deve-se considerar que aproximações assintóticas assumem que as variáveis aleatórias sejam independentes em cada ensaio e identicamente distribuídas (MOOD; GAYBILL; BOES, 1974).

A partir da aproximação Normal podem ser derivados alguns estimadores intervalares. Um deles é dado por (1) (FERREIRA, 2005).

$$IC_{1-\alpha}(p) : \hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} \quad (1)$$

em que \hat{p} é a proporção amostral de respostas corretas, \hat{q} é a proporção de respostas incorretas, ou seja, $1 - \hat{p}$ e $z_{\alpha/2}$ é o quantil da distribuição normal padrão.

Quando se utiliza uma distribuição Normal (contínua) para aproximar uma distribuição binomial (discreta), deve-se associar um intervalo ao valor discreto. Dessa forma, realiza-se a subtração e adição de 0,5 aos valores discretos, de maneira a incluir todos os valores possíveis do intervalo. Essa regra de adição/subtração de 0,5 foi introduzida por Yates, também conhecida como correção de Yates. A finalidade da correção de continuidade é melhorar a cobertura da probabilidade no intervalo de uma binomial (EMURA; LIAO, 2018).

Outro estimador intervalar possível da Normal, com o uso da correção de continuidade, equivale a (2) (FERREIRA, 2005).

$$IC_{1-\alpha}(p) : \frac{y \pm 0,5}{n} \pm \frac{z_{\alpha/2}}{\sqrt{n}} \sqrt{\frac{y \pm 0,5}{n} \left(1 - \frac{y \pm 0,5}{n}\right)} \quad (2)$$

em que y é o somatório do número de sucessos em n ensaios e a correção é feita com a soma ou subtração de 0,5 ao valor da variável.

Ao utilizar a propriedade de normalidade assintótica do estimador \hat{p} , tem-se que:

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \approx N(0, 1)$$

cuja afirmação probabilística fornece

$$P \left(-Z_{\alpha/2} \leq \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \leq Z_{\alpha/2} \right) \cong 1 - \alpha$$

Ao resolver a equação quadrática para p e com algumas manipulações algébricas, estabelece-se outro estimador intervalar Normal assintótico (3) (FERREIRA, 2005).

$$IC_{1-\alpha}(p) : \frac{y + \frac{z_{\alpha/2}^2}{2} \pm z_{\alpha/2} \sqrt{y - \frac{y^2}{n} + \frac{z_{\alpha/2}^2}{4}}}{n + z_{\alpha/2}^2} \quad (3)$$

para mesmos valores atribuídos de y e $z_{\alpha/2}$. Observa-se que, para os dois últimos estimadores intervalares baseados na Normal, atribui-se o número de sucessos obtidos em n ensaios, ou seja, $y = \sum_{i=1}^n X_i$.

Uma aproximação possível para a distribuição de \hat{p} se utiliza da distribuição F. Existe uma relação pouco conhecida entre as funções de distribuição binomial e distribuição F. Essa relação é exata e dada por (FERREIRA, 2005):

$$F_b(y; n, p) = F_F \left\{ \frac{(y+1)(1-p)}{(n-y)p}; \nu_1 = 2(n-y), \nu_2 = 2(y+1) \right\}$$

em que as funções de distribuição de probabilidade binomial e F são $F_b(y; p, n) = P(Y \leq y)$ e $F_F(f; \nu_1, \nu_2) = P(F \leq f)$. Sendo $f = \frac{(y+1)(1-p)}{(n-y)p}$, com graus de liberdade $\nu_1 = 2(n-y)$ e $\nu_2 = 2(y+1)$ (FERREIRA, 2005). Ao realizar as substituições e isolar o limite superior (p_S), tem-se (LEEMIS; TRIVEDI, 1996):

$$F_{\alpha/2} = \frac{(n-y)p_S}{(y+1)(1-p_S)} \quad \text{logo} \quad p_S = \frac{1}{1 + \frac{(n-y)}{(y+1)F_{\alpha/2}}}$$

Para igualar $F_{1-\alpha/2}$ ao lado direito da expressão de f , substitui-se y por $y-1$, para a garantia de cobertura de pelo menos $1-\alpha$ de probabilidade (FERREIRA, 2005). Obtendo-se o limite inferior (p_I) (LEEMIS; TRIVEDI, 1996):

$$F_{1-\alpha/2} = \frac{(n - (y - 1))p_I}{(y - 1 + 1)(1 - p_I)} \quad \text{que fornece} \quad p_I = \frac{1}{1 + \frac{(n - y + 1)}{yF_{1-\alpha}}}$$

Assim, constrói-se o intervalo exato para o parâmetro p e, a partir dessa abordagem, tem-se o estimador intervalar (4) para uma proporção binomial (LEEMIS; TRIVEDI, 1996).

$$IC_{1-\alpha}(p) : \left(\frac{1}{1 + \frac{(n - y + 1)}{yF_0}}; \frac{1}{1 + \frac{(n - y)}{(y + 1)F_1}} \right) \quad (4)$$

em que $F_0 = F_{1-\alpha/2; 2y, 2(n-y+1)}$ e $F_1 = F_{\alpha/2; 2(y+1), 2(n-y)}$ são quantis da distribuição F. Esse estimador intervalar tem a vantagem de ser exato, embora possua comprimento grande para valores pequenos de p e n .

Existe outra forma de inferir sobre p , dessa vez por meio de um teste sequencial, que é utilizado na inspeção individual de cada ensaio. Dessa forma, cada ensaio é analisado um a um e, com base nos resultados, toma-se a decisão de aceitar, rejeitar ou inspecionar uma nova unidade. O processo ocorre da seguinte maneira: são criadas duas linhas de decisão, sendo uma reta de não seleção (L_0) e outra reta de seleção (L_1), em que n é o número de observações. Com os valores previamente determinados das proporções p_0 e p_1 que se referem, respectivamente, à proporção máxima para não selecionar o candidato como provador e à proporção mínima para selecionar o candidato como provador, assim como os valores de α, β correspondem às respectivas probabilidades de se cometer os erros tipo I e II, obtém-se as linhas apresentadas em (5) (SHIROSE; MORI, 1996; WALD, 2013):

$$L_0 = -h_0 + sn; \quad L_1 = h_1 + sn \quad (5)$$

em que h_0, h_1 e s são dados abaixo:

$$h_0 = \frac{\log\left(\frac{1-\alpha}{\beta}\right)}{k}; \quad h_1 = \frac{\log\left(\frac{1-\beta}{\alpha}\right)}{k}$$

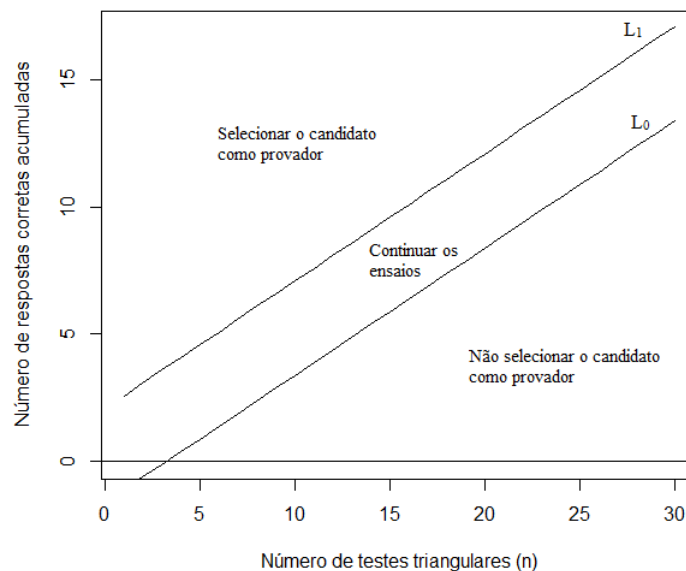
$$k = \log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right); \quad s = \frac{\log\left(\frac{1-p_0}{1-p_1}\right)}{k}$$

Nota-se que h_0 e h_1 são os interceptos das respectivas retas L_0 e L_1 , enquanto s representa a inclinação de ambas retas regressoras.

No processo de seleção de provadores, utilizando o teste sequencial triangular, gera-se a curva da seguinte maneira: se o primeiro teste obtiver resposta correta, então $(x, y) = (1, 1)$;

se a resposta for incorreta, $(x,y) = (1,0)$. A cada teste efetuado, acrescenta-se 1 a x , sendo que y recebe +1 caso a resposta seja correta ou +0, quando incorreta. A repetição do teste segue até que a curva gerada atinja uma das retas, em que toma-se a decisão de aceitar ou rejeitar o candidato a provador (SHIROSE; MORI, 1996). A Figura 2 mostra o gráfico do teste triangular sequencial para $p_0 = 1/3$, $p_1 = 2/3$, $\alpha = 0,05$ e $\beta = 0,10$. As definições dos erros tipo I e II, bem como sua relação com α e β são apresentadas detalhadamente no próximo tópico. Além disso, de acordo com os valores estipulados pelo pesquisador, o número médio de ensaios a serem realizados para se tomar uma decisão pode variar.

Figura 2 – Gráfico representativo do teste sequencial triangular para seleção de provadores.



Fonte: Adaptado de Shirose e Mori (1996, p. 51).

Nota-se que, na situação anterior, optou-se por utilizar $\alpha < \beta$, logo a probabilidade de selecionar um candidato sem acuidade sensorial é inferior à probabilidade de não selecionar um candidato com acuidade sensorial. A reta de aceitação fornece o número mínimo de respostas corretas que o candidato deve ter para ser selecionado, enquanto a reta de rejeição fornece o número máximo de respostas corretas para rejeitar um candidato. Caso o número de acertos esteja entre as duas retas, deve-se dar continuidade ao teste triangular (SHIROSE; MORI, 1996).

É possível também realizar uma aproximação da Poisson utilizando o quantil da qui-quadrado. Nesse caso, o estimador intervalar é obtido baseado na aproximação da Poisson para a Binomial, em que a variável Y é assintoticamente Poisson com parâmetro np_I . Portanto,

tem-se que o limite inferior p_I satisfaz (LEMIS; TRIVEDI, 1996)

$$\sum_{k=y}^n C_{n,k} p_I^k (1-p_I)^{n-k} = \frac{\alpha}{2}$$

cuja aproximação do limite inferior através da Poisson fornece:

$$\sum_{k=y}^{\infty} \frac{(np_I)^k e^{-np_I}}{k!} = \frac{\alpha}{2} \quad \text{ou} \quad 1 - \sum_{k=0}^{y-1} \frac{(np_I)^k e^{-np_I}}{k!} = \frac{\alpha}{2}$$

Observa-se que o lado esquerdo da última equação se trata de uma função distribuição acumulada da variável aleatória de Erlang com parâmetro de escala np_I e parâmetro de forma y (denotado por $E_{np_I,y}$) (LEEMIS; TRIVEDI, 1996). Consequentemente,

$$P(E_{np_I,y} \leq 1) = \frac{\alpha}{2}$$

em que $2np_I E_{np_I,y}$ é uma variável com distribuição χ^2 e $2y$ graus de liberdade, reduzindo a

$$P(\chi_{2y}^2 \leq 2np_I) = \frac{\alpha}{2}; \quad \text{ou} \quad p_I = \frac{\chi_{1-\alpha/2;2y}^2}{2n}$$

De forma similar, para o limite superior, obtém-se:

$$p_S = \frac{\chi_{\alpha/2;2(y+1)}^2}{2n}$$

Então, o estimador intervalar (6) é dado por (FERREIRA, 2005):

$$IC_{1-\alpha}(p) : \left(\frac{\chi_{1-\alpha/2;2y}^2}{2n}; \frac{\chi_{\alpha/2;2(y+1)}^2}{2n} \right) \quad (6)$$

sendo $\chi_{1-\alpha/2;2y}^2$ e $\chi_{\alpha/2;2(y+1)}^2$, respectivamente, os valores críticos da cauda esquerda e direita da distribuição qui-quadrado e, também, se referem aos limites inferior e superior.

De acordo com Leemis e Trivedi (1996), essa aproximação é melhor indicada quando p é pequeno. Além disso, quando a probabilidade de falhar é pequena e tem-se um número grande de replicações, a aproximação da Poisson possui melhor performance quando comparada à aproximação Normal.

Pode-se observar que cada estimador intervalar tem sua particularidade e condições em que são recomendados. Portanto, dependendo das condições em que são submetidos, esses

estimadores podem apresentar vantagens ou desvantagens em determinadas situações conforme sua adequação.

1.6 OS ERROS TIPO I, TIPO II E PODER

O erro tipo I refere-se a um resultado falso positivo do estudo de um efeito em particular (fenômeno em estudo) e ocorre quando há rejeição da hipótese nula, sendo ela verdadeira. O erro tipo II ocorre quando não rejeitamos a hipótese nula sendo que ela é falsa (LAWLESS; HEYMANN, 2010).

Para seleção de provadores, os erros tipo I e II se referem, respectivamente, à seleção do candidato sem acuidade sensorial e não selecionar o candidato com acuidade sensorial (SHIROSE; MORI, 1996).

A probabilidade de se cometer o erro tipo I é dada por α , enquanto a probabilidade de se cometer o erro tipo II equivale a β . É desejável cometer menor erro possível, então, o cientista pode controlar o erro tipo I, definindo o tamanho de α e minimizar o erro tipo II, através da maximização do poder ($1 - \beta$) (LAWLESS; HEYMANN, 2010). O poder refere-se à seleção do candidato com acuidade sensorial.

A avaliação das taxas de erro tipo I também permite definir se os testes são considerados conservadores ou liberais. Os testes são ditos conservadores quando as taxas de erro tipo I excedem negativamente e são ditos liberais quando as taxas de erro tipo I excedem positivamente (FELTOVICH, 2003).

1.7 A SIMULAÇÃO MONTE CARLO

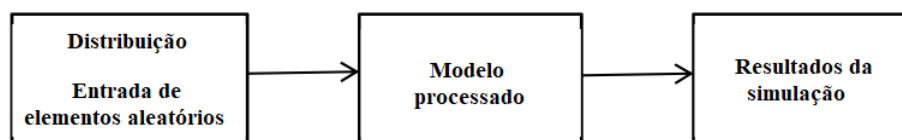
A simulação Monte Carlo (MC) recebeu esse nome como referência à cidade de Monte Carlo situada em Mônaco (França), famosa por jogos de cassinos. Essa simulação emprega técnicas de amostragem estatística no computador, em que são geradas variáveis aleatórias, para obter uma aproximação probabilística com resultados que indicam o desempenho do modelo desenvolvido (ADERIBIGBE, 2014).

Esse tipo de simulação utiliza modelos matemáticos que descrevem o comportamento

de um sistema durante longos períodos e incorporam a elementos estocásticos, no qual alguns ou todos os parâmetros têm um grau de variabilidade associado. Esse processo permite compreender sistemas complexos e solucioná-los, para tomar decisões. A simulação MC ocorre repetidamente, utilizando diferente conjunto de variáveis aleatórias (inputs), em cada etapa, fornecendo um conjunto de resultados possíveis de se obter (outputs) (BONATE, 2001).

A vantagem de se utilizar a simulação MC é que se trata de um método bastante flexível que pode ser facilmente desenvolvido de acordo com o que se busca na análise. Basicamente, MC implica na aplicação de um conjunto de variáveis aleatórias no modelo de estudo que processa e fornece resultados da simulação, como mostra o fluxograma da Figura 3 (ADERIBIGBE, 2014).

Figura 3 – Fluxograma representativo da simulação MC.



Fonte: Adaptado de Aderibigbe (2014, p.6).

Assim, a simulação MC forma um novo método científico entre o experimental e métodos numéricos, sendo denominado "experimentação numérica". Essa aleatoriedade faz com que essa simulação se assemelhe a experimentos reais (BEISBART; NORTON, 2012). Inclusive, para uma completa distribuição com acurácia razoável, deve-se realizar algumas centenas de milhares de simulações. O recomendável seria pelo menos 1000 simulações MC (KRAH; NIKOLIC; KORN, 2018). Através dessa simulação pode-se obter respostas quando não há solução analítica, sendo possível resolver problemas complexos de forma mais simples (BONATE, 2001).

1.8 REGRESSÃO SEGMENTADA

A regressão segmentada fornece uma estimativa do ponto de quebra e do erro padrão. Para alguns conjuntos de dados, esse tipo de regressão produz um ajuste visual satisfatório, por descrever bem os diferentes comportamentos dos dados em certas regiões, podendo ser linear até certo ponto e curvilínea em outro (ROBBINS; SAXTON; SOUTHERN, 2006).

Deve-se considerar que a função regressora é contínua no ponto de mudança (*change-point* ou *breakpoint*) τ , se satisfizer a equação 7 (LIU; QIAN, 2009):

$$\alpha_0 + \alpha_1\tau = \beta_0 + \beta_1\tau \quad (7)$$

para dois modelos lineares.

Então, a interseção da abscissa das duas retas é dada por:

$$\tau = \frac{(\beta_0 - \alpha_0)}{(\alpha_1 - \beta_1)}$$

Suponha que esses dois modelos lineares satisfaçam:

$$Y_i = \begin{cases} \alpha_0 + \alpha_1 i + e_i, & i = 1, \dots, r \\ \beta_0 + \beta_1 i + e_i, & i = r + 1, \dots, n \end{cases}$$

em que e_i é o erro.

Nota-se que τ encontra-se no intervalo $(r, r + 1)$, sendo r desconhecido. Portanto, deve-se encontrar a função de verossimilhança estimada de τ para se obter intervalos de confiança que possam ser associados (SOLOW, 1987).

Para encontrar o melhor modelo, são realizadas várias tentativas de regressões com diferentes pontos de mudança de inclinação, em que é obtida a soma de quadrados dos desvios para cada regressão. O ajuste do modelo de regressão segmentada é feito pelo método dos mínimos quadrados. Logo, a regressão linear que possui menor soma de quadrados dos erros é considerada a melhor (PORTZ; DIAS; CIRYNO, 2000).

A estimação das curvas e *breakpoints* são bem complexas. Uma abordagem mais simples e usual se dá por meio de algoritmos do tipo *grid-search*, em que uma grade com possíveis valores do parâmetro se ajusta a modelos lineares buscando aquele que fornece melhor ajuste ao modelo. O pacote *segmented* do software R Core Team (2019) auxilia nas estimativas pontuais e de linhas quebradas, denominadas *broken-lines*. Dessa forma, ele facilita a estimação e resumo dos modelos lineares segmentados, sendo que não há limite para o número de pontos de mudança existentes (MUGGEO, 2008).

REFERÊNCIAS

- ADERIBIGBE, A. **A term paper on Monte Carlo analysis / simulation**. University of Ibadan. 2014.
- ALVELOS, H. M. P. P. D. **Análise, Desenvolvimento e Teste de Métodos e Técnicas para Controle Estatístico em Análise Sensorial**. 2002. 253f. Tese (Doutorado em Ciências de Engenharia) - Faculdade de Engenharia, Universidade do Porto, Porto, 2002.
- ARAÚJO, F. B.; SILVA, P. H. A.; MINIM, V. P. R. Perfil sensorial e composição físico-química de cervejas provenientes de dois segmentos do mercado brasileiro. **Ciênc. Tecnol. Aliment.**, v. 23, n. 2, p. 121-128. 2003.
- BEISBART, C. ; NORTON. J. D. Why Monte Carlo simulations are inferences and not experiments. **Internatonal Studies in the Philosophy of Science**, v. 26, n. 4, p. 403-422. 2012.
- BI, J. **Sensory Discrimination Tests and Measurements**. Virginia: Wiley Blackwell, 2015.
- BONATE, P. L. A brief introduction to Monte Carlo simulation. **Clin Pharmacakinet**, v. 40, n. 1., p. 15-22. 2001.
- BROCKHOFF, P. B. Sensometrics for Food Quality Control. **Scandinavian Workshop on Imaging Food Quality**, Dinamarca, n. 15, p. 7-16. 2011.
- CALLEJO, M. J; VARGAS-KOSTIUK, M. E.; QUIJANO, M. R. Selection, training and validation process of a sensory panel for bread analysis: Influence of cultivar on the quality of breads made from common wheat and spelt wheat. **Journal of Cereal Science**, v. 61, p. 55-62. 2015.
- CARNEIRO, J. C. S. *et al.* Perfil sensorial e aceitabilidade de cultivares de feijão (*Phaseolus vulgaris L.*). **Ciênc. Tecnol. Aliment.**, v. 25, n.1, p. 18-24. 2005.
- CAYOT, N. Sensory quality of traditional foods. **Food Chemistry**, v. 101. p. 154-162. 2007.
- CHAMBERS IV, E.; WOLF, M. B. **Sensory testing methods**. 2nd ed. Lancaster: ASTM International, 2005.
- EMURA, T.; LIAO, Y. T. Critical review and comparison of continuity correction methods: The normal approximation to the binomial distribution. **Communications in Statistics - Simulation and Computation**, v. 47, n. 8, 2266-2285. 2018.
- FELTOVICH, N. Nonparametric Tests of Diferences in Medians: Comparison of the Wilcoxon-Mann-Whitney and Robust Rank-Order Tests. **Experimental Economics**, v. 6, p. 273-297. 2003.
- FERREIRA, D. F. **Estatística Básica**. Lavras: UFLA, 2005.

GRUNERT, K. G. Food quality and safety: consumer perception and demand. **European Review of Agricultural Economics**, v. 32, n. 3, p. 369–391. 2005.

HARRIS, E. H. D. B. *et al.* **Sensory Methods for Measuring Differences in Food Quality**. Washington: Agriculture Information Bulletin No. 34, 1951.

ISO 5492 (2008): **Sensory analysis – Vocabulary**, 2^a ed., International Organization for Standardization. Disponível em: <<https://www.iso.org/obp/ui/#iso:std:iso:5492:ed-2:v1:en>>. Acesso em: 09 mar. 2019.

JOHNSON, S.; BONSU, N. K.; MCSWEENEY, M. Triangle Taste Test and Sensory Evaluation: A Novel Application for Determining Supplement-Placebo Match in a Clinical Trial. **Journal of Food Technology and Nutritional Sciences**, v. 2, n. 1, p. 1-4. 2016.

KRAH, A. S.; NIKOLIC, Z.; KORN, R. A Least-Squares Monte Carlo Framework in Proxy Modeling of Life Insurance Companies. **Risks**, v. 6, p. 62, p. 1-26. 2018.

KUNERT, J.; MEYNER, M. On the Triangle Test with Replications. **Food Quality and Preference**. 1999. Disponível em: <https://www.researchgate.net/publication/2765334_On_the_Triangle_Test_with_Replications>. Acesso em: 2 out. 2019.

LAWLESS, H. T.; HEYMANN, H. **Sensory Evaluation of Food**. 2nd ed. New York: Springer, 2010.

LEE, H. S.; HOUT, D. V. Quantification of Sensory and Food Quality: The R-Index Analysis. **Journal of Food Science**, v. 74, n. 6, p. 57-64. 2009.

LEEMIS, L. M.; TRIVEDI, K. S. A Comparison of Approximate Interval Estimators for the Bernoulli Parameter. **The American Statistician**, v. 50, n. 1, p. 63-68. 1996.

LIU, Z.; QIAN, L. Change-point Estimation in a Segmented Linear Regression via Empirical Likelihood. **Communications in Statistics - Simulation and Computation**, v. 39, n. 1, p. 85-100. 2009.

LOVRIC, M. **International Encyclopedia of Statistical Science**. Springer-Verlag Berlin Heidelberg, 2011.

MINIM, V. P. R. *et al.* Análise descritiva: Comparação entre metodologias. **Rev. Inst. Latic. Cândido Tostes**, v. 65, n. 374, p. 41-48. 2010.

MOOD, A. M., GAYBILL, D. C., BOES, C. **Introduction of theory of statistics**. 3^a ed., McGraw-Hill. 1974.

MOSKOWITZ, H. R.; SILCHER, M. The applications of conjoint analysis and their possible uses in sensometrics. **Food quality and preference**, v. 17, p. 145-165. 2005.

MUGGEO, V. M. R. Segmented: An R Package to Fit Regression Models with Broken-Line

Relationships. **R News**, v. 8, n. 1, p. 20-25. 2008.

NÆS, T.; BROCKHOFF, P. B.; TOMIC, O. **Statistics for Sensory and Consumer Science**. London: Wiley. 2010.

OLIVEIRA, A. P. V.; BENASSI, M. T. Avaliação sensorial de pudins de chocolate com açúcar e dietéticos por perfil livre. **Ciênc. agrotec.**, v. 34, n. 1, p. 146-154. 2010.

PAES, J. B. **Métodos de diferença em Avaliação Sensorial de Alimentos e Bebidas**. 3ª ed. Viçosa: UFV, 2005.

PAIVA, E. F. F. **Análise sensorial dos cafés especiais do Estado de Minas Gerais**. 2005. 55f. Dissertação (Mestrado em Ciência dos Alimentos) - Universidade Federal de Lavras, Lavras, 2005.

PDST. **Sensory Analysis Teacher's Manual**. Dublin. 2017.

PORTZ, L.; DIAS, C. T. S.; CYRINO, J. E. P. Regressão segmentada como modelo na determinação de exigências nutricionais de peixes. **Scientia Agricola**, v. 57, n. 4, p. 601-607. 2000.

POSTE, L. M. *et al.* **Laboratory Methods for Sensory Analysis of Food**. Canada: Agriculture Canada Publication, 1991.

R CORE TEAM (Version 3.6.0) [Computer software]. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <https://www.R-project.org/>. Acesso em: 29 jun. 2019.

ROBBINS, K. R.; SAXTON, A. M.; SOUTHERN, L. L. Estimation of nutrient requirements using broken-line regression analysis. **J. Anim. Sci**, v. 84, n. E. Suppl., p. E155–E165, 2006.

ROCHA, I. F. O.; BOLINI, H. M. A. Passion fruit juice with different sweeteners: sensory profile by descriptive analysis and acceptance. **Food Science & Nutrition**, v. 3, n. 2, p. 129-139. 2015.

RODARTE, M. P. **Análise sensorial, química e perfil de constituintes voláteis de cafés especiais**. 2008. 147f. Tese (Doutorado em Ciência dos Alimentos), Universidade Federal de Lavras, Lavras, 2008.

SCHEID, G. A. **Avaliação Sensorial e Físico-Química de Salame Tipo Italiano com Diferentes Concentrações de Cravo da Índia (*Eugenia caryophyllus*)**. 2001. 94f. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) - Universidade Federal de Viçosa, Viçosa, 2001.

SHIROSE, I.; MORI, E. E. M. **Estatística aplicada à análise sensorial (Módulo 2)**. Campinas: ITAL - Manual Técnico nº 13, 1996.

SILVA, A. S. F. A. M. **Reavaliação e Formação de um Painel de Provedores na**

FRULACT. 2018. 101f. Dissertação (Mestrado em Biotecnologia) - Universidade Da Beira Interior, Covilhã, 2018.

SIMÕES, D. R.; WASZCZYNSKYJ, N.; WOSIACKI, G. Aromas em maçãs, suco e sidra: Revisão. **B.CEPPA**, Curitiba, v. 27, n. 1, p. 153-172. 2009.

SOLOW, A. R. Testing for Climate Change: **An Application of the Two-Phase Regression Model**. Manuscript, American Meteorological Society. 1987.

SPADOTI, L.M.; DORNELLAS, J.R.F.; ROIG, S.M. Avaliação sensorial de queijo prato obtido por modificações do processo tradicional de fabricação. **Ciência e Tecnologia de Alimentos**, p. 705-712, 2005.

TEIXEIRA, L.V. Análise sensorial na indústria de alimentos. **Rev. Inst. Latic. “Cândido Tostes”**, v. 64, n. 366, p. 12-21, 2009.

VERRUMA-BERNARDI, M. R.; DAMÁSIO, M. H. Análise descritiva de perfil livre em queijo mozzarella de leite de búfala. **Ciênc. Technol. Aliment.**, v. 24, n. 4, p. 536-542. 2004.

WALD, A. **Sequential Analysis**. New York: Dover Publications, 2013.

CAPÍTULO 2 - COMPARAÇÃO DE TESTES PARA A SELEÇÃO DE POTENCIAIS PROVADORES

RESUMO: A análise sensorial envolve a avaliação de atributos sensoriais de um produto através dos sentidos humanos. A formação de painéis eficientes é imprescindível para o controle de qualidade, sendo necessário executar etapas de seleção e treinamento. Na seleção, utilizam-se testes triangulares para averiguar a habilidade discriminativa dos candidatos, submetendo ao treinamento apenas aqueles com habilidades sensoriais aguçadas. Logo, deseja-se encontrar o melhor teste (em termos de menor taxa de erro tipo I e maior poder) para a seleção de provadores. Esse trabalho tem como objetivo fazer um estudo de estimadores intervalares para proporção, reescrevendo-os como testes para a seleção de provadores. Seis testes (baseados nas distribuições Normal e F, no teste Sequencial e aproximação da Poisson utilizando quantil χ^2) foram comparados. Por meio de um estudo de simulação Monte Carlo foram gerados 1000 provadores virtuais submetidos a testes triangulares ao longo de 5, 10, 15, ..., 100 ensaios. Estabeleceu-se uma probabilidade de acerto constante desses provadores ao longo do tempo, fixada em 0,33, 0,35, 0,40, 0,45 ..., 1,00, avaliando o desempenho dos testes em diferentes números de ensaios (n) e proporção de acertos (p). Foi possível notar que o teste sequencial apresentou menor taxa de erro tipo I e maior poder, para um número pequeno de ensaios ($n < 20$), enquanto o teste baseado na Normal obteve esse comportamento para valor de n igual ou superior a 20 ($n \geq 20$). Portanto, ambos os testes poderiam ser empregados na seleção de provadores, de acordo com o número de ensaios estipulado.

Palavras-chave: Simulação Monte Carlo. Testes triangulares. Sensometria.

FOOD SENSORY ANALYSIS: A COMPARISON OF TESTS FOR POTENTIAL TASTERS SELECTION

ABSTRACT: Sensory analysis involves a product sensory attributes evaluation through human senses. Efficient panels are essential for quality control, therefore it's necessary to perform selection and training steps. In selection, triangular tests are used to verify the discriminative skills of the candidates, subjecting to training only those with acute sensory skills. Thus, the goal is to determine the best test (in terms of lowest type I error rate and highest power) for tasters selection. This work aimed to make a proportion interval estimators study, rewriting them as tests for tasters selection. Six tests (based on Normal and F distributions, Sequential test and Poisson approximation using χ^2 quantile) were compared. One thousand virtual tasters were generated through a Monte Carlo simulation study and submitted to triangular tests over 5, 10, 15, ..., 100 trials. It was established a constant probability hit of these testers over time, set at 0,33, 0,35, 0,40, 0,45, ..., 1,00, evaluating the test performance according with the number of tests (n) and proportion of correct answers (p). Sequential test presented the lowest type I error rate and highest power for a small number of trials ($n < 20$), while the Normal-based test showed this same behavior for n value equal to or greater than 20 ($n \geq 20$). Therefore, both tests could be used to select tasters according to the number of tests established.

Keywords: Monte Carlo simulation. Triangular tests. Sensometrics.

2 INTRODUÇÃO

A qualidade de alimentos tem sido um fator importante nos últimos anos. Os consumidores estão se tornando cada vez mais exigentes, críticos e fragmentados quando se trata da escolha de alimentos, o que leva a situações em que se faz necessário produzir alimentos com um diferencial na qualidade (GRUNERT, 2005).

A análise sensorial é uma ciência que envolve a avaliação de atributos organolépticos de um produto através dos sentidos (ISO 5492, 2008). No setor alimentício, a análise sensorial tem grande importância na avaliação da aceitabilidade no mercado e qualidade do produto, sendo imprescindível para o controle de qualidade industrial (TEIXEIRA, 2009).

Para se obter resultados adequados, é importante selecionar uma equipe de pessoas, denominadas juízes ou degustadores, responsável por avaliar sensorialmente um produto, constituindo o painel de análise sensorial (TEIXEIRA, 2009). Os painéis desempenham um papel importante na ciência sensorial, a qual está intimamente ligada à sensometria. Enquanto a ciência sensorial lida com a percepção humana dos estímulos e a maneira como eles atuam, a sensometria é o campo da Estatística que analisa dados dessa ciência (BROCKHOFF, 2011).

A seleção e treinamento de juízes, com habilidades discriminativas, são etapas que permitem formar painéis pequenos e bem eficientes. O treinamento desenvolve a memória e sensibilidade de um indivíduo a fim de se obter medidas sensoriais consistentes, precisas e padronizadas. Sendo assim, membros do painel devem ser treinados de forma a desenvolver e estabelecer um vocabulário descritivo em conotação com o que for previamente estabelecido, desconsiderando suas preferências pessoais (POSTE *et al.*, 1991).

Para averiguar a habilidade discriminativa, utiliza-se bastante o teste triangular na seleção de provadores, em que se aplica uma série de ensaios para calcular a porcentagem correta de identificação das amostras diferentes. Além disso, esse teste pode ser empregado para avaliar se há alteração perceptível entre duas amostras submetidas a processamentos diferentes ou que passaram por alguma substituição de ingrediente (POSTE *et al.*, 1991). Portanto, além de ser relevante ao processo de seleção, o teste triangular tem aplicação no controle de qualidade.

Na seleção de provadores, há uma discussão a respeito dos critérios a serem utilizados (ALVELOS, 2002). Os critérios são bem diversificados, podendo variar a porcentagem requerida para seleção bem como o número de repetições do teste triangular. Diante da enorme importância dessas etapas, o teste estatístico empregado para decidir se o candidato deve ou não

passar para a fase de treinamento deve apresentar características ótimas de erro tipo I e poder. Ou seja, deve ser capaz de recusar candidatos com pouca habilidade discriminativa e acolher candidatos habilidosos.

Sabe-se que a matemática envolvida na estimação intervalar e no teste de hipóteses está intimamente relacionada. Ambas são grandes áreas da inferência estatística, em que é possível desenvolver métodos para testar hipóteses e aplicá-los em alguns problemas. Concomitantemente, a obtenção de estimadores intervalares fornece uma ideia de qual seria um ótimo intervalo. A estimação e decisão levam a uma afirmação sobre os parâmetros (MOOD; GAYBILL; BOES, 1974).

Sendo assim, a transformação de um estimador intervalar em teste é bastante corriqueira. No entanto, deseja-se estudar possíveis testes que permitam selecionar provadores de forma que sejam cometidas menores taxas de erros, a fim de recomendar o melhor. Então, este trabalho tem como objetivo, reescrever estimadores intervalares como testes, adaptados a seleção de provadores em testes triangulares, e compará-los por desempenho (erro tipo I e poder), elegendo assim o de melhor desempenho a fim de recomendá-lo.

3 METODOLOGIA

3.1 ESTABELECIMENTO DOS TESTES

Como o interesse do presente trabalho é a comparação de testes de hipóteses acerca da proporção de acertos, faz-se necessário reescrever convenientemente alguns estimadores intervalares em pontos críticos ao longo do número de ensaios (n). Os testes foram obtidos a partir de três estimadores intervalares da Normal, denominados TN1, TN2 e TN3; um da Distribuição F denominado TF; um do teste Sequencial ou TS; e um da aproximação da Poisson, utilizando o quantil χ^2 , chamado TP. Os estimadores intervalares para o parâmetro p de uma binomial, utilizados na obtenção desses testes, foram provenientes de Ferreira (2005) e o teste sequencial de Shirose; Mori (1996).

Também, deve-se levar em consideração que na utilização desses testes assume-se que as variáveis aleatórias são independentes em cada ensaio e possuem probabilidade constante de ocorrência de acerto ou fracasso. Portanto, a distribuição Binomial se trata de n ensaios independentes de Bernoulli com p constante.

Em cada ensaio do teste triangular são apresentadas duas amostras iguais e uma diferente. Então, pede-se ao avaliador que identifique a amostra diferente (POSTE *et al.*, 1991). A probabilidade de o avaliador selecionar corretamente a amostra é $1/3$ e de selecionar incorretamente é $2/3$ (PAES, 2005). Por isso, se em n ensaios o candidato acertar a uma proporção de $1/3$, significa que ele não está identificando corretamente a amostra e o acerto é ao acaso. No entanto, se ele acertar a uma proporção maior que $1/3$, ele está identificando corretamente a amostra diferente.

O par de hipóteses em estudo, para n ensaios triangulares é dado por (NÆS; BROCKHOFF; TOMIC, 2010):

$$\begin{cases} H_0 : p = \frac{1}{3} \\ H_a : p > \frac{1}{3} \end{cases}$$

Nota-se, pelo par de hipóteses, que há uma lateralidade no teste. Por construção unila-

teral à direita e assumindo $p = 1/3$ e $q = 2/3$ para TN1, tem-se Eq. 8:

$$PC_1(n) = \frac{1}{3} + z_{0,05} \sqrt{\frac{2}{9n}} \quad (8)$$

Essa expressão representa o ponto crítico, escrito como função do número de ensaios (n). Se a estatística de teste for maior que o ponto crítico, decide-se por rejeitar H_0 . Se for menor, por aceitar.

Sendo assim, para os demais testes baseados nos estimadores intervalares em que são atribuídos o número de sucessos em determinado número ensaios, realiza-se a substituição de y pelo valor esperado de sucessos em n ensaios, que equivale à esperança de uma Binomial, ou seja, $E(X) = np$. Dessa forma, para $y = n/3$ e $z_{0,05} = 1,64$, o TN2 fornece Eq. 9:

$$PC_2(n) = \frac{n + 1,5}{3n} + \frac{z_{0,05}}{\sqrt{n}} \sqrt{\left(\frac{1}{3} + \frac{0,5}{n}\right) \left(\frac{2}{3} + \frac{0,5}{n}\right)} \quad (9)$$

Impondo H_0 para obter o TN3, tem-se Eq. 10:

$$PC_3(n) = \frac{\frac{2n + 3z_{0,05}^2}{6} + z_{0,05} \sqrt{\frac{3n - n}{9} + 0,41}}{n + z_{0,05}^2} \quad (10)$$

Observa-se que o somatório do número de sucessos em n ensaios tem uma relação com H_0 , já que se pode assumir um valor esperado de acordo com a proporção mínima de acertos p , que equivale à proporção de acertos ao acaso $1/3$. Logo, y pode ser calculado pela multiplicação de n por $1/3$.

Para o TF, realizando as devidas substituições de y , a equação 11 se torna:

$$PC_4(n) = \frac{\left(\frac{n}{3} + 1\right) F_1}{\left(\frac{n}{3} + 1\right) F_1 + \left(n - \frac{n}{3}\right)}; \quad \text{com} \quad F_1 = F_{0,05; 2n/3 + 2, 4n/3} \quad (11)$$

Com relação ao TS, para que se possa trabalhar com proporção de acertos, as funções correspondentes às curvas de seleção (L_1) e não seleção (L_0) foram divididas por n , fornecendo as respectivas equações apresentadas em 12:

$$PC_{5,1}(n) = \frac{h_1}{n} + s; \quad PC_{5,0}(n) = \frac{-h_0}{n} + s \quad (12)$$

em que h_0 , h_1 e s equivalem, respectivamente, às equações dadas em 13 (Shirose; Mori, 1996).

$$h_0 = \frac{\log\left(\frac{1-\alpha}{\beta}\right)}{\log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)}; \quad h_1 = \frac{\log\left(\frac{1-\beta}{\alpha}\right)}{\log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)}; \quad s = \frac{\log\left(\frac{1-p_0}{1-p_1}\right)}{\log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)} \quad (13)$$

com $\alpha = \beta = 0,05$.

Na teoria sequencial, devem-se estabelecer valores de p_0 e p_1 (tal que $p_0 < p_1$), de modo que o candidato não seja selecionado sempre que $p \leq p_0$ e seja selecionado quando $p \geq p_1$ (WALD, 2013). Como na prática de seleção de provadores podem-se aplicar 4 testes triangulares, devendo o candidato acertar pelo menos 75%, conforme utilizado por Minim *et al.* (2010), Oliveira e Benassi (2010), decidiu-se por adotar $p_1 = 3/4$ e $p_0 = 1/3$.

Esse teste é composto por duas curvas em que a decisão só pode ser tomada ao atingir uma dessas curvas. Caso nenhuma delas seja atingida, deve-se dar prosseguimento ao teste triangular. Portanto, para essa abordagem em particular, foram consideradas ambas as curvas.

Por fim, para o TP, tem-se a equação 14:

$$PC_6(n) = \frac{\chi_{0,05;(\frac{2n}{3}+2)}^2}{2n} \quad (14)$$

Nota-se que em todos os testes definidos, os pontos críticos dependem apenas de n . Concomitantemente, para todos os testes é assumida a mesma estatística de teste, a saber Eq. 15:

$$p_{ic} = \frac{\sum_{i=1}^n x_i}{i}; \quad i = 1, \dots, n \quad (15)$$

em que p_{ic} é a proporção de acertos calculada no i -ésimo ensaio; e $x_i=0, 1$ é associado a um ensaio de Bernoulli, assumindo 0 quando fracasso (erro) ou 1 quando sucesso (acerto).

Na etapa posterior, foi realizado um estudo de simulação Monte Carlo para comparar os testes.

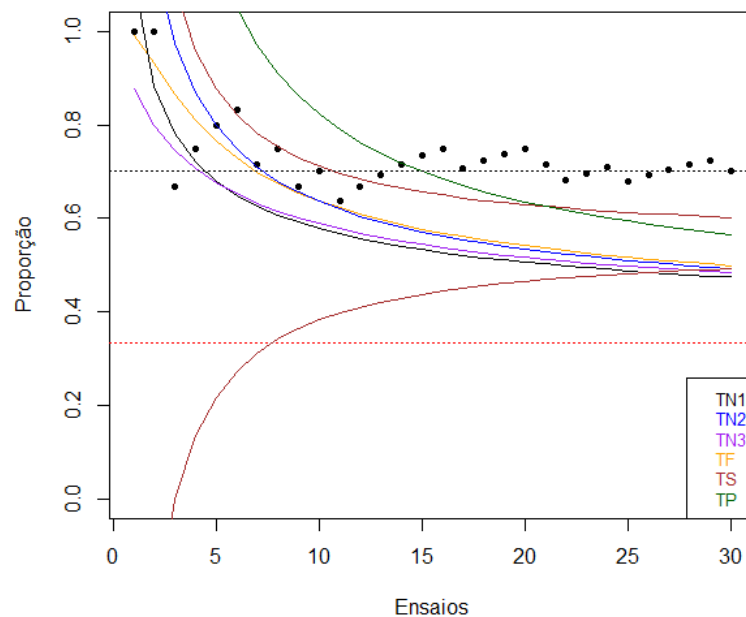
3.2 SIMULAÇÃO MONTE CARLO

Por meio de um estudo de simulação Monte Carlo, foram gerados 1000 provadores virtuais submetidos a testes triangulares discriminativos ao longo de 5, 10, 15, ..., 100 ensaios. Esses provadores foram gerados com probabilidade de acerto constante ao longo do tempo, fixada em 0,33, 0,35, 0,40, 0,45, ..., 1,00. Os testes foram assim comparados em termos de erro tipo I e poder, computando-se a proporção de vezes em que H_0 era rejeitada para cada valor do parâmetro p . O teste que teve menor erro tipo I e maior poder foi recomendado para um pequeno ou grande número de ensaios.

4 RESULTADOS E DISCUSSÃO

Os testes aqui comparados para proporção de uma binomial fornecem curvas com diferentes pontos críticos ao longo do número de ensaios. Para avaliar o comportamento desses testes e ilustrar o procedimento de simulação, foi apresentado na Figura 4 o gráfico de um provador virtual com uma proporção de acerto constante de 0,7.

Figura 4 – Gráfico dos testes aplicados a um provador virtual com 0,7 de acerto.



Fonte: Da autora.

A linha tracejada, em preto, destaca a aleatorização feita para um provador com 0,7 de chance de acerto, cujos pontos representam sua proporção de respostas corretas. Dessa forma, cada ponto é calculado a partir da estatística de teste, ou seja, cada ponto fornece a razão entre o número de acertos do avaliador em relação ao número de ensaios percorridos. Já a linha tracejada em vermelho destaca $1/3$.

Todos os testes obtidos a partir do limite superior dos estimadores intervalares apresentaram apenas uma curva que divide duas regiões de decisão. Em contrapartida, somente o teste sequencial (TS) apresentou duas curvas de decisão, sendo que a decisão é tomada somente ao atingir uma das curvas ou quando ultrapassar seus limites. Dessa forma, o TS não decide entre as duas curvas.

Com base na teoria da decisão e nas hipóteses de pesquisa, nota-se que os testes separam regiões críticas. As proporções acima da curva atingem uma região de rejeição de H_0 , enquanto

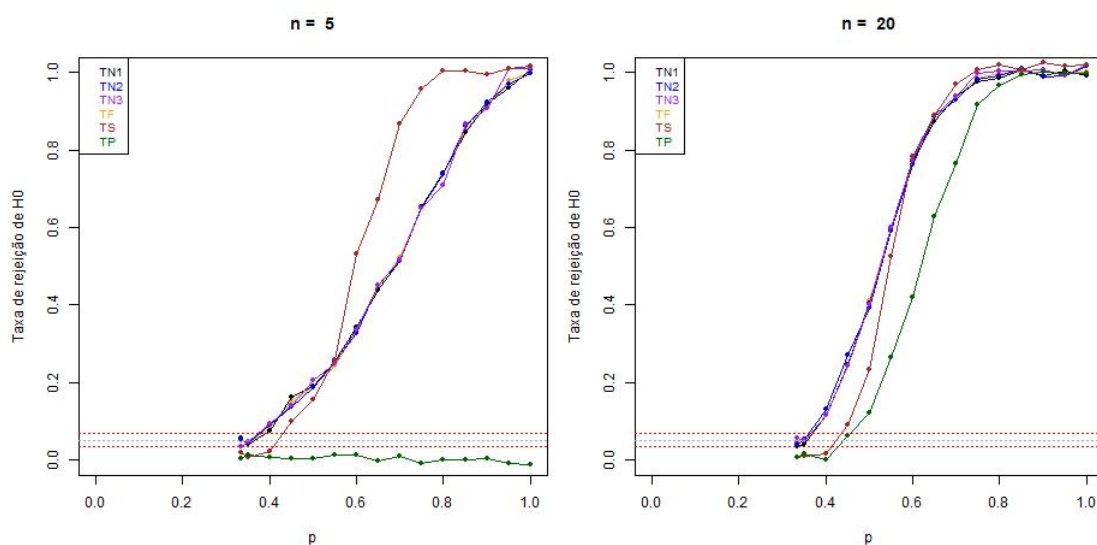
as proporções abaixo da curva atingem a região de aceitação de H_0 . Para o TS, em especial, a região crítica é dada acima da curva superior (rejeição de H_0) e abaixo da curva inferior (aceitação de H_0). Entre as curvas, deve-se apenas dar prosseguimento ao teste triangular.

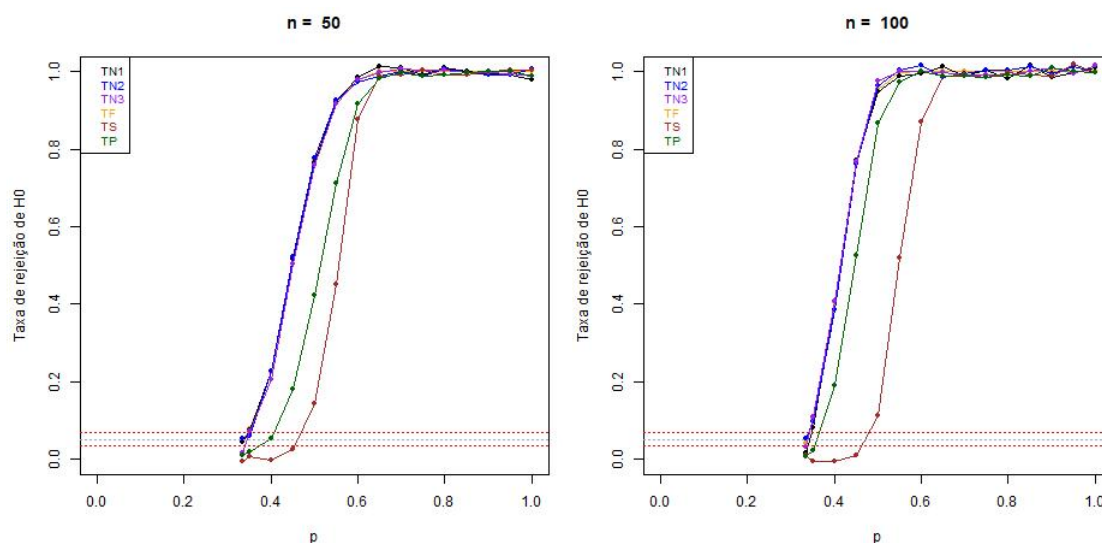
Observa-se que o TP é bastante exigente, portanto a proporção de acertos deve se manter bem elevada, durante os 7 primeiros ensaios, para ultrapassar a curva dessa função e o teste começar a detectar que o juiz está identificando diferença nas amostras. Também, é possível notar que essa exigência se mantém elevada até o 20º ensaio, comparada aos demais testes. Nos ensaios 21 e 22, o TP se torna igualmente exigente quando comparado ao TS que, no 23º ensaio, se torna o mais exigente, dentre os apresentados.

Com relação aos demais testes, apenas aqueles referidos como TN3 e TF detectam avaliadores que identificam diferenças nas amostras a partir do primeiro ensaio. Para avaliadores que acertam a uma proporção de 100%, os testes TN1, TN2 e TS, só fazem essa detecção a partir do 2º, 3º e 4º ensaios, respectivamente. É possível notar que no 5º e 6º ensaios, os testes denominados TN1 e TN3 se igualam e possuem comportamentos similares a partir de então, bem como o TN2 e TF se igualam no 9º e 10º ensaios.

Para análise da hipótese alternativa (H_1), em que foram realizadas simulações para as proporções de 0,35 a 1,00, foram construídos gráficos de poder, para todos os testes analisados. A Figura 5 mostra essas curvas para $n = 5, 20, 50$ e 100. Os demais gráficos são apresentados no Apêndice A.

Figura 5 – Gráfico do poder para 5, 20, 50 e 100 ensaios.





Fonte: Da autora.

Nota-se que todas as curvas do poder partem do intervalo do erro tipo I com o aumento do número de ensaios, exceto para os testes TS e TP. Estes partem de valores inferiores ao limite inferior. Pode-se dizer que os testes TS e TP são conservadores (possuem menor erro tipo I), embora o TP adquira alto poder e o TS adquira baixo poder, com o aumento de n . Quanto aos demais testes são ditos liberais (possuem maior erro tipo I), ainda dentro do limite desejado para o α escolhido, e maior poder. Harrison (2011) sugere que os testes podem ser ditos conservadores ou liberais, quando se realiza essa comparação aos limites de α . Todos os gráficos são apresentados no Apêndice A.

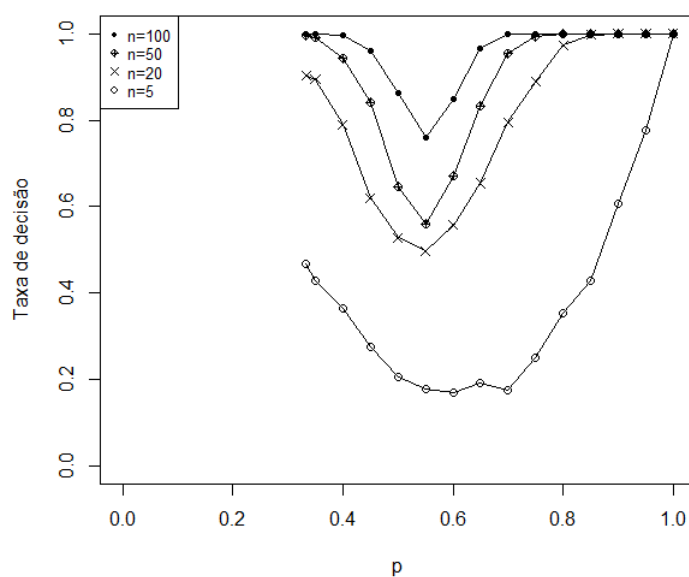
Observa-se que, em 5 ensaios, o TP possui poder nulo, já que um número muito pequeno de ensaios não é suficiente para que o teste decida rejeitar H_0 . Isso ocorre devido a alta exigência desse teste nos 7 primeiros ensaios. Concomitantemente, o TS apresenta poder elevado em poucos ensaios. No entanto, vale ressaltar que o TS não toma a decisão para todos os avaliadores simulados, pois há uma região de indecisão entre os limites do teste, em que apenas pede-se a continuidade do teste triangular até que a proporção atinja um dos limites. Inclusive, nota-se que o poder do TP somente supera o poder do TS com um grande número de ensaios, como pode ser observado nos gráficos obtidos para $n = 50$ e $n = 100$.

À medida que ocorre o aumento do número de ensaios, as curvas do poder de todos os testes tendem ao formato sigmóide. Ao analisar o comportamento dos testes em todos os ensaios, apresentados no Apêndice A, nota-se que o TS apresenta melhor comportamento até o décimo nono ensaio. No vigésimo ensaio, esse teste apresenta apenas uma pequena região de maior poder (entre $p = 0,6$ e $p = 0,8$), enquanto o TN3 adquire melhor comportamento

para os outros valores de p a partir do 20º adiante. No entanto, de maneira geral, pode-se dizer que as aproximações Normais são adequadas para 20 ensaios ou mais, o que é esperado em aproximações assintóticas. Apenas na ausência de qualquer ruído, observa-se que o TN3 possui menor taxa de erro tipo I e maior poder para n maior ou igual a 20, enquanto o TS possui menor erro tipo I e maior poder, para n inferior a 20.

Dentre os 1000 provadores simulados, o TS não toma uma decisão para todos os avaliadores virtuais, ele apenas decide para uma parcela desses avaliadores, já que alguns se mantêm na região em que deve-se dar prosseguimento ao teste triangular. No entanto, as taxas de rejeição e aceitação de H_0 podem ser calculadas a partir do número de decisões realizadas. Logo, construiu-se o gráfico da taxa de decisão, do TS, em relação a p para os ensaios 5, 20, 50 e 100, como mostra a Figura 6.

Figura 6 – Gráfico da taxa de decisão em relação a p , para diferentes números de ensaios (5, 20, 50 e 100), no TS.



Fonte: Da autora.

Nota-se que as taxas de decisão são elevadas inicialmente e reduzem a uma proporção de aproximadamente 50% ou 55%, retornando ao comportamento ascendente. Essa queda da taxa de decisão ocorre devido a essa região de proporções com aleatorizações inconclusivas. Ou seja, no TS, não se toma a decisão para juízes que acertam a uma taxa de 50%, enquanto aqueles que acertam a uma taxa bem inferior, pode-se tomar a decisão de não selecioná-los no processo, bem como, aqueles que acertam a uma taxa elevada, pode-se decidir selecioná-los no processo.

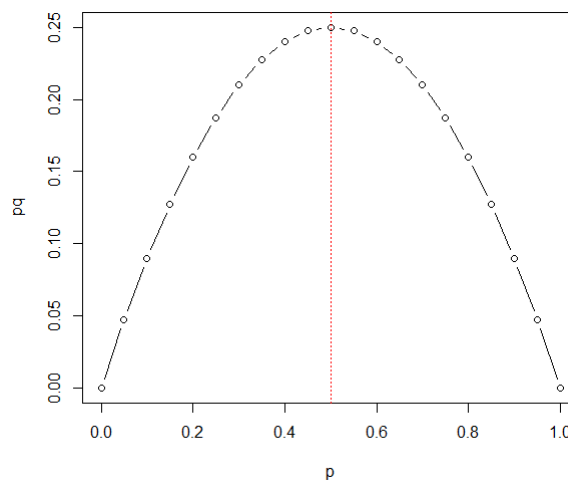
Observa-se, também, que as taxas de decisão são altas quando o número de ensaios é elevado. Logo, quanto maior o número de ensaios, maiores são as taxas de decisão, seja para selecionar ou não o candidato.

Segundo Miot (2011), o tamanho amostral é calculado por:

$$n = \left(\frac{z_\alpha}{\epsilon} \right)^2 \hat{p}\hat{q}$$

em que ϵ é a estimativa de erro para o parâmetro p , estipulado pelo pesquisador e z_α é tabelado. Logo, n depende de $\hat{p}\hat{q}$. A Figura 7 apresenta a variação de pq em relação a p .

Figura 7 – Gráfico da variação de pq em relação a p .



Fonte: Da autora.

Nota-se que a proporção $p = 0,5$ é a que exige maior tamanho amostral para ser estimada, o que também justifica a necessidade de maior número de ensaios para a tomada de decisão quando a proporção de acertos é 50%.

Quando um juiz acerta com uma proporção bem pequena, pode-se decidir aceitar H_0 e não selecioná-lo como provador, bem como aquele que acerta a uma alta proporção, pode-se tomar a decisão de rejeitar H_0 e selecioná-lo como provador. Entretanto, aqueles que acertam a uma proporção intermediária de 50%, por exemplo, não é possível tomar uma decisão, seja para selecioná-lo ou não. Para esses candidatos, geralmente o TS recomenda continuar a execução do teste triangular até que se tome uma decisão adequada.

Muitos autores trabalham com um número pequeno de ensaios na seleção de provadores, como Minim *et al.* (2010), Scheid (2001), Oliveira e Benassi (2010), que estipularam 4 ensaios. Nessas situações, recomenda-se o uso do TS.

5 CONCLUSÕES

A adaptação dos estimadores intervalares para proporção permitiu avaliar potenciais testes a serem empregados na seleção de provadores, com proporções de acertos constantes. Notou-se que, em todos os testes, as taxas de erro tipo I se aproximavam do intervalo do valor de significância estabelecido e a curva do poder apresentava característica sigmoideal, com o aumento dos números de ensaios. Entretanto, os testes TS e TP foram mais conservadores, enquanto os demais testes apresentaram características liberais.

A simulação permitiu gerar diferentes comportamentos de valores de proporções constantes ao longo dos ensaios. Através da menor taxa de erro tipo I cometido e maior poder, foi possível definir o melhor teste para valores grandes e pequenos de n . Dessa forma, para 20 ensaios ou mais, recomenda-se o TN3, enquanto para um número inferior a 20 ensaios, recomenda-se o TS. Todavia, deve-se ressaltar a adequação das aproximações Normais que, em geral, apresentaram curvas com alto poder para um número de ensaios maior ou igual a 20.

REFERÊNCIAS

- ALVELOS, H. M. P. P. D. **Análise, Desenvolvimento e Teste de Métodos e Técnicas para Controlo Estatístico em Análise Sensorial**. 2002. 253f. Tese (Doutorado em Ciências de Engenharia) - Faculdade de Engenharia, Universidade do Porto, Porto, 2002.
- BROCKHOFF, P. B. Sensometrics for Food Quality Control. **Scandinavian Workshop on Imaging Food Quality**, Dinamarca, n. 15, p. 7-16. 2011.
- FERREIRA, D. F. **Estatística Básica**. Lavras: UFLA, 2005.
- GRUNERT, K. G. Food quality and safety: consumer perception and demand. **European Review of Agricultural Economics**, v. 32, n. 3, p. 369–391. 2005.
- HARRISON, M. T. **Conservative Hypothesis Tests and Confidence Intervals using Importance Sampling**. Providence: Brown University. 2011.
- ISO 5492 (2008): **Sensory analysis – Vocabulary**, 2ª ed., International Organization for Standardization. Disponível em: <<https://www.iso.org/obp/ui/#iso:std:iso:5492:ed-2:v1:en>>. Acesso em: 09 mar. 2019.
- MINIM, V. P. R. *et al.* Análise descritiva: Comparação entre metodologias. **Rev. Inst. Latic. Cândido Tostes**, v. 65, n. 374, p. 41-48. 2010.
- MIOT, H. A. Tamanho da amostra em estudos clínicos e experimentais. **J. Vasc. Bras.**, v. 10, n. 4, p. 275-278. 2011.
- MOOD, A. M., GAYBILL, D. C., BOES, C. **Introduction of theory of statistics**. 3ª ed., McGraw-Hill. 1974.
- NÆS, T.; BROCKHOFF, P. B.; TOMIC, O. **Statistics for Sensory and Consumer Science**. London: Wiley. 2010.
- OLIVEIRA, A. P. V.; BENASSI, M. T. Avaliação sensorial de pudins de chocolate com açúcar e dietéticos por perfil livre. **Ciênc. agrotec.**, v. 34, n. 1, p. 146-154. 2010.
- PAES, J. B. **Métodos de diferença em Avaliação Sensorial de Alimentos e Bebidas**. 3ª ed. Viçosa: UFV, 2005.
- POSTE, L. M. *et al.* **Laboratory Methods for Sensory Analysis of Food**. Canada: Agriculture Canada Publication, 1991.
- SCHEID, G. A. **Avaliação Sensorial e Físico-Química de Salame Tipo Italiano com Diferentes Concentrações de Cravo da Índia (*Eugenia caryophyllus*)**. 2001. 94f. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) - Universidade Federal de Viçosa, Viçosa, 2001.
- SHIROSE, I.; MORI, E. E. M. **Estatística aplicada à análise sensorial (Módulo 2)**.

Campinas: ITAL - Manual Técnico nº 13, 1996.

TEIXEIRA, L.V. Análise sensorial na indústria de alimentos. **Rev. Inst. Latic. “Cândido Tostes”**, v. 64, n. 366, p. 12-21, 2009.

WALD, A. **Sequential Analysis**. New York: Dover Publications, 2013.

CAPÍTULO 3 - TESTES PARA PROPORÇÃO SOB QUEBRA DE PRESSUPOSIÇÃO: OS CASOS DA FADIGA E MEMÓRIA SENSORIAIS

RESUMO: A análise sensorial de alimentos é um procedimento experimental utilizado para medir a resposta de uma pessoa com relação a sensações produzidas pelos alimentos, através dos sentidos. No entanto, a nota dos provadores pode ser influenciada pela fadiga sensorial, ocorrida quando os provadores são submetidos a um grande número de ensaios e perdem a capacidade de discriminar os diferentes estímulos aos quais estão sujeitos, o que pode afetar a qualidade da análise. Concomitantemente, provadores podem desenvolver memória e sensibilidade sensorial através do treinamento, havendo um processo de aprendizagem. A fadiga sensorial provocará uma queda da proporção de acertos e o aprendizado ocasionará um crescimento da proporção de acertos. Dessa forma, deseja-se avaliar testes obtidos a partir de estimadores intervalares da distribuição Normal (TN1, TN2 e TN3), Distribuição F (TF), Teste Sequencial (TS) e aproximação da Poisson com quantil χ^2 (TP), através da simulação de provadores submetidos a testes triangulares que adquiriram fadiga ou aprendizado gerados por modelos lineares e não lineares, a fim de verificar o comportamento desses testes com a quebra de pressuposição de proporção constante. Para isso, foram efetuadas simulações de Monte Carlo para 1000 provadores, ao longo dos ensaios 5, 10, 15, ..., 100, que se cansam de forma linear e não linear a partir de um acerto inicial de 100% com queda até 1/3. Para o aprendizado, considerou-se um acerto inicial de 1/3 crescente até 100%. Observou-se que a quebra de pressuposição de p constante afeta bastante as taxas de erro tipo I e poder dos testes. As curvas de poder para os modelos linear e não linear da fadiga apresentaram características bem similares, com altas taxas de erro tipo I e baixo poder com o aumento de p , em que n é pequeno. Já as curvas de poder para os modelos linear e não linear do aprendizado foram crescentes com o aumento de p e n . O valor de n tem maior influência sobre as curvas de poder do que p , sendo que a estatística de teste foi outro fator que interferiu na eficiência dos testes devido a sua memória longa. Além disso, o TS apresentou menor poder para as funções aprendizado em um grande número de ensaios, adquirindo melhor comportamento quando n é pequeno e o TP adquiriu menor taxa de erro tipo I em todas as situações. Os testes TN1, TN2 e TF apresentaram simultaneamente menor taxa de erro tipo I e maior poder tanto para a função linear do aprendizado quanto não linear com o aumento de n . No entanto, não se recomenda o uso de nenhum dos testes para a fadiga sensorial.

Palavras-chave: Fadiga sensorial. Aprendizado sensorial. Sensometria.

PROPORTION TESTS UNDER PRESUPTION BREAK: THE CASES OF FATIGUE AND SENSORY MEMORY

ABSTRACT: Food sensory analysis is an experimental procedure used to measure a person's response to sensations produced by food through the senses. However, when they are subjected to a large number of trials, tasters' grades may be influenced by sensory fatigue, losing their ability to discriminate between different stimuli to which they are subjected, which may affect the quality of the analysis. Concomitantly, tasters can develop memory and sensory sensitivity through training, occurring a learning process. As sensory fatigue causes a decrease in the proportion of correct answers and learning causes an increase in the proportion of hits, it is desired to evaluate tests obtained from interval estimators from Normal distribution (TN1, TN2 and TN3), F Distribution (TF), Sequential Test (TS) and Poisson approximation with χ^2 quantile (TP), by simulating tasters submitted to triangular tests that developed tiredness or learning generated by linear and nonlinear models, in order to verify the behavior of these tests breaking constant proportion assumption. For this, Monte Carlo simulations were performed for 1000 tasters throughout the tests 5, 10, 15, ..., 100, which get tired linearly and nonlinearly from an initial hit of 100 % with drop up to 1/3. For learning, an initial hit of 1/3 increasing up to 100% was considered. Breaking p constant assumption affects a lot type I error rates and testing power. The power curves for linear and nonlinear models of fatigue showed similar characteristics, with high type I error rates and low power with the increase of p , when n is small. Power curves for linear and nonlinear models of learning, on the other hand, increased with the increase of p and n . The n value has more influence on power curves than p , besides, another factor that interfered the tests efficiency was test statistic, due to its long memory. Also, TS presented less power for the learning functions in a large number of trials, showing better behavior when n is small and TP presented lower type I error rate in all situations. The TN1, TN2 and TF tests presented lower type I error rate and higher power for both linear and nonlinear functions of learning when n increases. However, it's not recommended to use any of the tests for sensory fatigue.

Keywords: Sensory fatigue. Sensory learning. Sensometrics.

6 INTRODUÇÃO

A análise sensorial de alimentos é um procedimento experimental utilizado para medir a resposta de uma pessoa com relação a sensações produzidas pelos alimentos, através dos sentidos de visão, paladar, olfato, audição e tato. Esse procedimento é muito útil em alimentos e bebidas variados que produzem sensações como, por exemplo, sabor, aroma e textura. No entanto, a avaliação dos provadores pode ser influenciada pela fadiga sensorial, que ocorre quando os provadores são submetidos a uma grande quantidade de ensaios sensoriais de alimentos ou bebidas e perdem a sensibilidade aos estímulos aos quais foram sujeitos (ZWIMPFER, 2013).

As respostas humanas sobre os alimentos devem ser medidas de forma exata, a fim de minimizar efeitos que possam enviesar a análise de dados, tais como marca e quaisquer influências sobre a percepção do consumidor. A ideia é isolar as propriedades sensoriais, que são informações relevantes e bastante úteis para desenvolvedores, cientistas dos alimentos e gerentes (LAWLESS; HEYMANN, 2010).

A discriminação ou habilidade de diferenciar dois estímulos é fundamental nas respostas de análise sensorial. Inclusive, testes discriminativos são muitas vezes utilizados para coletar evidências de que um produto sofreu alteração (LAWLESS; HEYMANN, 2010). Dentre eles, destacam-se os testes triangulares, comumente utilizados na seleção de provadores através da aplicação de uma série de ensaios (POSTE *et al.*, 1991).

Em uma sessão teste, o número de amostras a serem apresentadas é uma função do tipo de produto a ser analisado e o fator psicológico dos provadores. O número mínimo depende do método do teste, no entanto, há uma preocupação com o máximo permissível, na maioria deles. A quantidade de amostras depende do quão rápido os provadores se tornam fatigados ou adaptados aos produtos. Quando o número de amostras ultrapassa um determinado ponto, os avaliadores podem se tornar fatigados e os resultados começam a apresentar menos discriminação (CHAMBERS IV; WOLF, 2005).

Existem testes em que não se pode realizar muitas repetições porque os sentidos se tornam rapidamente entorpecidos e, portanto, podem ser obtidas impressões erradas através das várias tentativas em avaliar as amostras. O tempo e a precisão funcionam em favor do pequeno número de amostras. Quanto mais forte o gosto e o odor de uma substância, menor o número de amostras que um indivíduo pode testar antes de descansar (HARRIS *et al.*, 1951).

Através dessas impressões erradas, torna-se visível o alcance da fadiga sensorial com a

queda da proporção de acertos. Adicionalmente, segundo Poste *et al.* (1991), uma pessoa pode desenvolver memória e sensibilidade quando submetida ao treinamento. Portanto, ao longo dos ensaios, o indivíduo pode se cansar e adquirir uma queda na proporção de acertos; ou aprender e adquirir um crescimento na proporção de acertos.

Como nessas duas situações a proporção paramétrica não se mantém constante, deseja-se avaliar alguns testes obtidos a partir de estimadores intervalares para proporção, em que provadores submetidos a ensaios triangulares foram simulados (adquirindo cansaço por meio de um modelo linear e não linear). De forma similar, deve-se levar em consideração o processo de aprendizagem com modelagem linear e não linear. O propósito da análise é verificar o comportamento desses testes com a quebra de pressuposição para proporção constante.

7 METODOLOGIA

7.1 TESTES ESTABELECIDOS

Tendo em vista que a fadiga sensorial pode causar a perda da credibilidade dos dados na análise sensorial, é importante identificá-la através de testes estatísticos e avaliar o impacto da quebra de pressuposição de p constante nesses testes. Concomitantemente, o aprendizado causa a quebra dessa pressuposição, devendo também ser avaliado. Para isso, foram realizadas adaptações de estimadores intervalares para proporção obtidos de Ferreira (2005) e do teste sequencial de Shirose; Mori (1996).

Os testes foram obtidos a partir de alguns estimadores intervalares, sendo três da Normal, um da Distribuição F, um proveniente do teste Sequencial e um oriundo da aproximação da distribuição Poisson, em que utiliza-se o quantil χ^2 . Cada um desses testes receberam nomes específicos sendo, respectivamente, TN1, TN2, TN3, TF, TS e TP.

Espera-se que o avaliador selecione a amostra correta com $1/3$ de chance de acerto ao acaso. Portanto, ao longo dos ensaios, as hipóteses de pesquisa são (NÆS; BROCKHOFF; TOMIC, 2010):

$$\begin{cases} H_0 : p = \frac{1}{3} \\ H_a : p > \frac{1}{3} \end{cases}$$

O TN1 obtido com as devidas substituições das proporções de sucesso ($p = 1/3$) e fracasso ($q = 2/3$) do teste triangular, fornece (16):

$$PC_1(n) = \frac{1}{3} + z_{0,05} \sqrt{\frac{2}{9n}} \quad (16)$$

em que $z_{0,05} = 1,64$, por se tratar de uma hipótese unilateral.

Já o TN2, em que foi realizada a substituição de y pelo valor esperado de respostas corretas $n/3$, corresponde a (17):

$$PC_2(n) = \frac{n + 1,5}{3n} + \frac{z_{0,05}}{\sqrt{n}} \sqrt{\left(\frac{1}{3} + \frac{0,5}{n}\right) \left(\frac{2}{3} + \frac{0,5}{n}\right)} \quad (17)$$

O TN3, para mesmos valores de y , é dado por (18):

$$PC_3(n) = \frac{\frac{2n + 3z_{0,05}^2}{6} + z_{0,05} \sqrt{\frac{3n - n}{9} + 0,41}}{n + z_{0,05}^2} \quad (18)$$

para mesmo valor de $z_{0,05}$.

O TF, adquirido pelas devidas substituições de y no limite superior do estimador intervalar da distribuição F com aproximação pela Binomial, equivale a (19):

$$PC_4(n) = \frac{\left(\frac{n}{3} + 1\right) F_1}{\left(\frac{n}{3} + 1\right) F_1 + \left(n - \frac{n}{3}\right)} \quad (19)$$

cujo quantil F_1 corresponde a $F_{0,05;2n/3+2,4n/3}$.

O TS fornece duas curvas apresentadas em (20). Esse teste em particular tem uma característica de apenas tomar uma decisão quando a proporção atinge uma das curvas, caso contrário, não toma decisão e requer a continuidade dos ensaios triangulares.

$$PC_{5,1}(n) = \frac{h_1}{n} + s; \quad PC_{5,0}(n) = \frac{-h_0}{n} + s \quad (20)$$

Nesse caso, tomou-se os limites superior e inferior do intervalo da Sequencial, dividido por n , já que é desejável trabalhar com proporção. Os valores de h_0 , h_1 e s são dadas pelas respectivas relações apresentadas em 21 (SHIROSE; MORI, 1996).

$$h_0 = \frac{\log\left(\frac{1 - \alpha}{\beta}\right)}{\log\left(\frac{p_1(1 - p_0)}{p_0(1 - p_1)}\right)}; \quad h_1 = \frac{\log\left(\frac{1 - \beta}{\alpha}\right)}{\log\left(\frac{p_1(1 - p_0)}{p_0(1 - p_1)}\right)}; \quad s = \frac{\log\left(\frac{1 - p_0}{1 - p_1}\right)}{\log\left(\frac{p_1(1 - p_0)}{p_0(1 - p_1)}\right)} \quad (21)$$

Assim, as proporções p_0 e p_1 foram estabelecidos como 1/3 e 3/4, que seriam, respectivamente, as taxas de acerto ao acaso e a taxa usual na pré-seleção de provedores. Na pré-seleção de provedores é comumente empregado uma série de quatro testes triangulares, em que a proporção de acerto deve ser pelo menos 75% (MINIM *et al.*, 2010). Portanto, utilizou-se o valor 3/4 como referência à proporção de acerto do provedor no teste triangular. Adicionalmente, tomou-se as respectivas taxas de erro tipo I e II, $\alpha = \beta = 0,05$.

E, por fim, para TP tem-se (22):

$$PC_6(n) = \frac{\chi_{0,05;(\frac{2n}{3}+2)}^2}{2n} \quad (22)$$

sendo adquirido pelo limite superior do estimador intervalar com aproximação da Poisson para Binomial e quantil $\chi_{0,05;(\frac{2n}{3}+2)}^2$.

A estatística de teste assumida para todos os testes é dada por $p_{ic} = \frac{\sum_{i=1}^n x_i}{i}$ com $i = 1, \dots, n$; em que p_{ic} refere-se à proporção de acertos no i -ésimo ensaio, portanto corresponde ao somatório do número de acertos acumulados em relação ao ensaio i .

Posteriormente, foram realizadas simulações para averiguar o melhor teste a ser indicado em termos de menor taxa de erro tipo II e maior poder.

7.2 OS MODELOS LINEAR E NÃO LINEAR PARA O CANSAÇO

Para simular o provador que cansa, devem-se analisar duas funções do cansaço: um modelo linear e outro não linear. Assim, para demonstrar o modelo linear do cansaço temos a proporção de acerto inicial do provador (p_1), que se refere à habilidade inicial específica de cada provador para acertar no teste. Portanto, este é um valor bastante variável. Para cada proporção seguinte p_2 e p_3 , há uma redução de valor de acordo com a queda do desempenho c que foi estimado:

$$\begin{aligned} p_2 &= p_1 - 1c \\ p_3 &= p_1 - 2c \\ &\dots \\ p_i &= p_1 - (i - 1)c \end{aligned}$$

Considerando que a proporção obtida no ponto de alcance da fadiga sensorial equivale a $1/3$, obtém-se a expressão abaixo:

$$p_1 - (m - 1)c = 1/3$$

Isolando c , tem-se:

$$c = (p_1 - 1/3) / (m - 1)$$

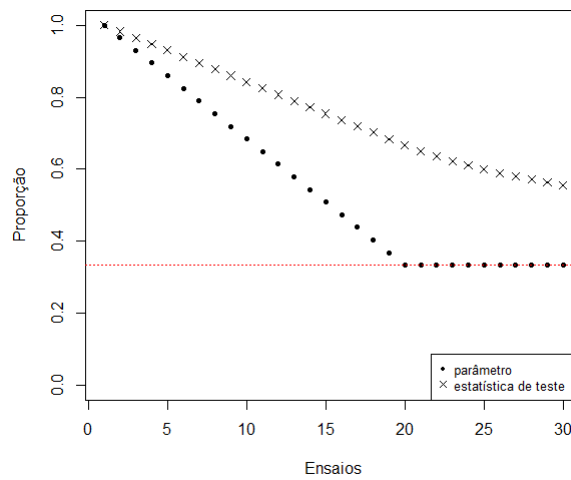
Logo, o modelo linear do cansaço é dado por (23):

$$p_i = p_1 - (i - 1) \frac{p_1 - 1/3}{m - 1} \quad (23)$$

em que i é o índice correspondente ao ensaio, p_i é a proporção de acertos nesse ensaio e m é o ensaio em que o provador se cansa, ou seja, atinge a proporção de 1/3.

A Figura 8 representa o gráfico do cansaço para o modelo linear (parâmetro), com uma proporção inicial de acerto de 100% e alcance da fadiga no vigésimo ensaio, juntamente com sua estatística de teste.

Figura 8 – Figura representativa do cansaço para o modelo linear.



Fonte: Da autora.

Para obtenção do modelo não linear do cansaço, foi realizada uma adaptação da função logística. A função logística é apresentada em (24).

$$f(x) = \frac{p_1}{1 + \exp^{-k(x-x_0)}} \quad (24)$$

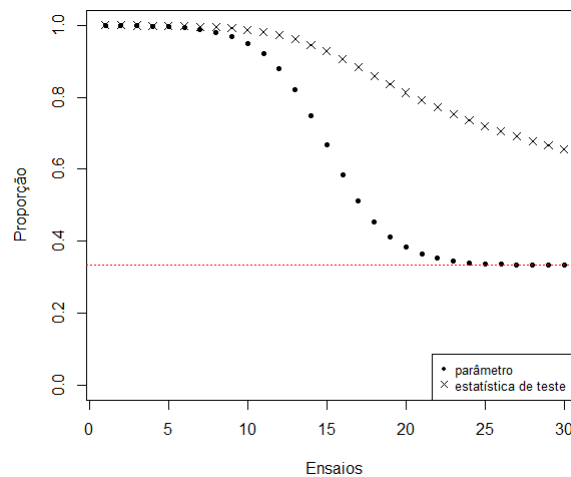
Tomando esta função como base para o comportamento em que o provador se cansa de forma não linear, devem-se fazer algumas alterações algébricas. Para torná-la descendente com queda em 1/3 e, portanto, parar na proporção final de 1/3, temos a função (25).

$$p_i = 1/3 + \frac{p_1 - 1/3}{1 + \exp^{k(i-x_0)}} \quad (25)$$

em que k é uma constante que corresponde à inclinação da curva que pode ser mais íngreme (k elevado), ou menos íngreme (k baixo).

O gráfico obtido para o modelo não linear do cansaço fornece a Figura 9, com um acerto inicial de 100% e a respectiva estatística de teste ao longo dos ensaios.

Figura 9 – Figura representativa do cansaço para o modelo não linear.



Fonte: Da autora.

Dessa forma, k estaria relacionado ao quão rápido o provador está se cansando, enquanto x_0 seria o ponto central dessa inclinação, que foi estipulada como metade do valor de n . Para a curva da função não linear, foram adotadas os mesmos valores de p_1 da função linear e k igual a 0,5.

7.3 OS MODELOS LINEAR E NÃO LINEAR PARA O APRENDIZADO

De forma similar às funções cansaço, as funções de aprendizado foram modeladas de forma linear e não linear. Dessa vez, a função deve representar um candidato que acerta a uma pequena taxa e que aumenta ao longo dos ensaios. A demonstração para o aprendizado linear ocorre de forma similar ao cansaço, considerando que o candidato possa atingir uma aprendizagem e adquirir uma taxa de acerto de 100%, obtendo-se:

$$p_1 + (m - 1)c = 1$$

em que m agora representa o ensaio em que o candidato alcança seu ápice de aprendizagem, p_1 é a proporção de acerto inicial, estipulado em 1/3, que equivale à taxa de acerto ao acaso, e c agora representa o quanto o candidato aprende. Logo, isolando c ,

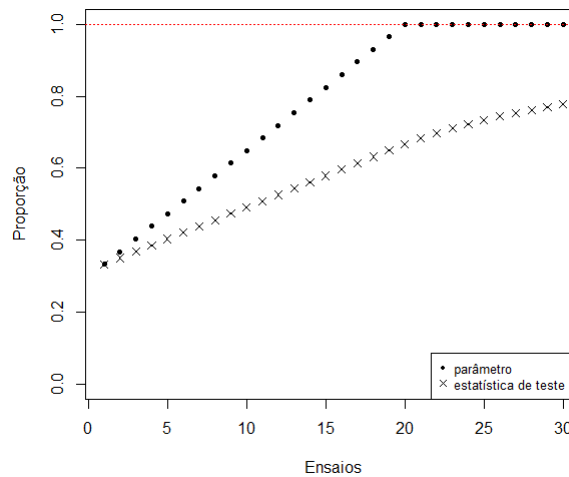
$$c = (1 - p_1) / (m - 1)$$

Sendo assim, a função de aprendizado linear equivale a (26):

$$p_i = p_1 + (i - 1) \frac{1 - p_1}{m - 1} \quad (26)$$

O gráfico obtido para a função aprendizado com modelo linear e alcance da proporção de acertos de 100% no vigésimo ensaio é representado junto à estatística de teste na Figura 10.

Figura 10 – Figura representativa do aprendizado para o modelo linear.



Fonte: Da autora.

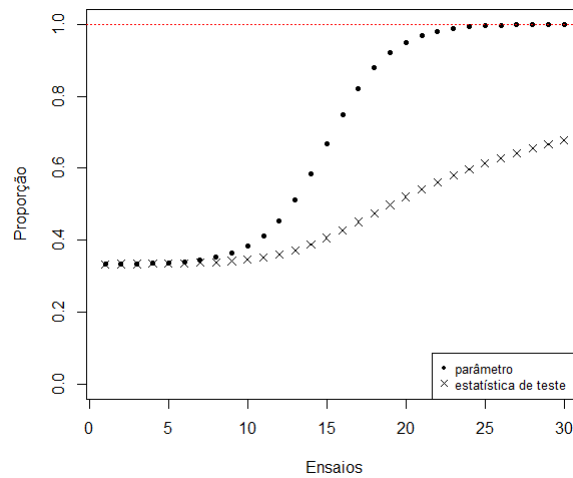
Para o modelo não linear de aprendizagem é desejável trabalhar com a função logística, de forma que ela também cresça de 1/3 a 1, obtendo-se (27):

$$p_i = p_1 + \frac{1 - p_1}{1 + \exp^{-k(i - x_0)}} \quad (27)$$

para k relacionado ao quão rápido o provador aprende, adotando-se valor igual a 0,5, enquanto x_0 também representa o ponto central da inclinação, sendo a metade do valor de n .

Dessa forma, obtém-se o gráfico da Figura 11.

Figura 11 – Figura representativa do aprendizado para o modelo não linear.



Fonte: Da autora.

Esse último gráfico possui o comportamento de uma função logística, que parte da proporção de acerto inicial de 1/3 até 100%.

7.4 DESEMPENHO DOS TESTES E SIMULAÇÃO

Para averiguar o desempenho dos testes para provedores que possuem um comportamento de cansaço linear e não linear, bem como o aprendizado, foram realizadas simulações de Monte Carlo. Dessa forma, tem-se duas funções utilizadas para a aleatorização correspondente ao cansaço e duas para o aprendizado, que geram proporções de acertos.

Para as curvas da função linear e não linear do cansaço, foi adotado valor de p_1 igual a 100%, enquanto para as curvas da função linear e não linear do aprendizado foi adotado p_1 igual a 1/3. O último ensaio de cada análise foi considerado o ponto em que se atinge a fadiga ou ápice de aprendizagem, dado por m .

Através das simulações efetuadas para 1000 provedores em diferentes números de ensaios variando entre 5, 10, 15, ..., 100, foram analisadas graficamente as curvas de poder para avaliar das taxas de rejeição de H_0 . Esse processo foi realizado a fim de determinar os testes de maior poder e menor erro tipo I.

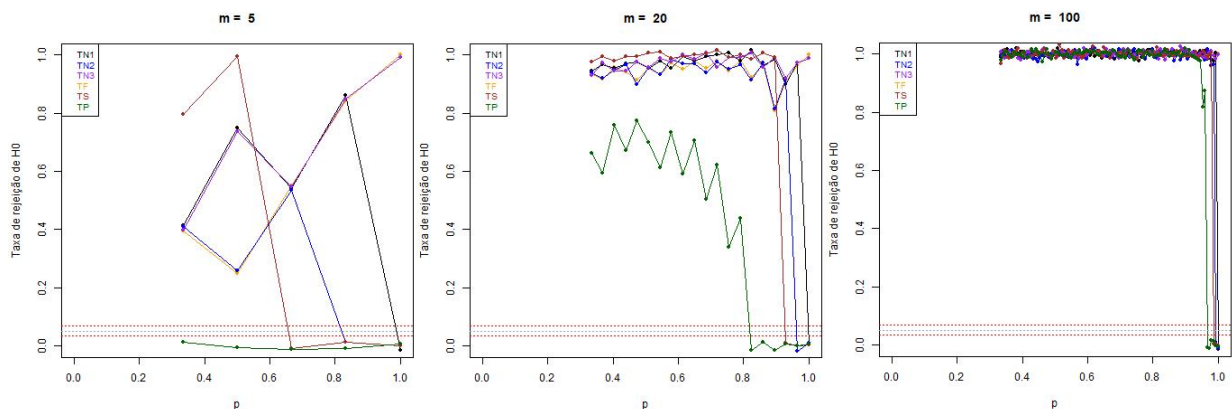
8 RESULTADOS E DISCUSSÃO

No processo de simulação para o cansaço e aprendizagem, os provedores podem ter habilidades diferentes e, portanto, taxas diferentes de acerto inicial. Nas simulações para o cansaço e aprendizagem foram consideradas uma proporção de acerto inicial de 100% e 1/3, respectivamente.

No estudo de simulação foi estabelecido um modelo para fadiga com descrições do cansaço através do modelo linear e não linear, bem como aprendizado, em que houve quebra da pressuposição de p constante (sendo não identicamente distribuídos), mas manteve-se a pressuposição de independência. Nesse caso, verificou-se o impacto no poder e nas taxas de erro tipo I dos mesmos testes, uma vez que essa pressuposição não foi atendida.

A Figura 12 representa as taxas de rejeição de H_0 para o modelo linear do cansaço nos ensaios 5, 20 e 100.

Figura 12 – Gráfico do poder para o modelo linear do cansaço.



Fonte: Da autora.

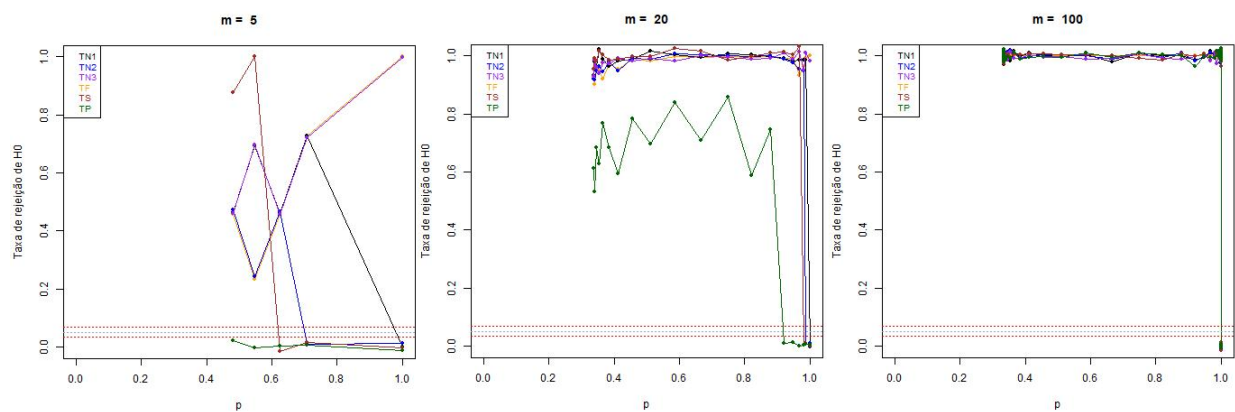
Observa-se que, com a quebra de pressuposição de p constante, os gráficos apresentam uma curva com alta rejeição de H_0 quando p é pequeno e n é elevado, mas baixa rejeição de H_0 quando p aumenta e n é reduzido. Nota-se que é esperado um alto poder nos primeiros ensaios e que esse poder seja reduzido nos últimos ensaios, já que a proporção de acertos é menor, porém o número de ensaios afeta bastante as curvas de poder. O alto valor de p está associado ao baixo valor de n nas funções cansaço, sendo que n tem maior influência sobre as curvas de poder do que p .

Os testes deveriam rejeitar H_0 com a alta proporção de acertos inicial e aceitar H_0 com

a posterior diminuição da proporção de acertos ao longo dos ensaios, para a análise do cansaço. No entanto, não é isso o que acontece, os testes não detectam a alta proporção inicial, portanto o poder é pequeno quando p é elevado. Isso ocorre devido a alta exigência dos testes e ao baixo valor de n , dessa forma, nos primeiros ensaios o poder é pequeno. À medida que aumenta o número de ensaios, essa exigência é reduzida e os testes passam a rejeitar H_0 e adquirir alto poder com menor proporção de acertos.

O mesmo acontece com as taxas de rejeição de H_0 nos ensaios 5, 20 e 100 para o modelo não linear do cansaço, apresentado na Figura 13.

Figura 13 – Gráfico do poder para modelo não linear do cansaço.



Fonte: Da autora.

O que diferencia o poder do modelo linear do cansaço para o não linear é que o primeiro possui uma queda gradual, enquanto o último possui uma queda mais lenta inicialmente e abrupta ao final. Isso ocorre devido ao comportamento das funções cansaço, ou seja, a função linear do cansaço tem apenas um ponto superior que reduz continuamente a cada ensaio, enquanto a função não linear do cansaço tem uma pequena região com altos valores da proporção de acerto, seguida de uma curvatura decrescente mais íngreme que leva a uma outra região com baixos valores da proporção de acerto.

Na situação em que foi estabelecido um modelo linear e não linear que descreveria a fadiga, observa-se que: quando é mantida a pressuposição de independência mas é quebrada a pressuposição de p constante em estudo de simulação, verifica-se o impacto no poder e erro tipo I desses mesmos testes devido ao não atendimento dessa pressuposição.

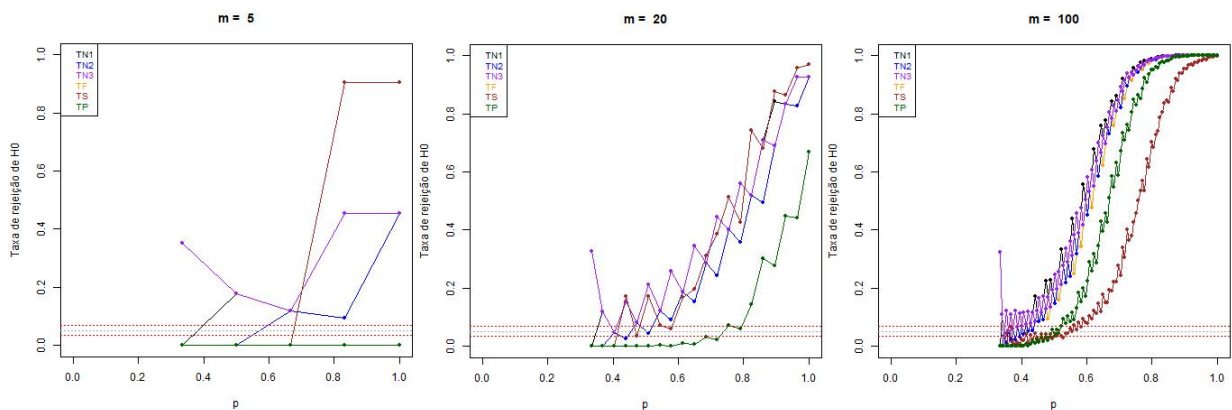
Segundo Kunert; Meyners (1999), assume-se que a segunda apresentação é randomizada independentemente da primeira, logo, qualquer que seja a estratégia do avaliador, sob hipótese

nula, a probabilidade da amostra diferente ser colocada na mesma posição que ele escolheu é de $1/3$ em todas as replicações, ou seja, as replicações são independentes e o resultado do segundo julgamento independe do primeiro resultado. Além do mais, a escolha é independente da ordem apresentada em cada ensaio, pois assume-se que os produtos são igualmente apresentados e não há nenhuma outra característica que interfira no processo como temperatura e aparência, por exemplo. Por isso, a pressuposição de independência foi mantida.

Para a análise do aprendizado em que também há quebra da pressuposição de p constante, mantendo a pressuposição de independência, verifica-se que a taxa de rejeição de H_0 é inferior quando p e n são pequenos, no entanto, há um aumento na taxa de rejeição de H_0 quando p e n aumentam.

A Figura 14 mostra as taxas de rejeição de H_0 dos mesmos valores de n para a aprendizagem linear.

Figura 14 – Gráfico do poder para o modelo linear do aprendizado.

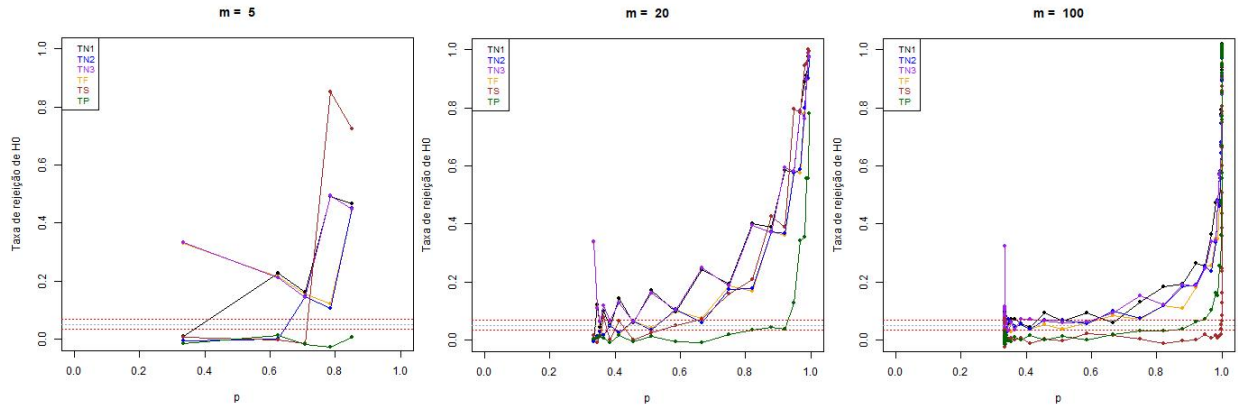


Fonte: Da autora.

Nota-se que, agora, os testes não rejeitam H_0 inicialmente, devido às baixas taxas da proporção de acerto e pela alta exigência dos testes. Entretanto, essa rejeição aumenta com o número de ensaios, alta proporção de acertos e menor exigência dos testes. A curva do poder passa a adquirir o formato sigmóide. Deve-se ressaltar que o TS tem alto poder quando são realizados poucos ensaios, mas adquire menor poder com o aumento do número de ensaios para as funções aprendizado.

Por fim, na avaliação das taxas de rejeição de H_0 para o modelo não linear da aprendizagem, tem-se a Figura 15.

Figura 15 – Gráfico do poder para o modelo não linear do aprendizado.



Fonte: Da autora.

Nota-se que a curva do poder para o modelo não linear de aprendizagem tem uma baixa rejeição de H_0 inicial que é elevada de forma mais lenta nos primeiros ensaios, adquirindo alto poder final. Isso também ocorre devido ao comportamento da função não linear de aprendizagem que possui uma pequena região com valores baixos da proporção de acerto, seguida de uma curvatura mais íngreme que leva a uma outra região com altos valores da proporção de acerto. Portanto, essa função atinge uma elevação súbita do poder ao final dos ensaios e não de forma gradual, como ocorre com a função aprendizagem linear. Verifica-se que, de forma similar, essa última função possui apenas um ponto inferior que cresce continuamente a cada ensaio, adquirindo crescimento gradual no poder.

Pode-se observar que, nas funções de aprendizado, n aumenta à medida que p também aumenta. Logo, as condições são ideais para que o poder seja alto, pois o alto valor de p está associado ao alto valor de n . Adicionalmente, quando p diminui, n também é reduzido, levando ao baixo poder dos testes nas funções aprendizado.

Com relação ao poder dos testes para as funções cansaço, se o número de ensaios não fosse grande quando a proporção aproximasse de $1/3$, o poder tenderia a cair. No entanto, a baixa proporção não é suficiente para causar a queda do poder pois, nessa situação, o número de ensaios é grande, o que o torna os testes poderosos quando p está próxima a $1/3$. Concomitantemente, quando a proporção é alta nas funções cansaço, o número de ensaios é muito pequeno, o que reduz drasticamente o poder. Nota-se, então, que o número de ensaios tem maior influência sobre as curvas de poder do que a proporção, o que justifica a necessidade do maior número de observações para a melhoria do poder dos testes.

Além disso, existe uma alta variância da proporção de acertos inicial e, antes da estabili-

zação acontecer, os testes percebem que a proporção de acertos está sendo alterada (diminuindo ou aumentando). No entanto, esse aumento ou diminuição demora a acontecer comparado à proporção real definida (paramétrica), até mesmo porque os resultados anteriores influenciam no acúmulo de proporção de acertos. Ou seja, quando há muitos acertos iniciais, mesmo que o candidato tenha uma alta probabilidade de errar no próximo ensaio, a proporção de acertos ainda tende a se manter alta. Ao passo que, se o candidato tem muitos erros iniciais, a proporção tende a se manter baixa, mesmo com um grande aumento da probabilidade de acertos no ensaio seguinte.

A estatística de teste adquire uma memória longa que, segundo VedoVatto (2014), ocorre quando a autocorrelação entre os dados possui um decaimento muito lento (ou infinito). Portanto, quando a proporção paramétrica é alterada, o estimador não reflete essa alteração na mesma velocidade, devido à influência dos estimadores obtidos nos ensaios anteriores com parâmetros diferentes.

Ao comparar as taxas de erro tipo I e poder para as funções cansaço, observa-se que todos os testes apresentam comportamento similares, sendo que o TP demora mais para atingir as mesmas taxas de rejeição de H_0 que os outros testes. Todos os gráficos do poder obtidos para o cansaço e aprendizado foram apresentados no Apêndice B.

Para os modelos linear e não linear do aprendizado, é possível notar que, a medida que aumenta o número de ensaios, a maioria dos testes passam a adquirir menores taxas de erro tipo I. Inclusive, o teste que apresentou menor taxa de erro tipo I, não só para as funções aprendizado mas também as funções cansaço, foi o TP. Já o TN3 apresentou altas taxas de erro tipo I nas funções de aprendizado devido ao seu baixo nível de exigência inicial e n pequeno.

Os testes que simultaneamente apresentam menor taxa de erro tipo I e maior poder foram o TN1, TN2 e TF, para um número de ensaios elevado das funções aprendizado. Nota-se que o poder do TS diminuiu com o aumento do número de ensaios, sendo este o menos poderoso. Entretanto, esse último teste ainda adquire melhor comportamento para pequeno número de ensaios. Para as funções cansaço, não se recomenda a utilização de nenhum dos testes.

9 CONCLUSÕES

Foi possível notar o quanto a quebra da pressuposição de p constante afeta bastante as taxas de erro tipo I e poder dos testes. De acordo com as características das funções, sendo crescente ou decrescente, as curvas de poder adquirem comportamentos opostos.

Para as curvas de cansaço ao longo das proporções de acerto, observa-se uma alta taxa de erro tipo I e baixo poder final, enquanto as curvas de aprendizagem adquirem menor taxa de erro tipo I e alto poder final. Inclusive, o número de ensaios (n) tem maior efeito nas curvas de poder do que a proporção (p).

Todos os testes adquirem comportamento similares para as funções cansaço, no entanto, o TP demora um pouco mais para atingir as mesmas taxas de rejeição de H_0 dos demais testes. Os testes TN1, TN2 e TF fornecem menor taxa de erro tipo I e maior poder tanto para o modelo linear do aprendizado quanto o não linear, para n elevado. Entretanto, TS adquire melhor comportamento quando n é pequeno, apresentando queda do poder com o aumento de n .

O fato da inclinação correspondente às funções não lineares no cansaço e aprendizagem ser mais íngreme afeta a velocidade de alcance das taxas finais de rejeição de H_0 . Já a função linear de aprendizagem apresenta crescimento gradual no poder, enquanto a função linear do cansaço apresenta uma queda um pouco menos brusca quando comparado ao modelo não linear cansaço. Deve-se ressaltar que a estatística de teste adquire uma memória longa, então, o estimador não reconhece a alteração do parâmetro na mesma velocidade. Os estimadores dos ensaios anteriores afetam os resultados posteriores, prejudicando a eficiência dos testes.

O TP foi o teste que apresentou menor taxa de erro tipo I em todas as situações. Enquanto o TN3 apresentou maiores taxas de erro tipo I para as curvas aprendizado, devido à baixa exigência inicial do teste e n pequeno.

Nota-se que, caso p não seja constante, os testes TN1, TN2 e TF podem ser recomendados para o aprendizado com n elevado, enquanto o TS pode ser recomendado para n pequeno. Já o caso da fadiga sensorial, não é aconselhado o uso de nenhum dos testes.

REFERÊNCIAS

CHAMBERS IV, E.; WOLF, M. B. **Sensory testing methods**. 2nd ed. Lancaster: ASTM International, 2005.

FERREIRA, D. F. **Estatística Básica**. Lavras: UFLA, 2005.

HARRIS, E. H. D. B. *et al.* **Sensory Methods for Measuring Differences in Food Quality**. Washington: Agriculture Information Bulletin No. 34, 1951.

KUNERT, J.; MEYNER, M. On the Triangle Test with Replications. **Food Quality and Preference**. 1999. Disponível em:
<https://www.researchgate.net/publication/2765334_On_the_Triangle_Test_with_Replications>. Acesso em: 2 out. 2019.

LAWLESS, H. T.; HEYMANN, H. **Sensory Evaluation of Food**. 2nd ed. New York: Springer, 2010.

MINIM, V. P. R. *et al.* Análise descritiva: Comparação entre metodologias. **Rev. Inst. Latic. Cândido Tostes**, v. 65, n. 374, p. 41-48. 2010.

NÆS, T.; BROCKHOFF, P. B.; TOMIC, O. **Statistics for Sensory and Consumer Science**. Wiley. 2010.

POSTE, L. M. *et al.* **Laboratory Methods for Sensory Analysis of Food**. Canada: Agriculture Canada Publication, 1991.

SHIROSE, I.; MORI, E. E. M. **Estatística aplicada à análise sensorial (Módulo 2)**. Campinas: ITAL - Manual Técnico nº 13, 1996.

VEDOVATTO, T. **Medidas de memória longa em séries temporais: Comparação de métodos de estimação do coeficiente de Hurst**. 2014. 97f. Dissertação (Mestrado em Ciências) - Universidade de Brasília, Brasília, 2014.

ZWIMPFER, T. **Sensory Evaluation Lab Report**. Nutrition 205, Section 3. 2013.

CAPÍTULO 4 - ANÁLISE SENSORIAL DE OLFATO DE CAFÉ: A REGRESSÃO SEGMENTADA AUXILIANDO OS TESTES DE SELEÇÃO DE JUÍZES

RESUMO: A análise sensorial está intimamente relacionada com a qualidade dos alimentos em que, através dos sentidos, pode-se medir, analisar e interpretar reações produzidas pelos mesmos. Os sentidos como o paladar e o olfato se destacam nas análises de café, sendo o aroma determinante na qualidade. O aroma do café pode variar de acordo com a espécie do grão, região de produção e procedimento pós-colheita, em que há classificações de suave a intenso com características frutadas, achocolatadas, florais e cítricas. Uma equipe de provadores, que saiba identificar e distinguir diferentes aromas, é essencial para avaliação do café. No processo de seleção de provadores são comumente utilizados ensaios triangulares. Entretanto, os dados obtidos por um avaliador podem adquirir uma alta variabilidade inicial, como qualquer estimador consistente e não viesado que obedece a lei dos grandes números, prejudicando o desempenho dos testes de seleção e fazendo com que eles tomem uma decisão errônea. Sendo assim, deseja-se saber o número necessário de ensaios para se tomar uma decisão correta, a partir do qual o teste poderia ser aplicado. Para isso, buscou-se uma ferramenta que auxiliasse na decisão do teste de forma que ele cometesse menor erro tipo I e II, decidindo apenas quando as proporções de acerto convergissem em algum ponto ou reduzissem sua variabilidade. O objetivo deste trabalho é realizar experimentos de análise sensorial de odor de café, submetendo candidatos a um grande número de ensaios triangulares e modelando os dados com o auxílio da regressão segmentada para posterior aplicação do teste sequencial (TS), a fim de determinar o número de ensaios necessários para se tomar uma decisão que cometa menor erro tipo I e II. Para isso, verificou-se que não houve quebra de pressuposição de independência pelo teste de Smith, estimaram-se os pontos de mudança utilizando o pacote *segmented* e tomou-se uma decisão pelo TS a partir do momento que observou uma certa estabilização ou redução de variabilidade. Através das etapas anteriores, foi possível selecionar 6% dos candidatos, outros 26% dariam prosseguimento ao teste triangular e 68% seriam dispensados. Para os candidatos que dariam prosseguimento ao teste, sugere-se que eles retornem à análise em um outro momento. Por fim, recomenda-se utilizar 15 ensaios na etapa de seleção da análise olfativa de café, tendo em vista que não houve mudança de comportamento de nenhum candidato após esse ensaio.

Palavras-chave: Regressão segmentada. Ponto de mudança. Café.

COFFEE OLFACTORY SENSORY ANALYSIS: THE REGRESSION SEGMENTED AIDING JUDGE SELECTION TESTS

ABSTRACT: Sensory analysis is closely related to food quality in which, through the senses, it is possible to measure, analyze and interpret reactions produced by them. Senses such as taste and smell stand out in coffee analysis, since the aroma determines quality. Coffee aroma may vary according to the bean species, production region and post-harvest procedure, in which there are classifications between mild to intense with fruity, chocolate, floral and even citrus characteristics. A team of tasters who can identify and distinguish different aromas is essential for coffee evaluation. Triangular tests are commonly used in tasters selection process. However, data obtained by an evaluator may acquire a high initial variability, like any consistent and unbiased bias estimator that obeys the law of large numbers, undermining the selection tests performance and leading to erroneous decisions. Therefore, it is desirable to determine the number of trials needed to make a correct decision, from which the test could be applied. For this, a tool assisting the test decision was sought, so it made the smallest error type I and II, deciding only when the correct proportions converged at some point or reduced its variability. The objective of this work was to perform coffee smell sensory analysis experiments, submitting candidates to a large number of triangular trials and modeling the data with segmented regression for subsequent application of sequential test (TS) in order to determine the number of trials needed to make a decision that makes the smallest error type I and II. For this, it was verified the assumption of independence by the Smith test and the change points were estimated using the segmented package. A decision was taken by the TS from the moment it was observed a certain stabilization or reduction of variability. Through the previous steps, it was possible to select 6% of the candidates, others 26% would proceed with the triangular test and 68% would be dismissed. For candidates who would continue the test, it is suggested that they return to the analysis at a later time. Finally, it is recommended to use 15 trials in selection stage of coffee olfactory analysis, considering there wasn't any change in candidates behavior after this trial.

Keywords: Segmented regression. Breakpoint. Coffee.

10 INTRODUÇÃO

A análise sensorial pode ser definida como uma disciplina científica utilizada para evocar, medir, analisar e interpretar reações provocadas por características de alimentos ou materiais, que são percebidas através dos sentidos de visão, olfato, paladar, tato e audição (POSTE *et al.*, 1991). Inclusive, a pesquisa sensorial, envolvendo alimentos, foca em como os sentidos reagem aos alimentos e como isso pode ser utilizado no controle de qualidade e desenvolvimento de novos produtos (BROCKHOFF, 2011).

A qualidade e segurança alimentar são tópicos importantes na indústria e na pesquisa. Essa percepção de qualidade e segurança está relacionada à demanda dos consumidores e escolha dos alimentos (GRUNERT, 2005). Logo, a análise sensorial tem grande importância na avaliação da aceitabilidade mercadológica e qualidade dos alimentos (TEIXEIRA, 2009).

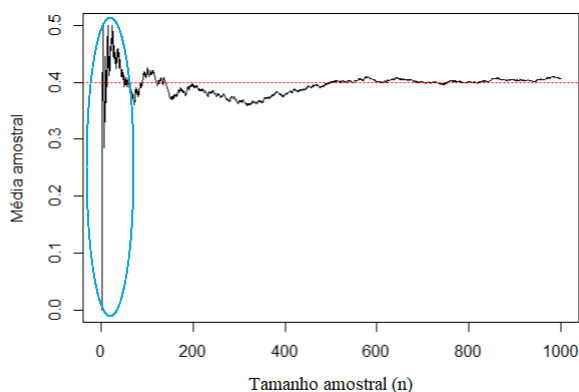
A qualidade do café é influenciada pela composição química dos grãos, bem como fatores relacionados ao cultivo, preparo, torra e conservação. O efeito da temperatura causa transformações químicas e físicas nos grãos, liberando compostos voláteis responsáveis pelo aroma e sabor do café (MODESTA *et al.*, 1999). O aroma do café pode variar de acordo com a espécie do grão, região de produção e procedimento pós-colheita (RODARTE, 2008). Ele pode ser classificado de suave a intenso, apresentando características frutadas, achocolatadas, florais e cítricas (PAIVA, 2005).

Os sentidos como o paladar e o olfato ganham destaque nas análises de café, sendo o aroma determinante na qualidade (RODARTE, 2008). Na avaliação sensorial para o controle de qualidade de torra e moagem, indica-se a formação de uma equipe de provadores que possuem habilidade de reconhecimento e distinção de diferentes aromas, sabores e gostos, que são selecionados e treinados especificamente para avaliar características do café (MODESTA *et al.*, 1999).

Para averiguar a habilidade discriminativa, uma série de ensaios triangulares é comumente empregada na seleção de provadores, em que calcula-se a porcentagem de acertos (POSTE *et al.*, 1991). Entretanto, sabe-se que, pela lei dos grandes números, a proporção de acertos estimada (\hat{p}) converge para a proporção paramétrica (p), à medida que aumenta o número de ensaios (DINOV; CHRISTOU; GOULD, 2009). Essa lei se aplica não só para \hat{p} , como também para outros estimadores. A Lei Fraca dos Grandes Números afirma que para variáveis aleatórias independentes e identicamente distribuídas, à medida que o tamanho amostral aumenta, a média

amostral tende a se aproximar da média populacional (ZIBETTI, 2019). A Figura 16 mostra um exemplo de simulação que represente essa lei fraca dos grandes números para uma Binomial.

Figura 16 – Exemplo de simulação para a lei fraca dos grandes números de uma Binomial.



Fonte: Adaptado de Zibetti (2019).

Então, o estimador adquire uma variabilidade cada vez menor e converge para um determinado valor, à medida que aumenta o tamanho amostral. No entanto, a alta variabilidade inicial da proporção pode prejudicar os testes na seleção de provedores, como mostra o Capítulo 3. Nesse capítulo observa-se que a alteração da proporção e consequente quebra de pressuposição prejudicam o desempenho dos testes.

Como a proporção de acertos tem uma distribuição Binomial, as pressuposições de independência entre observações e proporção constante devem ser consideradas. Outro fator que deve ser ressaltado é a alta variabilidade da proporção de acertos em poucos ensaios, pois ela pode ser interpretada pelo teste como uma variação da proporção paramétrica, o que de fato não ocorre e resulta em falsos positivos ou negativos. O comportamento ascendente ou descendente, que pode acontecer com um pequeno tamanho amostral, pode ser confundido como uma proporção de sucesso variável e prejudicar o teste em termos de taxa de erro tipo I e poder.

Se a proporção de acertos nos primeiros ensaios é bastante variável e fornece um falso comportamento crescente ou decrescente, deseja-se identificar a partir de qual ensaio \hat{p} adquire-se uma variabilidade relativamente pequena, em que pode-se aplicar o teste e adquirir um bom desempenho (com menor taxa de erro tipo I e maior poder), como apresentado no Capítulo 2. O que seria útil para que esse falso comportamento inicial não seja interpretado pelo teste como uma alteração do parâmetro. Para isso, é necessário buscar uma ferramenta que auxilie nessa detecção da estabilidade do estimador para fazer o uso dos testes.

Os modelos de regressão segmentada são modelos cuja relação entre a variável resposta (dependente) e a variável explicativa (independente) é representada por duas ou mais retas conectadas por valores desconhecidos. Esses valores podem ser referidos como ponto de mudança (*change points*), de interrupção ou de junção (MUGGEO, 2008). Através desse modelo, seria possível detectar o momento a partir do qual a proporção de acertos adquire comportamento constante.

Sendo assim, o objetivo deste trabalho foi fazer experimentos reais, com a coleta de dados de candidatos submetidos à análise de olfato de café em um grande número de ensaios triangulares, a fim de verificar o comportamento da proporção de acertos na prática e, por meio da regressão segmentada, estimar os pontos de mudança. A partir dessa estimativa, a estabilidade ou convergência da proporção de acertos seria melhor visualizada para definir o número de ensaios em que se deve aplicar o teste de seleção e adquirir ótimo desempenho.

11 METODOLOGIA

11.1 ANÁLISE SENSORIAL

No procedimento da análise sensorial, foram executados 20 ensaios triangulares com 31 indivíduos, membros do meio acadêmico. Nesse teste, os julgadores foram informados de que receberiam três amostras, sendo duas iguais e uma diferente. Eles deveriam analisá-las através do odor e indicar qual delas era a diferente. Também, foram previamente informados do número de ensaios a que seriam submetidos, recebendo as devidas orientações em todo o processo. A ficha de avaliação utilizada foi apresentada na Figura 17, adaptada ao exemplo de Paes (2005).

Figura 17 – Ficha de avaliação para teste triangular.

Ficha de avaliação

Laboratório de análise sensorial da UNIFAL

Nome: _____ Idade: _____

Produto: _____ Data: __/__/__

Duas das três amostras são idênticas. Por favor, através do odor, analise-as da esquerda para a direita e marque com um "X" o espaço correspondente ao código daquela que você julgar DIFERENTE.

Rep	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
546																				
765																				
389																				

Comentários:

Fonte: Adaptado de Paes (2005, p. 22).

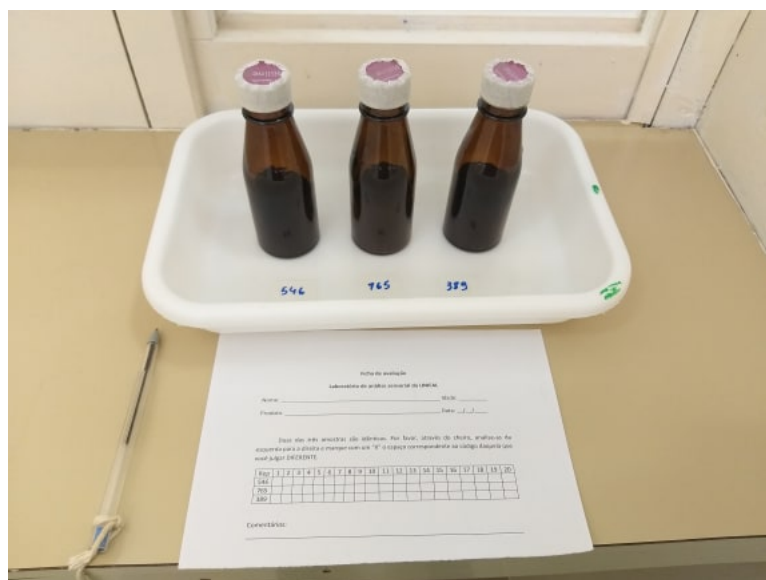
As amostras foram distribuídas em vidros de cor âmbar, para que a cor do produto não interferisse na análise. Então, o avaliador deveria apenas assinalar com "X" o espaço correspondente à amostra julgada diferente por ele, no ensaio determinado. O tempo de apresentação entre ensaios foi de aproximadamente 30 segundos.

A prática foi realizada no Laboratório de Análise Sensorial da Unifal-MG, unidade centro de Alfenas. Nessa análise, participaram candidatos não treinados, com idade entre 18 e 39

anos, sendo 8 do gênero masculino e 23 do gênero feminino, que assinaram o Termo de Consentimento Livre e Esclarecido (TCLE). O estudo foi aprovado pelo Comitê de Ética em Pesquisa da Unifal-MG sob Parecer nº 3.457.659.

Foram utilizadas duas marcas comerciais de café com descrição tradicional, produzidos na região de Guaxupé e Varginha, cujas respectivas torras são denominadas média e clássica. Devido à diferença de densidade dos pós de café, pesou-se 17 g do primeiro e 20 g do segundo, de forma a obter visualmente o mesmo volume. A Figura 18 mostra a aplicação do teste triangular nas cabines do laboratório de análise sensorial.

Figura 18 – Aplicação do teste triangular.



Fonte: Da autora.

A escolha das marcas de café foi baseada em pré-testes realizados, cuja diferença olfativa entre marcas era relativamente pequena. Não era desejável que a diferença fosse facilmente identificada, tal que todos os candidatos nunca errassem, nem mesmo que os produtos fossem muito similares de forma que nenhum candidato acertasse.

A partir dos dados obtidos, realizou-se a verificação da independência dos ensaios triangulares. Para avaliar se houve quebra de pressuposição de independência dos dados, utilizou-se o teste de Smith.

No teste de Smith, primeiramente, calcula-se o número de respostas corretas em cada ensaio e, então, o maior valor é utilizado para representar o número de respostas corretas no estudo. Com o auxílio da Tabela 4.3 de Lawless; Heymann (2010 p. 91), esse valor é comparado

com o número mínimo de respostas corretas para um determinado nível de significância (0,05) no teste triangular, cujo número de provadores foi 31. Se o valor for superior ao valor tabelado, os ensaios são declarados independentes, caso contrário, são ditos dependentes. Posteriormente, ajustou-se um modelo de regressão segmentada aos dados.

11.2 MODELAGEM COM REGRESSÃO SEGMENTADA

O pacote *segmented* do software R Core Team (2019) facilita a estimação e resumo de modelos lineares generalizados que possuem relações segmentadas. Para utilizá-lo, não há limite de variáveis ou pontos de mudança (*changeoints*; *breakpoints*). Em um amplo contexto, a ideia é estimar o modelo de linhas quebradas, também denominado *broken-lines*, a fim de fornecer estimativas pontuais de incerteza que são relevantes no modelo (MUGGEO, 2008).

Se dois modelos lineares satisfazem (28):

$$Y_i = \begin{cases} \alpha_0 + \alpha_1 i + e_i, & i = 1, \dots, r \\ \beta_0 + \beta_1 i + e_i, & i = r + 1, \dots, n \end{cases} \quad (28)$$

em que e_i é o erro, então, o *changeoint* τ deve estar no intervalo $(r, r+1)$, com r desconhecido, tal que $\alpha_0 + \alpha_1 \tau = \beta_0 + \beta_1 \tau$ (LIU; QIAN, 2009; SOLOW, 1987).

A função de verossimilhança estimada de τ fornece intervalos de confiança que possam ser associados (SOLOW, 1987). Para encontrar o melhor modelo são realizadas várias tentativas, cujo ajuste da regressão segmentada é feito pelo método dos mínimos quadrados. Logo, a regressão linear que possui menor soma de quadrados dos erros é considerada a melhor (PORTZ; DIAS; CIRYNO, 2000).

Para modelar os dados obtidos na análise sensorial, foram avaliadas as proporções de acertos em relação aos ensaios, com aplicação do pacote *segmented*. O pacote verifica se há ponto de mudança e fornece um valor estimado. Caso não seja encontrado nenhum *changeoint*, realiza-se a modelagem por regressão linear simples. Esse ajuste do modelo foi utilizado como uma ferramenta auxiliar para a tomada de decisão do teste sequencial. A estatística de teste é obtida pela razão entre o número de acertos e o número de ensaios percorridos, fornecendo a proporção para cada ensaio.

11.3 ANÁLISE DO TESTE SEQUENCIAL

Após calcular as proporções de acerto dos dados reais e efetuar as etapas anteriores, foi aplicado o teste sequencial (TS), conforme recomendação do Capítulo 2. Para obtenção do TS, foi realizada uma adaptação do Teste Sequencial de Shirose; Mori (1996). Na adaptação, ambos os limites foram divididos por n , para que se possa trabalhar com proporção. O TS tem a forma apresentada em (29):

$$PC_{5,1}(n) = \frac{h_1}{n} + s; \quad PC_{5,0}(n) = \frac{-h_0}{n} + s \quad (29)$$

Os interceptos h_0 , h_1 e a inclinação s são dados por (30) (SHIROSE; MORI, 1996).

$$h_0 = \frac{\log\left(\frac{1-\alpha}{\beta}\right)}{\log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)}; \quad h_1 = \frac{\log\left(\frac{1-\beta}{\alpha}\right)}{\log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)}; \quad s = \frac{\log\left(\frac{1-p_0}{1-p_1}\right)}{\log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)} \quad (30)$$

em que α e β são as respectivas taxas de erro tipo I e tipo II; p_0 e p_1 são as proporções definidas para não selecionar ou selecionar um provador.

Assim, foram estabelecidos os valores de $\alpha = \beta = 0,05$, $p_0 = 1/3$ e $p_1 = 3/4$. Para o valor de $p_1 = 3/4$ tomou-se a referencia da proporção mínima utilizada na prática para selecionar provadores em testes triangulares com 75% de acerto feita por Minim *et al.* (2010).

O TS possui uma região em que não se toma uma decisão e deve-se dar prosseguimento aos ensaios triangulares até a tomada de decisão. Ele só decide selecionar ou não um candidato quando a proporção atingir um dos limites.

Nesse teste, os candidatos foram avaliados com o auxílio da modelagem de regressão segmentada, em que foi possível tomar uma decisão entre selecioná-lo ou não, ou até mesmo continuar a aplicação do teste triangular.

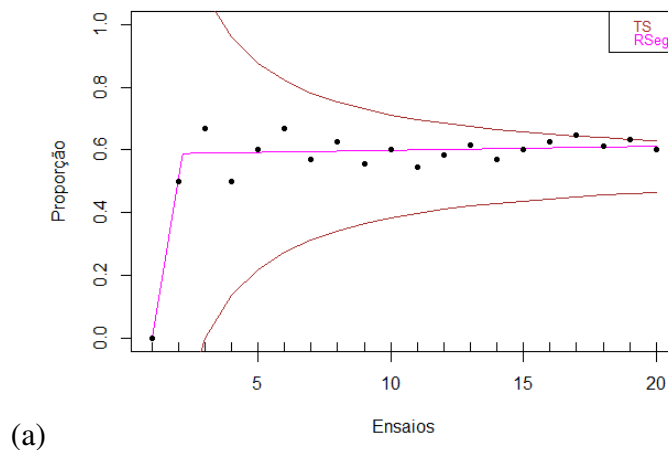
12 RESULTADOS E DISCUSSÃO

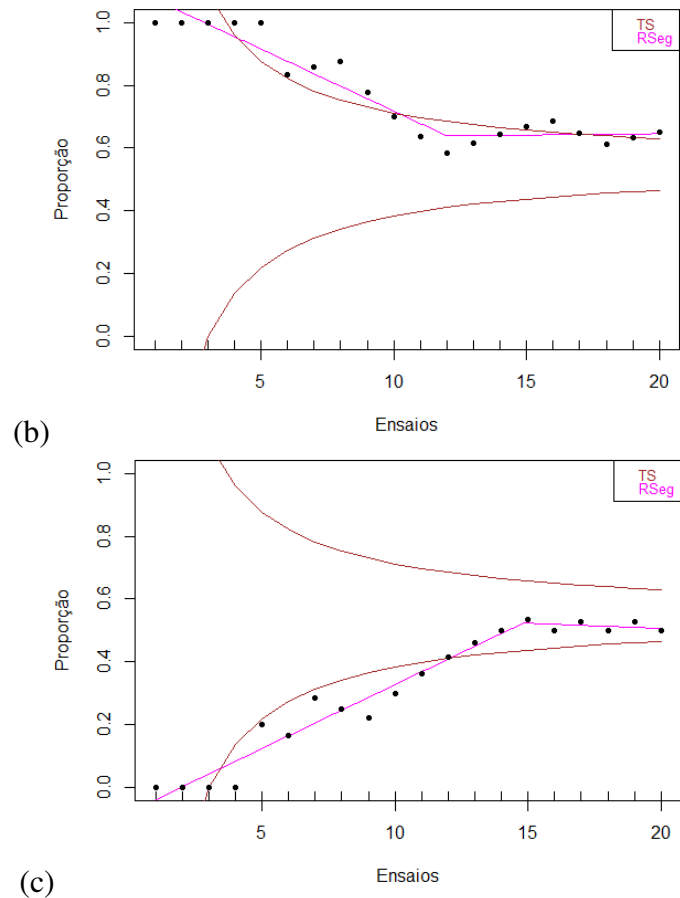
Com os dados obtidos pelo procedimento de análise sensorial, realizou-se a verificação de independência entre as replicações. Para isso, calculou-se o número de acertos em cada ensaio, obtendo-se o maior valor no segundo ensaio, equivalente a 19. Dessa forma, ao compará-lo ao valor tabelado de Lawless; Heymann (2010), para os 31 avaliadores submetidos a análise sensorial, notou-se que o valor foi superior, tanto para o nível de significância de 0,05 quanto para 0,01. Logo, pode-se dizer que os ensaios dos testes triangulares são independentes.

A regressão segmentada foi utilizada para verificar se houve ponto de mudança ou não, ao longo dos ensaios efetuados em cada candidato. Também, foi uma ferramenta auxiliar para a tomada de decisão do TS. O teste foi aplicado para cada pessoa a fim de avaliar se o indivíduo seria selecionado como provador, dispensado, ou até mesmo, se o teste triangular deveria ter prosseguido.

Pela análise dos dados, foi possível notar diferentes comportamentos: pessoas cuja taxa de acerto convergiu para um determinado valor desde os primeiros ensaios; pessoas que tiveram alta taxa de acerto inicial, mas que obtiveram uma queda da proporção de acertos ao longo dos ensaios; pessoas que tiveram baixa taxa de acerto inicial, que foi elevada ao longo dos ensaios. A Figura 19 representa os dados de alguns voluntários com aplicação do teste, para cada uma das respectivas situações anteriormente citadas.

Figura 19 – Teste aplicado em dados obtidos pela análise sensorial de três candidatos, com o auxílio da regressão segmentada.



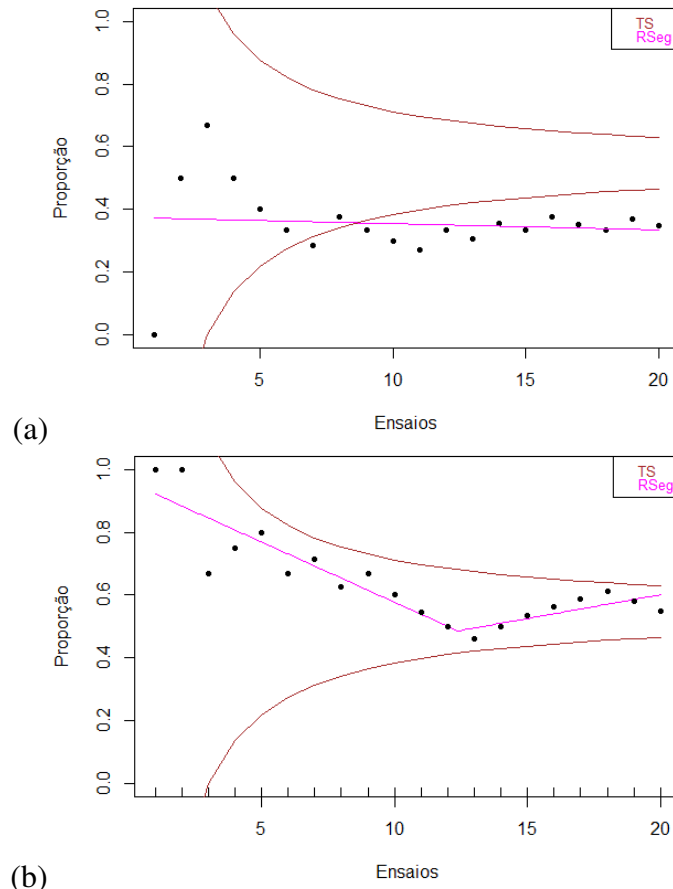


Fonte: Da autora.

Observa-se que, na primeira situação (a), o indivíduo apresentou 12 acertos em 20 ensaios e, portanto, sua taxa de acerto convergiu para 60%; a segunda pessoa (b), teve maior acerto inicial, no entanto apresentou um total de 13 acertos em 20 ensaios, adquirindo uma proporção convergente em 65%; por último, o terceiro indivíduo (c) apresentou baixo acerto inicial, que foi elevado e adquiriu uma proporção que converge em 50%, com um acerto de 10 ensaios dos 20 realizados.

Esses três candidatos obtiveram os respectivos *breakpoints* equivalentes a 2,18, 12 e 14,82. Portanto, cada candidato apresenta um ponto de mudança diferente, no entanto, convergem para um determinado valor após essa mudança. Houveram candidatos que não obtiveram ponto de mudança e apresentaram uma curva de regressão linear simples. Outros candidatos apresentaram uma curva inicial decrescente e outra crescente após seu *breakpoint* ou vice-versa. Acredita-se que, nessas últimas situações os candidatos ainda não atingiram sua habilidade real de acerto, mas tiveram sua variabilidade reduzida e estão tendendo para seu valor real. A Figura 20 mostra as situações anteriormente descritas.

Figura 20 – Gráficos obtidos para um candidato sem *breakpoint* e um candidato cuja proporção de acertos ainda não convergiu.



Fonte: Da autora.

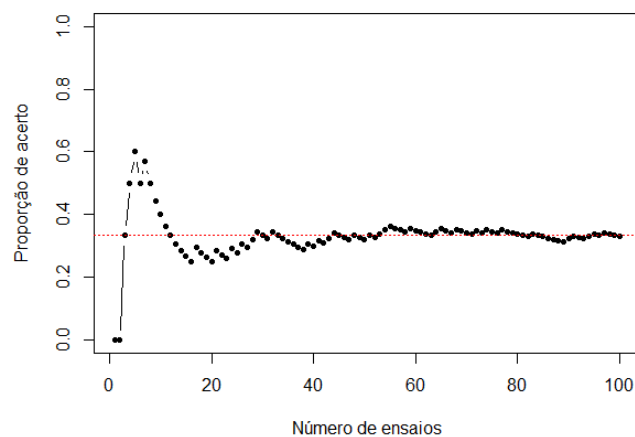
Não houve diferença de avaliação e taxa de acerto quanto ao gênero, mas houveram candidatos que aparentemente tiveram um pouco mais de facilidade durante a análise e outros que tiveram um pouco mais de dificuldade. O número de candidatos que não tiveram *breakpoint* foi equivalente a 4, enquanto 10 candidatos adquiriram característica constante e 17 apresentaram curva crescente ou decrescente.

A regressão segmentada é muito utilizada para avaliar níveis de exigência ou um ponto específico em que há mudança do comportamento dos dados. Portz; Dias; Ciryno (2000) utilizaram a regressão segmentada para encontrar o nível mínimo de um nutriente (proteína) que garantisse o máximo de desempenho para o crescimento de uma espécie de peixe, cuja resposta é fator importante para o custo-benefício da composição de rações. Solow (1987) utilizou a mesma ferramenta para encontrar o ponto de mudança climática ao longo do tempo, dada a importância em se detectar a alteração climática para quaisquer procedimento de interesse que a

utilize. De forma similar, a regressão segmentada permite avaliar a alteração do comportamento dos dados de cada candidato, dando a possibilidade de indicar um número de ensaios a partir do qual se pode tomar uma decisão, cometendo menor erro tipo I e II.

É possível notar que, para cada uma das situações, a proporção de acertos dos candidatos converge para um determinado valor ou reduzem sua variabilidade, fornecendo sua real habilidade ou tendência em identificar diferenças no produto correspondente. Inclusive, esse processo pode ser explicado pela lei dos grandes números, em que à medida que ocorre o aumento do número de repetições em um experimento, a frequência relativa (obtida empiricamente) se aproxima do valor real esperado. Segundo Dinov; Christou; Gould (2009), ao conduzir repetidamente um experimento (idêntico e independente), cuja probabilidade de ocorrência do evento é p em cada experimento, a razão da frequência daquele evento observado pelo total de repetições converge para p quando o número de experimentações aumenta. A Figura 21 mostra uma simulação de convergência para um candidato virtual que deveria acertar a uma proporção de $1/3$ de acerto.

Figura 21 – Simulação de candidato virtual com proporção de acertos convergente para $1/3$.



Fonte: Da autora.

No entanto, vale ressaltar que em muitos testes de seleção recomendam-se realizar quatro ensaios triangulares para selecionar ou dispensar candidatos a provadores, como foi recomendado por Minim *et al.* (2010), Scheid (2001), Oliveira e Benassi (2010). Através dessa análise, é possível verificar que esse número de ensaios é muito pequeno para fazer essa definição, sendo que muitos candidatos ainda não alcançaram a proporção referente a sua habilidade discriminativa e há uma alta variabilidade inicial. A aplicação dos testes para todos os candida-

tos e a análise de regressão foram apresentadas no Apêndice C.

Conforme é definido na lei dos grandes números, espera-se que os candidatos atinjam a proporção de acertos real, com o aumento do número de ensaios. Tendo em vista que, a partir do 15º ensaio não houve mudança de comportamento ou tendência de nenhum candidato, recomenda-se $n = 15$ para a etapa de seleção da análise olfativa de café. Foi possível notar que, com o auxílio da regressão segmentada e a aplicação do TS seriam selecionados 2 candidatos (6%), 8 iriam continuar a avaliação sensorial do olfato de café (26%) e 21 candidatos (68%) não seriam selecionados. Essa baixa porcentagem de indivíduos selecionados sugere que a avaliação do odor de café em testes triangulares é bastante rigorosa para que potenciais provadores sejam submetidos, sem nenhum treinamento.

Dessa forma, as taxas de erro tipo I e II seriam reduzidas. Logo, apenas os candidatos mais habilidosos seriam selecionados e aqueles que apresentaram menor habilidade seriam dispensados. Para os candidatos que dariam prosseguimento ao teste, sugere-se que eles retornem à análise em um outro momento.

Alvelos (2002) indica que, em caso de número insuficiente de aprovados, candidatos que não atingiram o nível exigido para serem selecionados tampouco o nível para não serem selecionados, sejam submetidos a outras repetições dos testes. Ou seja, candidatos que não atingiram o limite superior nem mesmo o limite inferior, mas obtiveram uma proporção de acertos entre esses pontos críticos, podem ter uma segunda chance de serem selecionados.

13 CONCLUSÕES

Em um determinado grupo de pessoas, há diversos comportamentos da proporção de acertos ao longo dos ensaios de testes triangulares. Entretanto, todos tendem a convergir ou reduzir sua variabilidade em um determinado número de ensaios. Pela análise gráfica, nota-se que um número muito pequeno de ensaios no processo seletivo é insuficiente para averiguar a real habilidade do candidato. Por isso, recomenda-se $n = 15$ para a etapa de seleção na análise olfativa de café, tendo em vista que não houve alteração do comportamento ou tendência de nenhum candidato a partir desse ensaio.

A independência dos ensaios foi verificada pelo teste de Smith e, assim, o teste sequencial foi avaliado para os dados reais, com o auxílio da regressão segmentada. Através da regressão segmentada foi possível identificar pontos de mudança, porém, não se pode dizer quando há aumento ou queda da proporção de acertos. Apenas, percebe-se que a variação inicial é maior do que a variação final e que todos estão tendendo a sua real habilidade. A regressão segmentada é uma grande ferramenta que auxilia o TS na tomada de decisão.

Sendo assim, sugere-se que, no processo de seleção de provadores, sejam aplicados quinze ensaios triangulares e também, ajustado um modelo de regressão segmentada que estime o *change point* para, a partir dessa estimação, aplicar o TS mantendo-se as pressuposições. Especificamente no presente estudo, através desse procedimento, 6% dos provadores seriam selecionados para o olfato de café e 26% dariam prosseguimento ao teste triangular para se tomar uma decisão e 68% seriam dispensados. Os candidatos que dariam prosseguimento ao teste seriam convocados para mais uma tentativa de seleção.

REFERÊNCIAS

- ALVELOS, H. M. P. P. D. **Análise, Desenvolvimento e Teste de Métodos e Técnicas para Controle Estatístico em Análise Sensorial**. 2002. 253f. Tese (Doutorado em Ciências de Engenharia) - Faculdade de Engenharia, Universidade do Porto, Porto, 2002.
- BROCKHOFF, P. B. Sensometrics for Food Quality Control. **Scandinavian Workshop on Imaging Food Quality**, Dinamarca, n. 15, p. 7-16. 2011.
- DINOV, I. D.; CHRISTOU, N.; GOULD, R. Law of Large Numbers: the Theory, Applications and Technology-based Education. **J Stat Educ.**, v. 17, n. 1, p. 1-19. 2009.
- FERREIRA, D. F. **Estatística Básica**. UFLA. 2005.
- GRUNERT, K. G. Food quality and safety: consumer perception and demand. **European Review of Agricultural Economics**, v. 32, n. 3, p. 369–391. 2005.
- LAWLESS, H. T.; HEYMANN, H. **Sensory Evaluation of Food**. 2nd ed. New York: Springer, 2010.
- LIU, Z.; QIAN, L. Changepoint Estimation in a Segmented Linear Regression via Empirical Likelihood. **Communications in Statistics - Simulation and Computation**, v. 39, n. 1, p. 85-100. 2009.
- MINIM, V. P. R. *et al.* Análise descritiva: Comparação entre metodologias. **Rev. Inst. Latic. Cândido Tostes**, v. 65, n. 374, p. 41-48. 2010.
- MODESTA, R. C. D. *et al.* **Manual para o desenvolvimento do perfil sensorial para bebida de café torrado e moído**. Rio de Janeiro: Embrapa Agroindústria de Alimentos, 1999.
- MUGGEO, V. M. R. Segmented: An R Package to Fit Regression Models with Broken-Line Relationships. **R News**, v. 8, n. 1, p. 20-25. 2008.
- NÆS, T.; BROCKHOFF, P. B.; TOMIC, O. **Statistics for Sensory and Consumer Science**. Wiley. 2010.
- OLIVEIRA, A. P. V.; BENASSI, M. T. Avaliação sensorial de pudins de chocolate com açúcar e dietéticos por perfil livre. **Ciênc. agrotec.**, v. 34, n. 1, p. 146-154. 2010.
- PAES, J. B. **Métodos de diferença em Avaliação Sensorial de Alimentos e Bebidas**. 3^a ed. Viçosa: UFV, 2005.
- PAIVA, E. F. F. **Análise sensorial dos cafés especiais do Estado de Minas Gerais**. 2005. 55f. Dissertação (Mestrado em Ciência dos Alimentos) - Universidade Federal de Lavras, Lavras, 2005.
- PORTZ, L.; DIAS, C. T. S.; CYRINO, J. E. P. Regressão segmentada como modelo na determinação de exigências nutricionais de peixes. **Scientia Agricola**, v. 57, n. 4, p. 601-607.

2000.

POSTE, L. M. *et al.* **Laboratory Methods for Sensory Analysis of Food**. Canada: Agriculture Canada Publication, 1991.

R CORE TEAM (Version 3.6.0) [Computer software]. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <https://www.R-project.org/>. Acesso em: 29 jun. 2019.

RODARTE, M. P. **Análise sensorial, química e perfil de constituintes voláteis de cafés especiais**. 2008. 147f. Tese (Doutorado em Ciência dos Alimentos), Universidade Federal de Lavras, Lavras, 2008.

SCHEID, G. A. **Avaliação Sensorial e Físico-Química de Salame Tipo Italiano com Diferentes Concentrações de Cravo da Índia (*Eugenia caryophyllus*)**. 2001. 94f. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) - Universidade Federal de Viçosa, Viçosa, 2001.

SHIROSE, I.; MORI, E. E. M. **Estatística aplicada à análise sensorial (Módulo 2)**. Campinas: ITAL - Manual Técnico nº 13, 1996.

SOLOW, A. R. Testing for Climate Change: **An Application of the Two-Phase Regression Model**. Manuscript, American Meteorological Society. 1987.

TEIXEIRA, L.V. Análise sensorial na indústria de alimentos. **Rev. Inst. Latic. “Cândido Tostes”**, v.64, n. 366, p. 12-21, 2009.

ZIBETTI, A. **Teoremas Limites**. Disponível em: <https://www.inf.ufsc.br/~andre.zibetti/probabilidade/teoremas-limite.html>. Acesso em: 17 dez. 2019.

CAPÍTULO 5 - CONSIDERAÇÕES FINAIS

Ao reescrever estimadores intervalares como testes, devem-se considerar as pressuposições e condições necessárias para recomendá-los. Esse estudo permitiu avaliar o comportamento desses testes quando as pressuposições são atendidas, ou até mesmo quando uma delas é quebrada (p constante).

Através das análises, foi possível recomendar o TS para $n < 20$ e TN3 quando $n \geq 20$, quando p é constante. Com a quebra dessa pressuposição, observa-se que o comportamento descendente das taxas de acerto no teste triangular gera uma queda do poder, já que o valor de n é pequeno quando a proporção é alta. O número de ensaios interfere mais nas curvas de poder do que a proporção. As taxas de acerto no teste triangular, para o comportamento ascendente, causam um aumento do poder simultâneo a n e p . Além do mais, a não linearidade das funções alteram a velocidade de alcance do poder final e a estatística de teste adquire memória longa que afeta a eficiência dos testes.

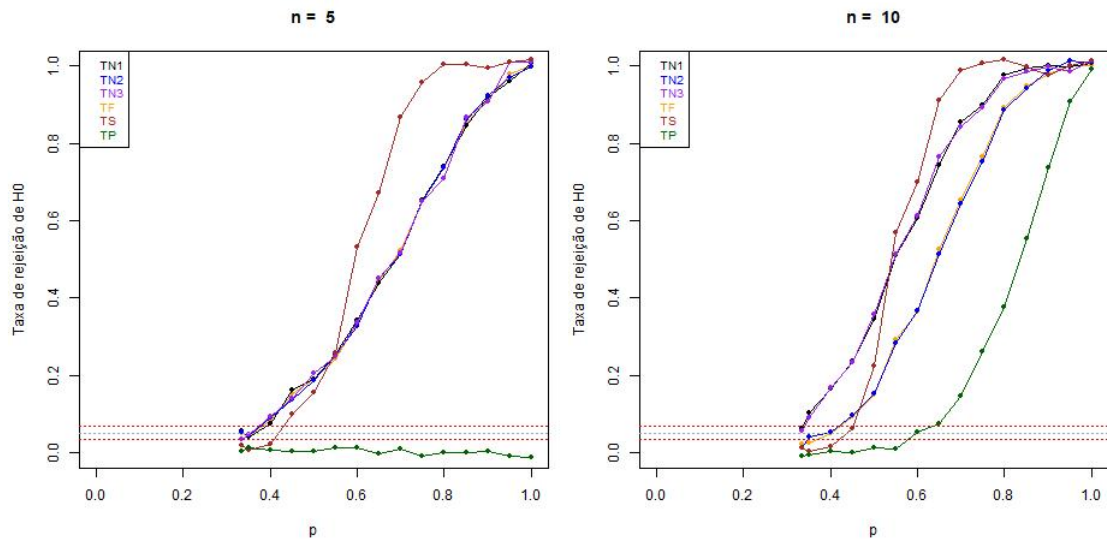
Na análise dos dados reais, obtidos através de ensaios triangulares de odor de café, foi possível notar diferentes comportamentos, que apresentaram regiões constantes em alguns momentos e até mesmo característica crescente ou decrescente em outras regiões. Também, deve-se ressaltar que os ensaios foram independentes. O ajuste de regressão segmentada auxiliou na tomada de decisão do TS, teste mais indicado ao número de ensaios efetuados. No entanto, como há alta variabilidade inicial, observou-se que o número de ensaios recomendado na literatura é muito pequeno para que o candidato atinja sua real habilidade. Para auxiliar na tomada de decisão a partir do momento em que a proporção converge ou reduz sua variabilidade, recomenda-se aplicar maior número de ensaios triangulares ($n = 15$) para análise olfativa de café e fazer o uso da regressão segmentada como ferramenta auxiliar do TS.

Para trabalhos futuros, sugere-se corrigir o efeito de memória longa da estatística de teste. Dessa forma, os testes iriam adquirir maior eficiência para refletir a mudança do parâmetro, tanto para o cansaço quanto para a aprendizagem.

APÊNDICES

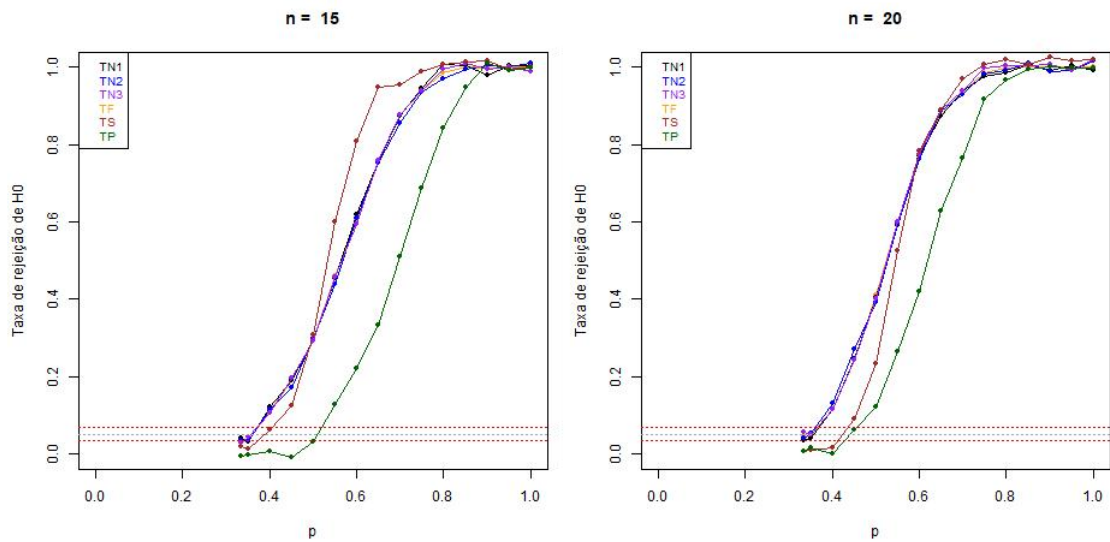
APÊNDICE A - Gráficos do poder para cada teste de 5 a 100 ensaios e proporção constante

Figura 22 – Gráfico das taxas de rejeição de H_0 para 5 e 10 ensaios.



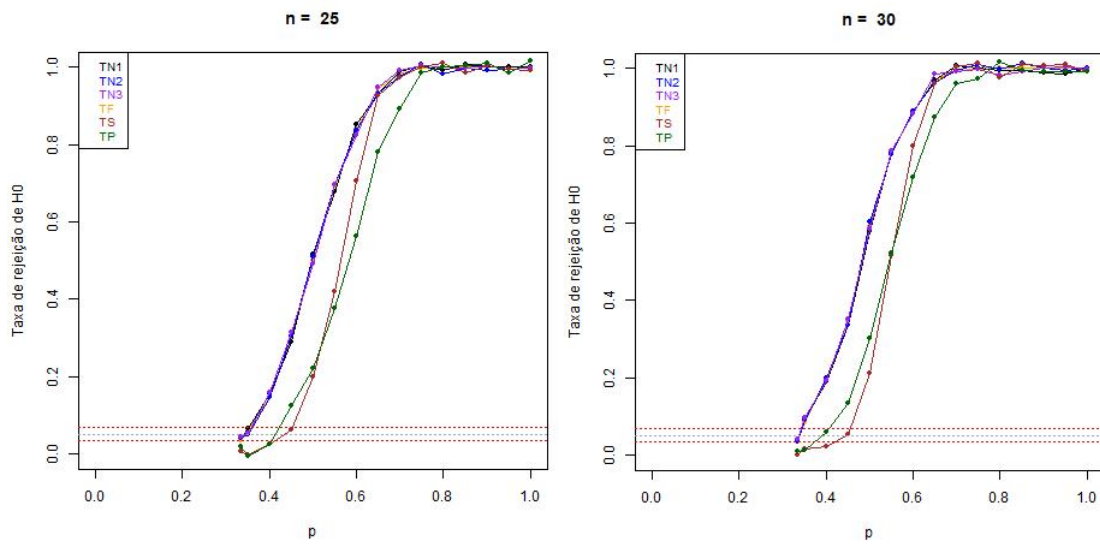
Fonte: Da autora.

Figura 23 – Gráfico das taxas de rejeição de H_0 para 15 e 20 ensaios.



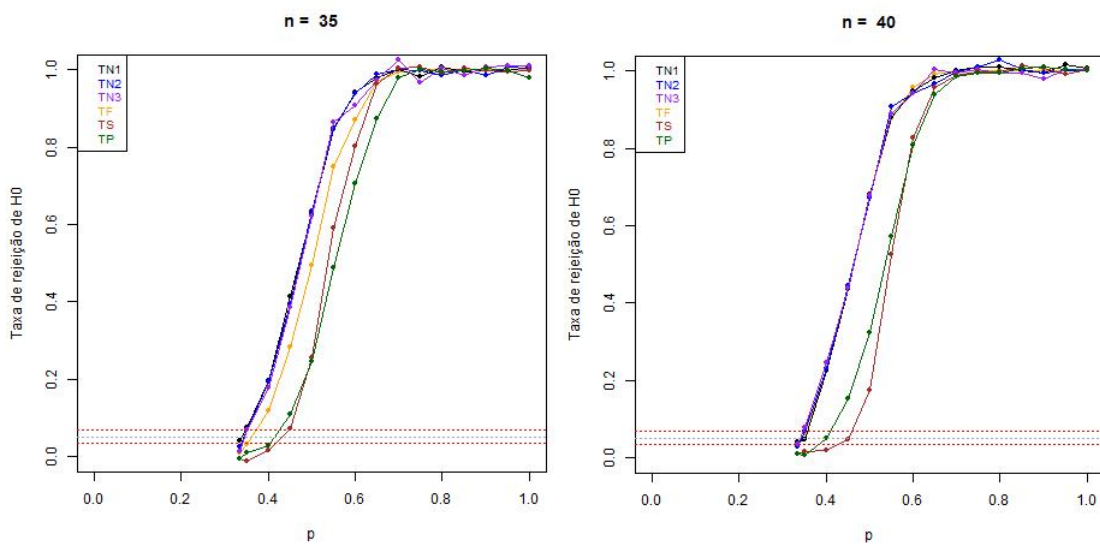
Fonte: Da autora.

Figura 24 – Gráfico das taxas de rejeição de H_0 para 25 e 30 ensaios.



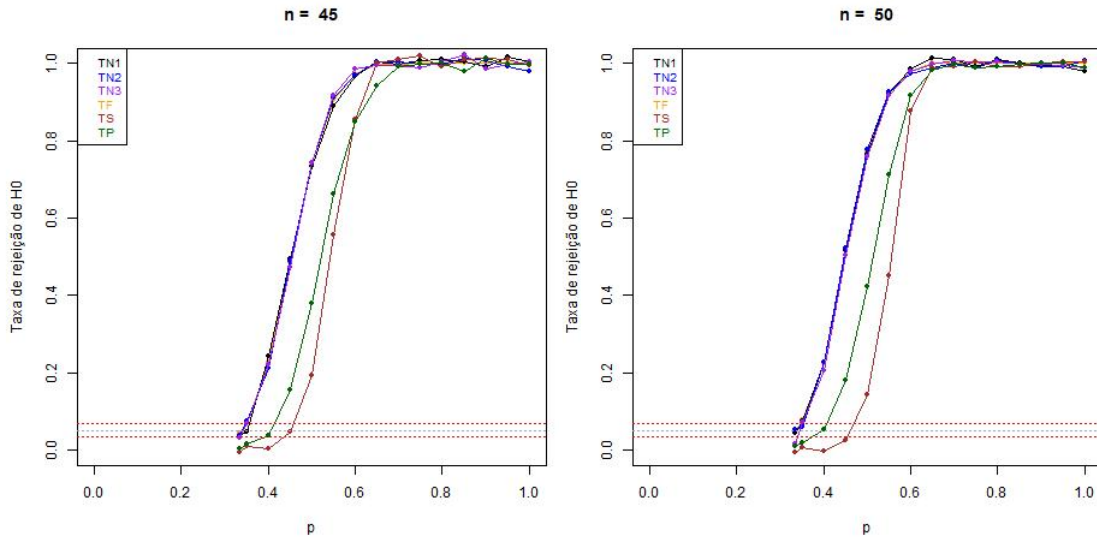
Fonte: Da autora.

Figura 25 – Gráfico das taxas de rejeição de H_0 para 35 e 40 ensaios.



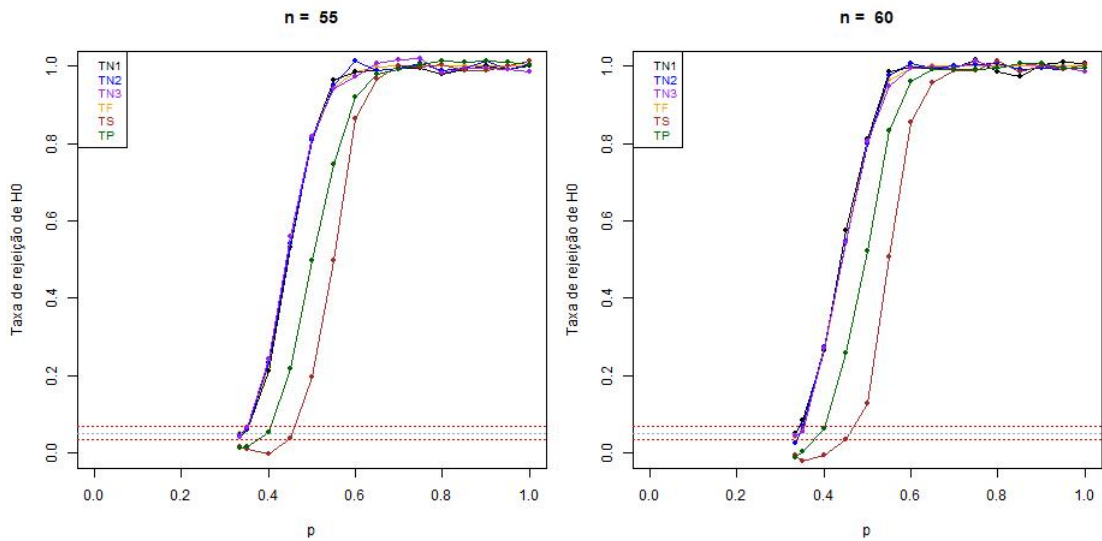
Fonte: Da autora.

Figura 26 – Gráfico das taxas de rejeição de H_0 para 45 e 50 ensaios.



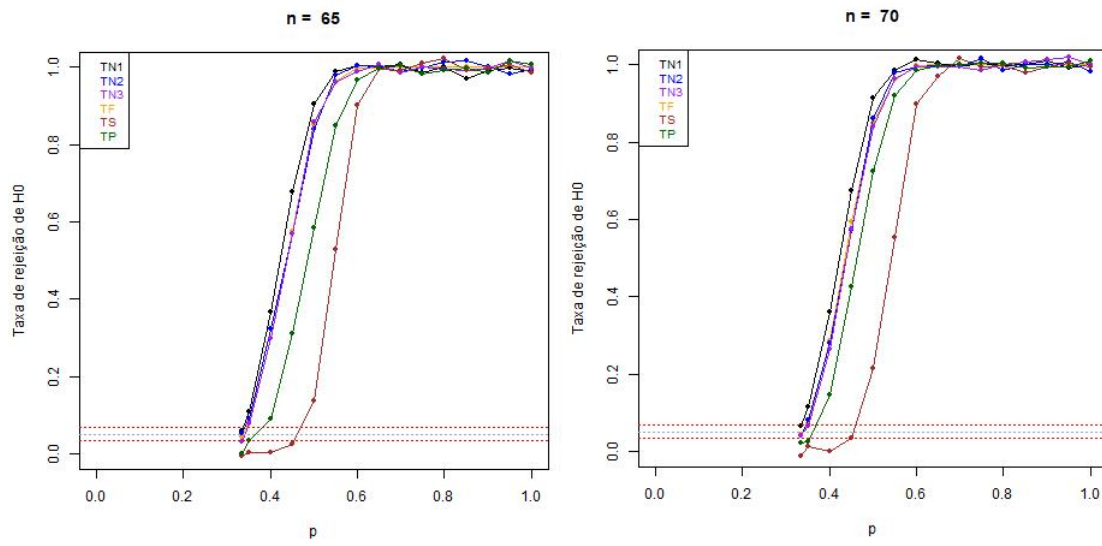
Fonte: Da autora.

Figura 27 – Gráfico das taxas de rejeição de H_0 para 55 e 60 ensaios.



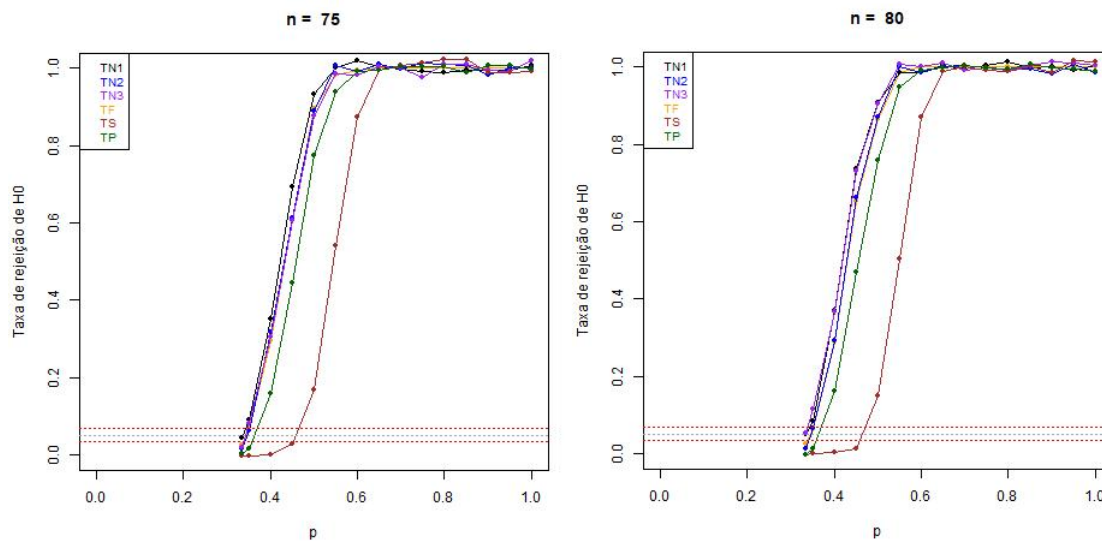
Fonte: Da autora.

Figura 28 – Gráfico das taxas de rejeição de H_0 para 65 e 70 ensaios.



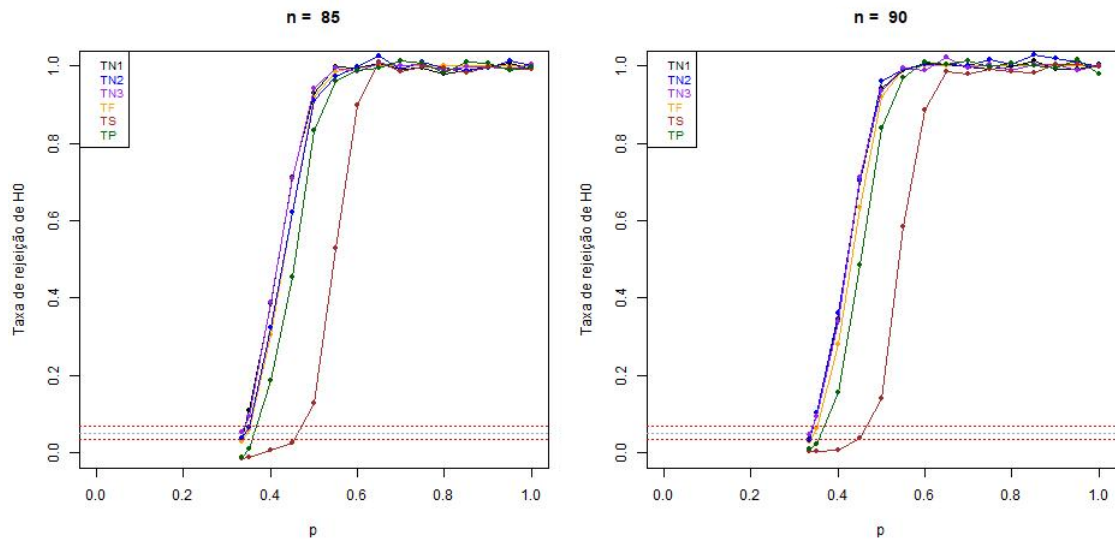
Fonte: Da autora.

Figura 29 – Gráfico das taxas de rejeição de H_0 para 75 e 80 ensaios.



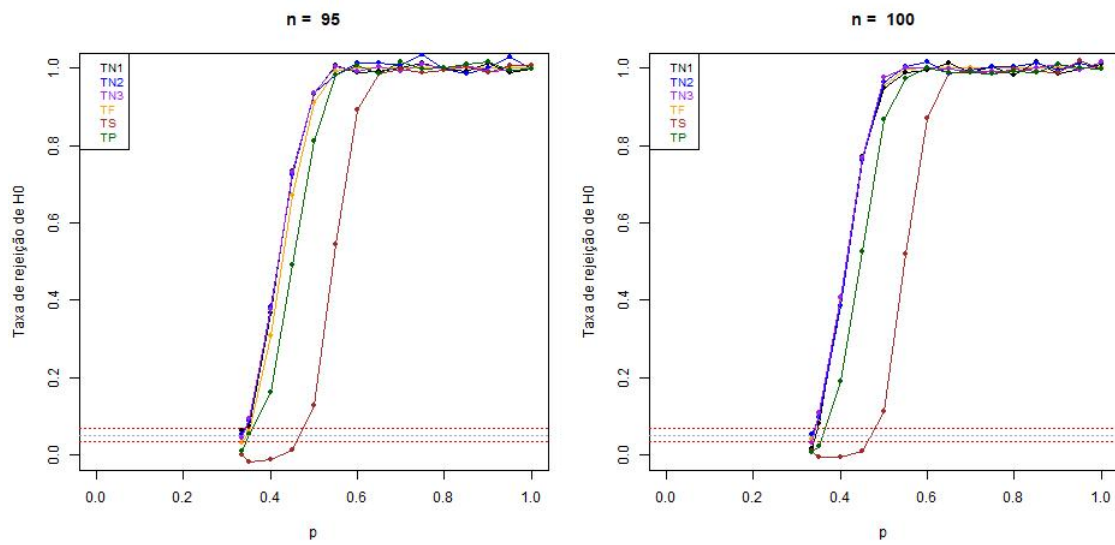
Fonte: Da autora.

Figura 30 – Gráfico das taxas de rejeição de H_0 para 85 e 90 ensaios.



Fonte: Da autora.

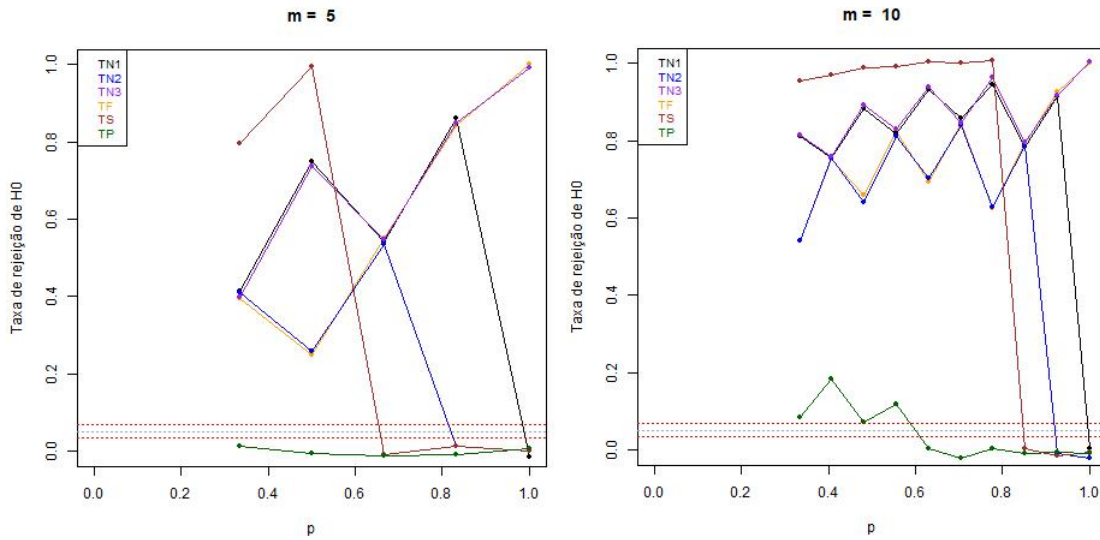
Figura 31 – Gráfico das taxas de rejeição de H_0 para 95 e 100 ensaios.



Fonte: Da autora.

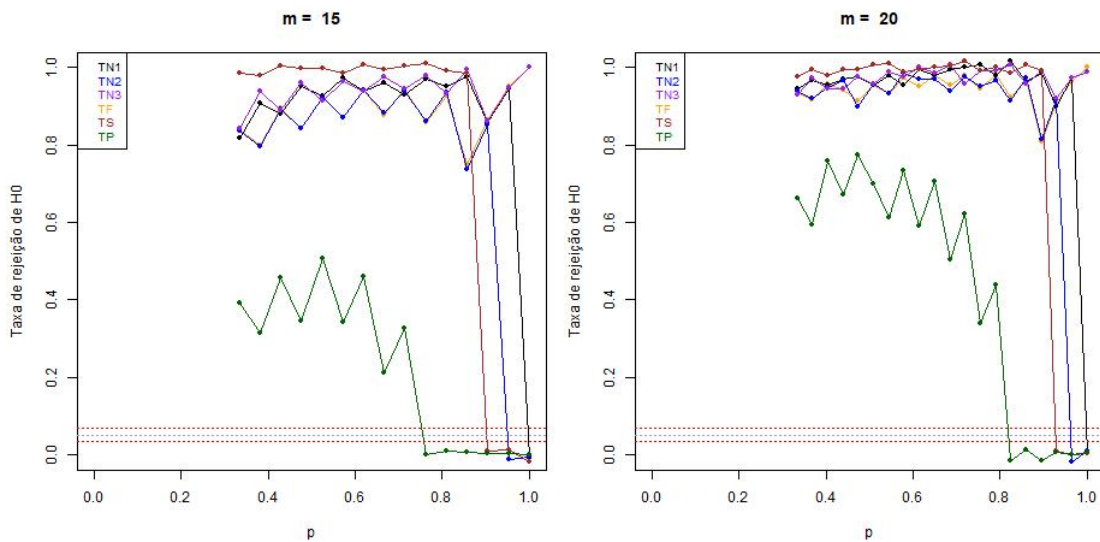
APÊNDICE B1 - Gráficos do poder para cada teste de 5 a 100 ensaios obtido para o cansaço linear

Figura 32 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 5 e 10 ensaios.



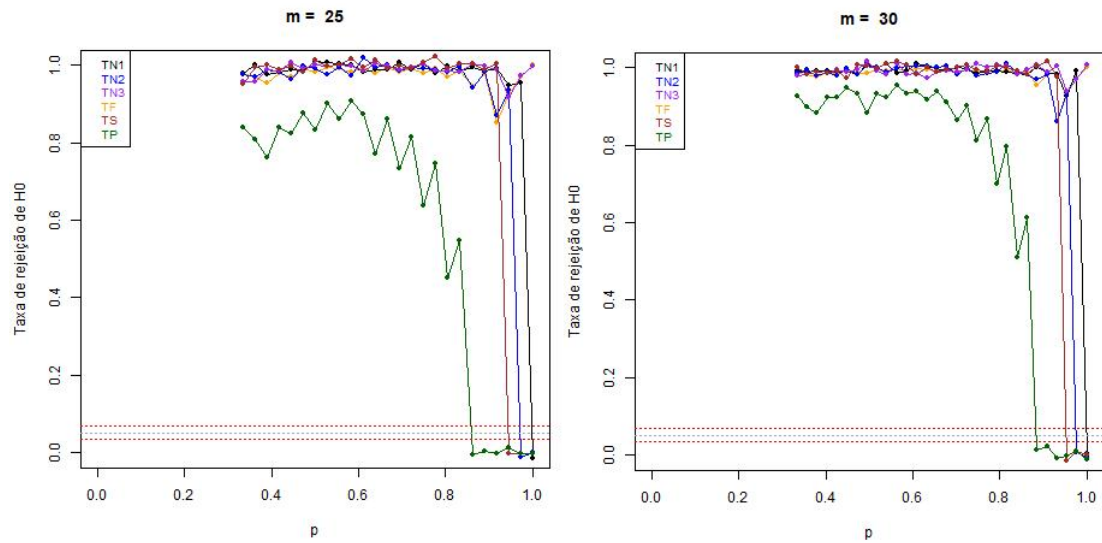
Fonte: Da autora.

Figura 33 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 15 e 20 ensaios.



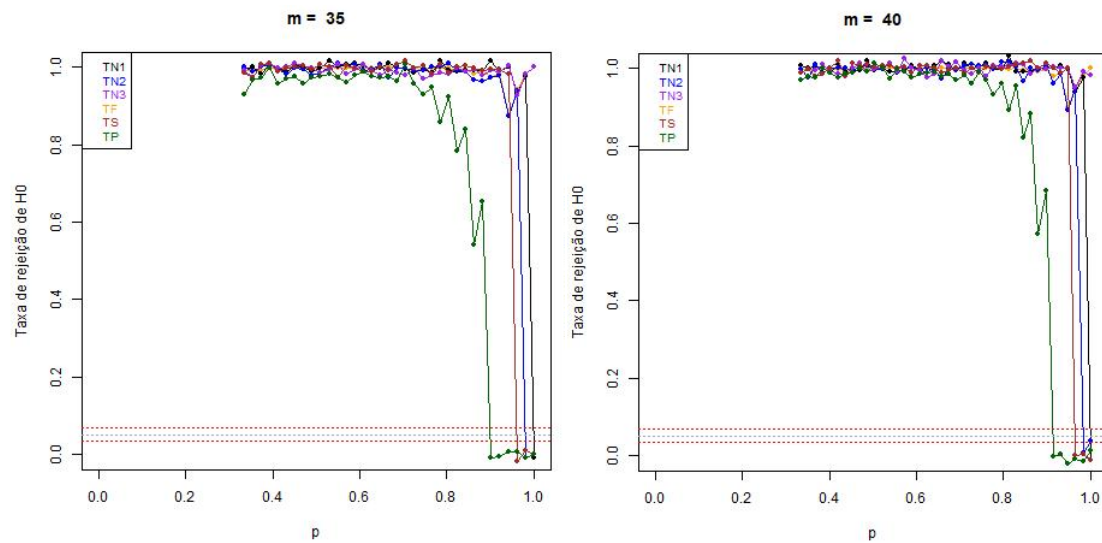
Fonte: Da autora.

Figura 34 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 25 e 30 ensaios.



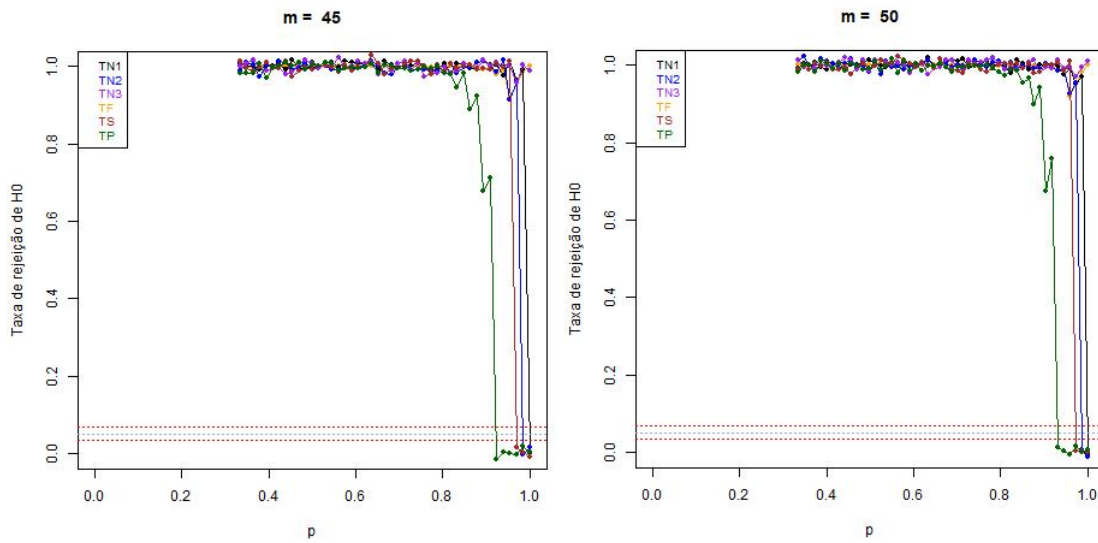
Fonte: Da autora.

Figura 35 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 35 e 40 ensaios.



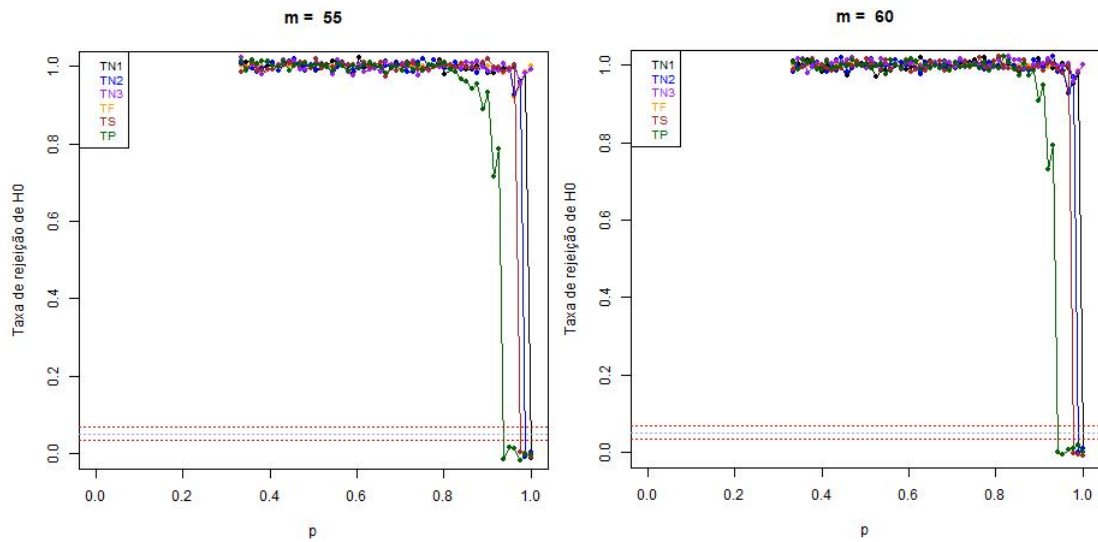
Fonte: Da autora.

Figura 36 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 45 e 50 ensaios.



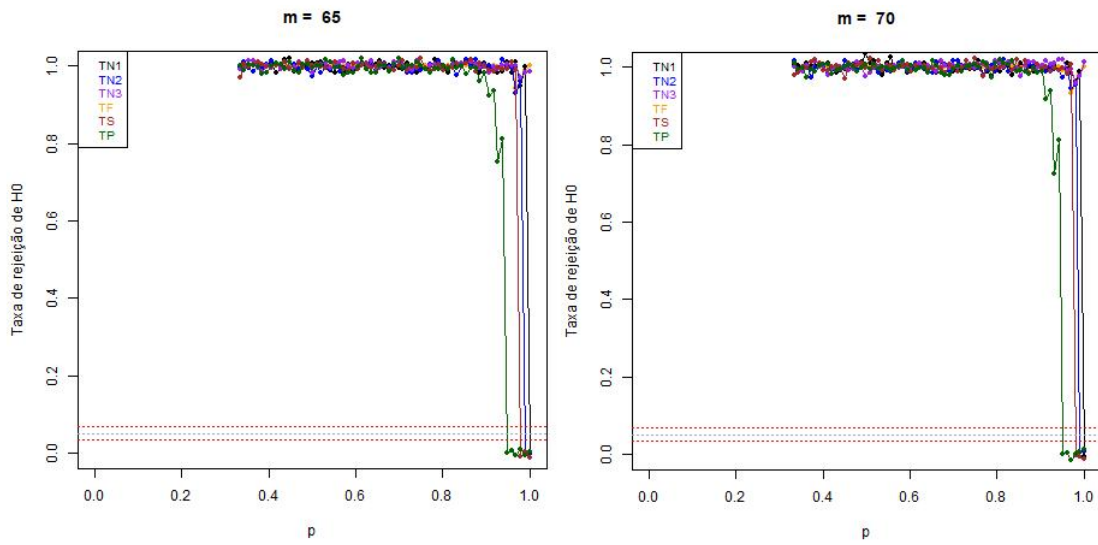
Fonte: Da autora.

Figura 37 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 55 e 60 ensaios.



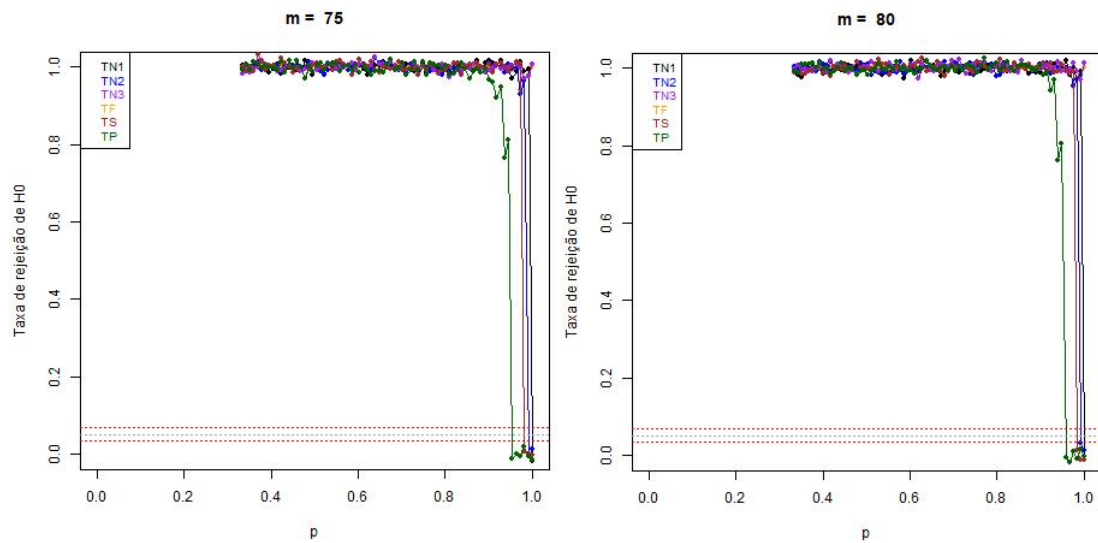
Fonte: Da autora.

Figura 38 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 65 e 70 ensaios.



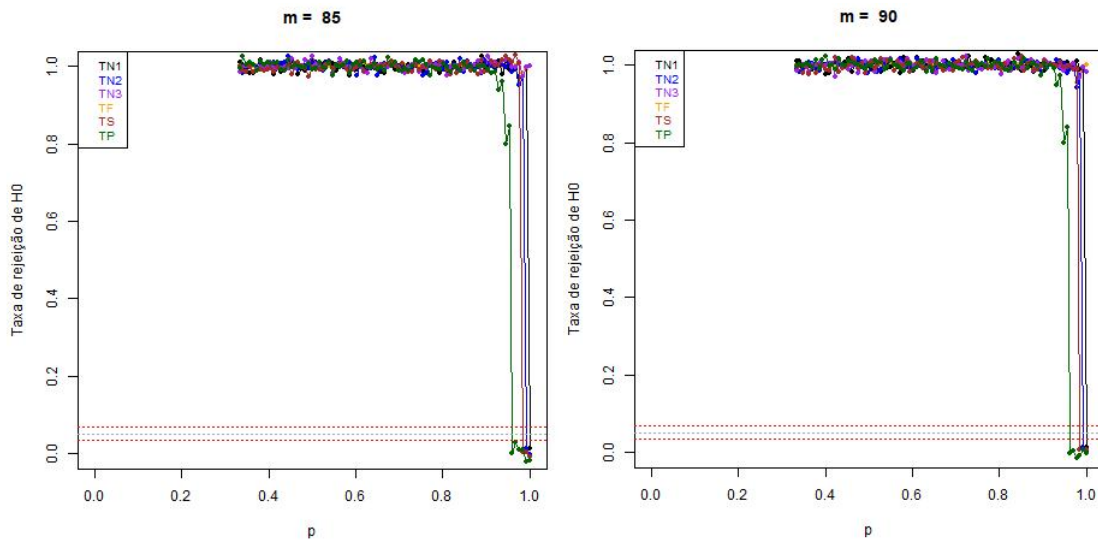
Fonte: Da autora.

Figura 39 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 75 e 80 ensaios.



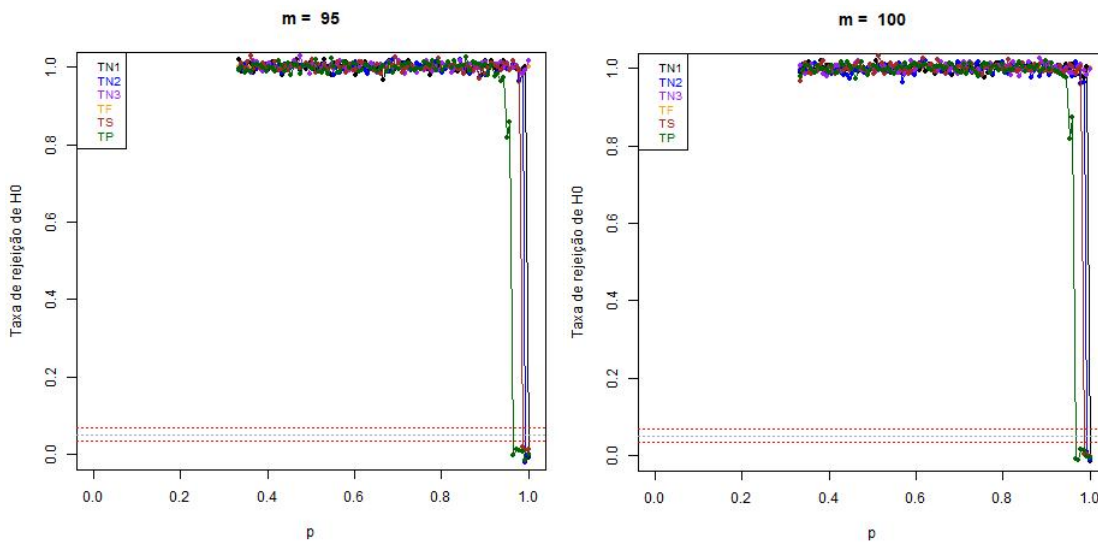
Fonte: Da autora.

Figura 40 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 85 e 90 ensaios.



Fonte: Da autora.

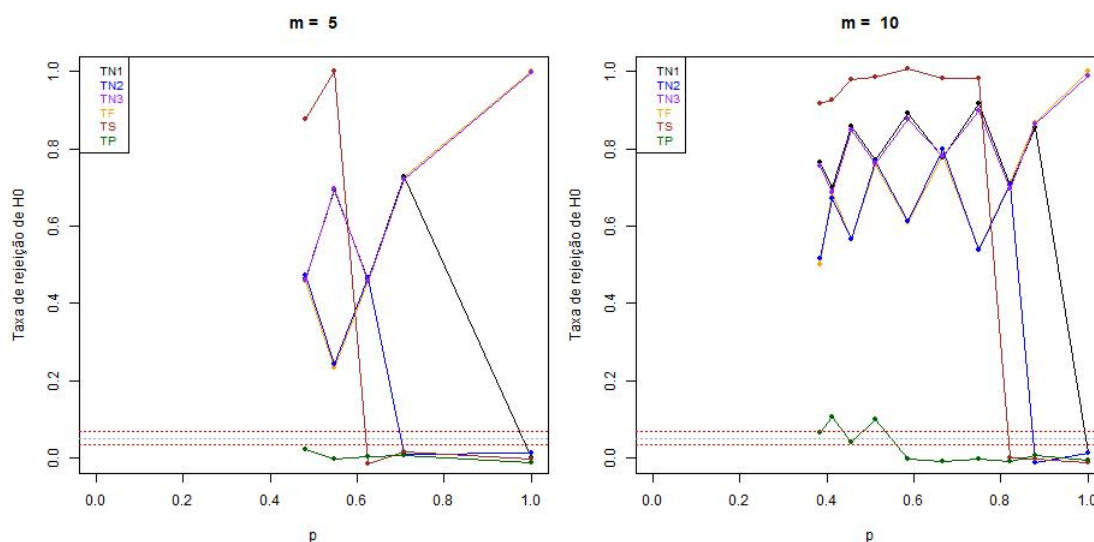
Figura 41 – Gráfico das taxas de rejeição de H_0 para o cansaço linear em 95 e 100 ensaios.



Fonte: Da autora.

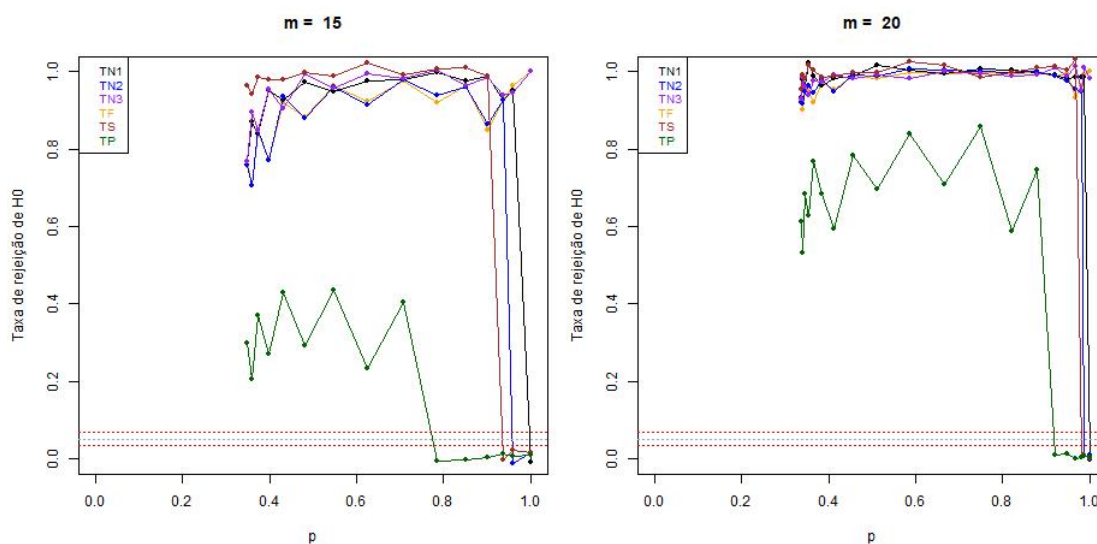
APÊNDICE B2 - Gráficos do poder para cada teste de 5 a 100 ensaios obtido para o cansaço não linear

Figura 42 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 5 e 10 ensaios.



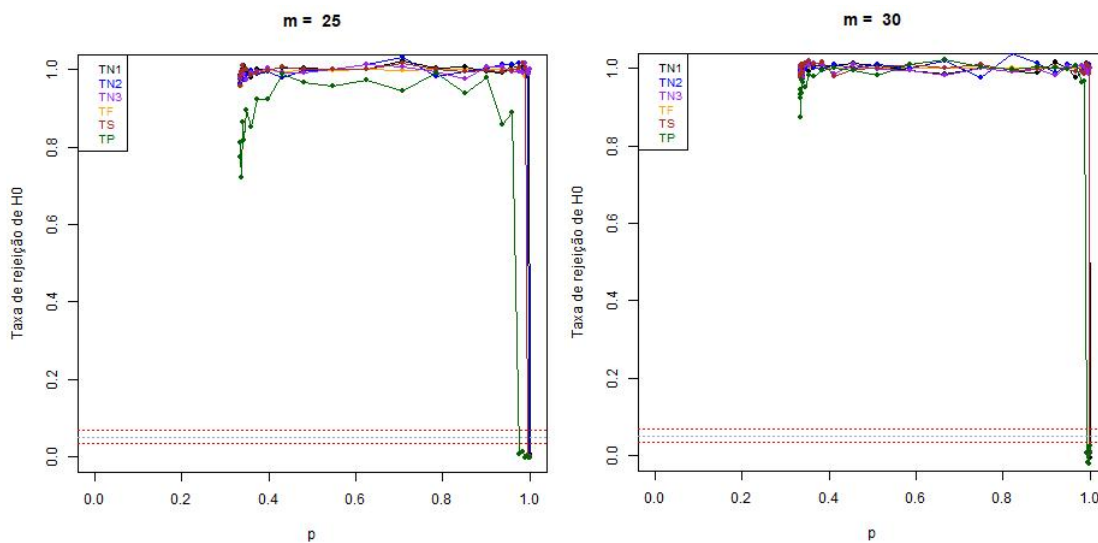
Fonte: Da autora.

Figura 43 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 15 e 20 ensaios.



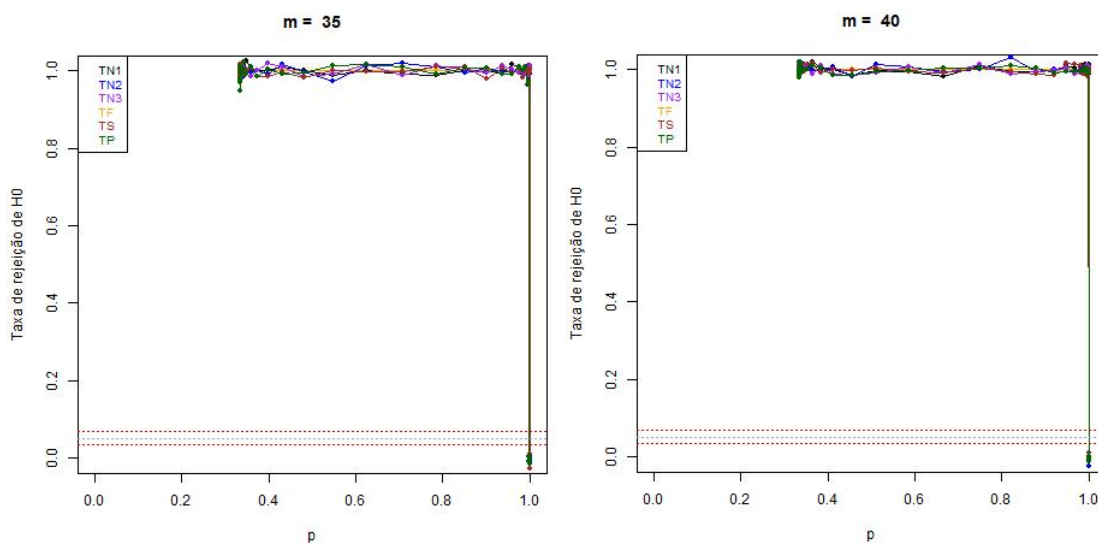
Fonte: Da autora.

Figura 44 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 25 e 30 ensaios.



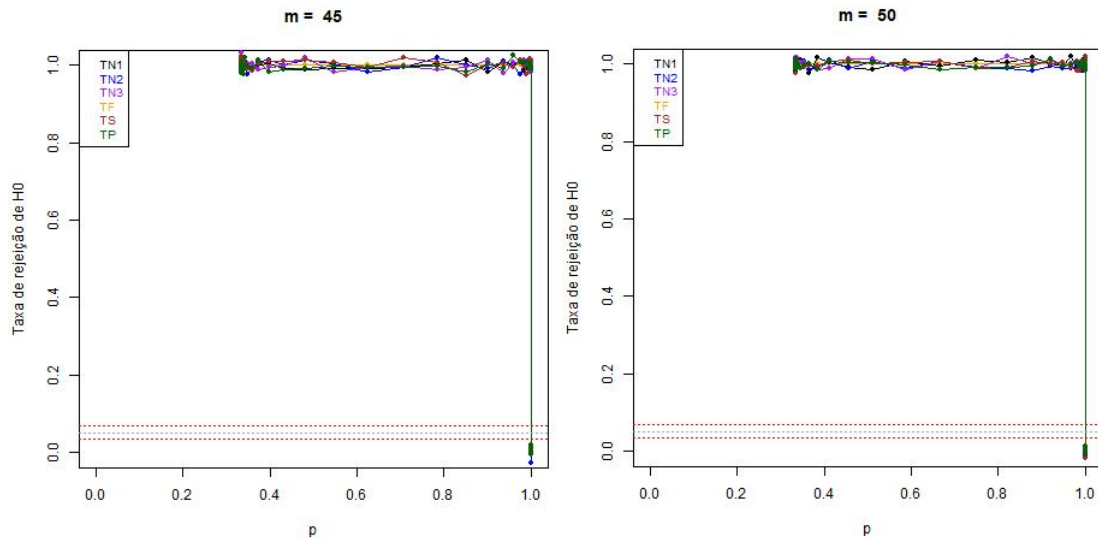
Fonte: Da autora.

Figura 45 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 35 e 40 ensaios.



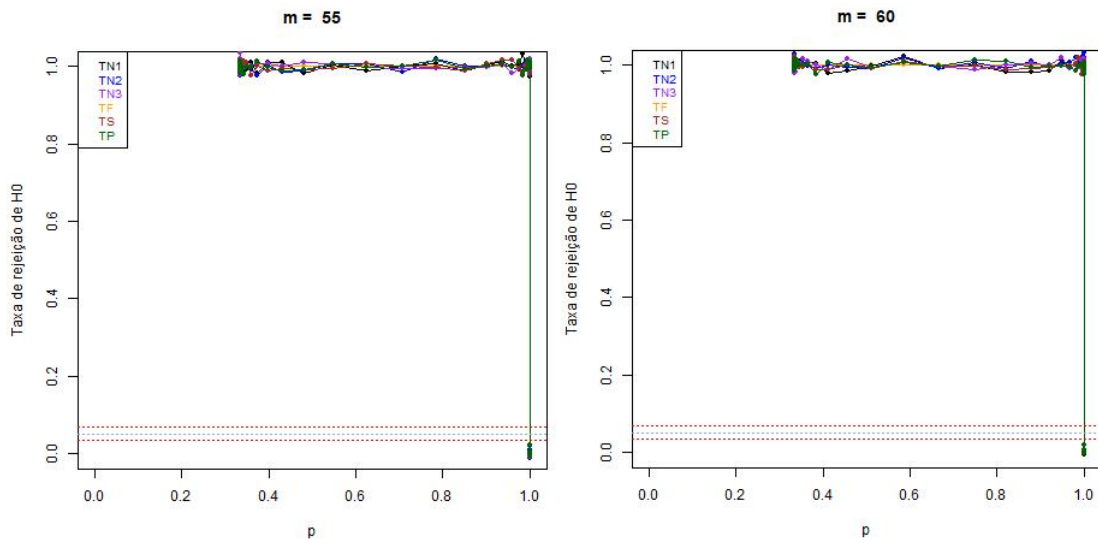
Fonte: Da autora.

Figura 46 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 45 e 50 ensaios.



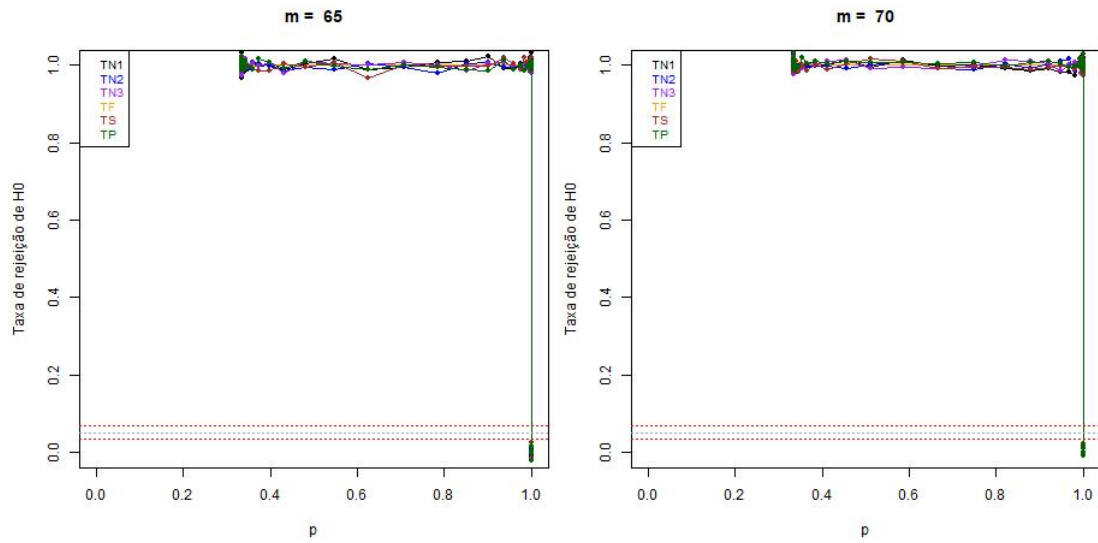
Fonte: Da autora.

Figura 47 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 55 e 60 ensaios.



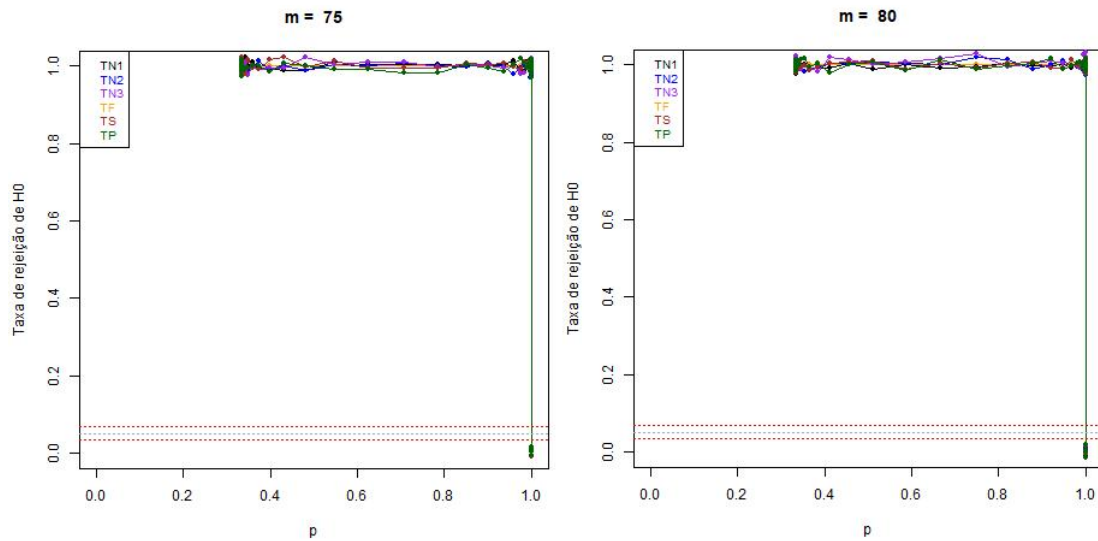
Fonte: Da autora.

Figura 48 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 65 e 70 ensaios.



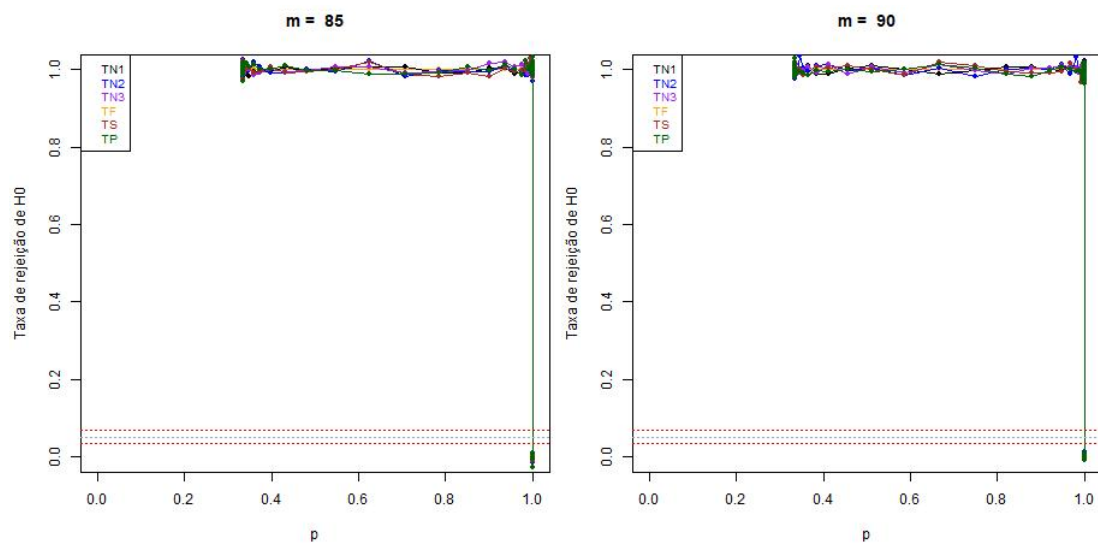
Fonte: Da autora.

Figura 49 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 75 e 80 ensaios.



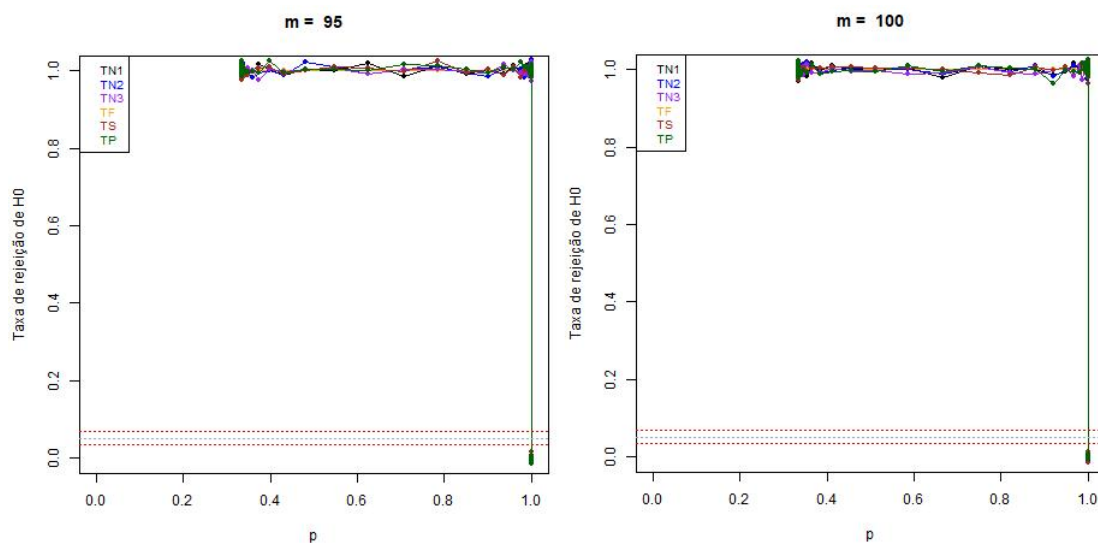
Fonte: Da autora.

Figura 50 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 85 e 90 ensaios.



Fonte: Da autora.

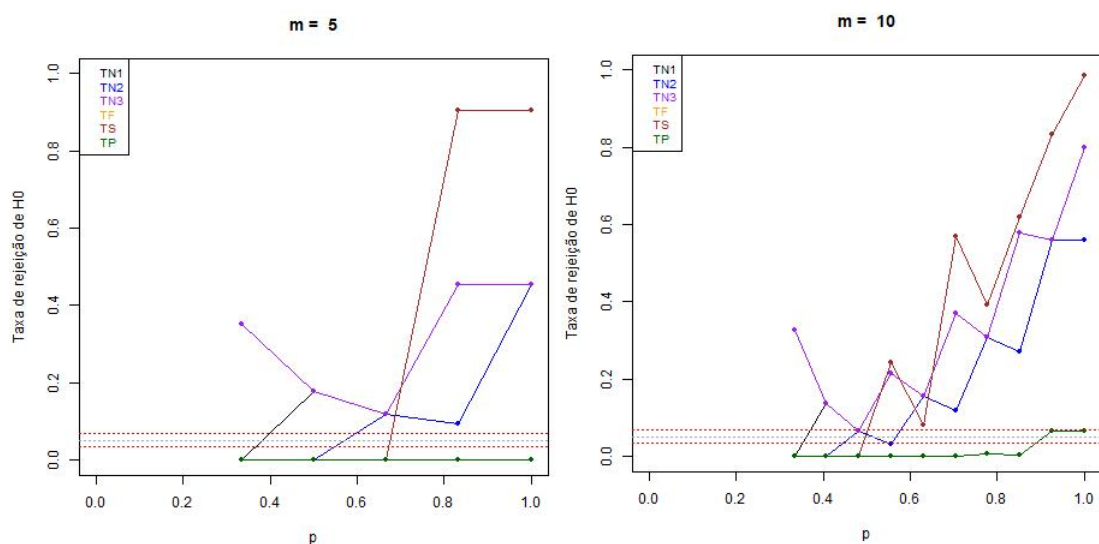
Figura 51 – Gráfico das taxas de rejeição de H_0 para o cansaço não linear em 95 e 100 ensaios.



Fonte: Da autora.

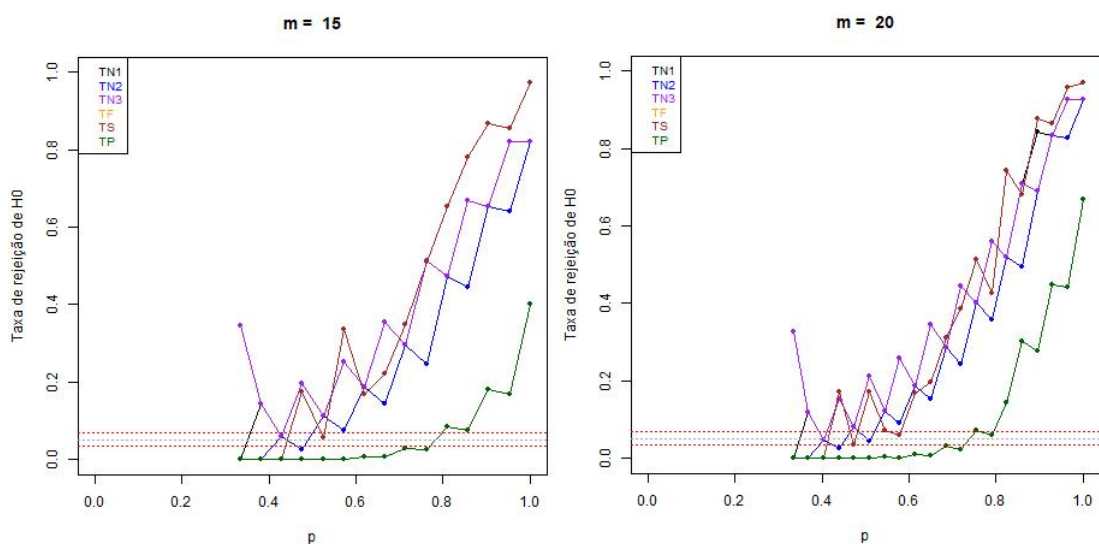
APÊNDICE B3 - Gráficos do poder para cada teste de 5 a 100 ensaios obtido para o aprendizado linear

Figura 52 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 5 e 10 ensaios.



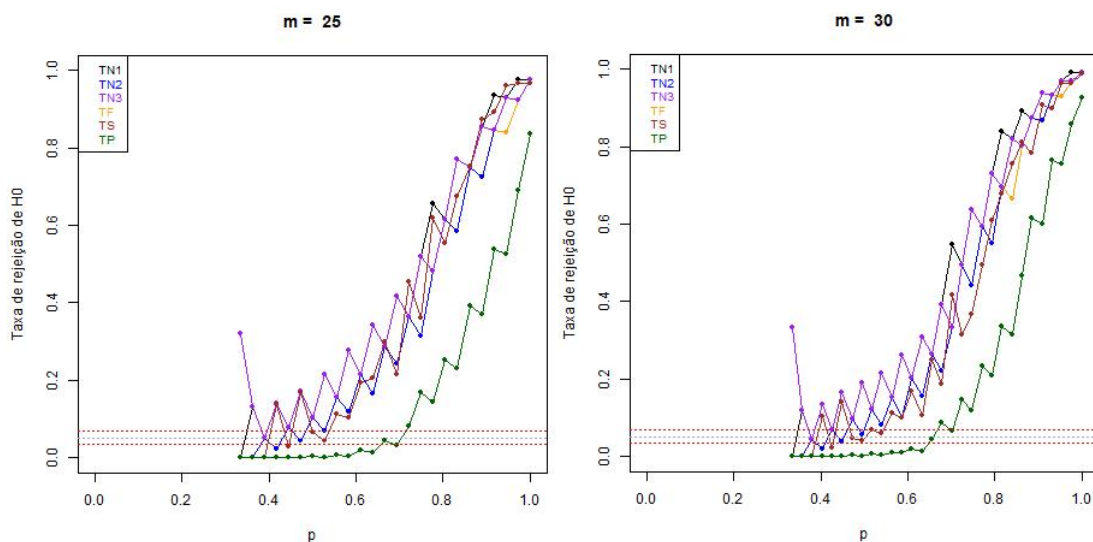
Fonte: Da autora.

Figura 53 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 15 e 20 ensaios.



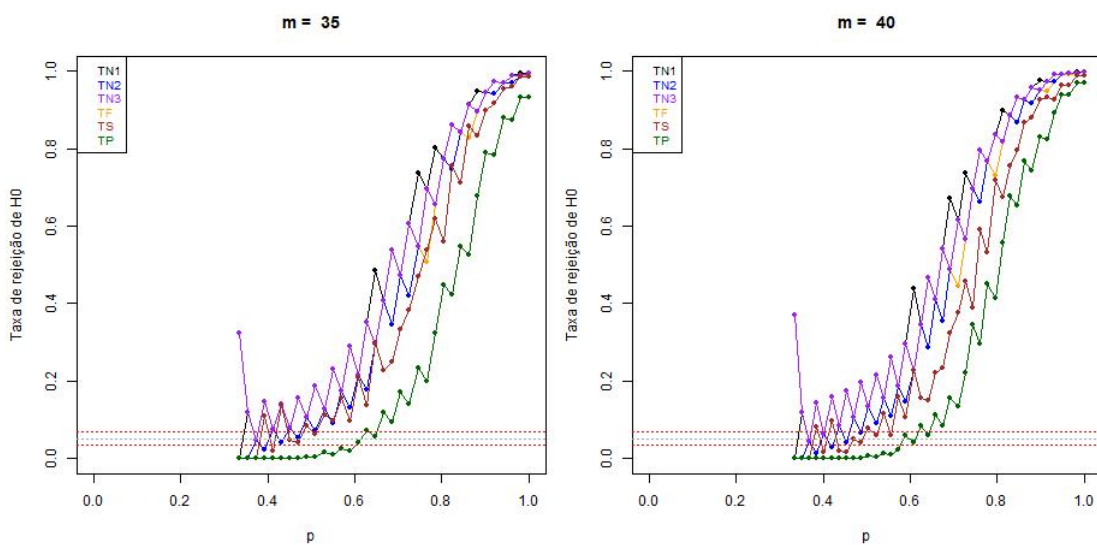
Fonte: Da autora.

Figura 54 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 25 e 30 ensaios.



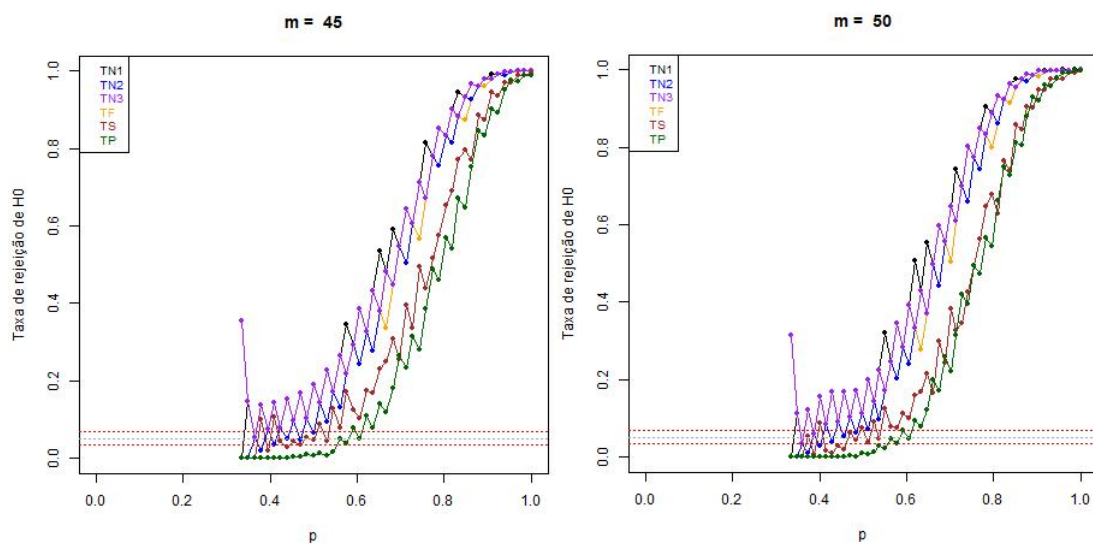
Fonte: Da autora.

Figura 55 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 35 e 40 ensaios.



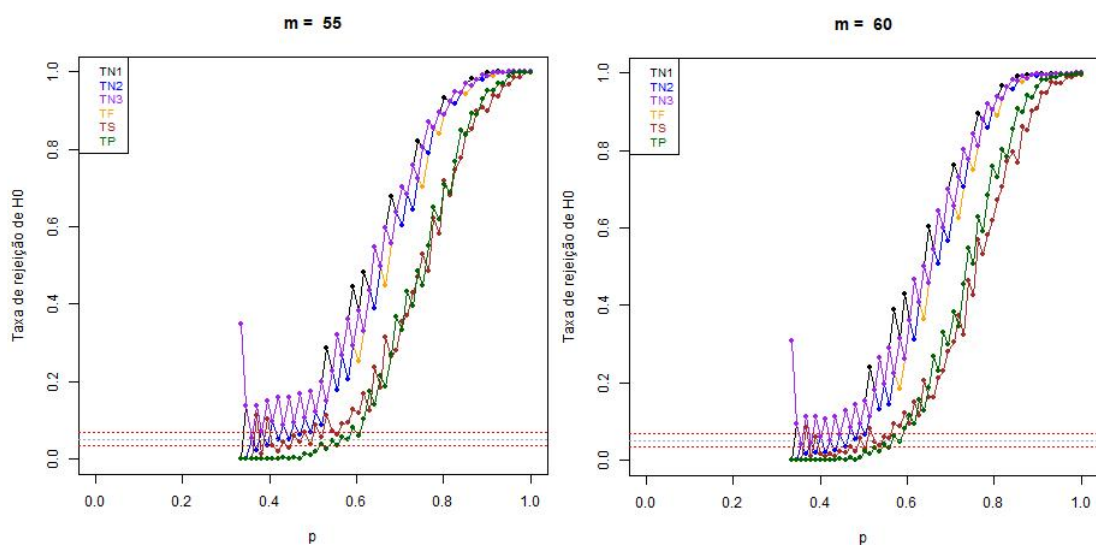
Fonte: Da autora.

Figura 56 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 45 e 50 ensaios.



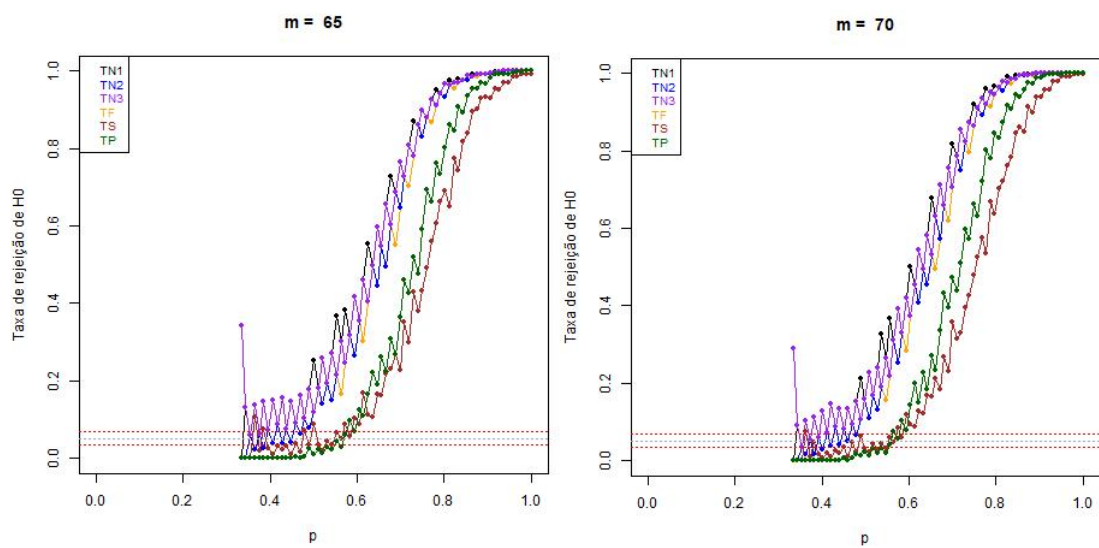
Fonte: Da autora.

Figura 57 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 55 e 60 ensaios.



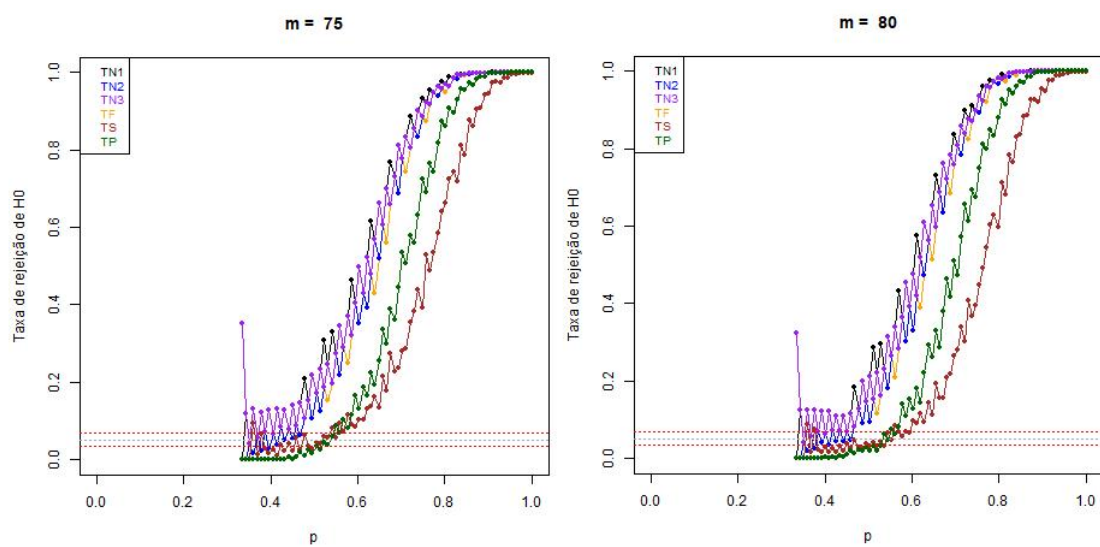
Fonte: Da autora.

Figura 58 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 65 e 70 ensaios.



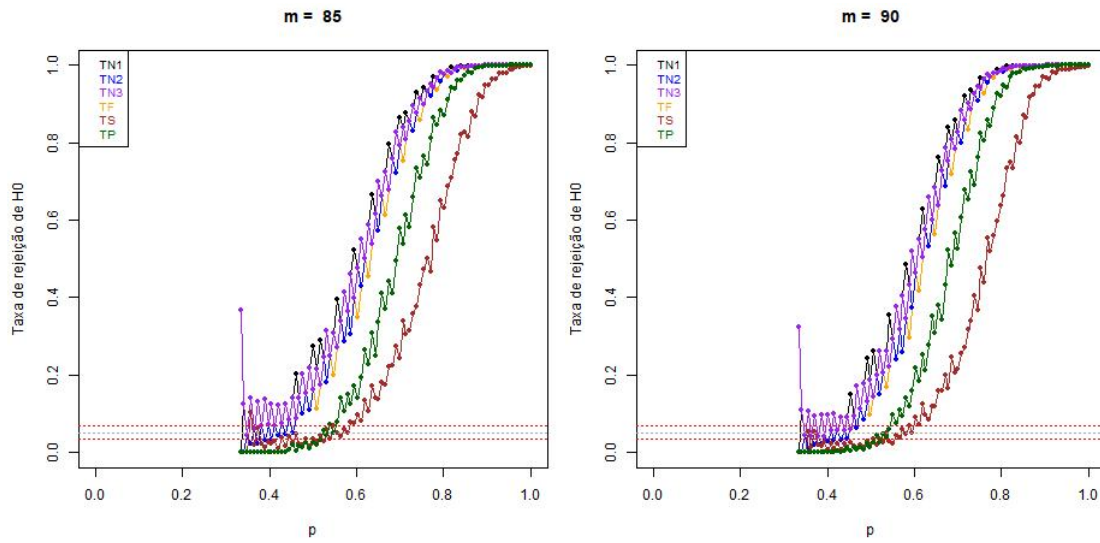
Fonte: Da autora.

Figura 59 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 75 e 80 ensaios.



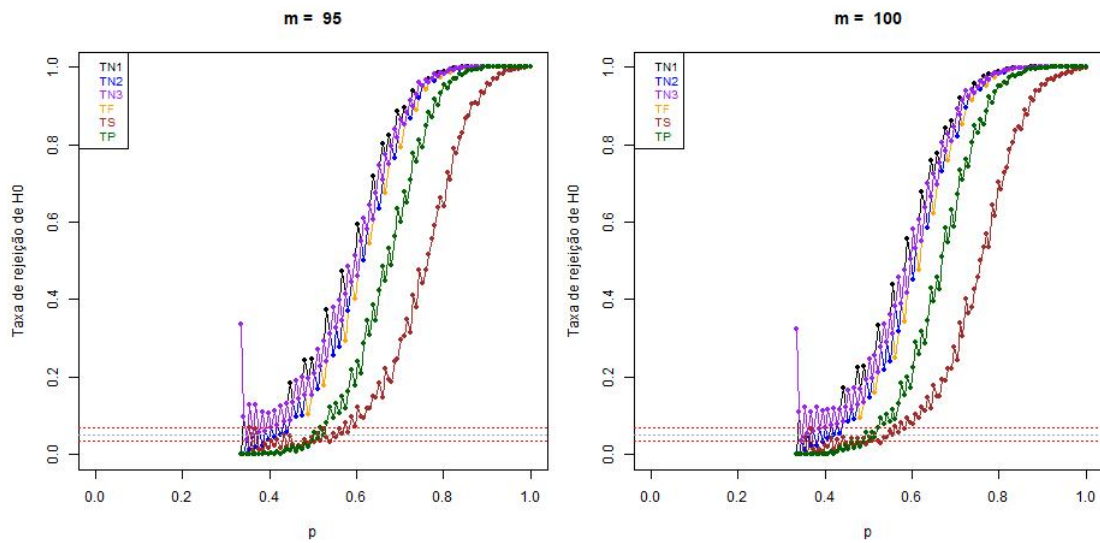
Fonte: Da autora.

Figura 60 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 85 e 90 ensaios.



Fonte: Da autora.

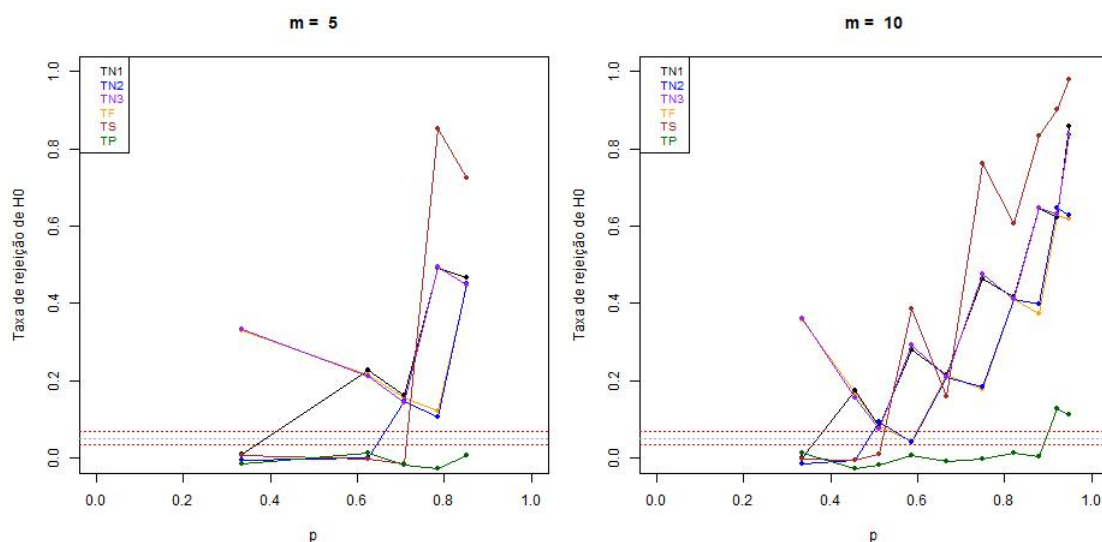
Figura 61 – Gráfico das taxas de rejeição de H_0 para o aprendizado linear em 95 e 100 ensaios.



Fonte: Da autora.

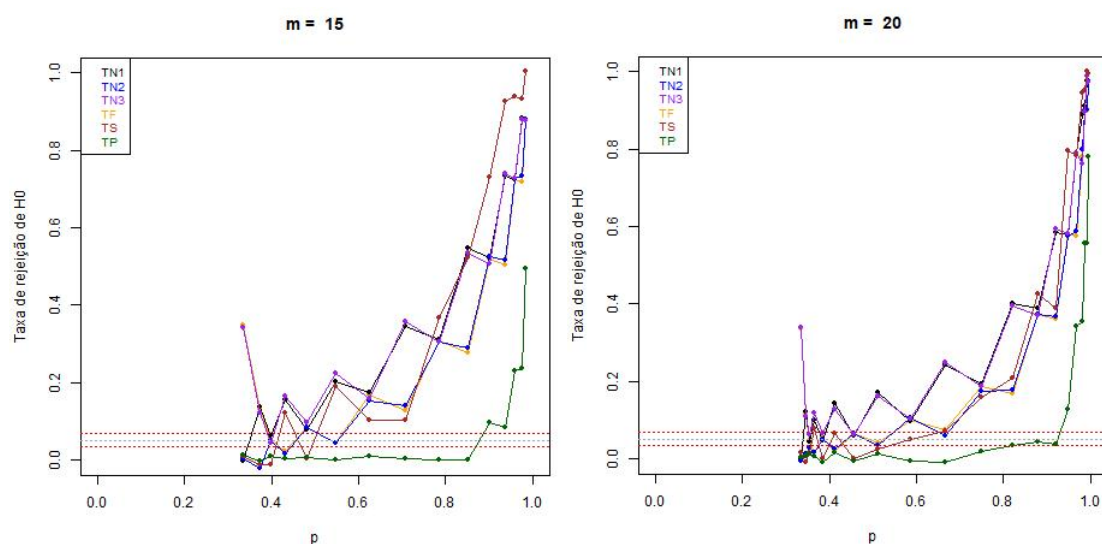
APÊNDICE B4 - Gráficos do poder para cada teste de 5 a 100 ensaios obtido para o aprendizado não linear

Figura 62 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 5 e 10 ensaios.



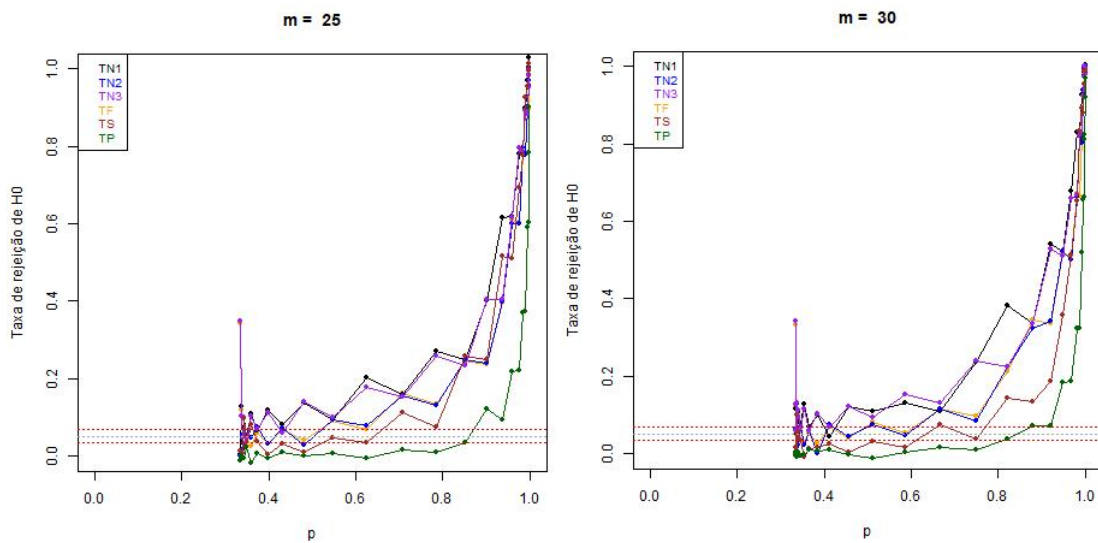
Fonte: Da autora.

Figura 63 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 15 e 20 ensaios.



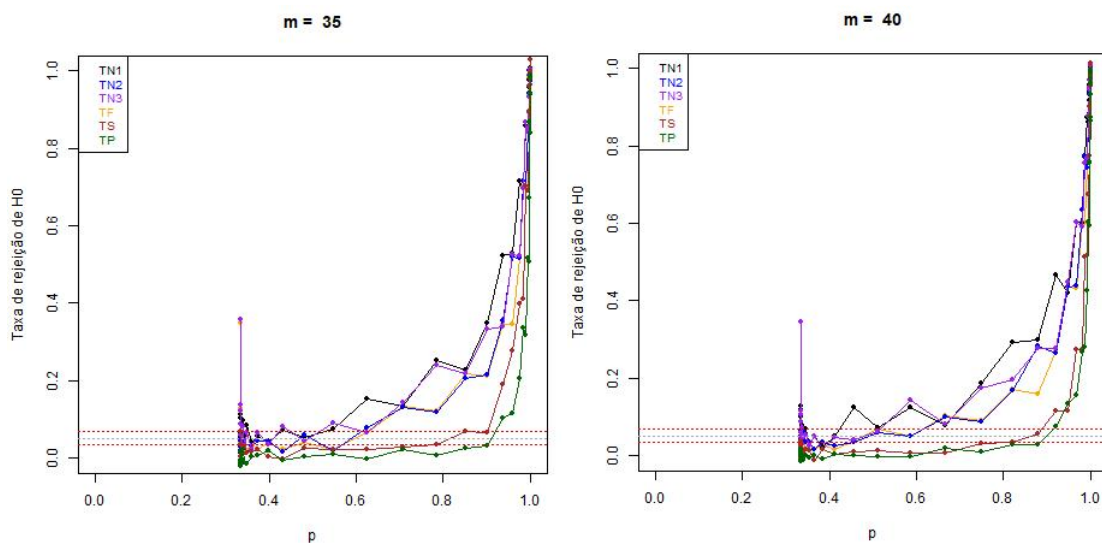
Fonte: Da autora.

Figura 64 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 25 e 30 ensaios.



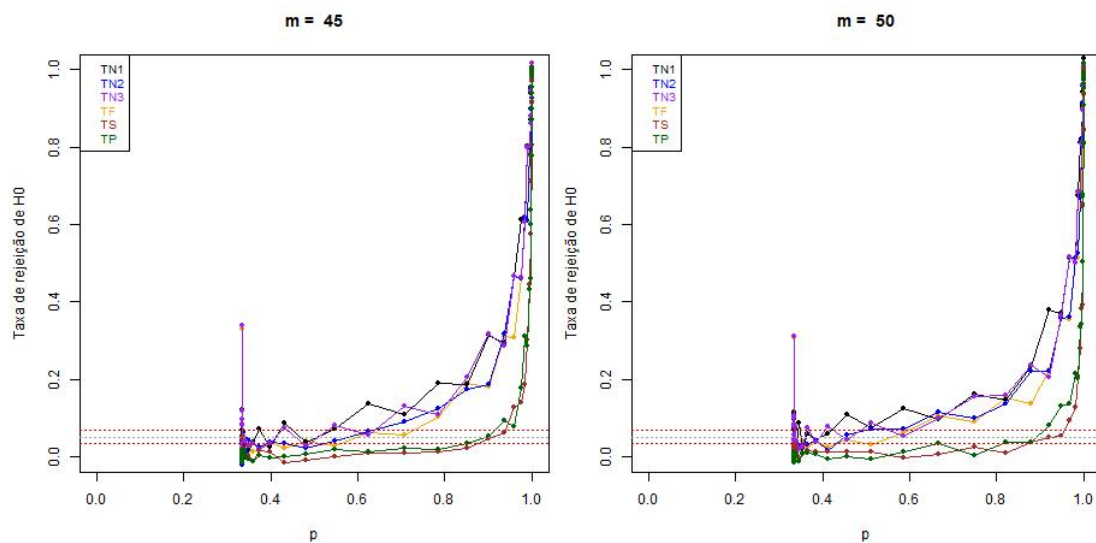
Fonte: Da autora.

Figura 65 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 35 e 40 ensaios.



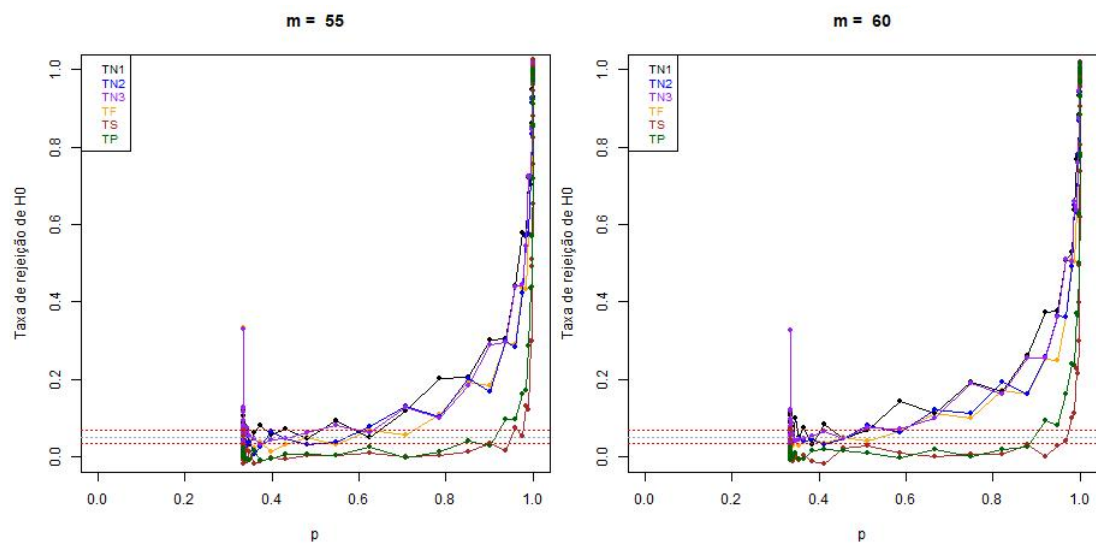
Fonte: Da autora.

Figura 66 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 45 e 50 ensaios.



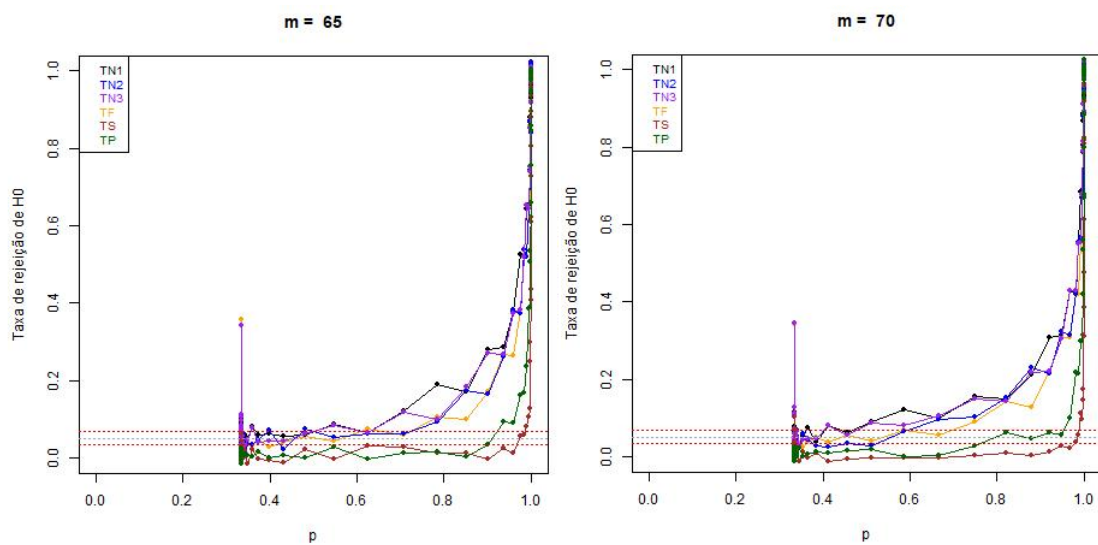
Fonte: Da autora.

Figura 67 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 55 e 60 ensaios.



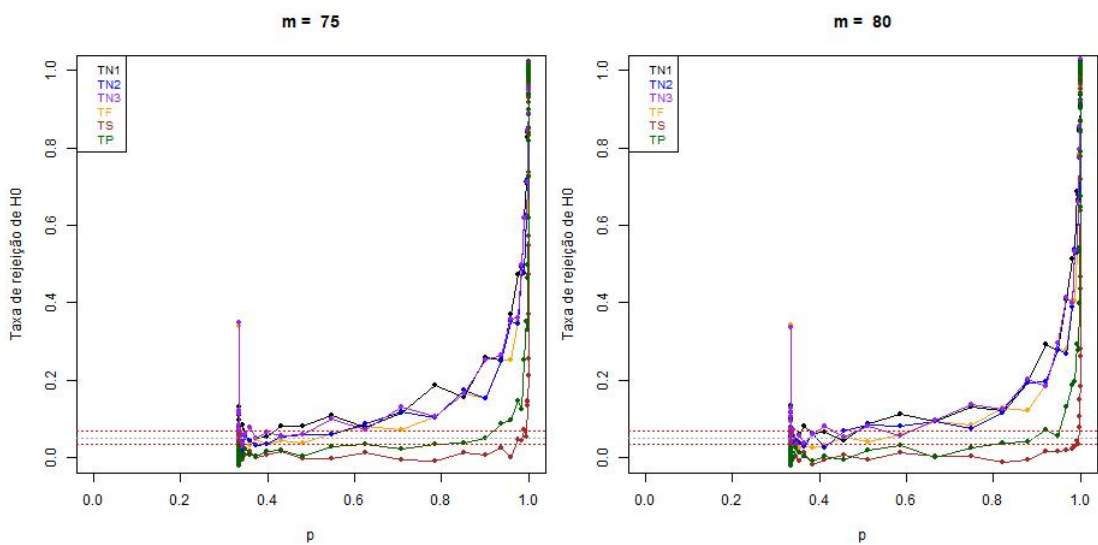
Fonte: Da autora.

Figura 68 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 65 e 70 ensaios.



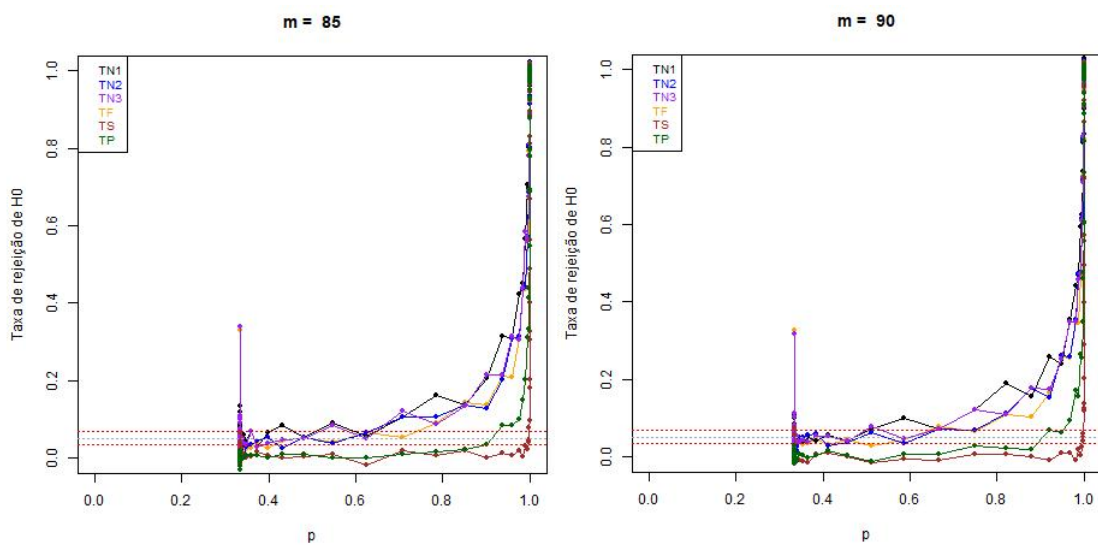
Fonte: Da autora.

Figura 69 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 75 e 80 ensaios.



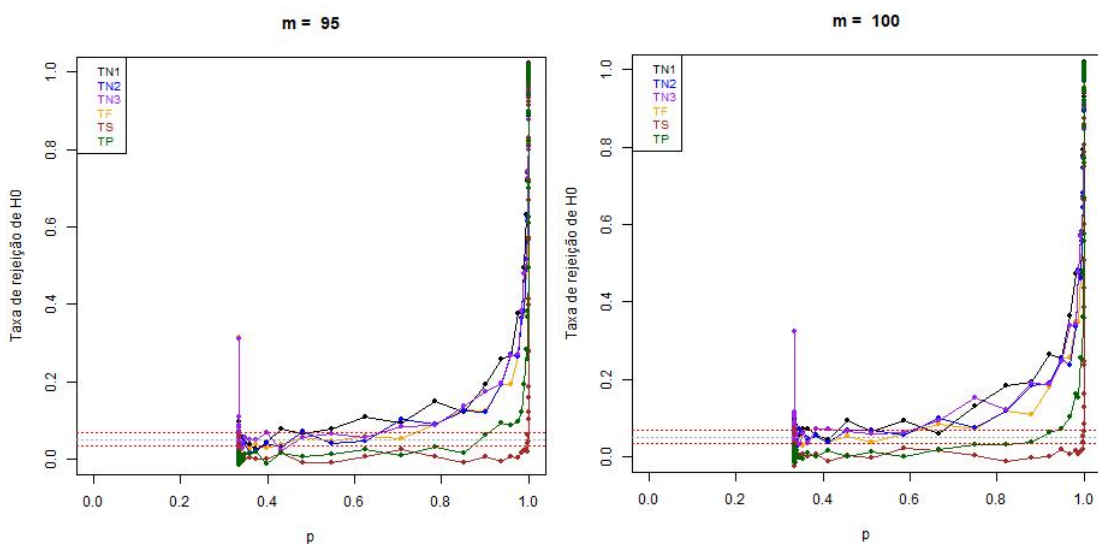
Fonte: Da autora.

Figura 70 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 85 e 90 ensaios.



Fonte: Da autora.

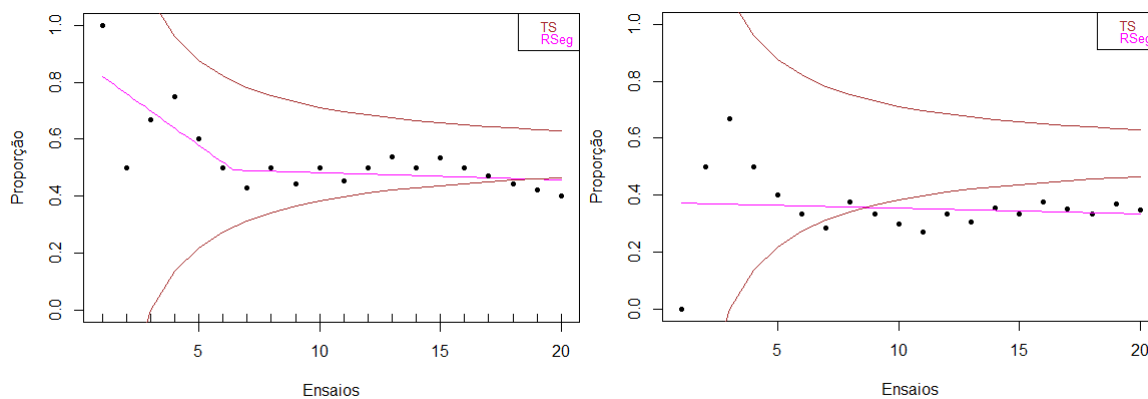
Figura 71 – Gráfico das taxas de rejeição de H_0 para o aprendizado não linear em 95 e 100 ensaios.



Fonte: Da autora.

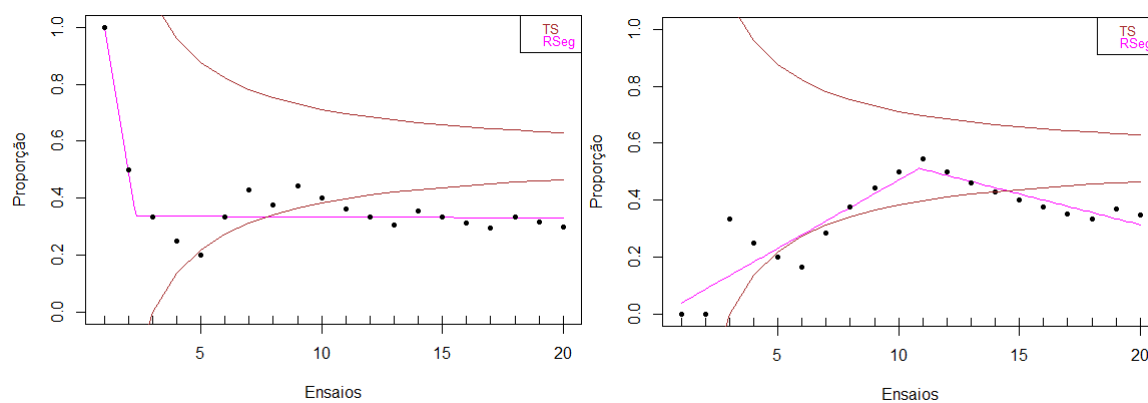
APÊNDICE C - Gráficos obtidos para análise de regressão segmentada e aplicação do teste sequencial em dados empíricos de análise sensorial

Figura 72 – Gráficos obtidos, respectivamente, para os candidatos 1 e 2.



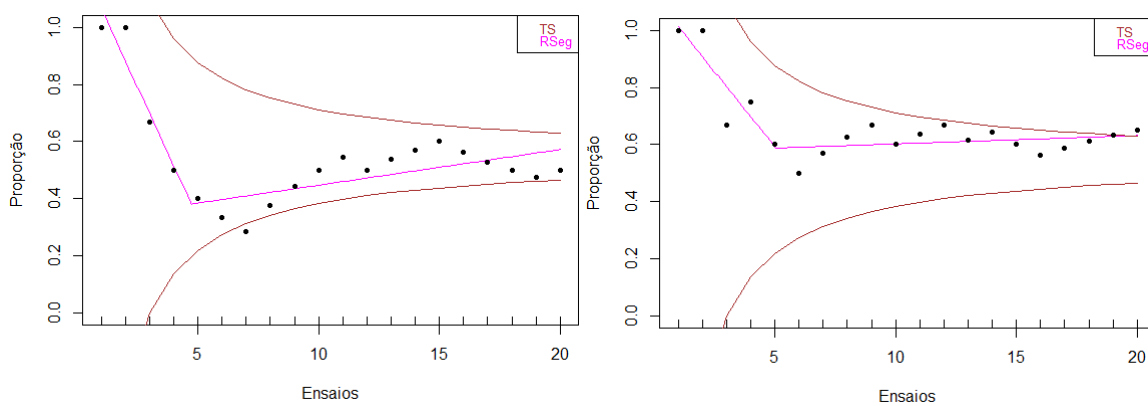
Fonte: Da autora.

Figura 73 – Gráficos obtidos, respectivamente, para os candidatos 3 e 4.



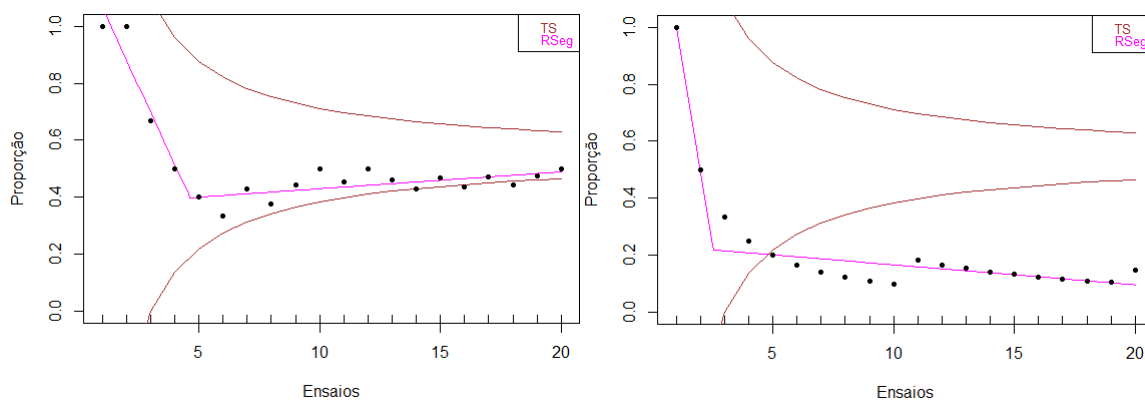
Fonte: Da autora.

Figura 74 – Gráficos obtidos, respectivamente, para os candidatos 5 e 6.



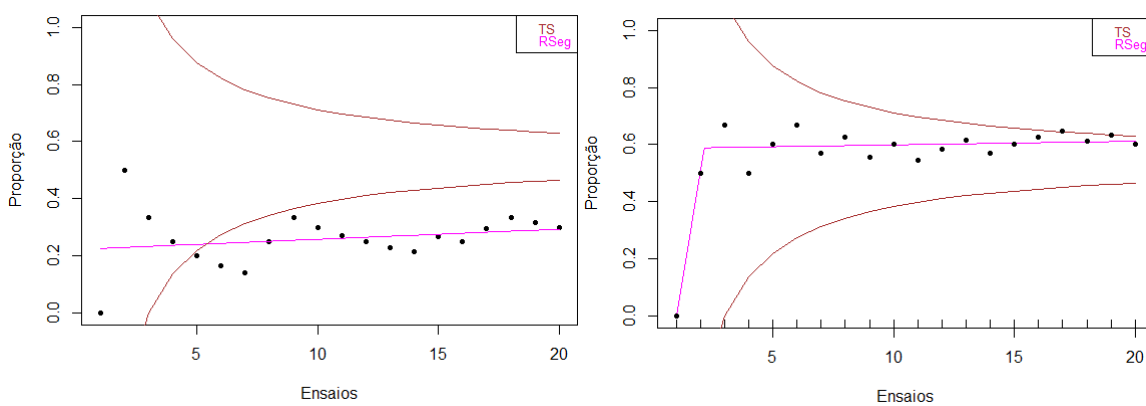
Fonte: Da autora.

Figura 75 – Gráficos obtidos, respectivamente, para os candidatos 7 e 8.



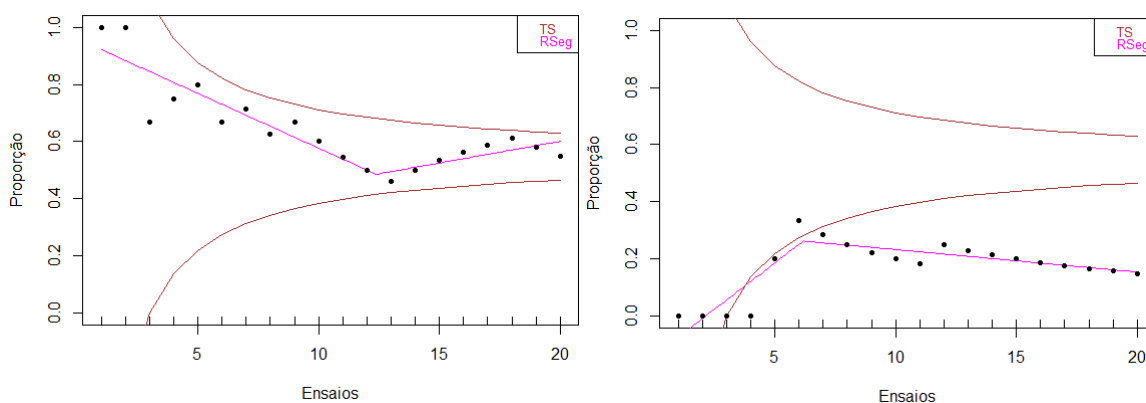
Fonte: Da autora.

Figura 76 – Gráficos obtidos, respectivamente, para os candidatos 9 e 10.



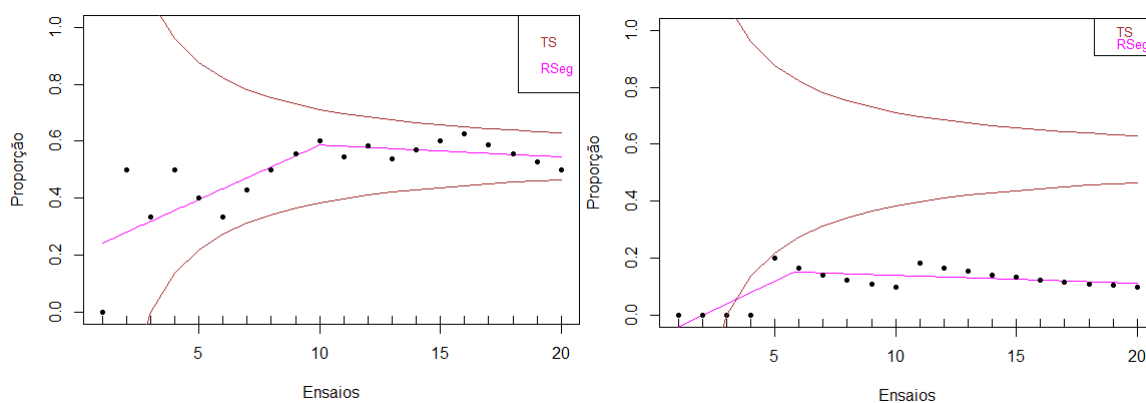
Fonte: Da autora.

Figura 77 – Gráficos obtidos, respectivamente, para os candidatos 11 e 12.



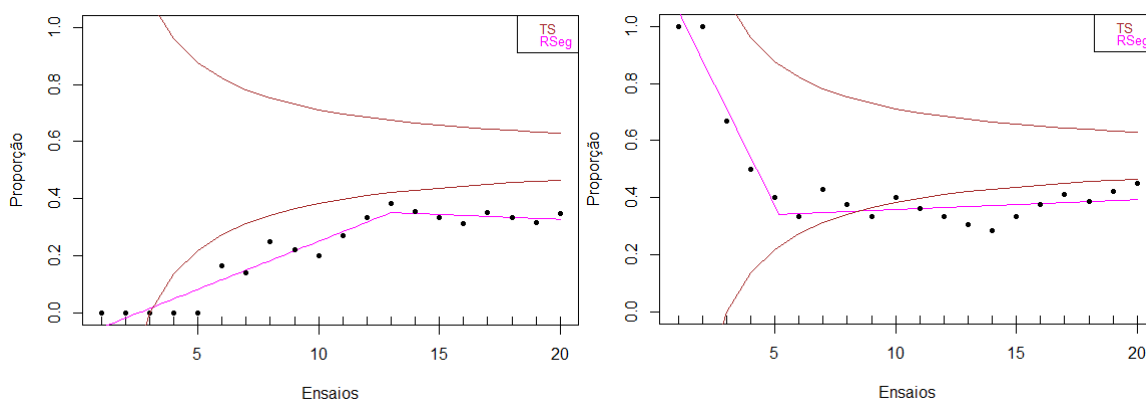
Fonte: Da autora.

Figura 78 – Gráficos obtidos, respectivamente, para os candidatos 13 e 14.



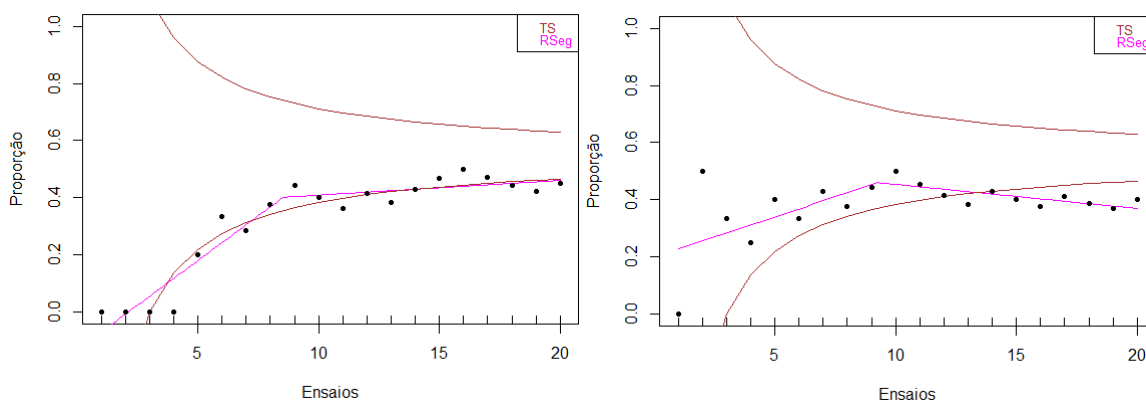
Fonte: Da autora.

Figura 79 – Gráficos obtidos, respectivamente, para os candidatos 15 e 16.



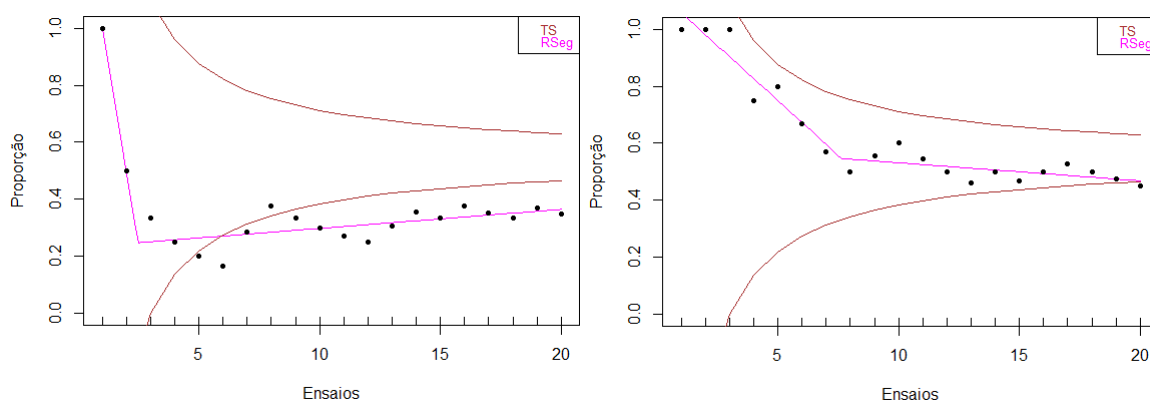
Fonte: Da autora.

Figura 80 – Gráficos obtidos, respectivamente, para os candidatos 17 e 18.



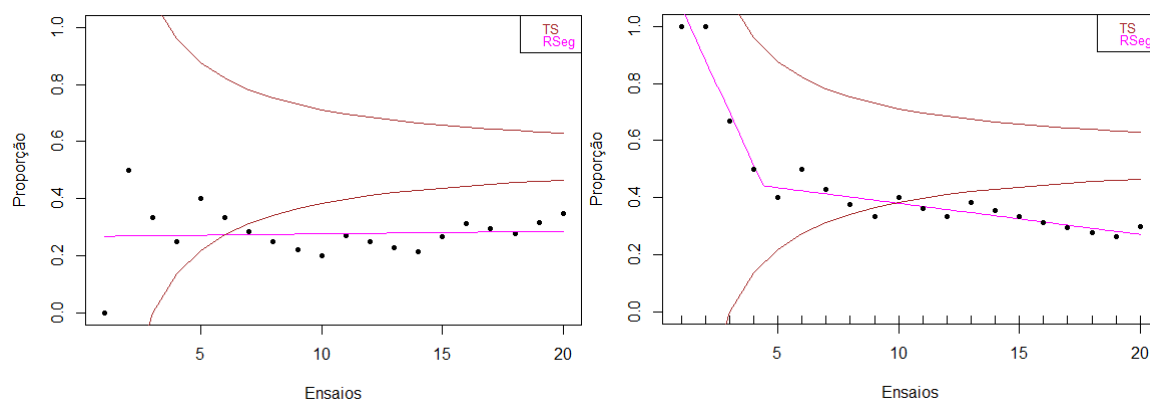
Fonte: Da autora.

Figura 81 – Gráficos obtidos, respectivamente, para os candidatos 19 e 20.



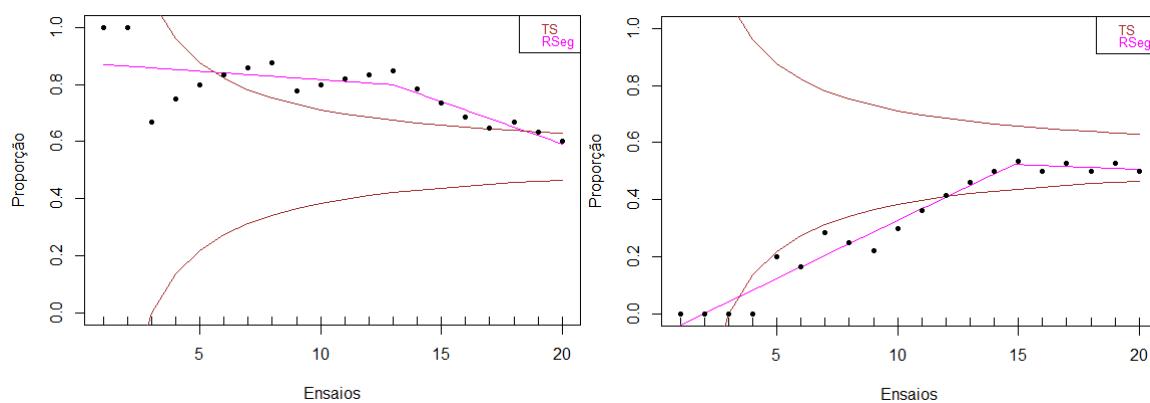
Fonte: Da autora.

Figura 82 – Gráficos obtidos, respectivamente, para os candidatos 21 e 22.



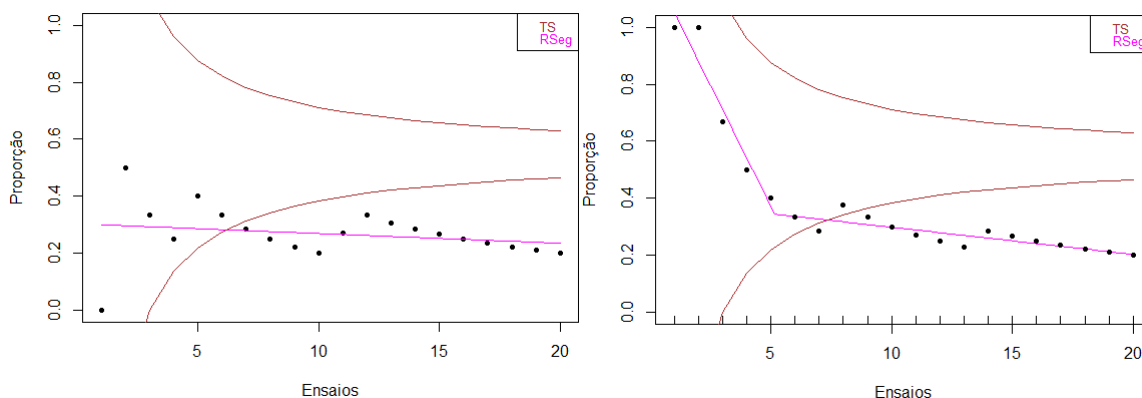
Fonte: Da autora.

Figura 83 – Gráficos obtidos, respectivamente, para os candidatos 23 e 24.



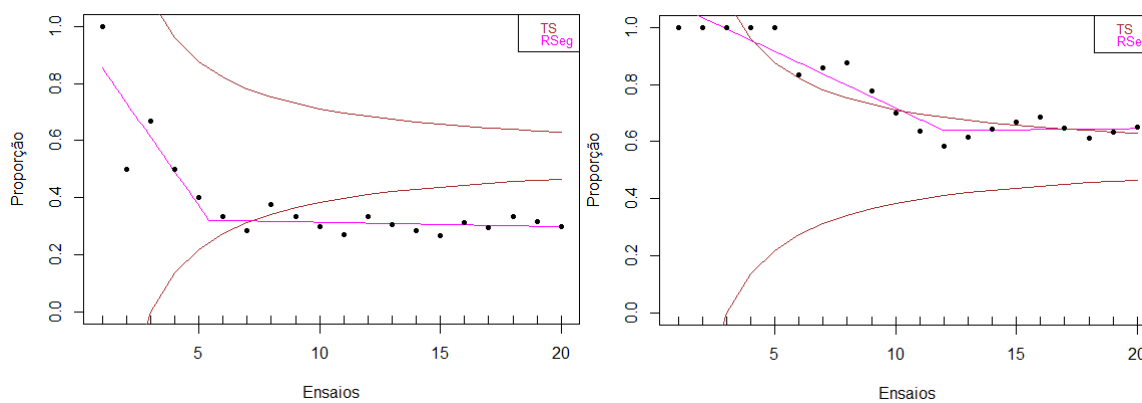
Fonte: Da autora.

Figura 84 – Gráficos obtidos, respectivamente, para os candidatos 25 e 26.



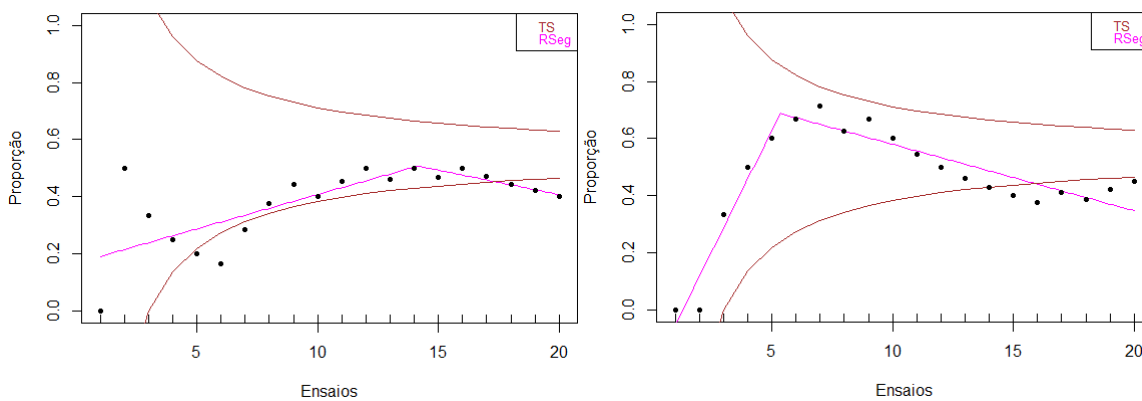
Fonte: Da autora.

Figura 85 – Gráficos obtidos, respectivamente, para os candidatos 27 e 28.



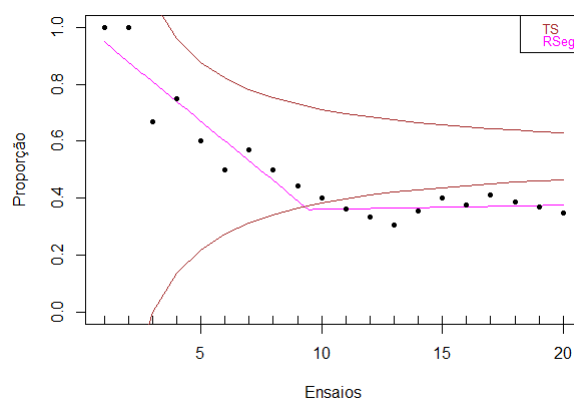
Fonte: Da autora.

Figura 86 – Gráficos obtidos, respectivamente, para os candidatos 29 e 30.



Fonte: Da autora.

Figura 87 – Gráfico obtido para o candidato 31.



Fonte: Da autora.