

Giovana Fumes Ghantous, Judite das Graças Lapa Guimarães, César Gonçalves
de Lima, José Eduardo Corrente

Análise Sensorial e testes estatísticos relacionados

Pirassununga - SP

FACULDADE DE ZOOTECNIA E ENGENHARIA DE ALIMENTOS (FZEA)
UNIVERSIDADE DE SÃO PAULO (USP)

2023

Sumário

Referências	10
1 Introdução	10
1.1 Histórico	10
1.2 Os sentidos humanos como instrumentos	11
1.3 Métodos	14
1.3.1 Principais métodos em análise sensorial	16
1.4 Ambiente de realização dos testes	17
1.5 Preparo das amostras	19
1.6 Avaliadores	20
2 Métodos Discriminativos	22
2.1 Introdução	22
2.2 Testes paramétricos	22
2.2.1 Teste de comparação múltipla e de diferença do controle	22
2.2.1.1 Teste F da Análise de Variância (ANOVA)	23
2.2.1.2 Teste de Tukey	25
2.2.1.3 Teste de Dunnett	26
2.2.2 Exemplos de aplicação	26
2.2.2.1 Resolução	37
2.2.2.2 Exemplo 3 - Teste de diferença do controle com análises repetidas	41
2.3 Teste não paramétricos	46
2.3.1 Teste triangular	46
2.3.1.1 Teste triangular não direcional	46
2.3.1.2 Exemplo	49
2.3.1.3 Resolução	50
2.3.1.4 anexo angular direcional	50
2.3.2 Exemplo de aplicação	51
2.3.3 Resolução	51
2.3.4 Teste duo-trio	52

2.3.4.1	Exemplo	54
2.3.4.2	Resolução pela abordagem binomial	54
2.3.4.3	Resolução pela abordagem qui-quadrado	56
2.3.4.4	Exemplo	57
2.3.4.5	Resolução	57
2.3.4.6	Resolução pelo método Qui-quadrado	57
2.3.5	Teste comparação pareada	58
2.3.5.1	Exemplo	59
2.3.5.2	Resolução pela abordagem binomial	59
2.3.5.3	Resolução pela abordagem qui-quadrado	60
2.3.5.4	Exemplo	60
2.3.5.5	Resolução	61
2.3.5.6	Exemplo	61
2.3.5.7	Resolução	62
2.3.5.8	Resolução pelo método do qui-quadrado	62
2.3.6	Teste de ordenação	62
2.3.6.1	Exemplo	64
2.3.6.2	Resolução	64
2.3.6.3	Exemplo	67
2.3.6.4	Resolução	69
2.3.6.5	Exemplo	71
2.3.6.6	Resolução	72
2.3.6.7	Exemplo	74
2.3.6.8	Resolução	75
2.4	Teste A ou NÃO A	75
2.4.1	Teste tetraédrico	77
2.4.2	Teste dois em cinco	78
2.5	Teorias sobre os testes discriminativos	78
2.5.1	Teoria da detecção de sinal	79
2.5.2	R-Index	79
3	MÉTODOS DESCRIPTIVOS	81
3.1	Introdução	81
3.2	Teste de escala	81
3.2.1	Estrutura	82
3.2.2	Posição	84
3.2.3	Polaridade	84
3.2.4	Quantidade de atributos avaliados	85

SUMÁRIO

3.2.5	Tipo de avaliação	85
3.3	Perfil de sabor	87
3.4	Perfil de textura	88
3.5	Perfil de escolha livre	93
3.5.1	Curiosidade - origem do termo Procrustes	94
3.6	Análise descritiva quantitativa (ADQ)	94
3.6.1	Técnicas estatísticas utilizadas na ADQ	96
3.6.2	Check-All-That-Apply (CATA)	97
3.6.3	Perfil Flash	99
4	MÉTODOS AFETIVOS	101
4.1	Introdução	101
4.1.1	Locais de Aplicação	102
4.2	Classificação dos testes afetivos	103
4.2.0.1	Testes afetivos qualitativos	103
4.2.0.2	Testes afetivos quantitativos	104
4.3	Testes de preferência	104
4.3.1	Testes de comparação pareada	104
4.3.1.1	Exemplo	105
4.3.2	Teste de Ordenação	106
4.3.2.1	Exemplo	106
4.3.3	Teste de Comparação Múltipla	107
4.3.3.1	Exemplo	108
4.4	Teste de aceitação	108
4.4.1	Escala Hedônica	108
4.4.2	Escala FACT	110
4.4.3	Escala JAR	111
4.5	Associação de Palavras na Análise Sensorial	114
4.5.1	Etapas da Associação de Palavras	114
4.5.2	Exemplo de aplicação	115
4.6	Características Não Sensoriais	115
4.6.1	Análise estatística	117
4.6.1.1	Teste t para amostras pareadas	117
4.6.1.2	Análise de risco (relative risk)	118
4.7	Pesquisa de mercado e a indústria de alimentos	119
4.7.1	Tipos de pesquisa	120
4.7.2	Coleta de dados	121
4.7.3	Níveis de mensuração	122

4.7.4	Estilo das questões	123
4.7.5	Considerações sobre a amostragem	124
4.7.5.1	Determinação do número da amostra	124
4.7.6	Exemplo de cálculo	125
4.8	Análise dos Dados	126
4.9	Mapa de Preferência	127
4.9.1	Análise de componentes principais (PCA)	128
4.9.2	Mapa de preferência interno	129
4.10	Redes Neurais Artificiais	130
4.10.1	Organização de um RNA	131
4.10.2	Etapas para resolução de problemas utilizando RNAs	135
4.11	Limiares Afetivos	140
4.11.1	Limiar de rejeição pelo consumidor (LRC)	140
4.11.1.1	Cálculo do limiar de rejeição do consumidor	140
4.11.2	Limiares hedônicos	141
4.11.2.1	Cálculo do limiar de aceitação comprometida	143
4.11.2.2	Cálculo do limiar de rejeição hedônica	144
5	Referências	146
6	ANEXOS	152
6.0.1	ANEXO A	153
6.0.2	ANEXO B	154
6.0.3	ANEXO C	156
6.0.4	ANEXO D	156
6.0.5	ANEXO E	157
6.1	ANEXO F	158
7	APÊNDICE PROGRAMAÇÕES	159
7.1	Problema 3 - Teste de comparação múltipla	159
7.2	Problema 15	161
7.3	Problema 14	164
7.3.1	Análise estatística dos resultados por Excel	167
7.3.2	Teste de tukey e análise gráfica	174

Capítulo 1

Introdução

A análise sensorial é de extrema importância para as indústrias de cosméticos, higiene, têxteis e principalmente de produtos alimentícios, uma vez que ela possui diversas finalidades, como desenvolvimento de novos produtos, controle de qualidade do produto, diminuição de desperdícios, estimativa do tempo de vida de prateleira, entre muitas outras. Por meio dela, é possível analisar, medir, evocar e interpretar as diferentes reações que os produtos podem causar nos seres humanos ao estimular os seus sentidos [2] [5]

A aplicação das metodologias de análise sensorial requer a utilização da estatística nas etapas de planejamento de experimentos, coleta, organização e análise de dados, possibilitando a interpretação dos mesmos.

Diante disso, este livro tem como principal objetivo explicar como a análise sensorial e a estatística estão relacionadas, descrevendo as principais metodologias usadas para coleta de dados da primeira e as técnicas para análise de tais dados pela segunda.

1.1 Histórico

O estudo sobre a análise sensorial dos alimentos iniciou-se no período antes de Cristo (A.C), no qual o conceito de classificação dos produtos alimentícios básicos, como peixe, vinho, trigo, entre outros, surge, mas foi em 1890, no início da industrialização, que o conceito de análise sensorial passou a ser desenvolvido, sendo aprimorado na segunda metade do século XX, muito influenciado por fatores relacionados à Segunda Guerra Mundial (1939 - 1945) quando diversos outros estudos passaram a ser realizados com o objetivo de desenvolverem alimentos mais nutritivos, e que tivessem uma boa aceitabilidade pelos soldados [12].

Segundo MINIM[5] e DUTCOSKY[2], o conceito de análise sensorial e sua aplicação nos alimentos foi evoluindo durante décadas, e pode ser dividida em quatro principais fases:

1º fase - Antes de 1940: chamada de período artesanal, na qual o proprietário de uma empresa era a pessoa encarregada de determinar se a qualidade sensorial de um produto estava

adequada ou não.

2° fase - Entre 1940 e 1950: ocorreu a expansão da indústria de alimentos. Neste período, a análise sensorial passou a ser utilizada dentro das fábricas por técnicos químicos e farmacêuticos, para garantir um maior controle na produção, além disso, passou-se a introduzir os conceitos de controle de processos e produto final, dando enfoque nos métodos químicos e instrumentais padronizados.

3° fase - Entre 1950 e 1970: a terceira fase foi marcada pelo entendimento de que o ser humano é imprescindível como ferramenta para a análise da qualidade sensorial. Com essa nova abordagem, vários avanços foram desenvolvidos, como o *Círculo de Kramer* e os sentidos humanos (órgãos sensoriais), ambos fundamentais para qualificação sensorial, sendo que o Círculo de Kramer definiu as propriedades/características de um alimento (atributos primários). Ademais, pesquisadores compreenderam que os seres humanos possuem habilidades naturais de diferenciação, comparação, quantificação dos atributos sensoriais e de obter dados confiáveis. Para que fosse possível utilizar tais habilidades, notaram a necessidade de padronizar o tratamento estatístico de dados, normalizando as condições e a forma que as perguntas eram realizadas. Também foram desenvolvidas pesquisas básicas sobre os estímulos sensoriais, ou seja, como sensações sensoriais são provocadas e como o ser humano as expressa. O final dessa fase destacou-se pela definição de análise sensorial como instrumento científico, além de utilizar métodos estatísticos para analisar e correlacionar medidas sensoriais e instrumentais

4° fase - Após 1970: marcada pelo entendimento do que é a qualidade sensorial de um produto, que nada mais é do que o resultado das interações entre produto e ser humano, ou seja, a sua percepção sobre ele. Portanto, para determinar a qualidade sensorial de um alimento, é necessário compreender o quê os estímulos sensoriais causam no ser humano, além de levar em consideração as condições psicológicas, fisiológicas e sociológicas do participante da análise sensorial, pois estão diretamente relacionadas com a avaliação do produto.

1.2 Os sentidos humanos como instrumentos

Olfato, audição, visão, paladar e tato são os cinco sentidos humanos, os quais são fundamentais para a realização de uma análise sensorial, pois por meio deles pode-se receber e interpretar uma vasta quantidade de estímulos, como gosto, odor, aparência e textura.

O sentido da visão se destaca, em relação aos outros, pois é através dele que ocorre o primeiro contato entre um indivíduo e o alimento, portanto, sua aparência pode influenciar indiretamente ou diretamente na decisão de aceitação ou rejeição do produto. Todas as características visuais do produto, como tamanho, cor, textura e forma são analisados primeiramente

por esse sentido, e essas informações tendem a induzir o consumidor a especulações com relação ao sabor, odor e textura do alimento[13] [8].

O sentido do olfato ocorre porque o nariz possui receptores sensoriais capazes de identificar substâncias voláteis que penetrem na cavidade nasal e atinjam o bulbo olfatório. A percepção do odor e aroma de um alimento ou produto está diretamente relacionada com a percepção do sabor, como veremos mais adiante. Existe uma diferença sutil entre os termos odor e aroma. O odor é característico de cada alimento, e é formado por substâncias voláteis que são inaladas antes do alimento ser inserido na boca. Em contrapartida, o aroma de um alimento só é liberado quando ele já está dentro da boca, e, posteriormente, encaminhado até os receptores do sistema olfativo. Vale ressaltar que o termo fragânciá é utilizado para os cheiros de cosméticos[12] [7].

A percepção do olfato está sujeito a fenômenos fisiológicos que podem afetar negativamente uma análise sensorial, tais como: adaptação, anosmia e hiposmia. A adaptação é um fenômeno muito constante nos seres humanos, ele causa a diminuição parcial ou total da capacidade de diferenciação consciente de odores, e na maioria das vezes é ocasionada pela repetição de um estímulo, acarretando na saturação dos receptores nasais. A anosmia é uma anormalidade rara, no qual a pessoa não possui qualquer tipo de percepção olfativa, e afeta diretamente a avaliação do odor e do sabor de um alimento. Já a hiposmia é recorrente em pessoas com idade mais avançada, ou que possuem uma inflamação nas narinas, e também é capaz de provocar uma diminuição na percepção olfativa [2].

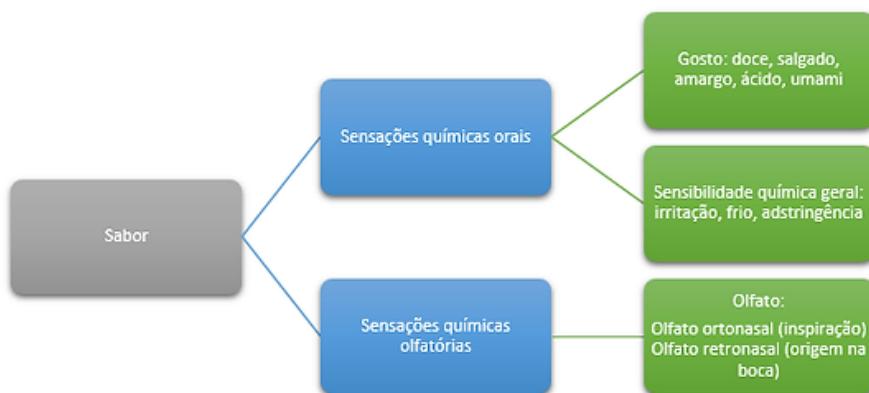
O paladar exerce um dos papéis mais importantes para a realização de uma análise sensorial, pois a língua humana possui diversos receptores que são capazes de identificar os cinco gostos, sendo eles: doce, salgado, ácido, amargo e umami (identificado como glutamato monossódio, conhecido popularmente como realçador de sabor). Entretanto, a língua não é a única estrutura responsável pela percepção dos gostos de alimentos, uma vez que o palato, as bochechas e o esôfago também possuem células receptoras que influenciam nessa percepção.

Além disso, a percepção do gosto pode variar de pessoa para pessoa, pois é influenciada pelas condições momentâneas, como sensação de fome ou saciedade, estado do humor e até mesmo por condições fisiológicas, como gripes ou resfriados.

O paladar também pode possuir anomalias como ageusia, quando o indivíduo não possui sensibilidade perante ao gosto, hipogeusia, em que ocorre a diminuição da percepção do gosto e disgeusia, no qual ocorre a distorção da percepção do gosto.

É preciso entender com clareza a diferença entre gosto e sabor dos alimentos. O gosto é a percepção do alimento sentida principalmente pela língua, como citado acima. Já o sabor é a união das percepções sentidas pelo paladar e olfato, uma vez que, durante a mastigação são liberadas algumas substâncias voláteis que geram sensações como formigamento e refrescância. Tais substâncias voláteis influenciam diretamente no sentido do olfato, enquanto as não voláteis, como a adstringência, auxiliam no paladar (Figura 1.2.1).

Figura 1.2.1: Classificação anatômica do sabor



Fonte: Adaptado de DUTCOSKY, 2019.

É importante ressaltar que o olfato exerce também uma grande influência na percepção do sabor de um alimento. Um exemplo simples de compreender é pensar em uma pessoa com resfriado, que possui congestão nasal, nesta condição a captação dos aromas liberados pelos alimentos nos receptores olfativos estará comprometida, e portanto a pessoa pode ter a sensação de que a comida está sem "gosto", apesar dos receptores do paladar estarem disponíveis para identificar os gostos normalmente[2].

A audição também possui uma contribuição importante para a análise sensorial, pois através dela é possível perceber os sons característicos da textura de cada alimento ao serem manuseados, mastigados e deglutidos. Isso ocorre porque a percepção auditiva possui um sistema complexo, onde as vibrações do ar entram no canal auditivo, causam vibrações no tímpano que são transmitidas para o ouvido médio, ocasionando movimentos no fluido em locais sensíveis do ouvido interno, e que são transmitidas por meio de impulsos nervosos para o cérebro, onde são interpretadas. É importante destacar que o reconhecimento do som de algum alimento é baseado unicamente nas experiências prévias do avaliador.

O quinto e último sentido é o tato, sentido que é possibilitado pela presença de diversas terminações nervosas presentes nas camadas da pele (epiderme, derme e tecido subcutâneo), e permitem a sensação de sensibilidade ao toque. Por meio do sentido do tato somos capazes de avaliar textura, peso, forma, temperatura, viscosidade e consistência dos alimentos. A percepção do tato se dá principalmente pela pele e mucosa que recobrem, respectivamente, as mãos e boca, sendo que a segunda engloba os lábios, língua, bochecha, palato, gengiva e dentes. É necessário compreender a importância dos dentes na análise sensorial de alimentos, pois o nervo dental possui várias ramificações responsáveis pela sensibilidade às pressões causadas pela mastigação, influenciando na percepção da textura. O tato e a audição se complementam na percepção da textura de um alimento, pois ao consumir um produto e ouvir os sons

característicos que ele produz, é gerada uma sensação de satisfação durante o consumo [2].

1.3 Métodos

Para a realização de uma análise sensorial confiável, ou seja, que possua o menor número de erros experimentais, é necessário compreender o que se deseja atingir com a análise: comparar amostras, descobrir se houve ou não alteração em alguma delas, entre outros. Para entender melhor os diferentes métodos utilizados, é importante saber o significado dos termos qualitativo, quantitativo e inferência.

Uma análise sensorial qualitativa busca analisar as qualidades dos alimentos, como: forma, cor, tamanho, textura, aroma e odor, gosto e sabor, propriedades mecânicas (viscosidade, elasticidade e dureza), interações entre pedaços (solto ou aglomerado), propriedades geométricas (arenoso, granuloso ou fibroso) e absorção, presença ou liberação de umidade e gordura do produto. Já uma análise quantitativa é utilizada para quantificar a intensidade de cada um dos atributos qualitativos no alimento[2].

No processo de inferência estatística são utilizados os dados resultantes de testes qualitativos e quantitativos, realizados em amostras. Utilizando ferramentas estatísticas adequadas, as conclusões obtidas nesses ensaios podem ser inferidas/generalizadas para todas as situações similares àquela que foi estudada. Dentre as ferramentas estatísticas usadas na análise sensorial, a mais comum é o teste de hipótese.

O teste de hipóteses é uma ferramenta que permite validar ou refutar uma hipótese formulada sobre algum parâmetro (populacional) de interesse, com base em resultados obtidos em amostras. São formuladas duas hipótese: H_0 , chamada de hipótese da nulidade e H_a , chamada de hipótese alternativa. A hipótese H_0 envolve, geralmente, uma igualdade e H_a , uma desigualdade (maior que, menor que ou diferente de) sobre o(s) parâmetro(s) de interesse(s).

Por exemplo: para testar que a média do produto A é superior à média do produto B, formulam-se as hipóteses:

$$H_0 : \text{média de A} = \text{média de B}$$

$$H_a : \text{média de A} > \text{média de B}$$

Uma estatística de teste adequada deve fornecer informações suficientes para rejeitarmos ou não a hipótese H_0 . Outras hipóteses alternativas poderiam ser de interesse, e isso caracteriza um teste ser uni ou bilateral (Tabela 1.3.1).

Tabela 1.3.1: Tipos de hipóteses

Hipótese nula (H_0)	Hipótese alternativa (H_a)
$\mu_a = \mu_b$	$\mu_a \neq \mu_b$ hipótese bilateral
	$\mu_a < \mu_b$ hipótese unilateral
	$\mu_a > \mu_b$ hipótese unilateral
	μ_a : média da população ou produto A μ_b : média da população ou produto B

Fonte: DUTCOSKY [2].

Em toda tomada de decisão estão envolvidos dois tipos de erros denominados genericamente de erros tipo I e tipo II. A Tabela 1.3.2 abaixo apresenta um resumo do mecanismo de ocorrência dos dois erros em um teste de hipóteses.

Tabela 1.3.2: Relação entre os erros tipo I e tipo II no processo de decisão

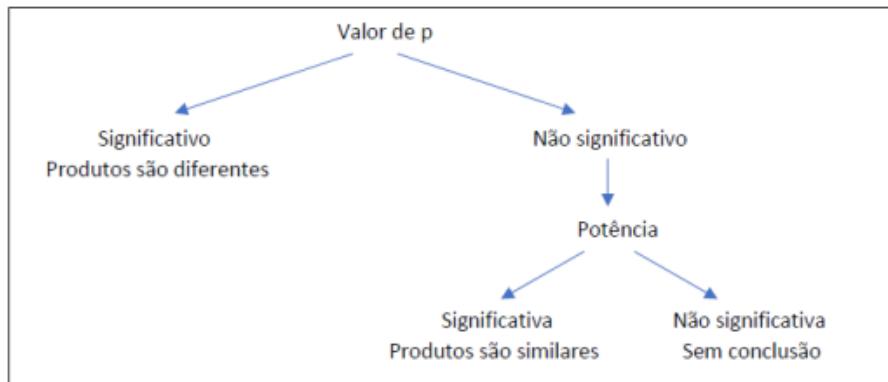
	Aceitar H_0	Rejeitar H_0
H_0 verdadeira	Decisão correta	Erro do tipo I (α)
H_0 falsa (ou H_a)	Erro do tipo II (β)	Decisão correta

Fonte: DUTCOSKY [2].

O erro tipo I, também denotado de nível de significância do teste ou pela letra grega alfa (α), ocorre quando se rejeita erroneamente a hipótese H_0 , ou seja, rejeita-se de modo equivocado a ideia de que as amostras são iguais. Já o erro tipo II, denotada pela letra grega beta (β), ocorre quando se aceita erroneamente a hipótese H_0 . Uma boa alternativa para diminuir a chance de cometer algum desses erros (α e β) consiste em utilizar amostra(s) de tamanho(s) grande(s), ou seja, que representem bem a população.

A probabilidade de tomar uma decisão correta é chamada de potência (ou poder) de um teste, é definida como $1 - \beta$, que representa a probabilidade de rejeitar H_0 quando ele realmente for falso. O poder de um teste depende do nível de significância (α) adotado, da diferença entre o real valor do parâmetro e do valor fixado em H_0 , da variância da população e também do tamanho da amostra utilizada. Um exemplo da tomada de decisões sobre tais erros, pode ser visto na avaliação de testes discriminativos da Figura 1.3.1.

Figura 1.3.1: Decisão de acordo com objetivos de análise dos testes discriminativos



Fonte: DUTCOSKY [2].

Na maioria dos testes, o nível de significância (α) costuma ser fixo no valor de 5% ou 1%. Os *softwares* estatísticos auxiliam na tomada de decisão sobre a rejeição ou não de H_0 apresentando o p-valor (denotado por Valor de p na Figura 1.3.1) ou nível descritivo do teste, que corresponde à probabilidade de ocorrência de valores mais extremos ao obtido com os dados amostrais. Se p-valor for superior ao nível de significância adotado, conclui-se que é alta a probabilidade de se rejeitar erroneamente H_0 , indicando que ela não deve ser rejeitada. Quando p-valor é inferior ao valor de alfa, conclui-se que é pequena a probabilidade de se rejeitar erroneamente H_0 , indicando que ela deve ser rejeitada.

1.3.1 Principais métodos em análise sensorial

Existem vários métodos que podem ser aplicados em uma análise sensorial de acordo com o objetivo que se deseja alcançar, os quais englobam em três principais categorias de métodos: discriminativos, descritivos e afetivos.

Os métodos discriminativos têm como objetivo verificar se existe ou não diferenças qualitativas (cor, sabor, textura, etc) e quantitativas (intensidade de cada qualidade) entre as amostras. Esses métodos são extremamente importantes e utilizados nas análises sensoriais para a realização de seleção, redução ou substituição de ingredientes, controle de qualidade, diminuição de custos, dentre outros. Com eles são possíveis os estudos para avaliar as semelhanças e diferenças entre dois ou mais produtos[2].

Os métodos discriminativos englobam dois tipos de testes:

- Não direcional: também chamado de teste diferenciação geral, é utilizado quando não se sabe qual a natureza da diferença, ou seja, quando se desconhece qual ou quais atributos do alimento podem ser afetados. Alguns exemplos são os testes triangular e duo-trio.

- Direcional: também chamado de teste diferença de atributos, é usado para comparar um atributo sensorial pré determinado, para saber se há ou não diferença ou similaridade nas amostras com relação a esse atributo. Alguns exemplos são os testes de comparação pareada e teste triangular direcional.

Já os métodos descritivos têm o objetivo de caracterizar as propriedades sensoriais de um alimento e descrever suas características de modo qualitativo (analisando cor, sabor, textura, cheiro entre outros) e quantitativo (analisando a intensidade de cada fator qualitativo), assim, é possível levantar as propriedades sensoriais do produto de modo detalhado. Devido ao nível técnico empregado nesses métodos, faz-se necessário a utilização de avaliadores treinados, que passam por etapas de recrutamento, seleção e capacitação, consequentemente, são métodos que demandam mais tempo e dinheiro quando comparados com outros. Exemplos de técnicas aplicadas nesse método são: o perfil de sabor, o método *Spectrum*, o perfil de textura, a análise descritiva quantitativa e o perfil livre [2].

Por fim, os métodos afetivos, também conhecidos como métodos subjetivos ou testes do consumidor, têm como objetivo avaliar o nível de aceitabilidade ou preferência de um produto pelo consumidor. É importante compreender que existe diferença entre os conceitos de preferência e aceitabilidade, pois o primeiro termo avalia os aspectos sensoriais que agregam ou não ao consumidor, já o segundo termo vai além, pois engloba questões que influenciam o consumidor no momento de compra (como por exemplo o preço do produto). Nesses métodos, não é necessário o uso de avaliadores treinados, eles são escolhidos conforme o público alvo determinado para o produto, levando em consideração fatores como: idade, sexo, frequência de consumo, localização geográfica, entre outros[13].

Nos métodos afetivos quantitativos, são analisadas respostas fornecidas por um grupo de consumidores que representem o público alvo, para determinar a aceitabilidade global do produto, já nos métodos qualitativos, o principal objetivo é buscar respostas a fenômenos específicos dos produtos, utilizando para isso uma maior interação entre os avaliadores e o entrevistador. Alguns dos testes afetivos mais utilizados são: ordenação de preferência, comparação pareada, *focus group*, comparações múltiplas de preferência, escalas hedônicas, escalas JAR (*Just about right*), entre outros[2].

1.4 Ambiente de realização dos testes

Para que os erros experimentais sejam minimizados durante a realização dos testes sensoriais, o ambiente em que for realizado os estudos deve seguir os requisitos normatizados pela *International Organization for Standardization* (ISO) 8589:2007, *American Society for Testing and Materials* (ASTM) e Associação Brasileira de Normas Técnicas (ABNT).

De acordo com as normas institucionais citadas acima, é necessário que o ambiente seja inodoro (podendo ser utilizado exaustores para evitar a contaminação do ambiente por odores externos), também deve ser calmo, silencioso, de fácil acesso, livre de aglomerações e confortável. Além disso, é preciso que esses ambientes possuam cabines individuais de dimensões adequadas segundo a legislação, para trazer maior conforto e individualidade aos avaliadores, prevenindo que interajam entre si ou com outras distrações. Também devem possuir paredes e móveis de tons neutros, iluminadas por luzes claras e sem sombras, dando preferência à luzes naturais, mas em alguns casos podem ser utilizadas luzes coloridas a fim de mascarar a cor das amostras, quando esse aspecto não está sendo avaliado[12] (Figura 1.4.1).

Figura 1.4.1: Exemplo de ambiente ideal

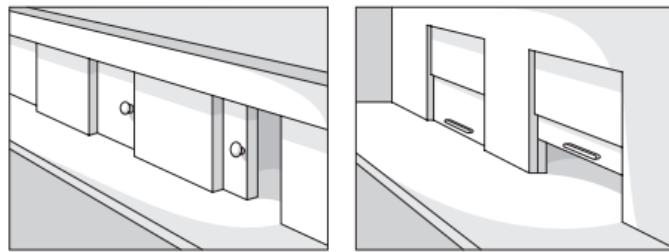


Fonte: Autoria própria.

Segundo a ISO 8589:2007, o ambiente em que são realizadas as análises sensoriais deve possuir uma sala de estocagem de amostras, outra sala para estocagem de materiais, um banheiro, um escritório, um ambiente de preparação de amostras separado da área dos avaliadores, cabines para avaliações individuais e mesas redondas para avaliações em grupo. As mesas utilizadas em grupo servem para a realização de reuniões, discussões, análises de perfis descritivos, padronização de procedimentos, entre outros.

Para que ocorra a passagem das amostras do local de preparo para as cabines individuais, devem ser projetadas aberturas com portas localizadas de tal forma que os participantes da análise sensorial não consigam observar a preparação e/ou codificação das amostras do experimento. Para isso, costuma-se utilizar pequenas portas deslizantes silenciosas com um sistema de sinalizador acoplado, com um interruptor interno, que quando ativado pelo participante ativará uma luz para a parte externa da cabine, onde a equipe de preparação identificará que existe um julgador na cabine, e também que ele terminou a avaliação de uma amostra[2] (Figura 1.4.2).

Figura 1.4.2: Exemplos de portinhais deslizantes para cabines de análise sensorial



Fonte: ISO 8589:2007.

1.5 Preparo das amostras

A preparação das amostras é realizada utilizando como base a padronização de quantidades e temperaturas ideias para a degustação delas pelo participante. Com relação a quantidade, as amostras deve ser porcionadas de modo que haja uma agilidade em seu preparo e uma mínima manipulação, para evitar possíveis contaminações. A atenção com as temperaturas das amostras deve ser reforçada, pois todas elas precisam ser entregues aos participantes com uma mesma temperatura, portanto, são realizados monitoramentos em todas as amostras antes de servi-las. Assim, as amostras devem estar uniformes entre si, tanto na questão de temperatura e quantidade, quanto nos vários outros fatores, como tamanho, tipo de corte, etc[2].

Cada produto em uma análise sensorial deve ser avaliado conforme os principais objetivos dos testes, portanto, os alimentos devem ser preparados conforme a necessidade de cada teste em particular, por exemplo, um produto deve ser fornecido na forma em que normalmente é consumido em casos de testes de aceitabilidade e preferência. Vale ressaltar que todas as amostras devem ser codificadas para evitar a escolha por tendência, ou associação a alguma empresa ou marca. Para a codificação geralmente são utilizados números com três dígitos. Ademais, a organização dos testes é essencial para que todas as amostras sejam fornecidas aos avaliadores em ordem casual, mas a ordem deve ser aleatória de maneira que todos os avaliadores realizem todos os testes a mesma quantidade de vezes.

Outra preocupação no preparo e distribuição das amostras é a fadiga sensorial que pode ocorrer no julgador, portanto, são considerados alguns aspectos para definir a quantidade máxima de amostras fornecidas por sessão. Os aspectos são: experiência do participante, quantidade de tempo e produto disponível para a realização da análise sensorial, tipo de teste que está sendo aplicado, tipo de produto, complexidade e intensidade da propriedade sensorial que está sendo avaliada[2].

Além disso, a fim de evitar o cansaço dos participantes, as análises devem ser realizadas

com no máximo 20 minutos de duração. Outra medida utilizada para evitar o cansaço e a interação entre amostras, é fornecer uma bandeja com água e biscoito de água e sal, para que o avaliador as consuma entre as amostras, evitando que os resquícios na boca interfiram na avaliação. Outro fator importante, é a recomendação de que o consumo das amostras deve ocorrer duas horas antes ou depois de uma refeição, para evitar interferência na percepção das características sensoriais das amostras.

1.6 Avaliadores

Os avaliadores, também chamados de julgadores, são os participantes que realizam a análise sensorial, e são fundamentais para a sua execução, pois através dos cinco sentidos humanos eles são capazes de avaliar as propriedades sensoriais de um determinado produto. Logo, a escolha adequada dos avaliadores é essencial para o bom desenvolvimento das análises, pois cada tipo de teste pode exigir um perfil de avaliador diferente.

Chamado de avaliador iniciado ou inexperiente, é aquele participante que não possui um treinamento e tem pouca experiência com análises sensoriais. Esse tipo de avaliador é muito utilizado em testes que possuem como objetivo descobrir a preferência do consumidor, pois assim, é possível examinar o comportamento do público alvo de determinado produto[2].

Já os avaliadores profissionais ou *experts*, são aqueles participantes que possuem uma grande experiência nas análises sensoriais, de modo que possuem habilidades e consistências nos resultados das avaliações. Esse grupo de avaliadores é muito utilizado em testes que visam estabelecer a qualidade das amostras, incluindo a quantificação da intensidade de diferentes atributos, portanto eles são treinados para que as respostas sejam ainda mais acertivas[2].

Em todas as análises sensoriais são realizados recrutamentos, os quais utilizam como base para a seleção dos avaliadores características, como sexo, faixa etária, classe social, fatores étnicos, localização geográfica, frequência de consumo, entre outros. Primeiramente, são recrutadas o dobro de pessoas desejado, sendo elas pertencentes à empresa que realiza análise (apenas aqueles que não tiveram nenhum contato com o produto que se deseja analisar), ou pessoas externas. Em seguida ocorre a etapa de entrevista, onde é esclarecido ao participante quais são os objetivos da análise, a importância de sua participação, o tempo necessário para a realização das análises, além disso, são questionados sobre sua responsabilidade, saúde, o quanto gosta do produto a ser avaliado, sua disposição para o estudo, seu apetite, entre muitos outros questionamentos. Aqueles que tiverem um bom desempenho na etapa de entrevista, serão encaminhados para a realização de testes sensoriais (identificação dos estímulos sensoriais) ou para uma seleção preliminar[2].

Ademais, em testes mais específicos, é realizada a seleção de participantes, sejam eles experientes ou não dependendo do tipo de teste. Essa etapa é importante para que os resultados

sejam o mais confiável possível, a fim de evitar os erros de decisão que podem comprometer a análise sensorial. Existem duas legislações que auxiliam nas etapas de seleção, treinamento, monitoramento e orientações de julgadores, sendo elas a ISO 8585-1 (1993) para os julgadores inexperientes, ISO 8586-2 (2008) para julgadores experientes[2].

De modo geral, costuma-se realizar quatro etapas para a seleção de provadores: entrevista, teste de reconhecimento de odores, teste de identificação de gostos e testes de diferença aplicadas ao produto.

- Entrevista: avalia a saúde do candidato, levando em conta intolerâncias, alergias, condições dentárias entre outros. Além disso, são levadas em consideração a seriedade do participante, sua vontade de realizar o projeto e sua curiosidade intelectual.
- Teste de reconhecimento de odores: costuma ser uma fase classificatória, onde são apresentados ao candidato entre 16 a 20 tipos de aromas comuns (como cebola, vinagre, café, entre outros), onde serão pontuadas a identificação correta, a associação e a caracterização do aroma. A ideia dessa etapa é que os candidatos sejam capazes de identificar ao menos 70% dos aromas que possuam um baixo grau de dificuldade.
- Teste de identificação de gostos: nessa etapa os candidatos recebem pequenas soluções previamente identificadas que correspondem aos gostos básicos (doce, salgado, ácido, amargo e umami). Para cada gosto existe uma quantidade diferente de concentração recomendada pela norma ISO 3972 (1991). Essa é uma etapa eliminatória em que o candidato deve identificar 100% dos gostos.
- Teste de diferença aplicada: essa etapa tem como objetivo avaliar se os candidatos possuem habilidades e percepções diferenciadas de acordo com o alimento avaliado. Por exemplo, um bom avaliador de cachaça não é necessariamente um bom avaliador de queijos.

Em algumas situações, aqueles que forem aprovados serão treinados, com duração de tempo variando de acordo com a complexidade do produto a ser analisado ou do teste a ser realizado. Nesse treinamento, os avaliadores passam a se familiarizar com os procedimentos dos testes, a aprender como observar características específicas, como sabor e textura, para então aprimorarem a memória e a capacidade sensitiva, resultando em análises mais precisas e consistentes.

Capítulo 2

Métodos Discriminativos

2.1 Introdução

De acordo com Dutcosky [2, p. 89], os métodos discriminativos têm como objetivo estabelecer diferenciação qualitativa e/ou quantitativa entre as amostras, desse modo, eles verificam se existe ou não uma diferença entre dois ou mais estímulos muito similares. Esta metodologia é usada para selecionar ingredientes de produtos, averiguar mudanças na formulação, processamento, embalagem, estocagem, redução de custos e estabilidade do controle de qualidade [2].

Nos métodos discriminativos diversos testes são utilizados, tais como: teste de comparação múltipla e de diferença de controle, teste triangular, teste duo-trio, teste de comparação pareada, teste de ordenação, teste tetraédrico e teste dois em cinco.

2.2 Testes paramétricos

Os testes paramétricos são aqueles baseados em modelos probabilísticos, sendo o modelo normal mais usualmente conhecido. A distribuição normal é definida com base nos parâmetros de média e de variância, sendo assim, os testes de hipóteses baseados neste modelo irão testar hipóteses sobre a média e a variabilidade dos dados provenientes de populações distintas. Nesta subseção, serão abordados os testes de comparação múltipla e de diferença do controle.

2.2.1 Teste de comparação múltipla e de diferença do controle

O teste de comparação múltipla e de diferença de controle são realizados com a finalidade de discriminar diferença entre várias amostras e o grau dessa diferença. Em especial, o teste de diferença de controle verifica se existe uma diferença significativa entre várias amostras e uma de referência (padrão) [2].

O teste de comparação múltipla apresenta ao avaliador amostras codificadas sem a presença de uma controle para comparação, julgando a intensidade percebida de determinado atributo por meio de escalas [16].

O teste de diferença de controle apresenta ao avaliador simultaneamente uma amostra referência codificada com as letras C ou R e as demais amostras codificadas, devendo prová-las comparando com a padrão e avaliar o grau de diferença entre as amostras codificadas e a controle usando uma escala, um exemplo é apresentado na figura 2.2.1. A escala pode ir de 1 (extremamente melhor que o controle) até 9 (extremamente pior que o controle) [2].

Professora Judite vai enviar exemplos sobre os dois tipos de testes e escalas para serem incorporadas aqui

Figura 2.2.1: Modelos de fichas para aplicação dos testes de comparação múltipla e de diferença com o controle.

A análise de seus resultados é feita por meio de uma estatística de teste F, proveniente de uma Análise de Variância (ANOVA) e um teste de comparação de médias é feito a posteriori no caso de diferenças significativas entre as amostras, sendo o teste de Dunnet usado para teste de diferença de controle e o teste de Tukey para comparações múltiplas [2].

2.2.1.1 Teste F da Análise de Variância (ANOVA)

A Análise de Variância (ANOVA) é uma técnica estatística de análise de dados experimentais que obtém uma estimativa exata e precisa do erro experimental (denominado resíduo) e verifica se há ou não diferença significativa entre as médias de tratamentos.

Para a realização da ANOVA, algumas pressuposições são assumidas como verdadeiras, como a aditividade dos fatores do estudo, a homogeneidade das variâncias, a independência e normalidade dos resíduos [1].

Cada indivíduo tem uma percepção sensorial e para controlar essa variedade de cada um, realiza-se agrupamento em blocos. Para um experimento em delineamento em blocos casualizados (DBC), o plano experimental baseia-se nos três princípios básicos da experimentação: repetição, casualização e controle local [2].

Para realizar a análise de variância em um delineamento em blocos completos casualizados, inicialmente calcula-se um fator de correção (C) dado pela Equação 2.2.1:

$$C = \frac{(\text{Total geral})^2}{\text{Total julgamentos}}, \quad (2.2.1)$$

em que Total geral = soma dos resultados de todos julgamentos e Total julgamentos = número de avaliadores (J) x número de amostras (I) [16].

Após, calcula-se as somas dos quadrados das amostras (SQAm), dos avaliadores (SQAv) e total (SQT), ou seja, a variação em relação as amostras, avaliadores e a variação total do experimento, respectivamente. As equações 2.3.4, 2.2.3 e 2.2.4 representam as fórmulas, respectivamente, para os cálculos:

$$SQAm = \frac{(Am\ 1)^2 + (Am\ 2)^2 + \dots + (Am\ J)^2}{I} - C, \quad (2.2.2)$$

$$SQAv = \frac{(Av\ 1)^2 + (Av\ 2)^2 + \dots + (Av\ I)^2}{J} - C, \quad (2.2.3)$$

$$SQT = \sum_{j=1}^I \sum_{i=1}^J valor_{ij}^2 - C, \quad (2.2.4)$$

Por meio das equações obtidas anteriormente, é possível calcular a variação devido ao erro experimental, ou seja, a soma dos quadrados dos resíduos (SQR), determinado por meio da equação 2.2.5 [16]:

$$SQR = SQT - (SQAm + SQAv). \quad (2.2.5)$$

Pode-se então calcular os valores dos quadrados médios das amostras (QMAM), avaliadores (QMAv) e resíduo (QMR), mostrados respectivamente pelas equações 2.2.6, 2.2.7, 2.2.8. Para isso, divide-se a soma dos quadrados resultantes pelo graus de liberdade (GL) de cada fator, no qual o número de graus de liberdade é determinado pelo número total de cada fator menos 1, ou seja, quando se tem J avaliadores, o número de graus de liberdade dos avaliadores = J - 1 [2].

$$QMAM = \frac{SQAm}{I - 1} \quad (2.2.6)$$

$$QMAv = \frac{SQAv}{J - 1} \quad (2.2.7)$$

$$QMR = \frac{SQR}{(I - 1)(J - 1)} \quad (2.2.8)$$

Por fim, para saber se há diferença significativa entre as amostras, calcula-se um valor F, determinado pela Equação 2.2.9:

$$F = \frac{QMA}{QMR}. \quad (2.2.9)$$

Para facilitar a análise dos resultados e realização dos cálculos são criadas tabelas, que seguem o modelo da Tabela 2.2.1, apresentada a seguir [2].

Tabela 2.2.1: Tabela modelo para realização da ANOVA.

Causas de variação (CV)	Graus de liberdade (GL)	Soma dos quadrados (SQ)	Quadrados médios (QM)	F
Amostra	I - 1	$\frac{1}{I} \sum_{j=1}^I Am_j^2 - C$	$\frac{SQAm}{I-1}$	$\frac{QMAm}{QMR}$
Avaliador	J - 1	$\frac{1}{J} \sum_{i=1}^J Av_i^2 - C$	$\frac{SQAv}{J-1}$	$\frac{QMAv}{QMR}$
Resíduo	(I - 1)(J - 1)	SQT - (SQAm + SQAv)	$\frac{SQR}{(I-1)(J-1)}$	
Total	IJ - 1	$\sum_{j=1}^I \sum_{i=1}^J valor_{ij}^2 - C$		

Por meio da estatística do teste F, avalia-se se há similaridade entre as amostras, ou seja, se os avaliadores não percebem diferenças sensoriais, ou se pelo menos duas amostras apresentam diferenças sensoriais significativas entre si.

O valor encontrado na equação 2.2.9 deve ser comparado com um valor de tabela da distribuição F, com $(I - 1, (I - 1)(J - 1))$ graus de liberdade, dado um nível de significância α . Se o valor encontrado for maior que o valor tabelado, não há homogeneidade, caso contrário, há homogeneidade. Outra forma de avaliação do resultado é dado pelo p-valor, se este for menor do que o nível α estabelecido, rejeita-se a hipótese de homogeneidade de percepção sensorial, caso contrário, não rejeita-se.

2.2.1.2 Teste de Tukey

Este teste de médias realiza a comparação de todas as amostras entre si. Para determinar se há diferença significativa entre elas, é necessário calcular a diferença mínima significativa (DMS_{Tukey}) entre as amostras analisadas, apresentada pela equação 2.2.10:

$$DMS_{Tukey} = q_\alpha \sqrt{\frac{QMR}{J}}, \quad (2.2.10)$$

em que q_α = valor resultante da tabela de Tukey, utilizando-se o número de amostras testadas e os graus de liberdade do resíduo; QMR = quadrado médio do resíduo, calculado na ANOVA; J = número de avaliadores [16].

A avaliação do teste pode ser feita comparando as diferenças entre as médias das amostras, comparando todas elas entre si e verificando se o valor é maior, menor ou igual que o valor DMS_{Tukey} . Uma amostra apresenta diferença significativa em relação a outra, se o valor calculado da diferença entre suas médias for maior que a diferença mínima significativa, caso contrário não há diferença entre tais amostras. Para um teste bilateral, a diferença deve ser considerada em módulo.

Pode-se também avaliar seus resultados com base no p-valor. Se o p-valor encontrado for menor que o nível de significância estabelecido, há diferença entre as médias estudadas, caso contrário, não há.

2.2.1.3 Teste de Dunnett

Neste teste, as médias das amostras são comparadas em relação a média da amostra controle, ou seja, as amostras não são comparadas entre si, somente com a amostra controle.

Para determinar se há diferença significativa entre uma amostra e um controle, é necessário calcular a diferença mínima significativa ($DMS_{Dunnett}$) dada pela equação 2.2.11:

$$DMS_{Dunnett} = D \sqrt{\frac{2(QMR)}{J}}, \quad (2.2.11)$$

em que D = valor tabelado da tabela Dunnett para teste unilateral ou bilateral, que considera o número de graus de liberdade da amostra e do resíduo; QMR = Quadrado médio do resíduo, obtido na ANOVA; J = número de avaliadores [2].

A avaliação do teste pode ser feita comparando as diferenças entre as médias e o valor de ($DMS_{Dunnett}$) ou com base no p-valor, com as mesmas regras já descritas para o teste de Tukey.

2.2.2 Exemplos de aplicação

Problema 1. Adaptado de Dutcosky [2, p. 218]. Uma comparação múltipla foi feita determinar a quantidade do conservante benzoato de sódio que poderia ser adicionada ao refrigerante de laranja e a avaliação foi realizada sobre o residual amargo no gosto. Os refrigerantes testados continham : 0,001%, 0,003%, 0,005% e 0,007% de benzoato de sódio. Cada grupo de prova continha quatro amostras codificadas. Vinte avaliadores fizeram a avaliação, em uma escala de 1 a 9, sendo 1 a menor intensidade de residual amargo no gosto e 9 a maior.

As amostras escolhidas para teste tinham como finalidade perceber o quanto a adição do conservante benzoato de sódio era perceptível em relação ao gosto amargo. O objetivo foi comparar as amostras para diferenciá-las em relação ao gosto amargo. Para os testes adotou-se $\alpha = 0,05$.

Tabela 2.2.2: Resultados obtidos no teste de comparação múltipla de refrigerantes.

Avaliadores	0,001%	0,003%	0,005%	0,007%	Total
Av1	3	2	4	6	15
Av2	2	4	5	7	18
Av3	3	4	6	8	21
Av4	6	5	3	6	20
Av5	2	3	5	9	19
Av6	2	4	5	8	19
Av7	3	2	4	7	16
Av8	5	5	6	8	24
Av9	3	2	5	8	18
Av10	4	3	3	9	19
Av11	3	4	5	8	20
Av12	4	5	5	9	23
Av13	3	5	5	8	21
Av14	4	6	5	9	24
Av15	4	5	6	9	24
Av16	5	4	4	8	21
Av17	4	3	5	8	20
Av18	5	6	6	9	26
Av19	4	4	6	9	23
Av20	3	4	5	9	21
Total	72	80	98	162	412
Média	3,6	4,0	4,9	8,1	

Fonte: Adaptado de Dutcosky [2, p.218].

Para comparação entre as amostras, os cálculos para a elaboração da tabela de análise de variância (ANOVA) foram realizados, dados por:

$$C = \frac{(\text{Total geral})^2}{\text{Total julgamentos}}$$

$$\begin{aligned} C &= \frac{(412)^2}{80} \\ C &= 2121,8; \end{aligned}$$

$$\text{SQAm} = \frac{(\text{Am } 1)^2 + (\text{Am } 2)^2 + \dots + (\text{Am } N)^2}{I} - C$$

$$\begin{aligned} \text{SQAm} &= \frac{(72)^2 + (80)^2 + (98)^2 + (162)^2}{20} - C \\ \text{SQAm} &= 2371,6 - 2121,8 \\ \text{SQAm} &= 249,8; \end{aligned}$$

$$\text{SQAv} = \frac{(\text{Av 1})^2 + (\text{Av 2})^2 + \dots + (\text{Av N})^2}{J} - C$$

$$\begin{aligned} \text{SQAv} &= \frac{[(15)^2 + (18)^2 + (21)^2 + (20)^2 + (19)^2 + \dots + (21)^2 + (20)^2 + (26)^2 + (23)^2 + (21)^2]}{4} - C \\ \text{SQAv} &= \frac{8638}{4} - 2121,8 \\ \text{SQAv} &= 2159,5 - 2121,8 \\ \text{SQAv} &= 37,7; \end{aligned}$$

$$\text{SQT} = \sum_{i=1}^J \text{valor}_{ij}^2 - C$$

$$\begin{aligned} \text{SQT} &= [(3)^2 + (2)^2 + (4)^2 + (6)^2 + (2)^2 + (4)^2 + (5)^2 + (7)^2 + \\ &+ (3)^2 + (4)^2 + (6)^2 + (8)^2 + (6)^2 + (5)^2 + (3)^2 + (6)^2 + \\ &+ (2)^2 + (3)^2 + (5)^2 + (9)^2 + (2)^2 + (4)^2 + (5)^2 + (8)^2 + \\ &+ (3)^2 + (2)^2 + (4)^2 + (7)^2 + (5)^2 + (5)^2 + (6)^2 + (8)^2 + \\ &+ (3)^2 + (2)^2 + (5)^2 + (8)^2 + (4)^2 + (3)^2 + (3)^2 + (9)^2 + \\ &+ (3)^2 + (4)^2 + (5)^2 + (8)^2 + (4)^2 + (5)^2 + (5)^2 + (9)^2 + \\ &+ (3)^2 + (5)^2 + (5)^2 + (8)^2 + (4)^2 + (6)^2 + (5)^2 + (9)^2 + \\ &+ (4)^2 + (5)^2 + (6)^2 + (9)^2 + (5)^2 + (4)^2 + (4)^2 + (8)^2 + \\ &+ (4)^2 + (3)^2 + (5)^2 + (8)^2 + (5)^2 + (6)^2 + (6)^2 + (9)^2 + \\ &+ (4)^2 + (4)^2 + (6)^2 + (9)^2 + (3)^2 + (4)^2 + (5)^2 + (9)^2] - C \\ \text{SQT} &= 2456 - 2121,8 \\ \text{SQT} &= 334,2; \end{aligned}$$

$$\begin{aligned} \text{SQR} &= \text{SQT} - (\text{SQAm} + \text{SQAv}) \\ \text{SQR} &= 334,2 - (249,8 + 37,7) \\ \text{SQR} &= 46,7. \end{aligned}$$

Os quadrados médios e o cálculo de F para a análise foram dados por:

$$\begin{array}{ll} \text{QMAm} = \frac{\text{SQAm}}{\text{GLAm}} & \text{QMR} = \frac{\text{SQR}}{\text{GLR}} \\ \text{QMAm} = \frac{249,8}{3} & \text{QMR} = \frac{46,7}{57} \\ \text{QMAm} = 83,27; & \text{QMR} = 0,82; \end{array}$$

$$\begin{array}{ll} \text{QMAv} = \frac{\text{SQAv}}{\text{GLAv}} & \text{F} = \frac{\text{QMAm}}{\text{QMR}} \\ \text{QMAv} = \frac{37,7}{19} & \text{F} = \frac{83,27}{0,82} \\ \text{QMAv} = 1,98; & \text{F} = 101,5. \end{array}$$

Monta-se a tabela de análise de variância, dada por:

Tabela 2.2.3: Análise de Variância para comparação das amostras de refrigerantes.

	CV	GL	SQ	QM	F
Amostras	3	249,8	83,27	101,5	
Avaliadores	19	37,7	1,98		
Resíduos	57	46,7	0,82		
Total	79	334,2			

Fonte: Autoria própria.

Para a avaliação do quadro da ANOVA, o valor tabelado da distribuição F deve ser considerado. Para o caso, $F_{tabelado}(3, 57)$ ao nível de 5% de significância, tem-se o valor na Tabela 2.2.4.

Tabela 2.2.4: Limites unilaterais de F em nível de 5% de probabilidade para o caso de F maior que 1.

$n_2 \setminus n_1$	1	2	3	4	5	6	8	12	24	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	238,9	243,9	249,0	254,3
2	18,51	19,00	254,3	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	19,16	9,12	9,01	8,94	8,84	8,74	8,64	8,53
:	:	:	:	:	:	:	:	:	:	:
40	4,08	3,23	2,93	2,61	2,45	2,34	2,18	2,00	1,79	1,51
60	4,00	3,15	2,92	2,52	2,37	2,25	2,10	1,92	1,70	1,39
:	:	:	:	:	:	:	:	:	:	:

Fonte: Adaptado de Dutcosky [2][p.477-478].

Assim, utilizou-se 3 graus de liberdade das amostras e 60 graus de liberdade do resíduo, por ser o valor mais próximo do requerido. Encontrou-se o valor de F tabelado, ao nível de 5% de significância de 2,92, então como o valor de F calculado foi maior que o valor de F tabelado, existindo diferença significativa entre as amostras.

Para a avaliação de quais amostras diferem entre si, o teste de Tukey foi utilizado. Para o cálculo do teste de Tukey, encontrou-se primeiramente o valor de DMS_{Tukey} , o qual depende da amplitude total estudentizada, do número de avaliadores e do quadrado médio do resíduo, obtido pela Análise de Variância (ANOVA). O valor da amplitude total (q) é tabelado e para ser encontrado requer o número de amostras (i) e o número de graus de liberdade do resíduo (GLres), dado por $q(4,57)$ (Tabela 2.2.5).

Tabela 2.2.5: Valores de amplitude total estudentizada (q), para uso no teste de Tukey.

GLres \ i	2	3	4	5	6	7	8	9	10	11
1	18,0	26,7	32,8	37,2	40,5	43,1	45,4	47,3	49,1	50,6
2	6,09	8,28	9,80	10,89	11,73	12,43	13,03	13,54	13,99	14,39
3	4,50	5,88	6,83	7,51	8,04	8,47	8,85	9,18	9,46	9,72
:	:	:	:	:	:	:	:	:	:	:
40	2,86	3,44	3,79	4,04	4,23	4,39	4,52	4,63	4,74	4,82
60	2,83	3,40	3,74	3,98	4,16	4,31	4,44	4,55	4,65	4,73
:	:	:	:	:	:	:	:	:	:	:

Adaptado de Dutcosky [?][p.483-484].

Pela Tabela 2.2.5 dos valores de amplitude estudentizada encontrou-se $q(4,60)=3,74$; utilizou-se o número de graus de liberdade de resíduo 60, por ser o valor mais próximo ao requerido. Assim, calculou-se o valor de DMS_{Tukey} , considerando $q=3,74$; $J=20$ e $QMR=0,82$:

$$DMS = q\sqrt{\frac{QMR}{J}}$$

$$DMS = 3,74\sqrt{\frac{0,82}{20}}$$

$$DMS = 0,76.$$

Com o valor de estabelecido, foi possível comparar as amostras entre si para verificação de diferença significativa.

$ 0,001\% - 0,003\% = 3,6 - 4,0 = 0,4 < DMS$	Não diferem entre si a 5% de significância,
$ 0,001\% - 0,005\% = 3,6 - 4,9 = 1,3 > DMS$	Diferem entre si a 5% de significância,
$ 0,001\% - 0,007\% = 3,6 - 8,1 = 4,5 > DMS$	Diferem entre si a 5% de significância,
$ 0,003\% - 0,005\% = 4,0 - 4,9 = 0,9 > DMS$	Diferem entre si a 5% de significância,
$ 0,003\% - 0,007\% = 4,0 - 8,1 = 4,1 > DMS$	Diferem entre si a 5% de significância,
$ 0,005\% - 0,007\% = 4,9 - 8,1 = 3,2 > DMS$	Diferem entre si a 5% de significância.

Tabela 2.2.6: Médias seguidas de letras diferentes diferem entre si pelo teste de Tukey a 5%.

Amostras	Médias
0,001%	3,6 c
0,003%	4,0 c
0,005%	4,9 b
0,007%	8,1 a

Fonte: Autoria própria.

As amostras de 0,001% e de 0,003% não apresentaram diferenças significativas em relação ao gosto residual amargo, com a menor avaliação quanto ao gosto amargo, e diferiram das amostras de 0,005% e de 0,007%, as quais também diferiram entre si, ao nível de 5% de significância.

Problema 2 (Adaptado de Dutcosky[p. 215-216]). [2]. Foi aplicado um teste de comparação múltipla com geleias dietéticas de kiwi para avaliar a interação do sabor da fruta e os diferentes adoçantes. As seguintes amostras foram analisadas em relação ao sabor:

309=C=geleia de kiwi com sacarose;

215=geleia de kiwi com frutose;

826=geleia de kiwi com sorbitol/aspartame;

437=geleia de kiwi com sorbitol/sacarina.

Sabe-se que a sacarose é a que apresenta melhores características de dulçor, sendo considerada padrão em todos os estudos de adoçantes e edulcorantes, por não apresentar efeitos indesejáveis de sabores (amargo, metálico, residuais), secura na boca, entre outros. Os extremos da escala utilizada foram:

1=extremamente melhor que o controle;

9=extremamente pior que o controle.

As amostras apresentadas aos avaliadores possuíam codificação de três numerações para evitar a indução de suposições. Para os testes adotou-se $\alpha = 5\%$.

Tabela 2.2.7: Respostas obtidas no teste de comparação de geleias.

	Av1	Av2	Av3	Av4	Av5	Av6	Av7	Av8	Av9	Av10	Total	Média
309	5	3	4	3	5	3	4	3	5	5	40	4,0
215	6	5	6	5	4	7	7	4	6	7	57	5,7
826	7	6	7	6	6	6	4	6	4	58	5,8	
437	8	7	9	8	7	8	8	9	8	8	80	8,0
Total	26	21	26	22	22	24	25	20	25	24	235	

Fonte: Adaptado de Dutcosky [?][p.216].

Para montagem do quadro de análise de variância, calculou-se: o fator de correção (C), soma de quadrados das amostras (SQAm), soma de quadrados dos avaliadores (SQAv), soma de quadrados total (SQT) e a soma de quadrados do resíduo (SQR).

$$C = \frac{(\text{Total geral})^2}{\text{Total julgamentos}},$$

$$\begin{aligned} C &= \frac{(235)^2}{40} \\ &= 1380,625; \end{aligned}$$

$$SQAm = \frac{(Am\ 1)^2 + (Am\ 2)^2 + \dots + (Am\ N)^2}{I} - C,$$

$$\begin{aligned} SQAm &= \frac{[(40)^2 + (57)^2 + (58)^2 + (80)^2]}{10} - C \\ &= 1461,3 - 1380,625 \\ &= 80,675; \end{aligned}$$

$$SQAv = \frac{(Av\ 1)^2 + (Av\ 2)^2 + \dots + (Av\ N)^2}{J} - C$$

$$\begin{aligned} SQav &= \frac{[(26)^2 + (21)^2 + (26)^2 + (22)^2 + (22)^2 + (24)^2 + (25)^2 + (20)^2 + (25)^2 + (24)^2]}{4} - C \\ &= \frac{5563}{4} - 1380,625 \\ &= 10,125; \end{aligned}$$

$$SQT = \sum_{i=1}^J valor_{ij}^2 - C$$

$$\begin{aligned} SQT &= [(5)^2 + (3)^2 + (4)^2 + (3)^2 + (5)^2 + (3)^2 + (4)^2 + (3)^2 + (5)^2 + (5)^2 + \\ &+ (6)^2 + (5)^2 + (6)^2 + (5)^2 + (4)^2 + (7)^2 + (7)^2 + (4)^2 + (6)^2 + (7)^2 + \\ &+ (7)^2 + (6)^2 + (7)^2 + (6)^2 + (6)^2 + (6)^2 + (4)^2 + (6)^2 + (4)^2 + \\ &+ (8)^2 + (7)^2 + (9)^2 + (8)^2 + (7)^2 + (8)^2 + (8)^2 + (9)^2 + (8)^2 + (8)^2] - FC \\ &= 1495 - 1380,625 \\ &= 114,375; \end{aligned}$$

$$\begin{aligned} SQR &= SQT - (SQAm + SQAv) \\ &= 114,375 - (80,675 + 10,125) \\ &= 23,575. \end{aligned}$$

Calcula-se os quadrados médios e o valor de F:

$$\begin{aligned} QMAM &= \frac{SQAm}{GLAm} \\ QMAM &= \frac{80,675}{3} \\ QMAM &= 26,892; \end{aligned}$$

$$\begin{aligned} QMAv &= \frac{SQAv}{GLAv} \\ QMAv &= \frac{10,125}{9} \\ QMAv &= 1,125; \end{aligned}$$

$$\begin{aligned} QMR &= \frac{SQR}{GLR} \\ QMR &= \frac{23,575}{27} \\ QMR &= 0,873; \end{aligned}$$

$$\begin{aligned} F &= \frac{QMAm}{QMR} \\ F &= \frac{26,892}{0,873} \\ F &= 30,80. \end{aligned}$$

Com esses cálculos, a tabela de análise de variância pode ser montada, dada por:

Tabela 2.2.8: Análise de Variância para comparação das geleias.

CV	GL	SQ	QM	F
Amostras	3	80,675	26,892	30,80
Avaliadores	9	10,125	1,125	
Resíduos	27	23,575	0,873	
TOTAL	39	114,375		

Tabela 2.2.9: *

Fonte: Autoria própria.

Para a avaliação do quadro da ANOVA, o valor tabelado da distribuição F deve ser considerado. Para o caso, $F_{tabelado}(3, 27)$ ao nível de 5% de significância, tem-se o valor na tabela 2.2.10.

Tabela 2.2.10: Limites unilaterais de F em nível de 5% de probabilidade para o caso de F maior que 1.

$n_2 \setminus n_1$	1	2	3	4	5	6	8	12	24	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	238,9	243,9	249,0	254,3
2	18,51	19,00	254,3	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	19,16	9,12	9,01	8,94	8,84	8,74	8,64	8,53
:	:	:	:	:	:	:	:	:	:	:
27	4,21	3,35	2,99	2,73	2,57	2,46	2,30	2,13	1,93	1,67
:	:	:	:	:	:	:	:	:	:	:

Fonte: Adaptado de Dutcosky [?][p. 477-478].

Assim, com 3 graus de liberdade das amostras e 27 graus de liberdade do resíduo, o valor de F tabelado, ao nível de 5% de significância é de 2,99. Como o valor de F foi maior que o valor F tabelado, pode-se concluir que há uma diferença significativa na percepção sensorial entre as amostras.

Para avaliação de quais amostras diferem da amostra padrão, o teste de Dunnett foi utilizado. Para o exemplo em questão, assumiu-se que os tratamentos sempre apresentam uma avaliação pior do que o grupo controle. Para o cálculo do teste de Dunnett, encontrou-se primeiramente o valor de D, que depende do número de graus de liberdade das amostras (GL=3) e do número graus de liberdade dos resíduos (GL=27), assim para o teste unilateral de Dunnett com 5% de significância, pela tabela 2.2.11 consultou-se o valor de D.

Tabela 2.2.11: Valores de D para teste unilateral de Dunnett ao nível de 5% de significância.

GL resíduo	GL amostras								
	1	2	3	4	5	6	7	8	9
5	2,02	2,44	2,68	2,85	2,98	3,08	3,16	3,24	3,30
6	1,94	2,34	2,56	2,71	2,83	2,92	3,00	3,07	3,12
7	1,89	2,27	2,48	2,62	2,73	2,82	2,89	2,95	3,01
:	:	:	:	:	:	:	:	:	:
24	1,71	2,01	2,17	2,28	2,36	2,43	2,48	2,53	2,57
30	1,70	1,99	2,15	2,25	2,33	2,40	2,45	2,50	2,54

Tabela 2.2.12: *

Fonte: Adaptado de Dutcosky [p. 481].

O valor mais próximo de D para resolução do exercício é de 2,17 ($D(3,24)$). Do quadro da ANOVA, QMR = 0,873 e J=10, então:

$$\begin{aligned}
 DMS_{Dunnnett} &= D \times \sqrt{\frac{2 \times QM_{res}}{J}} \\
 &= 2,17 \times \sqrt{\frac{2 \times 0,873}{10}} \\
 &= 2,17 \times 0,4178 \\
 &= 0,9067.
 \end{aligned}$$

A partir do cálculo da diferença mínima significativa (DMS) comparou-se as diferenças entre às médias estimadas dos tratamentos com a do controle.

- | | |
|---|---|
| $215 - 309 = 5,7 - 4,0 = 1,7 > DMS$
$826 - 309 = 5,8 - 4,0 = 1,8 > DMS$
$437 - 309 = 8,0 - 4,0 = 4,0 > DMS$ | Diferem entre si a 5% de significância,
Diferem entre si a 5% de significância,
Diferem entre si a 5% de significância. |
|---|---|

Dessa forma, é possível notar que ao nível de 5% de significância, que as geleias com frutose, sorbitol e aspartame, sorbitol e sacarina diferem da geleia com sacarose em relação ao gosto, e pode-se afirmar que são piores que a geleia elaborada com sacarose.

Problema 3. *Após realizar uma pesquisa de mercado, a equipe de Pesquisa e Desenvolvimento (P&D) de uma indústria cervejeira concluiu que um novo tipo de cerveja que se caracterize pelo sabor intenso e aroma de frutas pode ser bem aceita pelos consumidores de cerveja. Nas etapas iniciais de desenvolvimento da nova cerveja são produzidos 5 lotes de cerveja com a utilização de lúpulo de diferentes procedências e é realizada uma avaliação sensorial empregando-se o teste de comparação múltipla com 25 avaliadores treinados. No teste foi utilizada uma escala estruturada mista na qual 0=Sabor imperceptível e 8=Sabor forte. Os dados obtidos encontram-se na Tabela 2.2.2. Analise os dados e verifique se algum lote de lúpulo proporcionou maior sabor mais intenso à cerveja, use $\alpha = 5\%$.*

AVALIADOR	Lúpulo A	Lúpulo B	Lúpulo C	Lúpulo D	Lúpulo E
1	4	7	5	6	3
2	3	8	6	7	4
3	2	7	6	8	2
4	3	8	5	8	4
5	2	7	7	7	4
6	2	8	5	7	3
7	2	7	4	6	3
8	4	7	3	6	4
9	2	7	4	7	3
10	2	6	3	5	2
11	3	8	3	6	3
12	3	7	4	5	3
13	2	7	4	7	3
14	2	6	3	6	6
15	3	8	4	7	5
16	2	7	4	7	4
17	4	7	5	7	3
18	3	7	4	7	2
19	2	7	4	7	5
20	4	6	3	7	3
21	3	7	5	6	4
22	2	6	5	7	3
23	2	7	4	7	3
24	2	8	4	6	4
25	3	7	5	7	2
Média	2,64	7,08	4,36	6,64	3,4

2.2.2.1 Resolução

É possível calcular o quadro da ANOVA realizando os seguintes cálculos:

$$C = \frac{(\text{Total geral})^2}{\text{Total julgamentos}}$$

$$\begin{aligned} C &= \frac{(603)^2}{125} \\ C &= 2908,872; \end{aligned}$$

$$\text{SQAm} = \frac{(\text{Am } 1)^2 + (\text{Am } 2)^2 + \dots + (\text{Am } N)^2}{I} - C$$

$$\begin{aligned} \text{SQAm} &= \frac{(66)^2 + (177)^2 + (109)^2 + (166)^2 + (85)^2}{25} - C \\ \text{SQAm} &= 3293,88 - 2908,872 \\ \text{SQAm} &= 385,008; \end{aligned}$$

$$\text{SQAv} = \frac{(\text{Av 1})^2 + (\text{Av 2})^2 + \dots + (\text{Av N})^2}{J} - C$$

$$\begin{aligned} \text{SQAv} &= \frac{[(25)^2 + (28)^2 + (25)^2 + (28)^2 + (27)^2 + \dots + (25)^2 + (23)^2 + (23)^2 + (24)^2 + (24)^2]}{5} - C \\ \text{SQAv} &= \frac{14665}{5} - 2908.872 \\ \text{SQAv} &= 2931 - 2908,872 \\ \text{SQAv} &= 22,128; \end{aligned}$$

$$\text{SQT} = \sum_{i=1}^J \text{valor}_{ij}^2 - C$$

$$\begin{aligned} \text{SQT} &= [(4)^2 + (3)^2 + (2)^2 + (3)^2 + (2)^2 + (2)^2 + (2)^2 + (4)^2 + \dots \\ &+ (4)^2 + (3)^2 + (3)^2 + (4)^2 + (2)^2] - C \\ \text{SQT} &= 3381 - 2908,872 \\ \text{SQT} &= 472,128; \end{aligned}$$

$$\begin{aligned} \text{SQR} &= \text{SQT} - (\text{SQAm} + \text{SQAv}) \\ \text{SQR} &= 472,128 - (385,008 + 22,128) \\ \text{SQR} &= 64,992. \end{aligned}$$

Calcula-se os quadrados médios e o valor de F:

$$\begin{aligned} \text{QMAM} &= \frac{\text{SQAm}}{\text{GLAm}} \\ \text{QMAM} &= \frac{385,008}{4} \\ \text{QMAM} &= 96,252; \end{aligned}$$

$$\begin{aligned} QMAv &= \frac{SQAv}{GLAv} \\ QMAv &= \frac{22,128}{24} \\ QMAv &= 0,922; \end{aligned}$$

$$\begin{aligned} QMR &= \frac{SQR}{GLR} \\ QMR &= \frac{64,992}{96} \\ QMR &= 0,677; \end{aligned}$$

$$\begin{aligned} F &= \frac{QMAm}{QMR} \\ F &= \frac{96,252}{0,677} \\ F &= 142,1743. \end{aligned}$$

Com esses cálculos, a tabela de análise de variância pode ser montada, dada por:

Tabela 2.2.13: Análise de Variância para comparação dos lúpulos.

	CV	GL	SQ	QM	F
Avaliadores	24	22,128	0,922		
Lúpulo	4	385,008	96,252	142,1743	
Resíduos	96	64,992	0,677		
TOTAL	124	472,128			

Tabela 2.2.14: *

Fonte: Autoria própria.

Para a avaliação do quadro da ANOVA, o valor tabelado da distribuição F deve ser considerado. Para o caso, $F_{tabelado}(4, 96)$ ao nível de 5% de significância, tem-se o valor aproximado na tabela 2.2.15.

Tabela 2.2.15: Limites unilaterais de F em nível de 5% de probabilidade para o caso de F maior que 1.

$n_2 \setminus n_1$	1	2	3	4	5	6	8	12	24	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	238,9	243,9	249,0	254,3
2	18,51	19,00	254,3	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	19,16	9,12	9,01	8,94	8,84	8,74	8,64	8,53
:	:	:	:	:	:	:	:	:	:	:
120	3,92	3,07	2,84	2,45	2,29	2,17	2,02	1,83	1,61	1,25
:	:	:	:	:	:	:	:	:	:	:

Tabela 2.2.16: *

Fonte: Adaptado de Dutcosky [p. 477-478].

Assim, com 4 graus de liberdade das amostras e 120 graus de liberdade do resíduo, o valor de F tabelado, ao nível de 5% de significância é de 2,45. Como o valor de F calculado foi maior que o valor F tabelado, pode-se concluir que há uma diferença significativa na entre as amostras, sendo necessário a aplicação de um teste de médias para verificar quais amostras diferem entre si, sendo o utilizado para este caso o teste de Tukey. Temos que $q\alpha = 3,945$ ao observar a Tabela ??, $J = 5$ e $QM_{res} = 0,677$, então utilizando a equação 2.2.10.

$$DMS_{Tukey} = 3,945 \sqrt{\frac{0,677}{25}} \quad (2.2.12)$$

$$DMS_{Tukey} = 0,6492.$$

Fazendo a diferença entre as média:

- $|A - B| = |6,47 - 7,08| = 4,44 > DMS$ Diferem entre si a 5% de significância,
- $|A - C| = |2,64 - 4,36| = 1,72 > DMS$ Diferem entre si a 5% de significância,
- $|A - D| = |2,64 - 6,64| = 4 > DMS$ Diferem entre si a 5% de significância,
- $|A - E| = |2,64 - 3,4| = 0,74 > DMS$ Diferem entre si a 5% de significância,
- $|B - C| = |7,08 - 4,36| = 2,72 > DMS$ Diferem entre si a 5% de significância,
- $|B - D| = |7,08 - 6,64| = 0,44 < DMS$ Não diferem entre si a 5% de significância,
- $|B - E| = |7,08 - 3,4| = 3,68 > DMS$ Diferem entre si a 5% de significância,
- $|C - D| = |4,36 - 6,64| = 2,28 > DMS$ Diferem entre si a 5% de significância,
- $|D - E| = |6,64 - 3,4| = 3,24 > DMS$ Diferem entre si a 5% de significância,
- $|C - E| = |4,36 - 3,4| = 0,96 > DMS$ Diferem entre si a 5% de significância,

Organizando as médias com letras, onde médias seguidas de letras iguais possuem a mesma média a 5% de significância pelo teste de Tukey, temos:

$$A = 2,64 \text{ d}$$

$$B = 7,08 \text{ a}$$

$$C = 4,36 \text{ b}$$

$$D = 6,64 \text{ a}$$

$$E = 3,40 \text{ c}$$

Assim, os resultados permitem concluir que os lúpulos de procedência B e D conferem o sabor mais intenso às cervejas e não são diferenciados entre si. O lúpulo de procedência C confere sabor de intensidade intermediária e diferente de todos os outros, enquanto os lúpulos A e E não são diferenciados e conferem pouco sabor às cervejas. Considerando os objetivos do teste devem ser utilizados os lúpulos B e D nas próximas etapas de pesquisa para o desenvolvimento de cervejas com sabor intenso.

Para começar a resolver o problema, verifica-se se há diferença significativa entre as amostras realizando um Teste de Variância (ANOVA: fator duplo sem repetição), podendo ser facilmente realizado no excel (será explicado os procedimentos para realizar a ANOVA no excel no capítulo [7.3.1](#)).

2.2.2.2 Exemplo 3 - Teste de diferença do controle com análises repetidas

Problema 4. *Uma indústria de achocolatados pretende diminuir a utilização de sacarose em sua bebida láctea pronta para beber para diminuir o valor calórico da mesma e poder usar a alegação de menor valor calórico na embalagem mas não quer que os consumidores percebam diferença na intensidade do gosto doce. Para tanto a indústria fabricou 4 formulações de bebida láctea com diferentes reduções de sacarose: -2%, -4%, -7% e -10% e as comparou com a bebida láctea tradicional, sem redução de sacarose utilizando um Teste de Diferença do Controle. Vinte e cinco avaliadores avaliaram as bebidas 2 vezes utilizando uma escala que variava de 1=nenhuma diferença do controle em relação ao gosto doce a 9=muito diferente do controle em relação ao gosto doce. Os resultados estão expressos na Tabela [2.2.17](#). Verifique se há diferença entre as amostras e conclua sobre qual a máxima redução de sacarose possível.*

Tabela 2.2.17: Resultados do teste de diferença do controle entre bebidas lácteas com diferentes percentagens de redução de sacarose.

Avaliador	Bebida Trad	-2%	-4%	-7%	-10%
1	2	2	3	4	7
1	1	2	1	3	8
2	1	1	2	3	7
2	2	2	2	5	6
3	2	2	1	5	7
3	1	2	2	4	7
4	1	3	2	4	6
4	1	1	2	4	6
5	2	2	3	3	7
5	1	1	2	4	8
6	2	1	2	4	6
6	3	2	2	5	8
7	3	1	2	3	6
7	1	1	1	4	7
8	1	2	2	5	8
8	2	1	2	3	8
9	2	1	2	4	8
9	2	1	1	6	7
10	1	1	1	4	5
10	1	1	2	5	7
11	1	1	2	3	7
11	1	2	2	4	7
12	1	1	1	5	8
12	2	1	2	5	8
13	1	1	2	3	7
13	1	2	1	5	7
14	1	1	1	4	5
14	1	3	2	4	6
15	2	2	2	4	7
15	2	2	1	3	7
16	2	3	2	5	6
16	1	2	1	4	7
17	1	1	2	4	7
17	2	2	2	4	7
18	2	2	2	3	8
18	1	2	1	7	8
19	2	3	2	5	6
19	1	1	2	4	7

Avaliador	Bebida	Trad	-2%	-4%	-7%	-10%
20		1	1	2	6	7
21		2	2	1	5	7
21		3	2	2	6	8
22		2	1	2	6	7
22		1	1	1	5	7
23		2	1	2	5	6
23		1	2	1	6	7
24		2	3	2	6	7
24		2	1	2	5	7
25		1	2	1	5	6
25		2	2	2	6	7

Como deseja verificar se há diferença entre os tratamentos, é realizado um teste de variância (ANOVA), assim, temos:

$$C = \frac{(\text{Total geral})^2}{\text{Total julgamentos}}$$

$$\begin{aligned} C &= \frac{(816)^2}{250} \\ C &= 2663,424; \end{aligned}$$

$$SQT = \sum_i \sum_k \sum_j y_{ikj}^2 - C$$

$$\begin{aligned} SQT &= [(2)^2 + (2)^2 + (3)^2 \dots \\ &\quad + (6)^2 + (7)^2] - C \\ SQT &= 3958 - 2663,424 \\ SQT &= 1294,576; \end{aligned}$$

$$SQAm = \frac{(\text{Am 1})^2 + (\text{Am 2})^2 + \dots + (\text{Am N})^2}{I} - C$$

$$\begin{aligned} \text{SQAm} &= \frac{(77)^2 + (81)^2 + (87)^2 + (223)^2 + (348)^2}{50} - C \\ \text{SQAm} &= 3817,84 - 2663,424 \\ \text{SQAm} &= 1154,416; \end{aligned}$$

$$\begin{aligned} \text{SQAv} &= \frac{[(33)^2 + (31)^2 + (33)^2 + (30)^2 + \dots + (38)^2 + (33)^2 + (33)^2 + (37)^2 + (34)^2]}{10} - C \\ \text{SQAv} &= \frac{26788}{10} - 2663,424 \\ \text{SQAv} &= 2,678,8 - 2663,424 \\ \text{SQAv} &= 15,376; \end{aligned}$$

Calcula-se os quadrados médios e o valor de F:

$$\begin{aligned} \text{QMAm} &= \frac{\text{SQAm}}{\text{GLAm}} \\ \text{QMAm} &= \frac{1154,416}{4} \\ \text{QMAm} &= 288,604; \end{aligned}$$

$$\begin{aligned} \text{QMAd} &= \frac{\text{SQAv}}{\text{GLAv}} \\ \text{QMAd} &= \frac{15,376}{24} \\ \text{QMAd} &= 0,641; \end{aligned}$$

$$\begin{aligned} \text{QMR} &= \frac{\text{SQR}}{\text{GLR}} \\ \text{QMR} &= \\ \text{QMR} &= ; \end{aligned}$$

$$\begin{aligned} F &= \frac{\text{QMAm}}{\text{QMR}} \\ F &= \\ F &= . \end{aligned}$$

Tabela 2.2.18: Análise de Variância para comparação das reduções de sacarose.

CV	GL	SQ	QM	F
Avaliadores	24	15,376	0,641	—
Bebidas	4	1154,416	288,604	—
Interação	96	—	—	—
Resíduos	125	—	—	—
TOTAL	249	1294,576		

Tabela 2.2.19: *

Fonte: Autoria própria.

Para calcular ANOVA com fator duplo de repetição no excel basta seguir os passos do Capítulo 7.3.2.

Vemos que o p-valor dos Tratamentos (que representa as reduções de sacarose), é significativo podendo afirmar a 5% de significância que há diferença significativa entre as amostras, então pode ser aplicado um teste de média, que para esse caso que há um grupo controle é utilizado o teste de Dunnett, onde $D\alpha = 2,18$ (analizando a tabela ??), $QMR = 0,536$ e $J = 50$ Substituindo na equação 2.2.11, temos:

$$DMS_{Dunnett} = 2,18 \sqrt{\frac{2(0,536)}{50}} \quad (2.2.13)$$

$$DMS_{Dunnett} = 0,3192.$$

Fazendo a diferença entre as média:

$$\begin{aligned} |2\% - Tradicional| &= |1,62 - 1,54| = 0,12 < DMS \quad \text{Não diferem entre si a } 5\% \text{ de significância,} \\ |4\% - Tradicional| &= |1,74 - 1,54| = 0,2 < DMS \quad \text{Não diferem entre si a } 5\% \text{ de significância,} \\ |7\% - Tradicional| &= |4,46 - 1,54| = 2,92 > DMS \quad \text{Diferem entre si a } 5\% \text{ de significância,} \\ |10\% - Tradicional| &= |6,96 - 1,54| = 5,42 > DMS \quad \text{Diferem entre si a } 5\% \text{ de significância,} \end{aligned}$$

Os resultados permitem concluir que as bebidas lácteas com redução de 2% e 4% de sacarose não são diferenciadas da bebida láctea tradicional em relação ao gosto doce, enquanto as bebidas com redução de 7% e 10% são percebidas como diferentes da tradicional ao nível de significância de 5%. A empresa poderá reduzir o teor de sacarose na bebida láctea achocolatada em até 4% com baixo risco de que os consumidores percebam a redução do gosto doce.

QUESTÃO QUE PRECISA REVER COMO FAZ PELO RSTUDIO - professora vai conversar com a judite e com o cesar sobre ela

2.3 Teste não paramétricos

Os testes não paramétricos analisam as frequências de escolha das amostras. Neste capítulo serão abordados o Teste triangular, Teste duo-trio, Teste de comparação pareada, Teste de ordenação, Teste A ou Não A, Teste tetraédrico e Teste dois em cinco.

2.3.1 Teste triangular

2.3.1.1 Teste triangular não direcional

No teste triangular não direcional o avaliador não é informado sobre qual atributo está sendo avaliado e tem como objetivo verificar se existe uma diferença significativa entre duas amostras com tratamentos diferentes.

Para realização deste teste apresenta-se três amostras codificadas ao avaliador, das quais duas são iguais e uma diferente. As amostras devem ser provadas da esquerda para direita e o avaliador deve assinalar qual é a amostra diferente. As amostras devem ser oferecidas simultaneamente, com combinações apresentadas aleatoriamente, sendo possíveis AAB, ABB, BAA, BBA, ABA, e BAB. A probabilidade de acerto para este caso é de $\frac{1}{3}$ [12]. A Figura 2.3.1 mostra um exemplo de ficha para este método.

Figura 2.3.1: Modelo de ficha para aplicação do teste triangular.

Ficha de avaliação – Teste triangular																				
Avaliador:				Data: _____ / _____ / _____																
Em cada grupo de amostras apresentadas, duas são iguais e uma é diferente. Deguste cuidadosamente cada uma das amostras, na ordem em que são apresentadas, e faça um círculo em volta da amostra diferente																				
<table border="1"> <thead> <tr> <th>Grupo</th> <th colspan="3">Código da amostra</th> </tr> </thead> <tbody> <tr> <td>I</td> <td>928</td> <td>479</td> <td>110</td> </tr> <tr> <td>II</td> <td>171</td> <td>036</td> <td>245</td> </tr> <tr> <td>III</td> <td>352</td> <td>563</td> <td>684</td> </tr> </tbody> </table>					Grupo	Código da amostra			I	928	479	110	II	171	036	245	III	352	563	684
Grupo	Código da amostra																			
I	928	479	110																	
II	171	036	245																	
III	352	563	684																	
Comentários:	<hr/> <hr/>																			

Fonte: DUTCOSKY, 2019.

O número de avaliadores é escolhido de acordo com o nível de sensibilidade requerida para atingir os objetivos do teste, estando relacionada a proporção máxima permitida de discriminadores, erro tipo I (rejeita a hipótese nula) e erro do tipo II (acontece ao aceitar a hipótese nula quando esta é falsa).

A análise e interpretação dos dados são baseadas em estatísticas de testes definidas por duas formas: distribuições binomial ou qui-quadrado.

Para a primeira abordagem, baseada na distribuição binomial, os seguintes parâmetros devem ser definidos para a realização do teste: o risco alfa (α), o risco beta (β) e a proporção máxima permitida de discriminadores (P_d). A proporção máxima permitida de discriminadores (P_d) trata-se da porcentagem de avaliadores em uma população capazes de discriminar qual a amostra difere.

Com tais informações, usando a abordagem pela distribuição binomial, têm-se duas tabelas que podem ser consultadas para que os testes sejam conduzidos. A primeira tabela apresenta o número mínimo de avaliadores necessários para realização do teste, ou seja, o tamanho mínimo amostral (n). Para que este valor seja determinado, define-se os valores de P_d , α e β . A tabela 2.3.1 mostra o número mínimo de avaliadores necessários para execução deste teste.

Tabela 2.3.1: Número mínimo de avaliadores (n) para execução de um teste triangular.

α	$P_d=50\%$	β				
		0,20	0,10	0,05	0,01	0,001
0,20		7	12	16	25	36
0,10		12	15	20	30	43
0,05		16	20	23	35	48
0,01		25	30	35	47	62
0,001		36	43	48	62	81
0,20	$P_d=40\%$	12	17	25	36	55
0,10		17	25	30	46	67
0,05		23	30	40	57	79
0,01		35	47	56	76	102
0,001		55	68	76	102	130
0,20	$P_d=30\%$	20	28	39	64	97
0,10		30	43	54	81	119
0,05		40	53	66	98	136
0,01		62	82	97	131	181
0,001		93	120	138	181	233
:	:	:	:	:	:	:

Tabela 2.3.2: *

Fonte: Adaptado de Dutcosky [p. 451-452].

Após a definição do mínimo de avaliadores necessários (n), coleta-se os dados. Com as informações necessárias, consulta-se uma segunda tabela que informa o número mínimo de respostas corretas para estabelecer diferença significativa entre as amostras. Para o uso dessa tabela, precisa-se do número de participantes (onde espera-se que este valor seja igual ou maior ao tamanho mínimo amostral (n) estabelecido pela primeira tabela) e o valor de α , já definido anteriormente.

Tanto a distribuição binomial como a qui-quadrado comparam a proporcionalidade entre o número de respostas corretas sobre o número total dos julgamentos, consultando o valor na tabela mostrada abaixo. As amostras serão consideradas diferentes, se o valor calculado for igual ou superior ao valor indicado na tabela, ao nível de significância adotado e pode-se calcular o intervalo de confiança da proporção da população capaz de perceber essa diferença, mas se o número de respostas calculadas for menor do que o valor tabelado, as amostras são suficientemente similares [2].

Tabela 2.3.3: Número mínimo de respostas corretas necessárias para estabelecer diferença significativa entre as amostras para o teste triangular.

n	α				
	0,20	0,10	0,05	0,01	0,001
:	:	:	:	:	:
48	20	21	22	25	27
54	22	23	25	27	30
60	24	26	27	30	33
66	26	28	29	32	35
72	28	30	32	34	38
78	30	32	34	37	40
:	:	:	:	:	:

Tabela 2.3.4: *

Fonte: Adaptado de Dutcosky [p. 453-454] Dutcosky

Segundo [p.119] Dutcosky é possível calcular um intervalo para proporção de avaliadores que conseguem discriminar as amostras. Este cálculo é feito a partir da proporção de respostas corretas (P_c) calculada em cada teste, a partir da qual calcula-se a proporção de avaliadores que conseguem discriminar as amostras (P_d). Os intervalos de confiança são dados por:

$$L_s = P_d + z_\beta S_d,$$

$$L_i = P_d - z_\alpha S_d,$$

em que L_s é o limite superior de confiança, L_i é o limite inferior de confiança, P_d é a proporção de avaliadores que conseguem discriminar as amostras, S_d é o desvio padrão de P_d e z_α e z_β são os valores críticos da distribuição normal padrão, dados α e β . Assim, para o cálculo dos intervalos de confiança, primeiramente encontra-se os valores de P_c , P_d e S_d , dados por:

$$P_c = \frac{c}{n}$$

em que c é o número de respostas corretas e n o número de avaliadores e P_d e S_d são calculados

como:

$$P_d = 1,5P_c - 0,5;$$

$$S_d = 1,5\sqrt{\frac{P_c(1 - P_c)}{n}}.$$

Para a abordagem do qui-quadrado em um teste triangular, é utilizado como base a seguinte equação:

$$\chi^2 = \frac{(|O_A - E_A| - 0,5)^2}{E_A} + \frac{(|O_B - E_B| - 0,5)^2}{E_B}, \quad (2.3.1)$$

em que $O_A = X_A$ = número de respostas corretas observadas; $O_B = X_B$ = número de respostas incorretas observadas; $n = X_A + X_B$ total dos dados; p_A = probabilidade de responder corretamente a questão, dada por $\pi = p_A = \frac{1}{3}$ (por ser um teste triangular); E_A = número esperado de respostas corretas dado por $E_A = n \times \pi = n \times p_A = n \times \frac{1}{3} = \frac{X_A + X_B}{3}$; p_B = probabilidade de responder incorretamente a questão, dada por $1 - \pi = p_B = \frac{2}{3}$; E_B = número esperado de respostas incorretas; $E_B = n \times p_B = n \times \frac{2}{3} = \frac{2(X_A + X_B)}{3}$. Assim, a partir dos termos definidos, substituindo na equação 2.3.1, e supondo que $X_A > X_B$ ou o contrário, a estatística de teste dada pela qui-quadrado corrigido é expressa por

$$\chi^2 = \frac{(|4X_A - 2X_B| - 3)^2}{8n},$$

em que X_A é o número de respostas corretas, X_B é o número de respostas incorretas do teste e n o número de participantes do teste.

A rejeição da hipótese H_0 ocorre quando o valor calculado a partir dos dados observados for maior ou igual a um valor tabelado da distribuição de qui-quadrado com um grau de liberdade, com um nível de significância α adotado, caso contrário, a hipótese nula não é rejeitada.

2.3.1.2 Exemplo

Problema 5. Uma indústria de produtos lácteos encomendou uma pesquisa de mercado para descobrir o motivo de suas vendas de muçarela não crescerem, e descobriu que o sabor de seu produto era considerado muito suave e o gosto salgado muito fraco. Uma nova formulação foi desenvolvida com a utilização de um novo coalho e mais 10% de sal para comparação com a muçarela tradicional. Este é um teste de diferença, ou seja, deseja-se controlar o erro alfa, que significa o risco de rejeitar a hipótese nula, quando na verdade não existe a diferença (ou seja, as amostras são iguais e afirmamos que elas são diferentes). Como as alterações podem ter afetado mais que uma característica sensorial o técnico encarregado do projeto decidiu utilizar um teste triangular não direcional. Considerando que o erro α foi mantido em 0,05, o erro β em

0,20 e deseja-se que ao menos 40% da população (pD) perceba a diferença entre as formulações. Qual o número mínimo de avaliadores necessário? Supondo que foi utilizado o número mínimo de avaliadores no teste e que 10 deles acertaram qual era a amostra diferente, qual a conclusão alcançada?

2.3.1.3 Resolução

Para resolver o exercício, primeiramente é preciso deixar claro alguns dados fundamentais, como o α , β e pD (proporção de discriminadores). Observando o enunciado, temos que:

- $\alpha = 0,05$
- $\beta = 0,20$
- $pD = 40\%$

Utilizando a tabela 6.0.1, é possível observar que são necessários ao menos 23 julgadores para a realização adequada da análise sensorial. O segundo passo é analisar o número mínimo de respostas corretas em um teste triangular com 23 julgadores.

Segundo a tabela 6.0.4 o n° mínimo de respostas corretas é 12, maior do que aconteceu no exercício que registrou 10 respostas corretas. Portanto, como o valor tabelado foi maior que o registrado ($12 > 10$), pode-se concluir que não houve diferença entre as formulações.

Outro modo de se resolver é utilizando o método do qui-quadrado. Sabendo-se que o número de respostas corretas (X_a) = 10, é possível calcular o número de respostas erradas (X_b) = $23 - 10 = 13$. Substituindo os valores na fórmula, obtemos:

$$\chi^2 = \frac{(|4X_A - 2X_B| - 3)^2}{8n} = \frac{(|4 * 10 - 2 * 13| - 3)^2}{8 * 23} = 0,6576$$

Comparando com a tabela (VALORES CRÍTICOS DA DISTRIBUIÇÃO QUI-QUADRADOR), considerando graus de liberdade = 1 e $alpha = 0,05$, encontrou-se o valor de 3,841. Portanto, como o valor calculado é menor que o valor tabelado pode-se concluir que não existe diferença significativa, ao nível de 5% de significância, entre as formulações.

2.3.1.4 anexo angular direcional

Este teste também é conhecido como 3-AFC (*Alternative Forced Choice*) e impõe ao avaliador que determine entre três amostras codificadas, qual possui a maior ou menor intensidade de um atributo específico [2].

Para a análise de dados, o mesmo procedimento do anexo angular não direcional é usado, e a probabilidade de acerto da resposta também é de $\frac{1}{3}$ [2].

2.3.2 Exemplo de aplicação

Problema 6. Uma confeitoria famosa pelos seus pastéis de Belém quer aumentar a vida útil de seu produto que é comercializado em embalagens de 6 unidades e que tem no momento 7 dias de vida útil. Tal vida útil é limitada pela perda da crocância. A confeitoria decide testar um outro tipo de material de embalagem que apresenta menor permeabilidade ao vapor de água, mas que aumentará o custo de produção. A analista responsável pelo projeto decide fazer um teste TRIANGULAR AFC para avaliar se existe diferença na crocância dos pastéis de Belém após 10 dias de armazenamento dos produtos em embalagem tradicional e na nova. A empresa quer evitar o erro tipo I de concluir que existe diferença entre os produtos em função das embalagens quando na verdade não há, mas também não quer deixar de utilizar uma nova embalagem que possa aumentar a vida útil. Nesta situação, a analista decide equilibrar os erros tipo I e tipo II e utilizar valores de $\alpha = 0,10$, $\beta = 0,10$. Também estabelece que ao menos 40% da população (pD) deve perceber a diferença na crocância dos pastéis. Com esses parâmetros verificou-se na Tabela 6.0.1 que ao menos 25 avaliadores deveriam participar do teste e obteve-se 15 respostas corretas, em que os avaliadores indicaram que a amostra embalada na nova embalagem era mais crocante e que a amostra em embalagem tradicional era a menos crocante. Qual a conclusão do teste?

2.3.3 Resolução

Para o começo da resolução do problema, é necessário analisar e guardar os dados importantes que são dados no problema, como, o α , β e pD (proporção de discriminadores). Observando o enunciado, temos que:

- $\alpha = 0,10$
- $\beta = 0,10$
- $pD = 40\%$

Utilizando a tabela 6.0.4 verifica-se que nas condições do teste são necessárias 12 respostas corretas para se rejeitar a hipótese nula a 10% de significância e concluir que as amostras são diferentes. Como foram 15 respostas corretas pode-se concluir, a 10% de significância, que a nova embalagem mantém os pastéis de Belém com maior crocância após 10 dias de armazenamento, de modo que a confeitoria pode estabelecer um prazo de validade maior ao produto considerando esta característica sensorial.

Outro meio de resolver esse problema, é através do método do qui-quadrado. Sabendo-se que o nº de respostas corretas (X_a) foi 15, é possível calcular o nº de respostas erradas(X_b), que é igual a $25-15 = 10$. Substituindo na fórmula temos:

$$\chi^2 = \frac{(|4X_A - 2X_B| - 3)^2}{8n} = \frac{(|4 * 15 - 2 * 10| - 3)^2}{8 * 25},$$

$$\chi^2 = 6,845$$

Comparando com a tabela (VALORES CRÍTICOS DA DISTRIBUIÇÃO QUI-QUADRADOR), considerando graus de liberdade = 1 e $\alpha = 0,1$, encontou-se o valor de 2,706. Portanto, como o valor calculado é maior que o valor tabelado pode-se concluir que existe diferença significativa, ao nível de 10% de significância, entre as embalagens.

2.3.4 Teste duo-trio

O teste duo-trio é um teste não direcional simples e facilmente compreendido, no qual três amostras são fornecidas aos avaliadores, sendo uma delas definida como padrão, de acordo com o objetivo da análise. Este método é aplicado para saber se há ou não uma diferença sensorial entre dois produtos, seja ela causada por diferentes fatores, tais como alterações em formulação, processamentos, embalagem, estocagem, entre outros[2, 12].

Para a realização, apresenta-se ao avaliador três amostras, sendo uma padrão e as outras duas codificadas. Uma das amostras codificadas é igual a amostra padrão e a outra é diferente, pedindo-se para o avaliador identificar a amostra igual ao padrão [2]. Um modelo de ficha para aplicação do teste pode ser visto abaixo.

Figura 2.3.2: Modelo de ficha para aplicação do teste duo-trio.

Ficha de avaliação – Teste duo-trio		
Julgador:	Data: ___ / ___ / ___	
Tipo de amostra: Lenço facial dentro de um jarro de vidro		
Instruções:		
<ol style="list-style-type: none"> 1. Aspire cada amostra, iniciando pela amostra Referência. Remova a tampa por um período breve e inale através de aspirações curtas e sequenciais. 2. Depois, avalie as amostras codificadas da esquerda para a direita. Determine qual das duas amostras codificadas é igual à Referência. 3. Indique a amostra colocando um X no quadrado correspondente. 		
REFERÊNCIA	Código: _____ <input type="checkbox"/>	Código: _____ <input type="checkbox"/>
Comentários: _____ _____		

Fonte: DUTCOSKY, 2019.

A probabilidade de acerto ao acaso no teste é de $\frac{1}{2}$ pelo método da escolha forçada e

a ordem de apresentação das amostras pode ser AB e BA, além da apresentação da amostra padrão como referência para a discriminação (em que, P pode ser do tipo P=A, se o padrão for A ou P=B, se o padrão for B) [2].

O número de avaliadores para realizar o teste é definido de acordo com a sensibilidade requerida, que depende dos valores do risco α , risco β e da proporção máxima permitida de discriminadores (p_d). Assim, primeiramente define-se a sensibilidade do teste, e com base na distribuição binomial, utiliza-se uma tabela (mostrada abaixo) para encontrar o número de julgadores necessários para realizar o teste [2].

Tabela 2.3.5: Número mínimo de avaliadores (n) para execução de um teste duo-trio.

α	$P_d=50\%$	β								
		0,50	0,40	0,30	0,20	0,10	0,05	0,01	0,001	
0,40	$P_d=50\%$	2	4	4	6	10	14	27	41	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	
0,001		38	43	51	61	71	83	107	140	
0,40	$P_d=40\%$	4	4	6	8	14	25	41	70	
0,30		5	7	9	13	22	28	49	78	
0,20		5	10	12	19	30	39	60	94	
0,10		14	19	21	28	39	53	79	113	
0,05		18	23	30	37	53	67	93	132	
0,01		35	42	52	64	80	96	130	174	
0,001		61	71	81	95	117	135	176	228	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	

Tabela 2.3.6: *

Fonte: Adaptado de Dutcosky [p. 455-456].

Após aplicação do teste, a análise e interpretação dos resultados é baseada na distribuição binomial ou distribuição qui-quadrado, que determina se o teste apresenta diferenças significativas entre as amostras ou não.

Para os cálculos utilizando como base o qui-quadrado, a equação 2.3.1 também é utilizada como base, a única alteração está no valor da probabilidade de respostas corretas e incorretas, as quais serão dadas por $\pi = p_A = \frac{1}{2}$ e $1 - \pi = p_B = \frac{1}{2}$, respectivamente. Os demais termos utilizados na equação 2.3.1 são os mesmos descritos para o teste triangular, portanto, a equação final encontrada é:

$$\chi^2 = \frac{(|X_A - X_B| - 1)^2}{n},$$

em que X_A é o número de respostas corretas, X_B é o número de respostas incorretas do teste e n o número de participantes do teste.

Se o número de respostas corretas for maior ou igual ao valor tabelado, conclui-se que

existe diferença significativa, mas se o número de respostas corretas for menor que o valor tabelado, conclui-se que as amostras são suficientemente similares. Este teste só verifica se amostras testadas são diferentes, não avaliado se qual é a diferença ou se está é grande ou pequena [2].

Pode-se também calcular o intervalo de confiança da proporção da população que pode perceber a diferença entre as amostras [2].

2.3.4.1 Exemplo

Problema 7. (*DUTCOSKY, 2019*) Um fornecedor de equipamentos argumenta que seu novo equipamento que aromatiza o interior da caixa de lenço umedecido facial é superior à tecnologia atual, que aromatiza diretamente o lenço. O químico deseja confirmar esse argumento antes de realizar um teste com o equipamento.

2.3.4.2 Resolução pela abordagem binomial

Neste exercício deve-se determinar se os dois métodos de aromatização produzem alguma diferença na percepção da fragrância dos lenços após eles terem sido estocados no período de três meses.

O número de avaliadores deve proteger o químico de concluir falsamente que existe diferença, quando ela, na realidade, não existe. O analista sensorial propõe $\alpha=0,05$. Para manter o número de avaliadores num tamanho razoável, sugeriu-se $P_d=40\%$, com $\beta=0,20$.

Iniciou-se a resolução consultando a tabela do número mínimo de avaliadores para o teste duo-trio, representada a seguir como tabela 2.3.7.

Tabela 2.3.7: Número mínimo de avaliadores (n) para execução de um teste duo-trio.

α	$P_d=50\%$	β							
		0,50	0,40	0,30	0,20	0,10	0,05	0,01	0,001
0,40		2	4	4	6	10	14	27	41
:	:	:	:	:	:	:	:	:	:
0,001		38	43	51	61	71	83	107	140
0,40	$P_d=40\%$	4	4	6	8	14	25	41	70
0,30		5	7	9	13	22	28	49	78
0,20		5	10	12	19	30	39	60	94
0,10		14	19	21	28	39	53	79	113
0,05		18	23	30	37	53	67	93	132
0,01		35	42	52	64	80	96	130	174
0,001		61	71	81	95	117	135	176	228
:	:	:	:	:	:	:	:	:	:

Tabela 2.3.8: *

Fonte: Adaptado de Dutcosky [p. 455-456].

Pela tabela 2.3.7 o número mínimo de avaliadores para este teste são 37. Sabe-se que o número de avaliadores utilizados foi 40, estando assim dentro da sensibilidade requerida para o teste.

A descrição do procedimento do teste citou que suficientes quantidades de produtos foram preparadas com ambas tecnologias e os produtos foram armazenados, sob as mesmas condições, por um período de três meses. Sessenta amostras de lenço "A" com a fragrância aplicada diretamente sobre o lenço umedecido e sessenta amostras de lenço "B" com a fragrância aplicada dentro das caixas dos lenços foram preparadas para o teste. Dez repetições das quatro possibilidades de ordem de apresentação das amostras foram utilizadas: para $P=A$, AB e BA; e para $P=B$, AB e BA.

Os lenços foram pegos a partir do centro da caixa e cada lenço foi colocado em uma jarra de vidro selada, uma hora antes da análise. Esse procedimento permitiu que a fragrância migrasse para o *headspace* e o uso dos recipientes fechados reduziu a contaminação no ambiente das cabines.

Apenas 21 avaliadores de 40 julgadores conseguiram discriminar as amostras corretamente. Em seguida, consultou-se a tabela do número mínimo de respostas corretas, dada a seguir.

Tabela 2.3.9: Número mínimo de respostas corretas necessárias para estabelecer diferença significativa entre as amostras para o teste duo-trio.

n	α						
	0,40	0,30	0,20	0,10	0,05	0,01	0,001
2	2	2	-	-	-	-	-
3	3	3	3	-	-	-	-
4	3	4	4	4	-	-	-
5	4	4	4	5	5	-	-
6	4	5	5	6	6	-	-
:	:	:	:	:	:	:	:
40	22	23	24	25	26	28	31
:	:	:	:	:	:	:	:

Tabela 2.3.10: *

Fonte: Adaptado de Dutcosky [p. 457-458].

Como o número de respostas necessárias é 26 e apenas 21 avaliadores conseguiram discriminar as amostras, concluiu-se ao nível de 5% de significância que as amostras não diferem entre si. Assim, mesmo avaliadores treinados não conseguiram notar diferença em relação a nova tecnologia defendida pelo fornecedor de equipamentos, pois tanto os lenços aromatizados diretamente como os indiretamente apresentaram o mesmo resultado sensorial. Desta maneira, a troca de equipamento não trará nenhuma vantagem e portanto não há necessidade de ser implementada, uma vez que não fará diferença em relação ao odor dos lenços.

2.3.4.3 Resolução pela abordagem qui-quadrado

Como o número de respostas corretas obtidas neste teste foram 21 ($X_a=21$), então 19 avaliadores responderam errado ($X_b=19$). Desse modo, pode-se calcular:

Tabela 2.3.11: Número de respostas obtidas pelo teste comparação pareada.

Corretas (X_a)	Incorretas (X_b)	Total (n)
21	19	40

Tabela 2.3.12: *

Fonte: Autoria própria.

$$\chi^2 = \frac{((|X_a - X_b|) - 1)^2}{n} = \frac{((|21 - 19|) - 1)^2}{40} = 0,025.$$

Assim, o valor de χ^2 calculado foi de 0,025, comparando com o valor da tabela ??, o valor da distribuição para uma χ^2 com 1 grau de liberdade e 5% de nível de significância é de 3,841. Desse modo, como χ^2 calculado foi menor que o valor tabelado, concluiu-se ao nível de

5% de significância, que as amostras não diferem entre si, corroborando o resultado obtido pela abordagem binomial.

2.3.4.4 Exemplo

Problema 8. Um engenheiro do setor de controle de qualidade de uma indústria de condimentos quer verificar se um lote de molho de pimenta apresenta diferença sensorial perceptível em relação a um lote controle, pois houve um registro de elevação anormal da temperatura na unidade de concentração do molho. Considerando que a avaliação sensorial dos molhos é uma tarefa difícil devido a super-estimulação das papilas gustativas com a capsaicina, o engenheiro decide realizar um teste DUO-TRIO. O teste é de similaridade pois o profissional quer minimizar o risco de concluir que os molhos de pimenta são iguais, quando na verdade não são. Escolhendo risco alfa em 0,20, risco beta 0,05 e PD igual a 40% ele chega ao valor de 39 como o número mínimo de avaliadores necessário. O teste é realizado com 40 avaliadores e são obtidas 27 respostas corretas. Qual a conclusão obtida com o teste? O que o engenheiro de alimentos deve fazer com o lote de molho de pimenta?

2.3.4.5 Resolução

Para 40 respostas e alfa = 0,2 a Tabela N° MIN DE RESPOSTAS CORRETAS DUO-TRIO indica a necessidade de se obter ao menos 24 respostas corretas para que seja possível rejeitar a hipótese nula e concluir que as amostras são diferentes. No teste sensorial 27 avaliadores conseguiram indicar corretamente qual era a amostra igual ao padrão, portanto há diferença, ao nível de 20%, entre os molhos de pimenta. Considerando que a suspeita de que o sobreaquecimento do molho de pimenta causou alteração sensorial foi confirmada, o lote não deve ser comercializado.

2.3.4.6 Resolução pelo método Qui-quadrado

Considerando que o n° de respostas corretas (X_a) = 27, e o n° de respostas incorretas (X_b) = 40 - 27 = 13, basta substituir na fórmula do qui-quadrado, como pode-se observar abaixo.

$$\chi^2 = \frac{((|X_a - X_b|) - 1)^2}{n} = \frac{((|27 - 13|) - 1)^2}{40} = 4,225.$$

Em seguida, compara-se o valor calculado com o valor da tabela VALORES CRÍTICOS DE DISTRIBUIÇÃO QUI, considerando 1 grau de liberdade e 20% de nível de significância é de 1,642. Desse modo, como χ^2 calculado foi maior que o valor tabelado, concluiu-se que ao nível de 20% de significância, existe diferença entre as amostras.

2.3.5 Teste comparação pareada

Este método é um teste de diferença de atributos direcional, usado para identificar se existe uma diferença entre as amostras com relação a uma característica sensorial específica, sendo informado ao avaliador qual atributo sensorial está sendo avaliado [2]. Também é usado para treinar avaliadores e tem como vantagem não causar fadiga sensorial [12].

Para sua aplicação, são apresentadas duas amostras codificadas ao avaliador e solicita-se que o avaliador assinale qual delas apresenta o atributo em maior intensidade. A ordem das amostras é AB ou BA e a probabilidade de acerto ao acaso é $\frac{1}{2}$. A seguir é mostrado um modelo de ficha para este método.

Figura 2.3.3: Modelo de ficha para aplicação do teste de comparação pareada.

<u>Ficha de avaliação</u>	
Nome:	Data:
Instruções: Você está recebendo duas amostras codificadas de biscoito de polvilho. Por favor, avalie as amostras da esquerda para a direita quanto à dureza e assinale o código referente à amostra mais dura.	
261	843
Comentários: _____	

Fonte: ALVES, 2021.

Nos testes de comparação pareada há uma subdivisão entre teste unilateral (no qual se sabe a priori qual produto poderia apresentar maior intensidade do atributo sensorial - por exemplo, quando se compara dois sucos de laranja com diferentes concentrações de ácido cítrico quanto ao gosto ácido) e teste bilateral (quando não se sabe a priori qual produto pode apresentar maior intensidade do atributo - por exemplo, quando testa-se dois sucos de laranja provenientes de diferentes cultivares de laranja pêra quanto a acidez) [12].

A escolha do número de avaliadores dependerá da sensibilidade do teste, dependendo do risco α , risco β , se o teste é de natureza uni ou bilateral, da proporção máxima permitida de discriminadores (p_d) no teste unilateral ou de respostas para a amostra A e B no teste bilateral (P_{max}) [2].

A proporção máxima permitida de discriminadores é dada por:

$$P_d = 2P_c - 1;$$

A proporção de respostas corretas fica entre três faixas:

$P_{max} < 55\%$ representa um valor baixo

$55\% < P_{max} < 65\%$ valor de tamanho médio

$P_{max} \geq 65\%$ representa um valor grande

A análise dos resultados é feita por uma soma do número de julgadores que acharam a amostra A mais intensa e, o número de julgadores que acharam a amostra B mais intensa. Encontrado o maior número, usa-se uma tabela baseada na distribuição binomial. Deve-se definir se o teste é de natureza uni ou bilateral, pois para cada teste há um tipo de tabela que deve ser utilizada. Se o número de respostas for maior ou igual ao tabelado, conclui-se que as amostras diferem no atributo sensorial, mas, se o número de respostas for menor, o atributo específico é similar nas amostras [2].

2.3.5.1 Exemplo

Para realização do exemplo, utilizou-se dados cedidos pelo Laboratório Multiusuário de Análise Sensorial de Alimentos (LAMASA) da Universidade de São Paulo, *Campus de Pirassununga*.

Problema 9. *Um teste de comparação pareada unilateral foi realizado para diferenciação de iogurtes acidificados e iogurtes normais, no qual solicitou-se que os avaliadores assinalassem qual amostra era mais ácida. Esta análise é um exemplo de teste unilateral, uma vez que há o conhecimento a priori de qual a amostra pode apresentar maior acidez. O número de avaliadores neste teste foi 18 e a preparação e aplicação da análise seguiram todas as condições descritas neste trabalho. Dos participantes do teste, 15 responderam que a amostra de iogurtes acidificados era a mais ácida.*

2.3.5.2 Resolução pela abordagem binomial

A aplicação dos testes de análises sensoriais no Laboratório Multiusuário de Análise Sensorial de Alimentos (LAMASA) da Universidade de São Paulo, *Campus de Pirassununga*, não prevê antecipadamente o número de avaliadores necessários, devido a variação de participação dos alunos e funcionários.

Por isso conduziu-se as análises com os voluntários que se propuseram a participar, não sendo desse modo estabelecido o tamanho amostral considerando a probabilidade máxima de discriminantes e o risco beta. Desse modo, o $n=18$, trata-se do número de avaliadores que participaram espontaneamente do teste.

Consultando a tabela 2.3.13, para $n=18$ e $\alpha=0,05$, tem-se que são necessárias no mínimo 13 respostas corretas para estabelecer a diferença significativa entre as amostras. Como o número de acertos foi 15, então foi possível discriminar o iogurte com ácido em relação ao iogurte normal.

Tabela 2.3.13: Número mínimo de respostas corretas necessárias para estabelecer diferença significativa entre as amostras para o teste de comparação pareada unilateral.

n	α					
	0,50	0,20	0,10	0,05	0,01	0,001
:	:	:	:	:	:	:
16	9	11	12	12	14	15
17	9	11	12	13	14	16
18	10	12	13	13	15	16
19	10	12	13	14	15	17
20	11	13	14	15	16	18
:	:	:	:	:	:	:

Tabela 2.3.14: *

Fonte: Adaptado de Dutcosky [p. 465].

2.3.5.3 Resolução pela abordagem qui-quadrado

Como o número de respostas corretas obtidas neste teste foram 15 ($X_a=15$), então 3 avaliadores responderam errado ($X_b=3$). Desse modo, pode-se calcular:

Tabela 2.3.15: Número de respostas obtidas pelo teste comparação pareada.

Corretas (X_a)	Incorretas (X_b)	Total (n)
15	3	18

Tabela 2.3.16: *

Fonte: Autoria própria.

$$\chi^2 = \frac{((|X_a - X_b|) - 1)^2}{n} = \frac{((|15 - 3|) - 1)^2}{18} = 6,7222.$$

Assim, o valor de χ^2 calculado foi de 6,72222, comparando com o valor da tabela ??, o valor da distribuição para uma χ^2 com 1 grau de liberdade e 5% de nível de significância é de 3,841. Desse modo, como χ^2 calculado foi maior que o valor tabelado, concluiu-se ao nível de 5% de significância, que as amostras diferem entre si, corroborando o resultado obtido pela abordagem binomial.

2.3.5.4 Exemplo

Problema 10. Uma empresa de consultoria em Análise Sensorial prestou um serviço para uma empresa que produz sucos de laranja para verificar se sucos de laranja de dois cultivares diferentes e procedentes de duas fazendas ($F1$ e $F2$) são diferentes em relação a intensidade de sabor típico da fruta. Foi realizado um teste de comparação pareada. O teste é de similaridade,

já que os analistas querem minimizar o erro tipo II, de concluir que as amostras são iguais quando na verdade não são e o teste de hipóteses é bicaudal pois não se conhece as características dos cultivares em relação a intensidade do sabor típico de laranja. Os analistas definem então os valores de alfa = 0,1 e beta = 0,05 e PD=70% e por meio da tabela xxxx verificam a necessidade de fazer o teste com no mínimo 68 avaliadores. Ao realizar o teste com 76 avaliadores obtiveram 48 respostas que indicaram que a amostra da F2 apresentava maior intensidade de sabor de laranja. Qual a conclusão obtida com o teste?

2.3.5.5 Resolução

Para 76 respostas, a Tabela (N° MIN DE RESPOSTAS NECESSARIAS PARA BICAL-DAL) indica a necessidade de se obter ao menos 46 respostas corretas ao nível de alfa = 0,10 para que seja possível rejeitar a hipótese nula e concluir que as amostras de suco de laranja das fazendas F1 e F2 são diferentes em relação a intensidade de sabor. Considerando que 48 avaliadores indicaram o suco da fazenda 2 como sendo o que apresenta sabor mais intenso, há diferença significativa entre os cultivares de laranja.

Outro meio de resolver esse problema é utilizando o método do qui-quadrado, considerando que o nº de respostas corretas (X_a) = 48, e o nº de respostas incorretas (X_b) = 76 - 48 = 28, e substituindo na equação obtemos:

$$\chi^2 = \frac{((|X_a - X_b|) - 1)^2}{n} = \frac{((|48 - 28|) - 1)^2}{76} = 4,75.$$

Assim, o valor de χ^2 calculado foi de 4,75, comparando com o valor da tabela VALORES CRÍTICOS DE DISTRIBUIÇÃO QUI-QUADRADO, o valor da distribuição para uma χ^2 com 1 grau de liberdade e 1% de nível de significância é de 2,706. Desse modo, como χ^2 calculado foi maior que o valor tabelado, concluiu-se ao nível de 1% de significância, que as amostras diferem entre si.

2.3.5.6 Exemplo

Problema 11. A equipe de P & D de uma indústria de produtos de higiene pessoal está trabalhando no desenvolvimento de um creme dental que deve ser lançado com a alegação de produzir maior sensação de refrescância que o tradicional. A nova formulação foi elaborada com maior concentração de mentol e para se certificar de que houve alteração sensorial na percepção de refrescância a equipe decide fazer um teste de COMPARAÇÃO PAREADA no qual o creme dental novo e o tradicional serão comparados. Os técnicos querem se prevenir de lançar um produto que não apresente a característica alegada, portanto decidem por um teste de diferença no qual é controlado o erro tipo I, ou seja o risco de concluir que as amostras são diferentes quando na realidade não são. Este é um teste de hipótese monocaudal já que o produto novo,

com mais mentol deve ser o que irá apresentar maior sensação de refrescância. Utilizando a Tabela 6.0.6, com beta = 0,20, alfa = 0,05 e PD = 65% verificaram a necessidade de obter respostas de pelo menos 70 avaliadores. Considerando que 43 avaliadores assinalaram o creme dental novo como o sendo o que mais apresenta sensação de refrescância qual é a conclusão obtida no teste?

2.3.5.7 Resolução

Observando a tabela 6.0.8, utilizando nível de significância de 5%, vemos que ela indica a necessidade de se obter ao menos 42 respostas corretas para que seja possível rejeitar a hipótese nula e concluir que o creme dental com mais mentol produz uma sensação de refrescância. Considerando que 43 avaliadores indicaram o novo creme dental como mais refrescante, há diferença, há diferença ao nível de 5% e o creme dental novo poderá ser comercializado com a alegação de que produz maior sensação de refrescância.

2.3.5.8 Resolução pelo método do qui-quadrado

Observando os valores dados no exercício, temos que:

- $X_a = 43$
- $X_b = 27$
- $n = 70$

Então:

$$\chi^2 = \frac{((|X_a - X_b|) - 1)^2}{n} = \frac{((|43 - 27|) - 1)^2}{70} = 3,21.$$

Assim, o valor calculado foi de 3,21, comparando com o valor da tabela VALORES CRÍTICOS DE DISTRIBUIÇÃO QUI-QUADRADO, o valor da distribuição para uma 2 com 1 grau de liberdade e 5% de nível de significância é de 3,841. Desse modo, como o valor calculado foi maior que o valor tabelado, concluiu-se ao nível de 5% de significância, que as amostras diferem entre si.

Pelo método do qui-quadrado a resposta não bateu com a comparação de tabelas

2.3.6 Teste de ordenação

O princípio deste método consiste em apresentar três ou mais amostras casualizadas ao avaliador, solicitando-o ordenar de maneira crescente ou decrescente em relação a intensidade do

atributo sensorial em questão. Recomenda-se que os avaliadores façam uma primeira ordenação das amostras e depois a repitam, para verificar se foi feita de maneira correta[2]. A Figura 2.3.4 apresenta um modelo de ficha para este método.

Figura 2.3.4: Modelo de ficha para aplicação do teste de ordenação.

TESTE DE ORDENAÇÃO			
Avalie o odor de cada uma das amostras da esquerda para a direita e ordene-as em ordem crescente de odor oxidado.			
_____	_____	_____	_____
– oxidada			+ oxidada
Comentários: _____ _____ _____			

Fonte: DUTCOSKY, 2019.

O número de avaliadores necessários para a realização deste teste varia conforme o objetivo da análise.

Para determinação de avaliadores aptos a julgar as amostras, é possível calcular a concordância entre a ordem realizada pelo candidato e a prevista, por meio do coeficiente de correlação de *Spearman*, expresso pela Equação 2.3.2:

$$\rho = 1 - \frac{6 \sum D^2}{N(N^2 - 1)}, \quad (2.3.2)$$

Em que: ρ é a medida de correlação, D a diferença entre ordens correspondentes por amostra e N é o número de amostras analisadas.

Quanto mais próximo de 1 for esse valor, mais forte é a concordância. Um teste de hipótese pode ainda ser aplicado para avaliar se a correlação é significativa [2].

Para realizar um teste de ordenação no qual se tem uma ordem pré determinada, o teste de Page pode ser usado. Este teste pressupõe informações sobre a ordem dos produtos e sua estatística é dada por [17]:

$$L = S_1 + 2 * S_2 + 3 * S_3 + \dots + t * S_t$$

em que L é o coeficiente que será comparado a um valor crítico mínimo tabelado com o intuito de verificar se há diferença significativa entre as amostras, S significa a soma das ordenações do produto analisado, e t é o número de tratamentos.

Vale ressaltar que caso o valor final de L seja maior ou igual ao valor tabelado, significará que as amostras diferem entre si, já se o valor de L for menor que o tabelado, as amostras não diferem entre si.

2.3.6.1 Exemplo

Problema 12 (Adaptado de Dutcosky (p. 138-140)). *Deseja-se diminuir a quantidade de açúcar em um bolo para atrair outros nichos de consumidores, entretanto, deseja-se manter o gosto doce deste bolo. Para isso, um teste de ordenação foi aplicado para avaliar quanto de açúcar poderia ser reduzido sem alterar a qualidade do gosto do bolo. Atualmente a formulação contém 10% de açúcar. Testou-se com 7,5% e 5% de açúcar na formulação. No teste de ordenação solicitou-se que o bolo de gosto mais doce fosse ordenado em 1º lugar e o de gosto menos doce em último. O gosto dos bolos foram medidos através do paladar. O teste foi realizado por 48 julgadores, e seu resultado se encontra na tabela xxx. Analise os resultados e verifique se há ou não diferença significativa entre as amostras, pressupondo que é uma hipótese unilateral. Use $\alpha = 0,05$.*

Tabela 2.3.17: Somatório dos julgadores dos bolos.

Amostras	5%	7,5%	10%
Total 48 j	128	84	76

2.3.6.2 Resolução

Por ser uma hipótese unilateral utiliza-se o teste de Page, pois partimos do pressuposto que ao diminuir a quantidade de açúcar, perde-se o gosto adocicado do bolo:

$$L = S_3 + 2 * S_2 + 3 * S_1 + \dots + t * S_t$$

em que L é o coeficiente que será comparado a um valor crítico mínimo tabelado, S significa a soma das ordens atribuídas a uma amostras, e t é o número de amostras.

Segundo o enunciado, verifica-se que $t = 3$, $S_1 = 128$, $S_2 = 84$, $S_3 = 76$, então:

$$L = 76 + 2 * 84 + 3 * 128 = 628$$

O valor $L = 628$ deve ser comparado com os valores críticos tabelados especificamente do teste de Page dada na figura abaixo, no qual o número de julgadores e amostras são requeridos.

Figura 2.3.5: Valores críticos para o teste de Page aplicado ao teste de ordenação quando existe uma ordem prevista ou predeterminada.

Número de julgadores <i>j</i>	Número de amostras (ou produtos) <i>P</i>											
	Nível de significância $\alpha = 0,05$				Nível de significância $\alpha = 0,01$							
	3	4	5	6	7	8	3	4	5	6	7	8
7	91	189	338	550	835	1204	93	193	346	563	855	1232
8	104	214	384	625	950	1371	106	220	393	640	972	1401
9	116	240	431	701	1065	1537	119	246	441	717	1088	1569
10	128	266	477	777	1180	1703	131	272	487	793	1205	1736
11	141	292	523	852	1295	1868	144	298	534	869	1321	1905
12	153	317	570	928	1410	2035	156	324	584	946	1437	2072
13	165	343*	615*	1003*	1525*	2201*	169	350*	628*	1022*	1553*	2240*
14	178	368*	661*	1078*	1639*	2367*	181	376*	674*	1098*	1668*	2407*
15	190	394*	707*	1153*	1754*	2532*	194	402*	721*	1174*	1784*	2574*
16	202	420*	754*	1228*	1868*	2697*	206	427*	767*	1249*	1899*	2740*
17	215	445*	800*	1303*	1982*	2862*	218	453*	814*	1325*	2014*	2907*
18	227	471*	846*	1378*	2097*	3028*	231	479*	860*	1401*	2130*	3073*
19	239	496*	891*	1453*	2217*	3193*	243	505*	906*	1476*	2245*	3240*
20	251	522*	937*	1528*	2325*	3358*	256	531*	953*	1552*	2360*	3406*

Nota: Valores marcados com um asterisco* são valores críticos por aproximação utilizando-se a distribuição normal.

Figura 2.3.6: *

Fonte: ISO 8587: 2006.

Como a tabela acima não possui informações para 48 julgadores, utiliza-se uma fórmula para correção:

$$L' = \frac{12L - 3 * j * t * (t + 1)^2}{t * (t + 1) * \sqrt{j * (t - 1)}} \quad (2.3.3)$$

Substituindo os valores:

$$L' = \frac{12 * 628 - 3 * 48 * 3 * (3 + 1)^2}{3 * (3 + 1) * \sqrt{48 * (3 - 1)}} = 5,31 \quad (2.3.4)$$

Para um intervalo de confiança de 95%, o valor crítico da distribuição normal padrão é $z = 1,64$, portanto, L' é maior que o valor crítico ($5,31 > 1,64$), rejeitando-se a hipótese nula e concluindo que os julgadores identificaram diferenças entre as amostras, e que a ordem assumida a priori foi verificada.

Para comparar as amostras, pode-se consultar a tabela de Christensen

Tabela 2.3.18: Tabela de Christensen - Teste de ordenação.

Número de avaliadores	Número de amostras									
	3	4	5	6	7	8	9	10	11	12
43	18	23	29	34	39	44	50	55	60	66
44	18	24	29	34	40	45	50	56	61	66
45	19	24	29	35	40	46	51	56	61	67
46	19	24	30	35	40	46	51	57	62	67
47	19	24	30	35	41	46	52	57	63	68
48	19	25	30	36	41	47	52	58	64	69
49	19	25	30	36	42	48	54	59	65	70
50	19	25	31	37	42	48	54	59	65	70
:	:	:	:	:	:	:	:	:	:	:

Tabela 2.3.19: *

Adaptado de Dutcosky [p. 472].

Como observado na tabela, considerando um erro de 5%, a diferença crítica entre os totais de ordenação é de 19, portanto, todas as amostras que diferem entre si por um valor maior ou igual a 19 são significativamente diferentes.

Portanto, comparando as amostras obteve-se:

$$|5\% - 7,5\%| = |128 - 76| = 52 > 19 \quad \text{Diferem entre si a } 5\% \text{ de significância,}$$

$$|5\% - 10\%| = |128 - 84| = 44 > 19 \quad \text{Diferem entre si a } 5\% \text{ de significância,}$$

$$|7,5\% - 10\%| = |76 - 84| = 8 < 19 \quad \text{Não diferem entre si a } 5\% \text{ de significância.}$$

Assim, ao nível de 5% de significância tem-se que os bolos com 7,5% e 10% de açúcar possuem o mesmo gosto adocicado, enquanto o bolo com 5% de açúcar possui um gosto adocicado significativamente menor que os outros dois. Portanto, conclui-se que é possível reduzir a quantidade de açúcar de 10% para 7,5% da formulação inicial, mantendo o mesmo gosto do bolo.

Quando não existe uma ordem predeterminada das concentrações do atributo avaliado, há a necessidade de realizar uma primeira análise para comprovar a existência de diferença entre as amostras, para isso utiliza-se o teste de qui-quadrado, que segundo, trata-se de uma análise do teste F da estatística básica, aplicado às ordens das amostras por cada avaliador, por meio da Equação 3.6.1:

$$F_{teste} = \frac{12}{Av.t(t+1)}(S_1^2 + \dots + S_p^2) - 3Av(t+1), \quad (2.3.5)$$

em que: A_v é o número de avaliadores, t o número de tratamentos e S_i a soma das ordens atribuídas ao tratamento (amostra) i .

O teste de Friedman pode ser aplicado para análise de dados de testes realizados com um pequeno grupo de amostras, além de não precisar que as variâncias sejam homogêneas.

Os resultados podem ser avaliados por valores tabelados da distribuição qui-quadrado ou com base no p-valor. Se o p-valor encontrado for menor que o nível de significância estabelecido, há diferença entre as amostras estudadas, caso contrário, não há. Adicionalmente, testes para a comparação entre as somas das ordens atribuídas para cada amostra podem ser realizados [10].

2.3.6.3 Exemplo

Problema 13 (Adaptado de Dutcosky (p. 198-199)). *Um teste de ordenação foi utilizado para ordenar quatro amostras de óleo em ordem crescente de odor oxidado.*

1=amostra menos oxidada,

4=amostra mais oxidada.

A ficha de aplicação do teste está apresentada na figura 2.3.7 e os resultados na tabela 2.3.20. Analise os resultados e verifique se há diferença significativa entre as amostras. Aqui a pressuposição realizada é que não se sabe qual amostra tem o maior ou menor odor oxidado. Use $\alpha = 0,05$.

USAR FIGURA 2.3.4

Figura 2.3.7: Modelo de ficha para aplicação do teste de ordenação.

TESTE DE ORDENAÇÃO			
Avalie o odor de cada uma das amostras da esquerda para a direita e ordene-as em ordem crescente de odor oxidado.			
_____	_____	_____	_____
- oxidada			+ oxidada
Comentários:			
<hr/> <hr/> <hr/> <hr/>			

Figura 2.3.8: *
Fonte: Adaptado de Dutcosky [p. 198].

Tabela 2.3.20: Julgamentos obtidos no teste de ordenação para o odor oxidado de óleos.

Avaliador	Marca A	Marca B	Marca C	Marca D
Av1	3	1	2	4
Av2	1	3	2	4
Av3	2	1	3	4
Av4	4	2	1	3
Av5	1	2	3	4
Av6	2	1	3	4
Av7	2	1	3	4
Av8	4	1	2	3
Av9	3	1	2	4
Av10	3	1	2	4
Av11	4	1	2	3
Av12	2	1	3	4
Av13	2	1	3	4
Av14	3	1	2	4
Av15	3	1	2	4
Av16	2	1	3	4
Av17	3	1	4	2
Total	44	21	42	63

Tabela 2.3.21: *

Fonte: Adaptado de Dutcosky [p. 198].

2.3.6.4 Resolução

Para comparação das amostras, um teste de ordenação bilateral foi realizado, com base em:

$$F_{\text{teste}} = \frac{12}{Av \times t \times (t+1)} \times (S_1^2 + \dots + S_p^2) - 3 \times Av \times (t+1),$$

em que Av é o número de avaliadores, t é o número de amostras, e S_i a soma das ordens atribuídas a amostra i .

Segundo o enunciado, tem-se $Av = 17$, $t = 4$, $S_1 = 44$, $S_2 = 21$, $S_3 = 42$ e $S_4 = 63$, então:

$$\begin{aligned} F_{\text{teste}} &= \frac{12}{17 \times 4 \times (4+1)} ((44)^2 + (21)^2 + (42)^2 + (63)^2) - 3 \times 17 \times (4+1) \\ &= \frac{12}{340} \times (8110) - 255 \\ &= 286,235 - 255 \\ &= 31,235; \end{aligned}$$

Assim,

$$F_{\text{teste}} = 31,235.$$

O valor $F_{\text{teste}} = 31,235$ deve ser comparado com um valor tabelado de uma distribuição dada na tabela 2.3.22, na qual o valor tabelado da distribuição χ^2 com 3 grau de liberdade (ν) e 5% de nível de significância é de 7,815. Desse modo, como χ^2 calculado foi maior que o valor tabelado, conclui-se, ao nível de 5% de significância, que as amostras diferem entre si em relação ao odor oxidado.

Tabela 2.3.22: Valores críticos da distribuição qui-quadrado com ν graus de liberdade.

ν	Probabilidade (p)												
	0,990	0,980	0,975	0,950	0,900	0,800	0,700	0,500	0,300	0,200	0,100	0,050	0,040
1	0,000	0,001	0,001	0,004	0,016	0,064	0,148	0,455	1,074	1,642	2,706	3,841	4,218
2	0,020	0,040	0,051	0,103	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	6,438
3	0,115	0,185	0,216	0,352	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	8,311
4	0,297	0,429	0,484	0,711	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	10,026
:	:	:	:	:	:	:	:	:	:	:	:	:	:

Tabela 2.3.23: *

Fonte: Adaptado de Dutcosky [p. 460].

Para comparação entre as amostras, pode-se consultar a tabela de Christensen 2.3.24, para qual o número de avaliadores e o número de amostras são requeridos.

Tabela 2.3.24: Tabela de Christensen - Teste de ordenação.

Número de avaliadores	Número de amostras									
	3	4	5	6	7	8	9	10	11	12
12	10	12	15	18	21	23	26	29	32	35
13	10	13	16	19	22	24	27	30	33	38
14	10	13	16	19	22	25	28	31	34	37
15	11	14	17	20	23	26	29	32	36	39
16	11	14	17	21	24	27	30	33	37	40
17	11	15	18	21	25	28	31	34	38	41
18	12	15	19	22	25	29	32	35	39	42
19	12	16	19	23	26	29	33	36	40	44
20	12	16	20	23	27	30	34	37	41	45
:	:	:	:	:	:	:	:	:	:	:

Tabela 2.3.25: *

Adaptado de Dutcosky [p. 472].

Ao consultar a tabela, a diferença crítica entre os totais de ordenação a 5% é de 15. Desse modo, comparando as amostras, obteve-se:

- $|A - B| = |44 - 21| = 23 > 15$ Diferem entre si a 5% de significância,
 $|A - C| = |44 - 42| = 02 < 15$ Não diferem entre si a 5% de significância,
 $|A - D| = |44 - 63| = 19 > 15$ Diferem entre si a 5% de significância,
 $|B - C| = |21 - 42| = 21 > 15$ Diferem entre si a 5% de significância,
 $|B - D| = |21 - 63| = 42 > 15$ Diferem entre si a 5% de significância,
 $|C - D| = |42 - 63| = 21 > 15$ Diferem entre si a 5% de significância.

Que pode ser resumido em:

Amostras	Somas das ordens	
D	63	a
A	44	b
C	42	b
B	21	c

em que amostras seguidas de mesma letra não diferem entre si. Assim, ao nível de 5% de significância, tem-se que a amostra D foi a que apresentou o maior odor de oxidada, diferindo significativamente das demais e, a amostra B, a que apresentou o menor odor de oxidada, diferindo das demais. As amostras A e C não apresentaram diferenças entre si em relação ao odor oxidado, mas diferiram de D e B.

2.3.6.5 Exemplo

Problema 14. Uma indústria de conservas vegetais contratou uma empresa de assessoria em Análise Sensorial para verificar se a intensidade de sabor de morango da geléia de morango produzida pela indústria era similar a de outras 3 empresas líderes de venda no mercado. Um Teste de ordenação foi aplicado e vinte e quatro avaliadores analisaram as amostras, usando uma ficha em que a posição 1 = menos sabor de morango e a posição 4 = mais sabor de morango. Os resultados encontram-se na Tabela 2.3.27 abaixo. Verifique se há diferença significativa entre a geléia da indústria interessada e as das empresas líderes de venda.

Tabela 2.3.26: Resultados do teste de ordenação da intensidade de sabor de morango em geleias de morango.

Avaliadores	Empresa interessada	Empresa Líder A	Empresa Líder B	Empresa Líder C
1	2	3	1	4
2	1	3	2	4
3	3	1	2	4
4	1	3	2	4
5	4	3	1	2
6	4	3	2	1
7	3	4	1	2
8	2	3	1	4
9	3	2	1	4
10	3	4	2	1
11	3	2	1	4
12	4	2	1	3
13	3	2	1	4
14	4	2	1	3
15	3	1	2	4
16	2	3	1	4
17	4	2	1	3
18	3	4	1	2
19	3	2	1	4
20	4	2	1	3
21	4	2	1	3
22	4	2	1	3
23	3	2	1	4
24	3	2	1	4
Total amostra	73	59	30	78

Tabela 2.3.27: *
Fonte: Autoria Própria

2.3.6.6 Resolução

Como não se sabe qual amostra possuí o sabor mais ou menos intenso de morango, utilizando o teste F de Friedman. Substituindo os valores na fórmula, temos:

$$\begin{aligned}
 F_{\text{teste}} &= \frac{12}{24 \times 4 \times (4+1)} ((73)^2 + (59)^2 + (30)^2 + (78)^2) - 3 \times 24 \times (4+1) \\
 &= \frac{12}{480} \times (15794) - 216 \\
 &= 394,84 - 360 \\
 &= 34,84;
 \end{aligned}$$

Assim,

$$F_{\text{teste}} = 34,84.$$

Analizando a Tabela ??, é observado que o valor crítico para $F_{tab} = 7,74$, ou seja, F calculado \in a região crítica, ou seja, a 5% de significância, rejeita-se H_0 , onde H_0 é a hipótese de que as amostras são iguais, podendo concluir que há diferença significativa entre as amostras.

Como há diferença, é possível aplicar um valor crítico de Christensen para verificar quais amostras diferem entre si. Na Tabela 2.3.29, temos que se a diferença entre as médias for maior ou igual que 18, as amostras diferem entre si.

Tabela 2.3.28: Teste de ordenação - Tabela de Christensen

Número de julgadores	Número de amostras									
	3	4	5	6	7	8	9	10	11	
:	:	:	:	:	:	:	:	:	:	:
18	12	15	19	22	25	29	32	35	39	
19	12	16	19	23	26	29	33	36	40	
20	12	16	20	23	27	30	34	37	41	
21	13	16	20	24	27	31	35	38	42	
22	13	17	21	24	28	32	35	39	43	
23	13	17	21	25	29	33	36	40	44	
24	13	18	21	25	29	33	37	41	45	
:	:	:	:	:	:	:	:	:	:	

Tabela 2.3.29: *

Fonte: Adaptado de CHRISTENSEN et al., 2006.

Comparando as amostras, temos:

$$\text{Empresa interessada} - A = 73 - 59 = 14 < 18$$

$$\text{Empresa interessada} - B = 73 - 78 = 5 < 18$$

$$\text{Empresa interessada} - C = 73 - 30 = 43 > 18$$

Organizando com letras, onde medias seguidas de mesma letra não diferem entre si a 5% de significância no teste F de Friedman, temos:

Líder C 78 a

Empresa 73 ab

Líder A 59 b

Líder B 30 c

É possível resolver esse problema no programa Rstudio. O primeiro passo é carregar os pacotes necessários para a execução dos comandos.

O 2º passo é carregar o banco de dados, que seria a tabela do exercício. É possível digitar elemento por elemento ou importar ele diretamente do exel no formato csv. Nesse caso, foi

importado os dados da tabela de excel.

3º passo é reestruturar o banco de dados e renomear ele, de modo com que todas as empresas virem uma única variável, ou seja, transformando a tabela em 3 colunas: avaliadores, empresa e nota.

4º passo é transformar a variável avaliador em um fator

5º passo é a realização do teste de friedman, utilizando o comando `friedman.test`

Como o resultado do p-valor foi menor que 0,05, conclui-se que existe diferença significativa entre as empresas. Para descobrir qual das empresas está diferenciando, utiliza-se um post-hoc de bonferroni, o qual busca corrigir os valores de teste de hipótese.

Comparando os valores do p-valor da empresa interessada com as empresas líderas A,B e C, é possível concluir que existe diferença significativa entre a empresa interessada e a empresa B, uma vez que seu p-valor deu significamente pequeno, já as empresas A e C possuem p-valor próximo de 1, indicando que são semelhantes entre si.

2.3.6.7 Exemplo

Problema 15. *Sucos de caju com diferentes percentagens de adição de ácido cítrico foram utilizadas para selecionar avaliadores capazes de reconhecer diferenças na intensidade do gosto ácido da bebida por meio do Teste de ordenação. Na Tabela 2.3.30 estão os resultados de 6 pessoas; verifique quais avaliadores podem ser considerados aptos utilizando o coeficiente de correlação de Spearman.*

Tabela 2.3.30: Resultados de avaliadores que avaliaram a intensidade de doçura de sucos de caju em Teste de ordenação

Sucos	Ordenação	Av1	D2	Av2	D2	Av3	D2	Av4	D2	Av5	D2	Av6	D2						
1%	1	1	0	0	4	-39	1	0	0	2	-11	2	-11	3	-24				
2%	2	2	0	0	1	1	1	2	0	0	1	1	1	1	1				
4%	3	4	-11	3	0	0	3	0	0	4	-11	3	0	0	2	1	1		
6%	4	3	1	1	2	2	4	4	0	0	3	1	1	4	0	0	0		
8%	5	5	0	0	6	-11	5	0	0	5	0	0	5	0	0	5	0	0	
10%	6	6	0	0	5	1	1	6	0	0	6	0	0	6	0	0	6	0	0
		2		16		0		4		2		6							

2.3.6.8 Resolução

Como deseja-se determinar a concordância entre duas ordens (a realizada pelo candidato e a prevista), é necessário calcular o coeficiente de spearman, dado pela equação 2.3.2, assim, para cada avaliador, temos:

$$\rho_1 = \frac{1-12}{210} = 0,943$$

$$\rho_2 = \frac{1-96}{210} = 0,543$$

$$\rho_3 = \frac{1-0}{210} = 1$$

$$\rho_4 = \frac{1-24}{210} = 0,886$$

$$\rho_5 = \frac{1-12}{210} = 0,943$$

$$\rho_6 = \frac{1-36}{210} = 0,829$$

Com o resultado de cada cálculo e, levando em conta que quanto mais próximo de 1 a correlação, mais concordante são as respostas, pode-se concluir que o julgador 3 obteve o melhor julgamento, 1, 4, 5 e 6 obtiveram boas concordâncias e o avaliador 2 a pior concordância para discriminação do suco de caju com diferentes porcentagens de adição de ácido cítrico.

2.4 Teste A ou NÃO A

Este método é aplicado para avaliar amostras que possuem variações de aparência ou de gosto remanescente, sendo usado quando os testes duo-trio e triangular não são aplicáveis [2].

Para realizá-lo apresenta-se várias vezes a amostra A como referência para o avaliador até que ele possa reconhecê-la. Após, aleatoriamente apresenta-se amostras que podem ser "A" ou "NÃO A" para que identifique a amostra "A" [2].

Para isso, o julgador recebe um modelo de ficha de aplicação (ilustrado na figura 2.4.1) e pede-se que prove da esquerda para a direita verificando a familiarização com a amostra "A" [2].

Figura 2.4.1: Modelo de ficha para aplicação do teste A ou NÃO A.

TESTE “A” ou “NÃO A”		
Nome:	Data:	
Após familiarizar-se com a amostra “A”, prove cuidadosamente da esquerda para a direita as amostras codificadas. Após provar, anote quais amostras são “A”, bebendo água entre a prova de uma amostra e outra.		
Nº da amostra	“A”	“NÃO A”
-----	-----	-----
-----	-----	-----
-----	-----	-----
-----	-----	-----
Comentários: ----- -----		

Fonte: ALVES, 2021.

É recomendado uma quantidade de 20 avaliadores selecionados para realização do teste. A análise dos resultados é feita utilizando-se a distribuição qui-quadrada [2].

Problema 16. (*Adaptado de Dutcosky [p. 122-124]*) Um profissional de desenvolvimento de produtos está pesquisando adoçantes alternativos para bebida com 5% de açúcar em sua formulação. Testes preliminares de gosto estabeleceram 0,1% do novo adoçante como o nível equivalente a 5% de sacarose, mas também demonstraram que, se mais de uma amostra é apresentada ao mesmo tempo, a discriminação é prejudicada por causa do efeito residual da docura e de outros gostos e fatores de sensação bucal. O profissional deseja saber se as duas bebidas se distinguem pelo sabor. Vinte avaliadores recebem aleatoriamente cinco amostras “A”(bebida com adoçante) e cinco amostras “NÃO A”(bebida com sacarose), sendo-lhes recomendado que determinem quais amostras são “A”. Os resultados obtidos são representados na tabela 2.4.1.

Tabela 2.4.1: Julgamentos obtidos no teste "A"ou "NÃO A"de bebidas.

AMOSTRAS			
RESPOSTAS	A	NÃO A	TOTAL
A	60	35	95
NÃO A	40	65	105
TOTAL	100	100	200

Tabela 2.4.2: *

Fonte: Adaptado de Dutcosky [p. 123].

Para aplicar a fórmula, faz-se necessário calcular a frequência esperada para as amostras que são iguais e diferentes a amostra A, assim:

Frequência esperada para A:

$$\frac{100 \times 95}{200} = 47,5.$$

Frequência esperada para Não A:

$$\frac{105 \times 100}{200} = 52,5.$$

Com estes dados, calculou-se:

$$\begin{aligned}\chi^2 &= \frac{(60 - 47,5)^2}{47,5} + \frac{(35 - 47,5)^2}{47,5} + \frac{(40 - 52,5)^2}{52,5} + \frac{(65 - 52,5)^2}{52,5} \\ \chi^2 &= 12,53\end{aligned}$$

Para a análise dos resultados, consulta-se a tabela ?? da distribuição qui-quadrado, com 1 grau de liberdade e $\alpha = 0,050$, onde o valor de χ^2 tabelado é 3,841. Assim, como o valor calculado é maior que o valor tabelado, há diferença significativa entre a bebida adoçada com sacarose e a bebida adoçada com o novo adoçante.

2.4.1 Teste tetraédrico

Este método é aplicado para verificar diferença significativa entre duas amostras que receberam tratamentos diferentes [2].

O teste ainda está em desenvolvimento, mas, seu princípio é apresentar quatro amostras ao avaliador, em que duas pertencem ao tratamento A e as outras duas ao tratamento B. Então, solicita-se aos avaliadores que as agrupem as amostras duas a duas em dois conjuntos similares [2].

Esta metodologia é aplicável tanto em testes nos quais o atributo avaliado é revelado ao avaliador, como nos casos em que não há essa revelação [2].

Para a apresentação das amostras têm-se as seguintes combinações AABB, ABAB, ABBA, BABA, BAAB. Neste caso a probabilidade de acerto é de $\frac{1}{3}$, pois inicia-se pela primeira amostra, resultando em uma chance de um acerto em três amostras faltantes [2].

Para análise dos resultados utiliza-se os mesmos critérios do teste triangular, citados em 2.3.1.1. Se o número de respostas corretas for maior ou igual ao valor tabelado, há uma diferença significativa entre as amostras, mas se o número de respostas corretas for menor do que o valor tabelado, as amostras são suficientemente similares [2].

2.4.2 Teste dois em cinco

Este método também verifica se há uma diferença significativa entre duas amostras que recebem tratamentos diferentes.

Nele, cinco amostras codificadas são apresentadas aos avaliadores, sendo duas de um tipo e três de outro. Pede-se ao avaliador que identifique quais são as duas amostras que diferem das outras três remanescentes [2].

Quando o número de avaliadores for menor do que 20, a ordem de apresentação das amostras é selecionada ao acaso, com as combinações possíveis: AAABB, BBBAA, AABAB, BBABA, ABAAB, BABBA, BAAAB, ABBBA, AABBA, BBAAB, ABABA, BABAB, BABA, ABBAB, ABBAA, BAABB, BABAA, ABABB, BBAAA e AABBB [2].

A probabilidade de acerto ao acaso da primeira amostra de duas corretas é de $\frac{2}{5}$, e de acerto da segunda amostra correta é de $\frac{1}{4}$. Como estes dois eventos são dependentes entre si, ao multiplicá-los a probabilidade geral de acerto da resposta certa ao acaso é de $\frac{1}{10}$ [2]. por exemplo

2.5 Teorias sobre os testes discriminativos

Os testes discriminativos tem como principal objetivo determinar se existe ou não diferenças entre amostras facilmente confundíveis, e para se escolher o melhor teste em cada caso de estudo, é importante compreender as teorias básicas, como a moldagem Thurstoniana (D primo), a teoria de detecção de sinal e a abordagem R-Index. Essas três teorias permitem um melhor entendimento dos fatores que contribuem na validade dos testes de diferença, na sensibilidade e confiabilidade.

2.5.1 Teoria da detecção de sinal

A fundamentação teórica desta teoria é a mesma da modelagem thurstoniana, a principal diferença está no método utilizado em cada. Na detecção de sinal os testes de escolha forçada não são usados para estabilizar o critério de análise [2].

Nesta teoria, um diagrama representa o nível dos impulsos nervosos que o cérebro recebe, onde o eixo X é o tempo e o eixo Y a intensidade dos impulsos nervosos viajando no cérebro [2].

Todo conjunto de impulsos nervosos ocorre com alguns dos nervos disparando sem razão, e quando isso estes nervos disparam chama-se de "ruído". O cérebro separa o sinal do ruído desenhando uma linha virtual, se a intensidade do sinal estiver acima desta linha, será processado como sinal, mas se tiver abaixo será considerado como ruído [2].

Há muitos preconceitos e influências que podem afetar a resposta de uma pessoa, além da sensação real que ela experimenta. Por exemplo, eles podem esperar uma mudança na sensação e antecipar o nível em que algo será perceptível. Por outro lado, uma pessoa pode adotar uma postura muito conservadora e querer ter certeza de que algo é claramente perceptível antes de responder.

Em um experimento clássico de detecção de sinal, dois níveis de um estímulo devem ser avaliados, por exemplo, um estímulo de fundo ou em branco chamado de “Ruído”, um nível fraco, mas mais alto de intensidade do estímulo ou limite chamado de “sinal”. O observador teria que responder “Sim, acho que é o caso do sinal” ou “Não, acho que é o caso do ruído” (Semelhante ao teste A, não-A).

Quando o julgador decidir que a sensação é mais forte do que um certo nível, a resposta será “sinal” e se for mais fraca do que um certo nível, será dada uma resposta de “ruído”. Eles não sabem se é uma tentativa de sinal ou ruído, eles apenas respondem com base em quão forte é a sensação para eles. Ao longo de muitas tentativas as sensações de sinal e ruído são normalmente distribuídas com variância igual.

A escolha do avaliador não depende somente dos botões gustativos, depende também de como está o psicológico no momento da degustação, se a pessoa tiver passado por um dia ruim ou não está confiante, irá afetar na sua resposta. Por exemplo, é apresentado duas geleias de morango A e B para um julgador, sendo a A a mais doce e B a menos doce, nesse momento ele pode dizer com toda certeza de que a geleia A é a mais doce, porém, no outro dia o mesmo avaliador não consegue responder com certeza qual é a mais doce.

2.5.2 R-Index

O R-Index investiga os vários mecanismos da memória, sendo uma medida paramétrica e definido como a probabilidade de escolha do sinal em comparações pareadas como o ruído [2].

O teste é realizado apresentando ao avaliador dois produtos em ordem aleatória, onde um produto é o controle e outro reformulado com ingredientes de menor custo. O produto controle é o sinal e o reformulado o ruído [2].

Pede-se ao avaliador que responda se é o produto é o controle ou o reformulado e se tem certeza ou não. Com os resultados supõe-se que foram realizadas comparações pareadas de todos os sinais e ruídos [2].

Se forem obtidos 50% de acerto, o avaliador não conseguiu discriminar as duas amostras e se todas as respostas estiverem certas, significa uma discriminação perfeita [2].

Capítulo 3

MÉTODOS DESCRIPTIVOS

3.1 Introdução

Este capítulo irá apresentar os principais métodos descritivos utilizados em análises sensoriais, os quais possuem o objetivo geral de avaliar os aspectos qualitativos e quantitativos, caracterizando as propriedades sensoriais dos produtos, tal como apresentado na introdução.

3.2 Teste de escala

A utilização de escalas para a determinação da intensidade dos aspectos estudados em produtos alimentícios é muito comum, tal ferramenta permite a quantificação do grau de diferença entre amostra, do grau de aceitação do avaliador (em testes afetivos), e da intensidade de determinado atributo sensorial, tudo isso através de números ou palavras. Para a realização desse método, faz-se necessário a utilização de avaliadores treinados e habilidosos, além disso, as escalas podem ser de diversos tipos, e são classificadas de acordo com a sua estrutura, posição, polaridade, número de atributos analisados e tipo de avaliação [2]

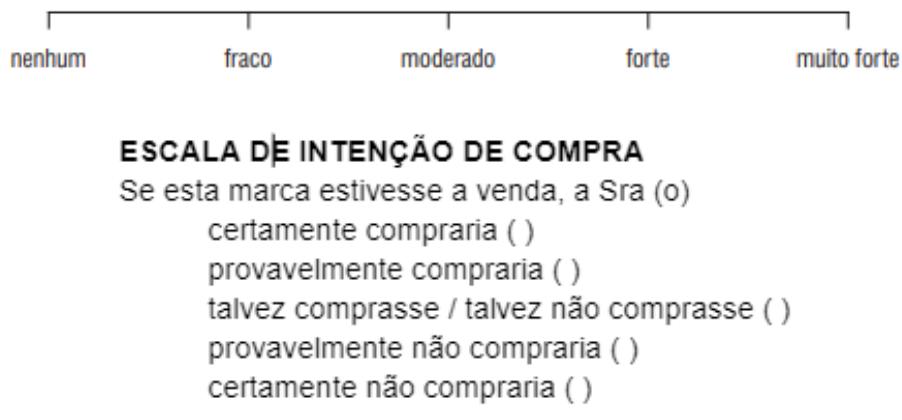
Os testes de escala dão a intensidade da sensação e também a direção das diferenças entre as amostras. Observando as escalas é possível enxergar o quanto uma amostra diferencia uma da outra e qual amostra apresenta maior intensidade do atributo sensorial que está sendo analisado, porém, exigem um maior treinamento e habilidade do avaliador, pois eles devem ser, no mínimo, consistentes e reproduzíveis [2]

É muito importante não utilizar termos redundantes nas escalas, evitando causar confusão e desconforto no avaliador, por exemplo, quando os julgadores estão avaliando um bife e devem responder quanto a percepção de maciez quanto a dureza da carne, seria muito mais claro decidir utilizar o termo "dureza" ou o termo "macia".

3.2.1 Estrutura

As classificadas com relação a sua estrutura são divididas em dois grupos: estruturadas e não estruturadas. As estruturadas, são escalas que possuem intervalos iguais entre suas categorias, podendo ser verbal/nominal, numérica, mista ou facial, onde no primeiro caso, são classificados por adjetivos utilizados no dia a dia, tais como fraco, forte, moderado, muito forte, entre outros. Já nas escalas estruturadas numéricas, os intervalos são divididos por uma sequência de números, enquanto a mista é uma escala que une tanto a verbal quanto a numérica. Para melhor compreensão, observe os exemplos de escalas verbais abaixo:

Figura 3.2.1: Exemplos de escalas verbais



Fonte: Adaptado de [2].

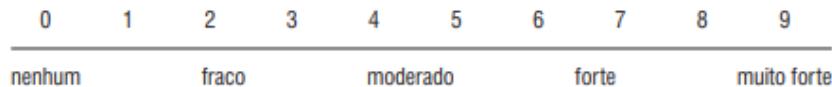
Algumas escalas apresentam maior vantagem que outras, como por exemplo, a escala numérica possui maior precisão, pois contém intervalos constantes, enquanto a escala verbal pode não ser muito vantajosa nesse sentido, por não ser possível definir bem os intervalos entre uma categoria e outra, ou seja, não há como afirmar com certeza que o intervalo entre "fraco" e "moderado" será igual ao intervalo entre "forte" e "muito forte".

Figura 3.2.2: Exemplos de escalas numérica



Fonte: [2].

Figura 3.2.3: Exemplos de escalas mistas



Fonte: [2].

As escalas faciais, são aquelas formadas por imagens, de tal forma que o avaliador deve marcar a imagem que melhor corresponde com a sua avaliação sensorial do produto. Em alguns casos, podem vir acompanhadas de números e palavras chaves, como ruim, bom, moderado, entre outras, para melhor compreensão.

Figura 3.2.4: Exemplos de escalas faciais



Fonte: [2].

Figura 3.2.5: Exemplos de escalas faciais

Marque a carinha que mais representa o que você achou do _____				
Detestei 1	Não gostei 2	Indiferente 3	Gostei 4	Adorei 5

Fonte:

Ao utilizar o método de escalas estruturadas, sugere-se que estas não sejam nem muito pequenas, pois fornecem pouca capacidade de discriminação, nem muito longas, pois aumentaria muito a variabilidade de respostas entre os avaliadores. Portanto, o ideal é que os pesquisadores identifiquem primeiro quantas categorias o participante é capaz de identificar, e diferenciar conscientemente no produto, para então estipular uma escala duas vezes mais longa. De modo geral, a maior parte dos pesquisadores utilizam escalas de nove categorias, mas podem variar entre cinco a quinze categorias [2].

Já as escalas não estruturadas são representadas por uma linha com apenas duas categorias em seus extremos que representam a intensidade do atributo avaliado, por exemplo os adjetivos "muito fraco" e "muito forte". Deste modo, o avaliador deve fazer uma marcação em algum ponto da reta que melhor represente a intensidade da sua sensação ao avaliar sensorialmente o produto, após isso, os dados serão avaliados pelos pesquisadores, medindo a distância da marcação a partir do extremo esquerdo (geralmente utiliza-se a unidade de medida em centímetros). Costuma-se utilizar linhas entre nove e 15 centímetros [2].

Figura 3.2.6: Exemplo de escala não estruturada



Fonte:[2].

A principal desvantagem da escala não estrutural, é a dificuldade que o participante pode vir a ter por não conseguir ser consciente das repetições de resposta, pois as posições em que as marcações foram feitas na reta não são facilmente lembradas quando comparadas a um número. Mas existem vantagem nesse método, tais como a diminuição dos erros psicológicos motivados pelo uso de número, afinal, por utilizar apenas adjetivos, o avaliador não é induzido a escolher números que possua maior afeição, outra vantagem com esta escala, é que a intensidade pode ser percebida facilmente por poder ser marcada em qualquer ponto da reta [2].

3.2.2 Posição

Em relação a posição das escalas, estas podem ser tanto verticais quanto horizontais, a depender do tipo de escala a ser utilizada, por exemplo, escalas não estruturadas costumam ser lineares, ou seja, horizontais, como demonstrado nas imagens abaixo.

3.2.3 Polaridade

É subdividida em duas vertentes: unipolar e bipolar. A unipolar diz respeito as escalas que avaliam a intensidade de apenas um atributo sensorial, como acidez, picância, entre outros. Já as bipolares, são as escalas que avaliam diferentes aspectos de um mesmo atributo sensorial simultaneamente [2]. Observe os exemplos abaixo.

Figura 3.2.7: Exemplo de escala unipolar

0	1	2	3	4	5	6	7	8	9
nenhum	fraco								muito forte

Fonte:[2].

Figura 3.2.8: Exemplo de escala bipolar

- 1 extremamente mole
- 2 muito mole
- 3 moderadamente mole
- 4 levemente mole
- 5 nem mole nem firme
- 6 levemente firme
- 7 moderadamente firme
- 8 muito firme
- 9 extremamente firme

Fonte:[2].

3.2.4 Quantidade de atributos avaliados

As escalas também variam conforme a quantidade de atributos que são utilizados na análise sensorial. Chamados de escala simples, são aqueles que analisam apenas um atributo, enquanto as escalas compostas avaliam dois ou mais atributos ao mesmo tempo, como aroma e sabor por exemplo, onde ao fim o avaliador deve fornecer uma única nota ao produto avaliado. Entretanto, a maior parte dos pesquisadores, atualmente, não recomendam a utilização desse tipo de escala, pois podem ocorrer erros uma vez que amostras com atributos sensoriais distintos entre si, podem receber uma mesma classificação final.

3.2.5 Tipo de avaliação

As escalas também podem ser sub divididas pelo tipo de avaliação, tais como: escala de intensidade, qualidade, hedônica e de magnitude. Nas escalas de intensidade são avaliados as características gerais, tais como aroma e sabor, ou características específicas das amostras, como aroma floral, rugosidade aparente, entre outros. devido a este fator específico, é necessário utilizar avaliadores treinados ou fornecer um treinamento adequado a eles.

Já as escalas de qualidade possui um papel fundamental no estabelecimento de padrões para que as avaliações sejam realizadas, assim, avaliam a qualidade sensorial de determinado alimento, além disso, também requerem avaliadores previamente treinados. Com relação as escalas hedônicas, estas buscam averiguar a aceitabilidade de um produto pelo público alvo, ou seja, descobrir se o avaliador gostou ou não de determinada amostra, este tipo de escala é utilizado de amplamente e frequentemente, sem precisar de avaliadores previamente treinados.

Figura 3.2.9: Exemplo de escala de qualidade



Fonte:[2].

Figura 3.2.10: Exemplo de escala hedônica

Indique utilizando a escala abaixo o quanto você gostou do suco de laranja.

- 9 – Gostei muitíssimo (ou adorei)
- 8 – Gostei muito
- 7 – Gostei moderadamente
- 6 – Gostei regularmente
- 5 – Nem gostei nem desgostei
- 4 – Desgostei regularmente
- 3 – Desgostei moderadamente
- 2 – Desgostei muito
- 1 – Desgostei extremamente (ou detestei)

Fonte: Autoria própria.

Com relação as escalas de magnitudes, o avaliador recebe previamente um valor de referência a intensidade do estímulo com relação ao atributo avaliado, e o compara com a amostra recebida. Para melhor compreensão, veja o seguinte exemplo: O avaliador recebe um biscoito de referência que possui o valor 25 atribuído a crocância. Após isso, o avaliador recebe amostras, previamente codificados, onde deve avaliar a crocância com relação à primeira amostra, de tal forma que, caso uma das amostras codificadas tenha o dobro da crocância da primeira amostra, o avaliador deve dar nota 50, já se a amostra for menos crocante que a primeira, a nota deve ser menor que 25, e assim por diante. Dessa forma, o avaliador consegue construir sua própria escala, sem possuir limites ou extremos pré determinados.

Figura 3.2.11: Exemplo de ficha de aplicação de um modelo de escala de magnitude

Prove primeiramente o biscoito referência (R) e assinale um valor para a sua maciez. A seguir, avalie a crocância das amostras codificadas com relação à referência. Se a crocância da amostra codificada for metade da amostra referência, assinale um valor equivalente à metade do valor dado à referência. Se a crocância da amostra codificada for duas vezes maior que a referência, dê o valor equivalente a duas vezes o valor assinalado para a referência.

Amostras	Crocância
R	<hr/>
824	<hr/>
591	<hr/>
407	<hr/>

Fonte: Adaptado de [2].

3.3 Perfil de sabor

Essa técnica foi desenvolvida por Arthur D. Little no final dos anos 1940 e início dos anos 1950, usada pela primeira vez com o objetivo de descrever sistemas complexos de sabor, através da medição dos efeitos do glutomato monossódio. Basicamente, o perfil de sabor tem como principal objetivo avaliar e caracterizar os aromas e sabores de amostras de alimentos, usando como base a ISO 6564:1985. [32] [2]

Esse método avalia as amostras de modo qualitativo e semi quantitativo, onde os avaliadores devem mensurar cinco características do sabor, tais como: impressões gerais do aroma e seu impacto, fatores perceptíveis do aroma e sabor, intensidade dos fatores, ordem de percepção dos fatores, e o sabor residual. Após isso, os avaliadores realizam discussões em grupo para que seja possível obter um único resultado final, por conta disso, o perfil de sabor pode ser considerado uma técnica de consenso. [2]. [32]

Para melhor compreensão, considere o seguinte exemplo de uma análise descritiva de um refrigerante de limão elaborada por Dutcosky:

1) Impacto do aroma e fatores perceptíveis: o primeiro impacto predomina o aroma “casca” do limão, característico do siciliano, com leves notas de laranja que resultam no aroma adocicado posterior.

2) Ordem de sensações na boca: 1º sensação de frescor / 2º sensação do gosto doce acentuado / 3º sensação de borbulhamento agradável / 4º gosto ácido, mascarado inicialmente pelo gosto muito doce.

3) Sabor residual: nota anestesiante, gosto levemente amargo com o doce predominando, leve sensação de adstringência.

4) Impressão global: refrigerante com sabor de limão típico, harmônico, bem equilibrado, apenas o doce está destoando no equilíbrio, mas está agradável.

Para que esse método seja bem sucedido, é necessário a utilização de avaliadores treinados que possuam disponibilidade de longo prazo, com número variando entre cinco e oito pessoas, pois o treinamento pode levar meses, e após treinados, podem ser requisitados por anos para analisar determinados alimentos. De acordo com Harry T. não é incomum encontrar julgadores de perfil de sabor que tenham servido no mesmo painel por mais de 10 anos. Por esses motivos, esse método leva tempo, esforço e dinheiro.

Existem alguns fatores que trazem riscos nesse método de teste, como o fato do número de julgadores ser pequeno e seu treinamento ser muito longo, outro fator é que no perfil de sabor não é aplicado uma análise estatística nos resultados obtidos. Além disso, dentro dos grupos dos avaliadores, um ou mais membros que possuirem personalidades fortes podem influenciar os demais e afetar nas respostas obtidas [2].

3.4 Perfil de textura

Criada na década de 1960 e posteriormente modificada por diversos especialistas sensoriais, o Perfil de textura é uma técnica que permite a avaliação de todas as características de textura de um produto, onde seus conceitos foram baseados inicialmente no perfil de sabor. [32]

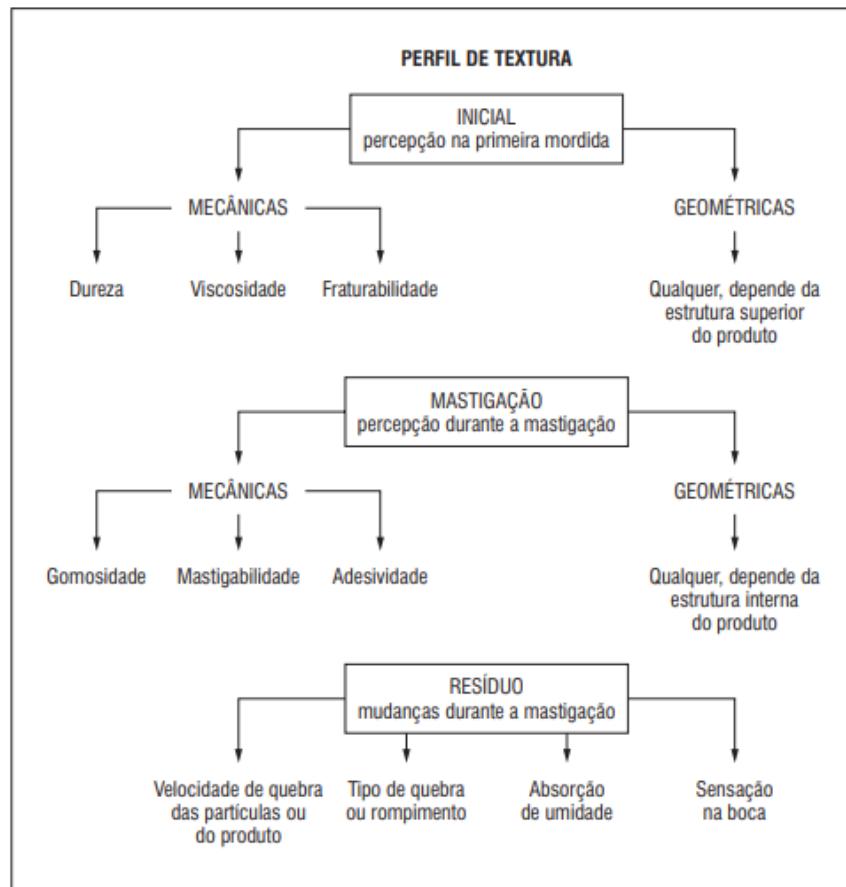
Assim, essa técnica busca obter uma análise completa da textura das amostras analisadas, e as manifestações sensoriais da estrutura interna dos produtos quanto a reação à força e propriedades táteis. A reação à força diz respeito às propriedades mecânicas das texturas, tais como viscosidade, firmeza/dureza, elasticidade, entre outros, que podem ser notadas utilizando os músculos da língua, mão, dedos ou lábio. Já as propriedades táteis tratam sobre as medidas de partículas geométricas, que podem relacionar forma e tamanho das partículas ou forma e orientação, tais como arenoso, cristalino, entre outros, além disso, também trata sobre as propriedades de umidade e gordura, todas essas características podem ser sentidas pelos nervos táticos presentes na superfície da pele da mão, lábios ou língua [2].

Para a realização deste teste, é necessário um clara compreensão das características de diferentes texturas em alimentos, para isso, eles são divididos em alguns tipos. As características mecânicas de uma textura é subdividida em dois grupos: o primário, o qual inclui a dureza, elasticidade, adesividade, coesividade e viscosidade, e o secundário, que retrata sobre a fragilidade, gomosidade e mastigabilidade. A textura também pode ser dividida com relação às suas características geométricas, teor de umidade e gordura, onde o primeiro, respectivamente, busca relacionar a forma e tamanho das partículas, sendo as grandes aquelas que são ásperas, arenosas, granulosas, parecidas com um pó ou que são pulverulentas. Já o teor de umidade retrata o quanto a textura é molhada, seca, úmida, suculenta ou encharcada, e o teor de gorduras

diz sobre a oleosidade do alimento [2].

O método do perfil de textura segue o procedimento recomendado na ISO 11036 (1994), demonstrado abaixo.

Figura 3.4.1: Procedimento para avaliação da textura sugerido por BRANDT; SKINNER; COLEMAN, 1963



Fonte: [2].

O método tradicional do teste de perfil de textura costuma utilizar uma escala com nove pontos para a maioria dos atributos de textura, como dureza, viscosidade, entre outros, fornecendo também exemplos de produtos comuns que representam cada ponto da escala [32]. Verifique abaixo as escalas de referências mais utilizadas.

Figura 3.4.2: Escalas padrão sugeridas por BRANDT; SKINNER; COLEMAN, 1963

Valor da escala	n. de mastigação	Produto	Marca ou tipo	Tamanho da amostra temperatura
Dureza				
1	queijo cremoso	Philadelphia, Kraft	cubos de 1,5cm/7 – 13°C	
2	clara de ovo	cozimento: 5"	1,5 cm/temperatura ambiente	
3	salsichas	grandes, Schneiders	pedaços de 1,5 cm/ 10 – 18°C	
4	queijo	Mild Cheddar, Kraft	cubos de 1,5cm/10 – 18°C	
5	azeitonas	Stuffed, tamanho "Queen", McLaren's	1 azeitona, remoção de pimentão/10 – 18°C	
6	amendoim	tipo em lata, Cocktail Planters	1 noz/temperatura ambiente	
7	cenouras	cruas, frescas	pedaços de 1,5 cm/ temperatura ambiente	
8	amêndoas	sem branquear, McNair	1 noz/temperatura ambiente	
9	bala dura	Dryden & Palmer	ambiente	
Viscosidade				
1	água		2 mL/temperatura ambiente	
2	creme leve	Sealtest, Neilson	2 mL/7 – 13°C	
3	creme pesado	Sealtest, Neilson	2 mL/7 – 13°C	
4	leite evaporado	Carnation	2 mL/7 – 13°C	
5	mel	Camp	2 mL/7 – 13°C	
6	xarope de chocolate	Hershey	2 mL/7 – 13°C	
7	misturar: 125 mL de maionese e 30 mL de creme pesado	Hellmann's Sealtest, Neilson	2 mL/7 – 13°C	
8	leite condensado	Eaglebrand, Sweetened	2 mL/7 – 13°C	

Fonte: [2].

Figura 3.4.3: 2º parte das escalas padrão sugeridas por BRANDT; SKINNER; COLEMAN, 1963

Valor da escala	n. de mastigação	Produto	Marca ou tipo	Tamanho da amostra temperatura
Fraturabilidade				
1	bolo de milho	Betty Crocker Mix	cubo de 1,5 cm/temp. ambiente	
2	suspiro	Primo	cubo de 1,5 cm/temp. ambiente	
3	Cream Craker	Nabisco ou Paulins	quadrado de 1,5 cm/temp. ambiente	
4	torrada Melba	plano, pedaço retangular do interior/old London	quadrado de 1,5 cm/temp. ambiente	
5	biscoito de farelinho de trigo	Weston	quadrado de 1,5 cm/temp. ambiente	
6	biscoito de gengibre	Christies	quadrado de 1,5 cm/temp. ambiente	
7	pé-de-moleque	Mc Cormick	ped. peq. de 1,5 cm/temp. ambiente	
Gomosidade				
1	40% pasta de farinha	Gold Medal ou Robin Hood	15 mL/temperatura ambiente	
2	45% pasta de farinha	Idem à anterior	15 mL/temperatura ambiente	
3	50% pasta de farinha	Idem à anterior	15 mL/temperatura ambiente	
4	55% pasta de farinha	Idem à anterior	15 mL/temperatura ambiente	
5	60% pasta de farinha	Idem à anterior	15 mL/temperatura ambiente	
Mastigabilidade				
1 10,3	pão de centeio	corte central (fresco)	cubo de 1,5 cm/temp. amb.	
2 17,1	salsicha	grande, sem a pele	pedaço de 1,5 cm/ 10 – 21°C	

Fonte: [2].

Figura 3.4.4: 3º e última parte das escalas padrão sugeridas por BRANDT; SKINNER; COLEMAN, 1963

Valor da escala	n. de mastigação	Produto	Marca ou tipo	Tamanho da amostra temperatura
Mastigabilidade				
3	25,0	bala de goma	Chuckle	1 pedaço/temp. ambiente
4	31,8	bife	1,5 de grossura, grelhado (10') cada lado	quadrado - 1,5 cm/ 60 - 85°C
5	33,6	bala "licorice"	Laura Secord Mc Cornick	1/2 pedaço
6	37,3	caramelos	Mc Cornick	1 pedaço/temp. ambiente
7	56,7	bisc. Tootsie Rolls	peq., Tootsie Rool Co.	1 pedaço/temp. ambiente
Adesividade				
1	margarina	Fleischmann	2 mL/7 - 13°C	
2	queijo velveeta	Kraft	1,5 cm/7 - 13°C	
3	queijo cremoso	Philadelphia	2 mL/7 - 13°C	
4	mashmallow	Kraft	2 mL/7 - 13°C	
5	pasta de amendoim	Smooth, Skippy	2 mL/7 - 13°C	

Fonte: [2].

Em todas as escalas, os alimentos vão aumentando de intensidade, como é possível observar na escala de dureza, onde os diferentes alimentos (requeijão/queijo cremoso, clara de ovo cozida, salsichas, queijo, azeitonas, amendoim, cenoura crua, amêndoas e balas duras) aumentam de intensidade do requeijão para o doce. No entanto, esses produtos alternadamente podem cisalhar, quebrar ou comprimir quando a força de compressão é aplicada. Assim, ao usar a escala de referência de ,dureza, os painelistas devem entender que, embora todos esses produtos variem em uma dimensão específica e definível, a dureza, eles não necessariamente reagem da mesma maneira a uma força compressiva aplicada [32].

O teste de perfil de textura é um método que requer um treinamento dos avaliadores, os quais precisam ser aptos a identificar sem dificuldade todas as características das texturas em alimentos. Para isso, cada característica de textura possui um método para treinar os avaliadores, mas a base é a mesma para todos: são apresentados a eles cinco amostras aleatorizadas, onde se pede que coloquem em ordem crescente da característica analisada, na análise de dureza os avaliadores devem posicionar as amostras em ordem crescente de dureza, em vis-

cosidade na ordem crescente de viscosidade, e na análise da escala geométrica pedi-se para cada candidato associar uma determinada descrição geométrica a cada um, por exemplo, granuloso, fibroso, aerado, floculento ou cheio de bolinhas (também com cinco amostras aleatoriamente selecionada).

Com isso, ao fim das 15 amostras, é necessário que o avaliador tenha tido um acerto de no mínimo 12 amostras para poder participar dos testes de perfil de textura. Estes procedimentos devem ser feitos em uma mesa redonda, e é necessário a seleção de 6 a 8 avaliadores com treinamento específico para o alimento que serão destinadas as análises sensoriais em questão. Essa metodologia é capaz de analisar o grau das características desde a primeira mastigação até a fase residual (quando já ocorreu a deglutição do produto), mas vale ressaltar que deve ser adaptada para cada tipo de produto [2].

3.5 Perfil de escolha livre

O método de Perfil Livre, também chamado de (Free-Choice Profiling), possui esse nome devido a sua metodologia inovadora de conduzir uma análise descriptiva: do início ao fim da análise, o avaliador possui total liberdade para utilizar os termos descriptivos que desejar, na quantidade que desejar.

Em outras palavras, no lugar de treinar os julgadores de modo extensivo utilizando um vocabulário pré definido para o produto, no método de perfil livre, cada membro pode criar sua própria lista de termos descriptivos, ou seja, sua própria terminologia. Isso não descarta a necessidade de treinamento, uma vez que os participantes devem utilizar de forma consistente as próprias terminologias criadas. Além disso, cada avaliador pode avaliar o produto do modo que desejar, seja tocando, provando ou cheirando, analisando a cor, brilho ou outros estímulos que o avaliador julgar importante [32]

A ideia central desse método é que, apesar de cada julgador utilizar uma linguagem diferente na descrição do produto, todos devem descrever os mesmos produtos no mesmo ambiente de análise sensorial. Vale ressaltar que o uso desse método só foi possível por conta do desenvolvimento da Análise Procrustes Generalizada (Generalised Procrustes Analysis – GPA), por Gower (1975), pois através dessa técnica estatística é possível analisar os dados obtidos através dos métodos do perfil livre.

Para a realização desse testes, são utilizados entre 10 a 15 julgadores, e o processo da análise sensorial deve ser realizada em três etapas: 1º - são apresentadas as amostras a serem avaliadas (três repetições serão feitas por amostra), 2º - cada julgador deve elaborar sua própria lista de definições (seua terminologia própria) os quais são apresentados as amostras que serão avaliadas, e 3º - são confeccionadas fichas de avaliação individuais com base das definições escolhidas, com uma escala de intensidade para cada termo.

Os resultados então são analisados pela Análise procrustes Generalizada (GPA), a qual busca aproximar as respostas de cada julgador a uma resposta média, ou consenso, maximizando as similaridades geométricas.

Basicamente, cada resultado é considerado uma coordenada em um espaço multidimensional, onde posteriormente são interpretadas usando os diferentes termos descritivos. Os resultados permitem detectar os julgadores que apresentam problemas de diferenças de percepção e de falta de repetibilidade. Assim, caso a resposta de um determinado avaliador seja muito diferente dos demais ou falte com repetibilidade, é possível detectá-lo e eliminar seus dados [32].

3.5.1 Curiosidade - origem do termo Procrustes

O termo vem da mitologia grega, o qual conta a história de Procusto, um assaltante e torturador que oferecia hospedagem aos viagentes locais que transitavam entre a região de Atenas e Eleusis. Caso o viajante fosse menor do que a cama oferecida, Procusto esticava a pessoa até que coubesse exatamente na cama, já se o viajante fosse maior que ela, Procusto multilava as partes que sobravam para que coubesse nas medidas exatas da cama [33].

A metodologia denominada de Procrustes, possui esse nome por tentar "forçar" as medidas para que elas se encaixem em um consenso, pois o objetivo desse método é tornar as medidas/respostas consensuais através de mudança de escala, rotação e/ou translação no espaço multidimensional.

3.6 Análise descritiva quantitativa (ADQ)

A análise descritiva quantitativa (ADQ) foi desenvolvida em 1974 por Stone e Sidel da Tragon Corporation, e tem como objetivo avaliar todos os atributos sensoriais presentes nos alimentos das análises sensoriais, tais como textura, aroma, sabor e aparência. Esse método foi normatizado pela NBR 14140 (1998), onde utiliza escalas não estruturadas com dois extremos que indicam a intensidade da característica avaliada, costumam ter entre nove e 15 centímetros. Para que a ADQ seja aplicada de modo adequado, é necessário a utilização de cinco etapas, as quais são:

- 1º Etapa - Seleção dos avaliadores:

Nessa etapa ocorre a seleção dos avaliadores para a montagem final da equipe, a qual deve possuir, no mínimo, 10 a 12 avaliadores treinados, portanto é iniciado a avaliação com três vezes o número de avaliadores desejado. [12] [2]

- 2º Etapa - Levantamento dos descritores/desenvolvimento da terminologia:

Os julgadores avaliam sensorialmente o produto e verbalizam as sensações percebidas, discutindo-as em grupo com a ajuda do líder da equipe. O objetivo principal desta etapa é o levantamento dos descritores que caracterizam os produtos[2]

Trata-se do aspecto qualitativo em que são estabelecidos as características das amostras que devem ser analisadas e as referências utilizadas. Os atributos estudados devem ser observados sensorialmente de acordo com as verdadeiras características da amostra como cor, forma, tamanho. Dessa forma, o aspecto qualitativo está relacionado com a aparência, aroma, sabor, textura oral sendo que, para cosméticos, também é avaliado os atributos de sensações da pele e, para tecidos, a textura ao tato [12]. É possível utilizar dois métodos para desenvolver uma lista de termos descritivos:

- Primeiro método (método tradicional) :

É baseado em oferecer o produto a cada julgador e pedir para que ele faça uma observação completa dos atributos sensoriais que aquele produto possui e, em seguida, é aberto uma discussão com o acompanhamento de um líder. Os atributos mais utilizados irão compor a ficha de análise do produto que está sendo avaliado[2]

- Segundo método (método de rede ou grid) :

O segundo método é composto pela apresentação das amostras aos pares ao avaliador, pedindo que o julgador liste as diferenças e similaridades entre elas. Por exemplo, se há três amostras, A, B e C, será apresentado ao julgador, no método de grid, para que seja feita a lista de termos descritivos comparando-se A x B, Bx C e A x C. Os termos descritivos que mais aparecem para dizer as diferenças compõem a ficha de análise.

- 3° Etapa - Treinamento:

Essa etapa é realizada com os próprios produtos e com os materiais de referência, podendo ser adicionado alguns sistemas modelos, por exemplo: pode-se adicionar um determinado aromatizante a uma das amostras com o objetivo de uma determinada nota aromática ser melhor percebida no produto pelos avaliadores, que poderão memorizá-la melhor[2].

Depois de feito o treinamento, é recomendado fazer uma nova seleção dos julgadores, a fim de determinar os avaliadores com melhor desempenho em discriminar, que possuem uma boa reproduzibilidade e possuem resultados consistentes.

- 4°Etapa - Teste sensorial:

Depois de realizado o treinamento e a seleção final dos julgadores, já é possível realizar os testes sensoriais em condições controladas que garantam a individualidade dos avaliadores e os outros requisitos necessários.

- 5° Etapa - Análise dos resultados:

Os dados são avaliados por Análise de Variância (ANOVA) e um teste de média, sendo mais utilizado o teste de Tukey, para comparação das amostras. Os resultados do teste são representados graficamente com uma forma típica chamada gráfico-aranha (spider-web), no qual possui a intensidade média de cada atributo, sendo o ponto central o zero.

A metodologia do ADQ ajudou muito na evolução dos métodos descritivos Perfil de Sabor e Perfil de Textura, trazendo até mesmo uma nova abordagem metodológica. A abordagem do ADQ possibilitou a avaliação de todas as propriedades sensoriais de um produto de uma vez só, a qualificação dos avaliadores, a possibilidade de usar uma equipe menor, sem a necessidade de que fossem experientes ou experts no produto. Além disso esse procedimento não sofre a influência direta do líder, garantindo a análise individual do produto e, o mais importante, possibilitou uma medida quantitativa e reproduzível dos descriptores do produto[2]

3.6.1 Técnicas estatísticas utilizadas na ADQ

O primeiro método mais utilizado é a análise de variância (ANOVA), pois como foi explicado anteriormente, através dela é possível realizar a comparação entre pelo menos 2 médias e descobrir se elas são ou não diferentes entre si. Costuma-se realizar, em conjunto da ANOVA, o delineamento em blocos casualizados (DBC), também chamado de delineamento em blocos completos casualizados, o qual possui três princípios básicos: repetição, casualização e controle local. O uso do DBC é essencial, uma vez que cada avaliador, ao provar a amostra, tende a ter uma percepção diferente do produto avaliador, e essa variabilidade deve ser levada em consideração na análise de resultados. O modelo estatístico de uma DBC pode ser representado pela seguinte equação:

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \text{ para } i = 1, 2, \dots, a, j = 1, 2, \dots, b, \quad (3.6.1)$$

onde Y_{ij} representa o valor observado na parcela que recebeu o tratamento i no bloco j , o μ é uma constante geral, o τ_i é o efeito do tratamento i aplicado na parcela, β_j é o efeito do bloco j aplicado a parcela, e ϵ_{ij} represesta o erro experimental.

O quadro da ANOVA, nesse caso, pode ser representado abaixo, onde o teste F é utilizado para inferência a respeito dos tratamentos.

Figura 3.6.1: ANOVA para um delineamento em blocos casualizado (DBC)

Causa de Variação	Soma de Quadrados (SQ)	Graus de Liberdade	Quadrado Médio (QM)	Teste F
Tratamentos	$SQ_{Tratamentos}$	$a - 1$	$QM_{Tratamentos}$	$F_0 = \frac{QM_{Tratamentos}}{QM_E}$
Blocos	SQ_{Blocos}	$b - 1$	QM_{Blocos}	
Erros	SQ_E	$(a - 1)(b - 1)$	QM_E	
Total	SQ_T	$N - 1$		

Em alguns casos, o interesse da análise sensorial é avaliar mais de um fator dentro do experimento, para isso utiliza-se o método de esquema fatorial.

Cada subdivisão de um fator é denominada de nível do fator, e os tratamentos nos experimentos fatoriais consistem de todas as combinações possíveis entre os diversos fatores nos seus diferentes níveis.

Para compreender melhor o conceito de esquema fatorial, imagine a seguinte situação: uma fábrica de doces deseja combinar duas variedades de farinhas com três diferentes tipos de corantes. Portanto, esse é um esquema fatorial 2×3 , onde existem dois fatores: Farinhas (F) e Corantes (C), sendo que o fator farinha possui dois níveis (F_1 e F_2), e o fator corante possui três níveis (C_1, C_2 e C_3). Portanto, pode-se analisar ao todo seis combinações diferentes: $F_1C_1, F_1C_2, F_1C_3, F_2C_1, F_2C_2$ e F_2C_3 .

3.6.2 Check-All-That-Apply (CATA)

A metodologia Check-all-that-apply (CATA), pode ser traduzida do inglês para "marque tudo que se aplique", e é uma técnica descritiva amplamente aplicada, pois não requer avaliadores previamente treinados, e através dela é possível distinguir quais as características do alimento os consumidores conseguem detectar, algo essencial para a criação de novos produtos, uma vez que a compreensão dos atributos sensoriais podem determinar o sucesso ou fracasso de uma inovação [34] [35].

Para a realização do CATA, os avaliadores não precisam ser necessariamente treinados, uma vez que esse método requer uma instrução mínima, sendo facilmente executado e concluído de modo rápido. No CATA, é apresentado aos avaliadores uma lista que contém todos os descriptores de um determinado produto, onde eles possuem o objetivo de selecionar aqueles que, de acordo com a sua percepção, estão presentes no alimento [36], para entender observe o exemplo abaixo.

Figura 3.6.2: Exemplo de lista para o método CATA

Marque todas as palavras que descrevem o produto que você acabou de avaliar

- Amargo
- Artificial
- Azedo
- Doce
- Natural
- Fresco

A lista de termos para descrição do produto, pode ser elaborada utilizando um painel de avaliadores previamente treinados, ou, por um conjunto de consumidores que testarem o produto. Entretanto, caso os consumidores criem seus próprios termos através da livre escolha, é importante ter em mente que a análise se tornará mais demorada, devido a interpretação subjetiva e a combinação de termos parecidos que deve ser realizada [37] [35] [38].

Uma das principais vantagens do método CATA é a espontaneidade das respostas fornecidas e sua simplicidade de realização pelos consumidores, quando comparada com a classificação de intensidade. Entretanto, essa ausência de medição de intensidade dos atributos acaba fornecendo um grau de discriminação menor quando comparado com outras técnicas. Outra dedvantagem é o número mínimo indicado de avaliadores, que é de 100 pessoas, para que tenha relevância estatística [35] [37].

Para que seja possível realizar uma análise adequada das respostas, costuma-se codificá-las em números binários, onde as palavras marcadas pelo avaliador são representadas por 1 e as não assinaladas por 0. Esses resultados normalmente são resumidos em tabelas de frequência, onde os atributos que possuírem maior porcentagem são considerados como presentes no produto, mas não é possível tirar conclusões sobre a intensidade percebida. Já com as palavras que tiverem baixa porcentagem, não é possível concluir exatamente o motivo de não terem sido assinaladas, pode ser que o atributo não estivesse presente no produto, ou que o avaliador não tinha certeza, ou não era adequado para o produto. É muito importante levar isso em consideração no momento de interpretação dos dados e conclusão [34]

Uma análise mais aprofundada pode ser feita usando o teste Q de Cochran, o qual certifica se, no mínimo, três conjuntos com frequências/proporções similares possuem diferenças significativas entre si, em outras palavras, determinará se para cada atributo, os produtos avaliados têm um padrão de resposta diferente [39].

Esse teste se adapta a dados em escala nominal ou na forma de ordinal dicotomizada, assim, é possível verificar se os itens da avaliação diferem entre si em relação ao grau de dificuldade através, por exemplo, da análise de dados que apresentam respostas do tipo “aprovado-reprovado” (pass-fail) para k itens e n elementos. Ademais, é possível ter um item para análise e realizar a comparação de n indivíduos para k condições diferentes, em que cada pessoa está sujeita a uma condição distinta [39].

Os resultados podem ser apresentados em uma tabela de dupla entrada com l linhas e k colunas, tornando possível avaliar a hipótese que estebelece que cada coluna apresenta a mesma proporção de respostas. A partir disso, se a hipótese nula for considerada verdadeira, isso é, se os “sucessos” ou “falhas” são distribuídos de forma aleatória na tabela e apresentam uma grande quantidade de linhas, o valor Q tem a seguinte fórmula

$$Q = \frac{k(k-1) \sum_{j=1}^k (G_j - G)^2}{k \sum_{i=1}^n L_i - \sum_{i=1}^n L_i^2},$$

Existe também uma expressão equivalente, que pode facilitar os cálculos, sendo representada pela equação abaixo

$$Q = \frac{(k-1)[(k \sum_{j=1}^k G_j^2) - (\sum_{j=1}^k G_j)^2]}{k \sum_{i=1}^n L_i - \sum_{i=1}^n L_i^2},$$

Se o valor Q encontrado através do cálculo não for menor a um certo valor especificado para um nível de significância

Para a análise estatística, tabelas de contingência e gráficos de barras são amplamente usados para apresentação dos resultados referentes a contagem ou porcentagem do número de avaliadores que assinalaram um determinado atributo [36].

3.6.3 Perfil Flash

O Perfil Flash (Flash Profile), proposto por Sieffermann (2000), é um método descritivo que proporciona acesso mais rápido de classificação e descrição de um conjunto de produtos em relação ao método descritivo tradicional [40].

Esse método consiste na rápida julgação de um produto, evitando a fadiga e influência de pensamentos muito profundos, sendo acompanhado de uma escala de intensidade, que é o único aspecto quantitativo do teste (a ordem de intensidade percebida nos produtos).

É uma técnica baseada no teste de Perfil Livre, na qual é feito uma lista individual de

atributos, para depois ser realizado o teste com o uso de uma ficha de ordenação e por fim, avaliação das amostras por ordenação.

Capítulo 4

MÉTODOS AFETIVOS

4.1 Introdução

Também conhecidos como testes de consumidores, esse método é o mais utilizado quando se deseja obter a preferência ou aceitação do consumidor com relação a algum produto. Para entender melhor os objetivos desses testes, é necessário ter clareza da diferença entre preferência e aceitação. A preferência pode ser definida como a expressão de gostar, a escolha de uma amostra ao invés da outra, e/ou o contínuo psicológico (percepção do que é agradável até o que é desagradável) do afetivo por meio do qual uma escolha é feita. Já a aceitabilidade pode ser definida como uma experiência caracterizada por uma atitude positiva, e/ou de hábito de comprar ou consumir um determinado alimento[2].

Os métodos afetivos buscam verificar o posicionamento do produto dentro do mercado e avaliar o seu potencial, servem para desenvolver novos produtos e otimizar formulações, averiguar qual é o produto mais aceito e, ou, preferido do público alvo. Além disso, esse método também pode ser utilizado para eliminar produtos de baixa qualidade[5].

Para a realização de um bom teste afetivo, são necessários participantes que sejam consumidores habituais ou potenciais do produto que será analisado, não sendo recomendado pessoas que de algum modo estão relacionados com o produto testado, pois podem analisar com base em um viés pessoal. Os participantes não são escolhidos com base nas suas habilidades sensoriais, mas sim por serem pertencentes do público alvo que o produto é destinado. Existem alguns critérios que devem ser levados em consideração na seleção dos participantes nos testes afetivos: faixa etária (o qual varia com base no público alvo do produto), a frequência de consumo (uma boa equipe é composta, preferencialmente, por indivíduos que possuam uma elevada frequência de consumo do produto avaliado, ou similares), além de sexo, localização geográfica, religião, educação, estado civil e poder aquisitivo, uma vez que todos esses fatores influenciam diretamente e/ou indiretamente no gosto pessoal e frequência de consumo dos produtos[5].

4.1.1 Locais de Aplicação

O ambiente de aplicação de qualquer tipo de teste pode afetar no resultado final de uma análise sensorial, no caso dos testes afetivos, eles podem ser realizados em três locais distintos: laboratório, locais centrais e domicílios.

O laboratório é o ambiente mais utilizado para a realização dos testes afetivos, sendo muito útil para seleção preliminar de amostras que possam vir a ser futuros novos produtos. São realizados em cabines individuais e entre 25 a 50 avaliadores, mas vale ressaltar que em análises envolvendo mais de dois produtos, o recomendado é utilizar painéis com 50 consumidores, aumentando a significância estatística. Os testes em laboratórios costumam ser mais baratos, rápidos e facilitam um maior controle das condições do teste, entretanto, o preparo do produto pode ser diferente do doméstico, o produto pode ser apresentado para o consumidor de forma e quantidade diferente de como é apresentado casualmente, não sendo similar a quantidade que consome, e também, os resultados não podem ser generalizados para o grupo consumidor, por apresentar um pequeno número de consumidores.

Já os testes realizados em locais centrais, ocorrem em ambientes abertos, com grande movimento de pessoas, como shopping, mercados, escolas, entre outros. Esses locais são muito utilizados na realização de testes de consumidores, e também para a pesquisa de marketing do produto, necessitando de ao menos 100 consumidores por produto. Esses consumidores podem ser aleatórios ou previamente pré-selecionados, desde que ambos se enquadrem nos pré-requisitos definidos no teste. Testes realizados nesse ambiente conseguem produzir resultados com validade assegurada, uma vez que são avaliados pelos próprios consumidores, além de que a porcentagem de retorno de respostas é alta, quando comparados com os testes em domicílio. Outra vantagem, é que qualquer dúvida sobre a avaliação ou sobre o produto pode ser esclarecida prontamente pelo avaliador. Entretanto, esse ambiente gera um maior custo, uma menor quantidade de informações retiradas, já que o número de questionamento que pode ser feito aos consumidores é limitado, e também, existe pouco controle das condições de teste, pouca individualidade, menor concentração dos consumidores etc.

Os testes em domicílio são realizados na própria residência do consumidor, levando em consideração a opinião de todos os membros da família, sem restringir a interação e influência da opinião de cada um. Testes realizados nesse ambiente normalmente avaliam apenas dois produtos, por se tratar de um estilo de teste demorado onde o primeiro produto seria avaliado por quatro a sete dias, para assim, iniciar a avaliação do segundo produto. Caso resolva avaliar mais produtos, deve levar em conta o risco de desistência, menor retorno etc. Os produtos enviados devem ser padronizados de forma que a forma, embalagem, tamanho sejam iguais e devem ser enumerados com 3 números aleatórios, contendo na embalagem a instrução clara e completa de como se preparar o produto.

As principais vantagens de testes realizados em domicílios é que a forma em que o produto

é enviado para a análise, é exatamente a forma que iria para o consumidor em condições reais de uso, possibilitando uma análise real de que forma o produto seria avaliado caso estivesse no mercado de vendas, tendo um feedback da embalagem, características específicas do produto e que preço estariam dispostos a pagar no produto. Além disso, não terá somente a primeira impressão do produto, e sim como seria avaliado com o seu uso contínuo, permitindo também realizar planos de amostragem estatística da população. Já as desvantagens enfrentadas são os riscos de alguns questionários não serem respondidos, o maior tempo de espera pelos resultados, o maior custo e também o número de amostras limitado de amostras avaliadas.

4.2 Classificação dos testes afetivos

Os testes afetivos podem ser divididos nos teste qualitativos e quantitativos, onde os qualitativos buscam uma análise mais observatória e trazer aprendizados sobre o produtos analisado, já os quantitativos são testes que produzem dados que são analisados por testes estatísticos.

4.2.0.1 Testes afetivos qualitativos

Esse tipo de teste não necessita de um grande número de avaliadores, visando ter respostas subjetivas em relação ao produto, sendo muito importante ter um contato entre entrevistador/moderador. São utilizados para saber a primeira impressão do consumidor com o protótipo do produto, buscando saber a terminologia que os consumidores irão utilizar na descrição dos atributos sensoriais da amostra analisada.

Os tipos de testes utilizados são:

- Grupo de foco:

É um teste no qual são selecionados de seis a nove consumidores com um critério específico, sendo necessário levar em conta o consumo frequente do produto, idade etc. é baseado em um grupo no qual é controlado por um moderador, que apresenta o tópico de interesse a guia os consumidores para facilitar a discussão utilizando dinâmicas em grupo, buscando retirar toda informação de valor da reunião, sendo realizado, no mínimo, três sessões.

- Equipes de foco:

Nesse teste é reutilizado o mesmo grupo de foco duas ou três vezes mais, possibilitando mais discussões com o grupo, sendo solicitado que utilizem o produto em casa para relatarem a experiência do seu uso contínuo.

- Entrevistas individuais:

Esse teste é realizado individualmente, buscando entender profundamente as opiniões de cada consumidor, sendo entrevistado normalmente 50 consumidores sucessivamente.

4.2.0.2 Testes afetivos quantitativos

São realizados para avaliar a preferência, gostos e opiniões dos consumidores em relação às características sensoriais dos produtos. Esses consumidores devem representar a população-alvo, esclarecendo a preferência global ou aceitação em relação a um ou mais produtos, podendo também medir respostas dos julgadores analisando atributos específicos, como a espessura do caldo de uma sopa, cor do caldo da sopa etc.

Esses testes podem ser divididos em duas grandes categorias, testes de preferência e testes de aceitação. Resumidamente, os testes de preferência tem como objetivo avaliar a preferência de um julgador quanto outros dois ou mais produtos e os testes de aceitação é avaliar o quanto um grupo de consumidores gostam ou desgostam de um produto, buscando sempre saber se o avaliador compraria ou não o produto.

4.3 Testes de preferência

Esse tipo de teste é utilizado para saber somente qual amostra o consumidor prefere, o que não possibilita saber se ele gosta ou não do produto, sendo utilizado nas empresas para melhorar o produto, desenvolver novos produtos etc. Os testes de preferência mais utilizados são os teste de comparação pareada, ordenação e comparação múltipla.

4.3.1 Testes de comparação pareada

Como o teste de comparação pareada dos testes discriminativos, duas amostras são apresentadas para o consumidor, mas dessa vez, ao invés de ser pedido para dizer qual amostra é diferente é pedido para anotar na ficha de avaliação qual amostra ele prefere. As amostras devem ser enumeradas com três dígitos e apresentadas de forma casualizada e balanceada entre os avaliadores, podendo ter dois tipos de combinação AB e BA, sendo que a equipe de consumidores deve receber igual número dessas duas ordens. Um exemplo de ficha de comparação pareada de preferência está na Figura 4.4.1.

Figura 4.3.1: Exemplo de ficha de para o teste de comparação pareada de preferência

Nome: _____	
Data: _____	
<p>Por favor, prove as amostras que estão codificadas da esquerda para a direita e circule o código da amostra de sua preferência. Entre as avaliações das amostras, enxague a boca com água e aguarde 30 segundos.</p> <p>— —</p>	
Comentários: _____	

A análise estatística dos dados desse teste é baseada em análises de tabelas, sendo necessário anotar o número de julgadores que realizaram o teste e o número de julgadores que preferiram cada amostra, identificando qual amostra foi indicada pela maioria como a preferida.

Após esses dados serem recolhidos, é consultada a Tabela 6.1.2 para determinar quantas respostas de preferência positiva são necessárias para dizer que a amostra é significativamente preferente. Após essa informação ser obtida, é feita a comparação da quantidade de respostas mínimas com a quantidade de respostas obtidas, sendo que:

- Se o número de respostas para a amostra indicada como preferência for maior ou igual ao número mínimo indicado pela tabela, é possível dizer que ela é significativamente preferente.
- Se o número de respostas para a amostra indicada como preferência for menor que o número mínimo indicado pela tabela, não é possível dizer que ela é significativamente preferente, ou seja, não há preferência significativa entre as amostras.

4.3.1.1 Exemplo

Problema 17. *Uma fábrica de brownies está em dúvida sobre o lançamento de um novo sabor e realiza um teste de preferência de comparação pareada para saber se os consumidores preferem brownies de paçoca ou de cream cheese. O teste é realizado com 60 julgadores e foi obtido 43 respostas para o brownie de paçoca como preferência, e 17 para o brownie de cream cheese. O brownie de paçoca é realmente o que possui maior preferência a 5% de significância?*

Para solucionar esse tipo de problema, basta analisar a tabela 6.1.2, assim, observando a 5% de significância, com 60 julgadores, é observado que é necessário 39 respostas positivas para preferência para dizer que uma amostra é significativamente preferível, ou seja, a 5% de significância o brownie de paçoca é o sabor preferido, quando comparado com o brownie de cream cheese.

4.3.2 Teste de Ordenação

O método de ordenação, além de dizer qual amostra os consumidores preferem, também ordena em ordem de preferência três ou mais amostras, podendo dizer qual o consumidor menos prefere e qual o consumidor prefere em um nível mediano. O consumidor recebe as amostras codificadas com três dígitos, de maneira casualizada e balanceada, pedindo que as ordene de acordo com sua preferência, marcando a resposta na ficha de avaliação. É um método rápido e capaz de avaliar várias amostras em uma mesma sessão, tendo como limite recomendado de no máximo seis amostras, um exemplo de ficha de um teste de ordenação está na 4.3.2.

Figura 4.3.2: Exemplo de ficha de para o teste de ordenação

Nome : _____		
Data : _____		
Das três amostras codificadas que foram entregues a você, ordene-as em ordem decrescente de preferência:		
_____	_____	_____
(maior preferência)	_____	(menor preferência)
Comentários: _____		

Após o teste ser realizado, os dados são organizados em uma tabela de dupla entrada para cada tratamento (amostra) e para cada avaliador (consumidor), possibilitando realizar a soma de ordens para cada tratamento, para então, realizar o teste de Friedman ou de Kramer e se necessário, é realizado um teste de comparações múltiplas, possibilitando estabelecer diferenças entre os tratamentos quanto à intensidade da preferência pelos consumidores, podendo ser utilizado o Método LSD ou o Método de Christensen.

4.3.2.1 Exemplo

Exemplo a ser adicionado

4.3.3 Teste de Comparação Múltipla

Esse tipo de teste é realizado quando se deseja avaliar mais de duas amostras ao mesmo tempo, podendo ser avaliado três ou mais amostras. É apresentado para o julgador em cabines individuais a amostra de referência, indicado com a letra R, junto com as outras amostras que serão avaliadas, seguindo a regra de codificação com três dígitos.

É pedido para os avaliadores indicarem as amostras quanto a sua preferência, classificando como “igual preferência”, “mais preferida” ou “menos preferida” que a referência e depois, indicar o grau de preferência, um exemplo de ficha para esse tipo de teste está na Figura 4.3.3.

Figura 4.3.3: Exemplo de ficha de para o teste de comparação múltipla

Nome: _____		
Data: _____		
Pedimos que, por favor, prove a amostra de referência (indicada com R) e as amostras codificadas. Compare as amostras codificadas com a amostra de referência, dizendo de acordo com seu julgamento, se ela é de igual preferência, mais preferida ou menos preferida. Logo em seguida, classifique a intensidade da preferência com o apoio da escala apresentada. Entre as avaliações das amostras, enxágue a boca com água e aguarde 30 segundos.		
Código da amostra:		
Mais preferida que R	_____	_____
Preferida igual a R	_____	_____
Menos preferida que R	_____	_____
Intensidade da preferência		
Nenhuma	_____	_____
Pequena	_____	_____
Moderada	_____	_____
Grande	_____	_____
Extrema	_____	_____
Comentários: _____		

Figura 4.3.4: *

Fonte: Adaptado de [5]

Após obter as respostas dos avaliadores, é organizado os dados e transformadas em valores numéricos, como mostrado na Tabela 4.3.1. Esses valores numéricos (escores) são organizados em uma tabela de dupla entrada, para assim, ser realizado um teste de variância

Tabela 4.3.1: Critérios utilizados para transformação das respostas dos avaliadores em valores numéricos (escores)

Classificação da amostra		Escores
Mais preferida que R	Extrema	9 ou 4
	Grande	8 ou 3
	Moderada	7 ou 2
	Pequena	6 ou 1
Preferida igual a R	Nenhuma	5 ou 0
Menos preferida que R	Pequena	4 ou -1
	Moderada	3 ou -2
	Grande	2 ou -3
	Extrema	1 ou 4

Tabela 4.3.2: *

Fonte: [5]

Vale notar que uma análise desse tipo transforma dados como extremo, grande, moderada em uma grandeza numérica, e por conta disso, estatisticamente não se torna algo de grande confiabilidade para uma análise de ANOVA, mas é a forma que é comumente utilizada.

4.3.3.1 Exemplo

Exemplo a ser adicionado

4.4 Teste de aceitação

Os testes de aceitação possuem como principal objetivo avaliar o quanto os provadores gostaram ou não de um determinado produto, para isso, são utilizadas escalas não balanceadas e balanceadas, sendo que a segunda é a mais empregada por ser considerada mais discriminativa e questionadora, uma vez que apresenta termos igualmente espaçados e números iguais de categorias positivas e negativas. De modo geral, existem diversas escalas que podem ser utilizadas nos testes de aceitação, porém as mais utilizadas são a escala hedônica, a FACT e a escala do ideal.

4.4.1 Escala Hedônica

A escala hedônica, como apresentada no capítulo 3, serve para se avaliar a aceitação global de um determinado produto, ou também avaliar a aceitação de diversos atributos do produto, tais como cor, aroma, sabor, etc. Ela costuma ser muito utilizada devido a facilidade de compreensão pelos consumidores, independente do seu tipo: verbal, facial e não estruturada. De modo geral, o consumidor recebe as amostras codificadas com três números aleatórios, e

pede-se que seja avaliado cada amostra utilizando a escala apresentada, com isso, os resultados são organizados e transformados em valores numéricos (no caso de escalas faciais ou verbais). Observe abaixo um exemplo de como transformar essas escalas em números.

Figura 4.4.1: Exemplo de transformação das classificações da escala em valores numéricos

CLASSIFICAÇÃO NA ESCALA	ESCORES
Gostei extremamente	9 ou 4
Gostei muito	8 ou 3
Gostei moderadamente	7 ou 2
Gostei ligeiramente	6 ou 1
Indefinidamente	5 ou 0
...Desgostei ligeiramente	4 ou -1
Desgostei moderadamente	3 ou -2
Desgostei muito	2 ou -3
Desgostei extremamente	1 ou -4

Para a realização da análise de resultados das escalas hedônicas são realizados dois métodos, o primeiro é a distribuição de frequência dos valores hedônicos, por meio de histogramas, os quais facilitam a visualização da segmentação dos valores hedônicos de cada amostra, deixando claro os níveis de aceitação e de rejeição e possibilitando uma comparação de desempenho entre duas ou mais amostras

A segunda forma de analisar os resultados, é através da análise de variância (ANOVA), onde é possível desprezar as individualidades e assumir que todos apresentam o mesmo tipo de comportamento, e também, caso sejam comparados mais de duas amostras é possível aplicar um teste de comparação de médias.

Problema 18. *DEU ERRADO Buscando comparar a aceitação do público com relação a três diferentes marcas de salsichas (A, B e C), foi realizado um teste utilizando uma escala hedônica de nove pontos e um total de 108 participantes. Os resultados da pesquisa se encontram na tabela abaixo. Qual a conclusão oferecida pelos testes?*

Avaliador	Amostras			
	A	B	C	Total
1	6	8	6	20
2	5	8	7	20
3	5	7	6	18
4	5	8	5	18
5	7	7	4	18
Total	28	38	28	94
Média	5,6	7,6	5,6	18,8

4.4.2 Escala FACT

A escala FACT, conhecida também como escala de atitude ou de intenção de compra, tem como objetivo medir o grau de aceitabilidade de um produto a partir das atitudes do consumidor com relação a frequência em que ele estaria disposto a consumir o produto em um determinado período.

Essa escala costuma ser dividida em nove categorias, sendo muito utilizada para a realização de testes de aceitação onde o público alvo não está familiarizado com o produto. Ela também tende a ser mais sensível que a escala hedônica, pois busca medir uma atitude ao invés do interesse afetivo, para compreender melhor, observe abaixo o exemplo de escala FACT.

Figura 4.4.2: Modelo de ficha de avaliação para o teste de aceitação utilizando a escala de atitude (FACT)

Nome _____

Data _____

Sexo: M() F()

Idade: _____ anos

Por favor, prove a amostra servida e marque a resposta que melhor corresponde ao seu julgamento (atitude).

Código da amostra: _____

() Comeria isso sempre que tivesse oportunidade.
 () Comeria isso muito frequentemente.
 () Comeria isso frequentemente.
 () Gosto disso e comeria de vez em quando.
 () Comeria isso se tivesse acessível, mas não me esforçaria para isso.
 () Não gosto disso, mas comeria ocasionalmente.
 () Raramente comeria isso.
 () Só comeria isso se não pudesse escolher outro alimento.
 () Só comeria isso se fosse forçado (a).

Comentários: _____

Fonte: [5]

As palavras ou frases utilizadas na montagem da escala FACT estão separadas por escalas previamente definidas, com base nas atitudes de consumo, esquematizadas em ordem sucessiva de atitudes, buscando facilitar a decisão de resposta do consumidor.

As palavras/frases da escala são convertidas em valores numéricos para que seja possível realizar a análise de resultados, utilizando o mesmo conceito de conversão apresentado na imagem 4.4.1. Caso seja avaliado apenas um produto pela escala, é possível expressar o resultado final pela média das amostras, entretanto, se a intenção do teste for comparar a aceitação de duas ou mais amostras, o uso da ANOVA é o método recomendado.

Problema 19. ADICIONAR EXEMPLO

4.4.3 Escala JAR

A escala do ideal, também chamada de Just Right Scales (JRS), é utilizada quando o objetivo da análise sensorial é otimizar a intensidade de um determinado atributo no produto, como maciez, docura, entre outros, otimizar as variáveis de processamento (como tempo de torragem, de cozimento, etc) e também otimizar o nível de adição de um determinado ingrediente. Portanto, de modo geral, essas escalas são utilizadas para determinar o grau de intensidade de uma característica sensorial específica no produto.

As escalas JAR podem ou não ser balanceadas, onde as mais utilizadas atualmente são aquelas entre 5 a 7 categorias. Para que seja possível analisar seus resultados, existem dois meios: o primeiro é calcular a porcentagem de avaliadores que respondeu cada categoria específica da amostra avaliada, ou comparar a distribuição das respostas das amostras com as de uma amostrapadrão pelo teste do qui-quadrado. O segundo meio, é analisar os resultados estabelecendo um calor mínimo de respostas para a categoria "ideal", sendo o valor de 70% o adequado.

Figura 4.4.3: Exemplo de modelo de ficha de avaliação para o teste de aceitação utilizando a escala do ideal (JRS)

Nome _____
 Data _____
 Sexo: M () F () Idade: _____ anos
 Por favor, prove a amostra servida e marque a resposta que melhor corresponde ao quanto próximo do ideal encontra-se cada um dos atributos.

Código da amostra: _____					
Doçura	()	()	()	()	()
	Pouco		Ideal		Muito
Acidez	()	()	()	()	()
	Pouco		Ideal		Muito
Amargor	()	()	()	()	()
	Pouco		Ideal		Muito

Fonte: Adaptado de [5]

Problema 20. (*Adaptado de Minim*) Uma indústria em São Paulo está desenvolvendo uma nova formulação de balas de morango e deseja avaliar a aceitação dela em relação ao sabor de morango. Para isso, 50 consumidores avaliaram a aceitação da bala de morango, utilizando a escala do ideal de sete categorias. Tanto o modelo da tabela JAR 4.4.4 quanto os resultados dos julgamentos dos consumidores 4.4.5 estão apresentados abaixo

Figura 4.4.4: Modelo de ficha de avaliação para o teste de aceitação utilizando a escala do ideal (JRS) para avaliar o atributo sensorial sabor de morango

Nome _____
 Data _____
 Sexo: M () F () Idade: _____ anos
 Por favor, prove a amostra de bala de morango servida e marque a resposta que melhor corresponde ao quanto próximo do ideal encontra-se a intensidade do sabor morango.
 Código da amostra: _____
 7 - sabor de morango muito mais intenso que o ideal
 6 - sabor de morango razoavelmente mais intenso que o ideal
 5 - sabor de morango levemente mais intenso que o ideal
 4 - sabor de morango de intensidade quase ideal
 3 - sabor de morango levemente mais fraco que o ideal
 2 - sabor de morango moderadamente mais fraco que o ideal
 1 - sabor de morango muito mais fraco que o ideal
 Valor: _____
 Comentários: _____

Fonte: Adaptado de [5]

Figura 4.4.5: Resultados do teste de aceitação da amostra de bala de morango utilizando a escala JRS

CATEGORIA	PORCENTAGEM (%) DAS RESPOSTAS DOS CONSUMIDORES EM CADA CATEGORIA
1	0
2	0
3	0
4	10
5	50
6	30
7	10

Fonte: Adaptado de [5]

Para observar melhor a análise de resultados, é possível montar um histograma com as porcentagens de respostas para cada categoria. Através dele, nota-se que apenas 10% dos avaliadores consideraram a intensidade do sabor de morango próximo do ideal, os outros 90% consideraram ele mais intenso que o ideal, dentre os quais 50% dos avaliadores considerou o

sabor levemente mais intenso que o ideal e 30% considerou razoavelmente mais intenso. Assim, a indústria deverá reformular a bala de morango, diminuindo a concentração da essência de morango.

4.5 Associação de Palavras na Análise Sensorial

O método da associação de palavras é um método qualitativo, muito utilizado na psicologia e sociologia, que se baseia na suposição de que quando um indivíduo é exposto a um estímulo e é pedido que ele associe a ideias vindas na mente, não há restrições quanto às definições mentais do estímulo em sua mente, diferente de entrevistas fechadas e questionários, que possuem resultados mais tendenciosos.

A metodologia da associação de palavras avalia desde a expectativa do consumidor ao ter a intenção de comprar um produto até sua aceitação, visto que o sucesso de um produto depende da combinação desses dois fatores, o momento em que o consumidor compra, e a da confirmação da expectativa. Por avaliar esses fatores, possui uma grande aplicação na área de desenvolvimento de produtos e Marketing, conseguindo extrair diversos impactos do produto no consumidor.

A utilização e aplicação desse método para que se obtenha bons resultados, avaliando a geração da expectativa e confirmação da expectativa em relação a um determinado produto, é combinado com metodologias de análise sensorial, possibilitando que se desenvolvam abordagens integradas, capazes de avaliar os atributos do produto e suas interações na influência do comportamento do consumidor.

4.5.1 Etapas da Associação de Palavras

Esse método é composto por quatro etapas, sendo elas:

(i) Definição do Estímulo a ser estudado

É possível utilizar diversos estímulos para a aplicação desse método, como imagens, palavras, ou descrições de interesse ao estudo, possibilitando que o consumidor seja livre para responder e interpretar o estímulo de acordo com suas experiências e conhecimentos.

(ii) Recrutamento dos Participante

A amostragem utilizada na Associação de palavras normalmente é por conveniência, sendo um método de análise não probabilístico, utilizado para se obter resultados grosseiros, sem se atentar aos custos ou tempo necessário, sendo recomendado que se escolha participantes com critérios que andam de acordo com o tipo de pesquisa, como frequência de consumo, idade, buscando sempre atingir o público alvo do produto em estudo, e por não se tratar de um método probabilístico, o pesquisador pode definir o número de avaliadores necessários para seu estudo.

(iii) Condução da Técnica

É realizado entrevistas individuais, onde é pedido ao consumidor que descreva as quatro primeiras imagens, pensamentos, associações ou sentimentos que apareçam em sua mente ao se deparar com o estímulo apresentado. Os estímulos devem ser apresentados de forma monódica e aleatória, fazendo com que não haja influência da ordem em que são apresentados os estímulos.

A abordagem desse teste pode ser em locais públicos, como praças, supermercados, shoppings etc.

(iv) Análise dos Dados

Os termos mais utilizados são agrupados de acordo com sua similaridade, sendo feito por meio do método de triangulação, já que se trata de uma avaliação qualitativa. O método de triangulação fornece um meio consideravelmente eficaz de avaliar o grau de convergência, bem como explicita as divergências entre os dados obtidos [5].

Após as categorias serem formadas, é verificado a concordância entre elas por três pesquisadores, sendo que somente as categorias mencionadas por mais de 10% dos avaliadores serão consideradas para análise, sendo que para essas contagens serem feitas é considerado o número total de avaliadores que utilizam aquela categoria para descrever os estímulos, não o total de avaliadores.

As frequências obtidas são organizadas em tabelas de contingência, cruzando as variáveis (categorias formadas) com os estímulos que foram avaliados (amostras), para então, aplicar-se o teste qui-quadrado, possibilitando verificar se há ou não associação dos estímulos com as categorias levantadas no estudo.

4.5.2 Exemplo de aplicação

Exemplo a ser adicionado

4.6 Características Não Sensoriais

Todos os produtos possuem características intrínsecas, onde é possível analisar com cuidado utilizando a análise sensorial, sendo características provenientes de todos os compostos presentes no alimento, garantindo sua qualidade sensorial. Porém, os consumidores também são influenciados por outros fatores que englobam o produto para finalizar sua escolha de compra, aceitação e preferência, como conceitos individuais que foram desenvolvidos durante sua vida, preço, embalagem, entre outros, esses sendo chamados os fatores não sensoriais. O consumidor

também associa o alimento a uma marca, sua embalagem, rótulo, fazendo com que o produto passe por um grande processo de avaliação antes mesmo de sua avaliação sensorial.

Observando esses fatores a serem avaliados, é visto que a análise sensorial deve ser combinada com métodos modernos de pesquisa de marketing, possibilitando analisar a influência dos fatores extrínsecos na escolha do consumidor, ou até mesmo a interação dos fatores intrínsecos e extrínsecos, utilizando métodos estatísticos de análise para que seja interpretado de maneira objetiva a relação entre a análise sensorial e o comportamento do consumidor ao ter sua primeira interação com o produto.

A expectativa tem uma grande influência até mesmo no sabor do produto, podendo piorar ou melhorar a experiência do consumidor antes de provar o alimento. Se o consumidor possui uma alta expectativa no produto, as chances de compra são muito maiores do que quando se tem uma expectativa negativa, sendo que é algo gerado por características extrínsecas do produto.

Existem dois grandes grupos que dividem as características não sensoriais dos produtos, sendo elas advindas das próprias características do consumidor e as características do alimento, sendo elas listadas a seguir:

Características não sensoriais do consumidor:

- (a) Conceitos éticos, culturais e religiosos
- (b) Preocupações nutricionais (bem-estar)
- (c) Idade
- (d) Sexo
- (e) Renda
- (f) Fatores sociais e influências

Características não sensorias do alimento:

- (a) Preço
- (b) Origem e tecnologia utilizada
- (c) Segurança e valor nutritivo
- (d) Marca
- (e) Embalagem

(f) Praticidade

O efeito da expectativa na aceitação de alimentos podem ser avaliados com testes sensoriais que consigam envolver todas as sessões de aceitação (teste cego, teste da característica não sensorial e teste com informações sobre a característica não sensorial).

4.6.1 Análise estatística

Para começar a estudar a influência das características não sensoriais em um alimento, é realizado os três testes sensoriais que englobam os processos de aceitação citados anteriormente e descritos a seguir.

- **Teste cego (sessão 1)**

Nessa sessão não é apresentado nenhuma característica não sensorial para o julgador, sendo que cada amostra é codificada com números de três dígitos aleatórios.

- **Teste da característica não sensorial (sessão 2)**

Aqui é apresentado ao consumidor as informações não sensoriais do produto anterior.

- **Teste com informação (sessão 3)**

Agora é apresentado para o avaliador o alimento junto com sua característica não sensorial e pedido que julgue a amostra, anotando sua aceitação na ficha disponibilizada contendo normalmente uma escala hedônica de nove pontos.

Os dados obtidos em cada teste podem ser avaliados utilizando a ANOVA (Análise de Variância), explicada no Capítulo 2.2.1.1 e caso a diferença seja significativa pode ser combinado com um teste de comparação múltipla. Para amostras pareadas é possível utilizar um teste t, comparando a sessão 1 com a 2, sessão 3 com a 2 e a sessão 3 com a 1. Também há outros métodos simples que podem ser utilizados, como a Análise de risco relativo (relative risk) e alternativamente, o método da Regressão de Probit, esses métodos serão explicados a seguir.

4.6.1.1 Teste t para amostras pareadas

Para realizar a análise de dados utilizando o teste t, é calculado todas as diferenças entre as notas da escala hedônica fornecidas por avaliador, sendo calculado para as amostras no teste da característica não sensorial e no teste cego (sessão 2 com a sessão 1), sessão 3 com a sessão 2 e, em seguida, sessão 3 com a sessão 1, assim, é englobado todas as 3 sessões no teste. Ao realizar esse teste, está sendo verificado a diferença entre as notas de cada um dos pares de sessões, buscando enxergar se há diferença entre as notas ou não, tendo a seguinte hipótese:

$$H_0 : \bar{D} = 0$$

$$H_a : \bar{D} \neq 0$$

Então, na hipótese nula (H_0) temos que a diferença das médias é igual a zero, ou seja, as médias são iguais, e na hipótese alternativa (H_a) temos que a diferença é diferente de zero, ou seja, são diferentes, portanto:

$$t = \frac{\bar{d}_i - \bar{D}}{\frac{s(d_i)}{\sqrt{n}}} \quad (4.6.1)$$

onde:

\bar{d}_i = média das diferenças da amostra i entre as sessões avaliadas (amostral)

\bar{D} = média das diferenças da população (sob H_0 , assume o valor 0)

$s(d_i)$ = desvio-padrão das diferenças amostrais.

n = número de avaliações por amostra

O valor de t será considerado significativo (onde rejeita H_0 se for igual ou superior ao valor t tabelado, para um determinado nível de significância α (bilateral), considerando um número $n-1$ de graus de liberdade.

Analizando os resultados desse teste, é possível concluir se há ou não influência das características não sensoriais na sua aceitação ou a expectativa gerada por elas na percepção do consumidor.

4.6.1.2 Análise de risco (relative risk)

Esse teste é realizado utilizando os dados das sessões 1 e 3, onde tem como objetivo comparar a probabilidade de se obter uma nota menor ou igual a 5 nesses dois testes, pois nesta análise o 5 é considerado como uma nota ruim, sendo a categoria "indeferente".

Para realizar essa análise os dados são organizados em uma tabela de frequência de dupla entrada, contendo as respostas das respostas boas (maior ou igual a 6) e ruins (menor ou igual a 5) da sessão 1 e das respostas boas e ruins da sessão 3. Com essa tabela será calculado o risco de obter uma resposta em duas condições diferentes, por exemplo, o risco de se obter uma resposta ruim na sessão 1 e o risco de se obter uma resposta ruim na sessão 3.

Para melhor entendimento, observe a tabela a seguir (Tabela 4.6.1).

Tabela 4.6.1: Tabela de frequências de notas boas e ruins obtidas na sessão 1 e 2

Teste	Resposta		
	Boa	Ruim	Total
Sessão 1	t_{11}	t_{12}	T1
Sessão 3	t_{21}	t_{22}	T2
Total	T.1	T.2	

Observando essa tabela, temos que o risco de se obter uma resposta ruim na sessão 1 é:

$$R_1 = \frac{t_{12}}{N}$$

Já na sessão 3 seria:

$$R_2 = \frac{t_{22}}{N}$$

Agora para calcular o risco relativo, dado por RR, é pela seguinte equação:

$$RR = \frac{R_1}{R_2} = \frac{t_{12}}{t_{22}} \quad (4.6.2)$$

Nesse teste, as hipóteses a serem analisadas são:

$$H_0 : RR = 1$$

$$H_a : RR \neq 1$$

Se a hipótese nula não for rejeitada, é concluído que os riscos de se obter uma nota ruim são iguais nas duas sessões, ou seja, não há influência dos fatores extrínsecos na nota do produto avaliado. Caso a hipótese alternativa seja válida, ou seja, rejeita-se H_0 , se o valor RR for maior que 1, o risco de se obter uma nota ruim é maior no teste cego, ou seja, há um impacto das características extrínsecas na nota do avaliador, nesse caso, aumentando a nota (melhorando as características do produto). Caso seja menor que 1, o risco de notas ruins é menor no teste cego, concluindo que as características não sensoriais impactam negativamente na nota do avaliador.

4.7 Pesquisa de mercado e a indústria de alimentos

A pesquisa de mercado no geral é muito importante para qualquer área que envolva a venda de um produto, para que seja possível enxergar qual é a necessidade do consumidor, auxiliar no desenvolvimento de novos produtos etc. Utilizando a pesquisa de mercado, é possível

ver o desejo, comportamento ou necessidade dos consumidores, possibilitando visualizar qual será a aceitação de um produto.

Basicamente, a pesquisa de mercado busca gerar informações que serão usadas para identificar possíveis oportunidades ou até mesmo ameaças dentro do mercado, dando informações necessárias para as questões levantadas e formulando qual método de coleta de dados poderá ser utilizado.

Para que tal pesquisa seja realizada, é importante ter concretizado qual é a pergunta que deverá ser respondida, o problema que será resolvido, sendo um importante passo a ser definido. Após o problema ser definido, irá ser buscado a solução para esse problema, o que será feito para solucioná-lo.

4.7.1 Tipos de pesquisa

Há três tipos de pesquisas, descritiva, exploratória e causal/experimental, sendo que uma pesquisa pode englobar vários tipos, mas não é recomendado que analise todas juntamente.

1. Pesquisa exploratória

Esse tipo de análise, é feito para iniciar a investigação de um problema, sendo uma pesquisa mais informal, exigindo flexibilidade e criatividade, buscando conhecer melhor seus concorrentes e uma possível expansão comercial do seu produto. Como coleta de dados, pode ser utilizado entrevistas pessoais ou questionários, não precisando ser uma amostra representativa

2. Pesquisa descritiva

Essa pesquisa é muito utilizada para descrever algo, como qual é o tipo de consumidor que compra seu produto, a idade, gênero, possibilitando que a comunicação com seu cliente seja aprimorado e aproximado, atendendo suas expectativas, melhorando o marketing, ou até mesmo saber se é necessário expandir seu consumo para outro possível público.

Ressalta-se que o objetivo principal é descrever e não explicar o fenômeno ou objeto estudado; entretanto, a comparação dos resultados constitui passo intermediário entre a lógica da explicação e descrição [5].

Um dos métodos de coleta de dado dessa pesquisa é um método qualitativo, como o grupo focal, já apresentado no Capítulo 4.2, ou até mesmo métodos quantitativos, utilizando questionários.

3. Pesquisa casual

Estabelece uma relação de causa e efeito entre as variáveis de estudo. É apropriada quando se deseja estabelecer quais variáveis são as causas (variáveis independentes) e quais são os efeitos (variáveis dependentes) de um fenômeno [5].

Utilizando esse tipo de pesquisa, é possível analisar a reação do consumidor com a mudança de um ingrediente na formulação do produto, ou até mesmo de uma informação no rótulo, visando buscar a resposta se essas mudanças causaram um impacto positivo ou negativo na venda do produto.

4.7.2 Coleta de dados

Para uma pesquisa de mercado, há três tipos de métodos de coleta de dados possíveis, sendo o método de observação, entrevista e interativo, sendo que o seu objetivo final definirá qual deverá ser utilizado para se obter os melhores resultados.

A técnica de observação deverá ser utilizada quando o objetivo é levantar informações preliminares do consumidor, por exemplo, pesquisas que dizem qual marca de suco é escolhida pelo consumidor no dia a dia, observando como o cliente se comporta no supermercado durante sua compra.

Essa observação direta do pesquisador com o comportamento do consumidor é uma vantagem desse tipo de pesquisa, possibilitando também que pessoas ou crianças com dificuldade em se expressar participem com melhores resultados na pesquisa, podendo ser feito de forma disfarçada, fazendo com que a artificialidade diminua. Porém, como se trata de um método que se tem mais proximidade entre o pesquisador e o consumidor, é necessário maturidade para que não haja influências da subjetividade por parte do pesquisador, também é necessário um grande número de observações para levar esses dados para uma representatividade geral, além disso, esse tipo de pesquisa não é capaz de detectar o que motiva um consumidor a comprar o produto.

O método de entrevista, consiste no contato do pesquisador e o consumidor, podendo ser realizada até mesmo por telefone, sendo uma forma mais rápida e prática, mas ao mesmo tempo, não se tem uma visualização do consumidor, fazendo com que as expressões corporais e faciais não sejam levadas em conta.

As entrevistas também podem ser realizadas por correspondência ou pessoalmente. Por correspondência, seria enviado para diversos consumidores questionários para serem respondidos, e como a entrevista por telefone, possui o problema de não ter a visualização do consumidor e não possui a capacidade de auxiliar o consumidor na entrevista. Porém, tem a vantagem de se poder enviar as correspondências para diversos lugares, sem precisar de um número grande de respostas, levando em conta que quem responde os questionários são os próprios consumidores. Como não se tem um retorno total dos questionários enviados, é necessário enviar uma quantidade maior, para se ter um retorno representativo.

As entrevistas pessoais são as mais utilizadas, possibilitando que o pesquisador tenha um contato direto com o entrevistado, fazendo com que seja possível instruí-lo e ter a visão de seu comportamento físico, buscando observar se suas respostas não contradizem sua linguagem corporal. Porém, é necessário o deslocamento do pesquisador, um grande grupo amostral, eleboração de roteiros, padronização de questionário e das entrevistas, para que o erro amostral seja o menor possível,

Finalizando os métodos de coleta de dados, será falado sobre o método interativo, onde será apresentado no computador o questionário com suas instruções, sendo um método agil e prático, porém não há controle da amostra.

Após o pesquisador definir qual será o estilo de coleta de dados, deverá ser elaborado os roteiros e questionários cuidadosamente para que evitar o máximo possível resposta tendenciosas, dúvidas no questionário, perguntas desnecessárias, buscando deixar tudo o mais claro possível, não cansando o entrevistado. Sendo assim, temos que ter em mente que um questionário tem os seguintes objetivos:

- (i) Transformar a informação que se deseja receber em perguntas.
- (ii) Minimizar o tédio e cansaço do entrevistado, o motivando a dar respostas de qualidade.
- (iii) Evitar erros causados por perguntas mal formuladas e também ter questionários homogêneos.

A formação de cada pergunta do questionário é de extrema importância para o resultado final do questionário, pois a qualidade da resposta depende muito da clareza da questão.

4.7.3 Níveis de mensuração

Para cada tipo de informação coletada, um tipo de escala diferente é utilizado, pois informações que são coletadas de forma inadequada não são capazes de se obter um resultado minimamente preciso. Sendo assim, as variáveis são divididas em quatro níveis diferentes, as qualitativas (nominais e ordinais) e as quantitativas (intervalares e de razão).

Tipos de variáveis:

- **Nominais**

Não definem intensidade, apenas se algo existe ou não.

- **Ordinais**

Definem a existência de uma variável, e as ordena de acordo com sua intensidade.

- **intervalares**

Nesse tipo de variável normalmente é utilizado números discretos, e permitem mensurar a existência de um fenômeno.

- **Razão**

Mensura a existência de um certo fenômeno, e ainda possibilita a realização de certas comparações.

Por meio dessas informações, é importante ter conhecimento de que o questionário elaborado deverá levar em conta qual é o tipo de variável que está sendo analisada. Resumindo, variáveis nominais servem para estratificações ou tipificações, enquanto as variáveis ordinais observam as dependências entre fenômenos. As intervalares e de razão permitem medir a intensidade de um fenômeno.

4.7.4 Estilo das questões

Existem diversos tipos de questões que podem ser inseridas nos questionários, e cada tipo irá variar, também, o estilo da resposta do entrevistado. Os estilos que serão abordados são: questões abertas, fechadas, semiabertas e com declarações.

- (a) **Questões Abertas**

Dão ao entrevistado uma maior liberdade de expressão, por conta disso, é mais trabalhosa e pode haver a possibilidade de se ter respostas que não agregam para a pesquisa.

- (b) **Questões Fechadas**

São limitados a escolher uma alternativa e podem ser divididos da seguinte forma:

- (a) Dicotômicas: Apenas duas opções de respostas, como: gosto ou não gosto.
- (b) Encadeadas: Uma pergunta depende da resposta da outra.
- (c) De ordem de preferência: normalmente organizadas em números limitados, que deve ser respondido de acordo com sua preferência.
- (d) Escala Likert: Se indica o grau de concordância ou discordância.
- (e) Escala de intenção de compra: Qual a chance de se obter o produto

- (c) **Questões Semiabertas**

Junção de uma pergunta aberta e uma fechada.

- (d) **Declarações**

Irá determinar o quanto o entrevistado concorda ou discorda com uma afirmação.

Além do estilo de cada questão, é importante considerar que as questões devem ser disponibilizadas de forma clara, evitando questões duplas (duas perguntas em uma só questão), evitar questões irrelevantes etc.

4.7.5 Considerações sobre a amostragem

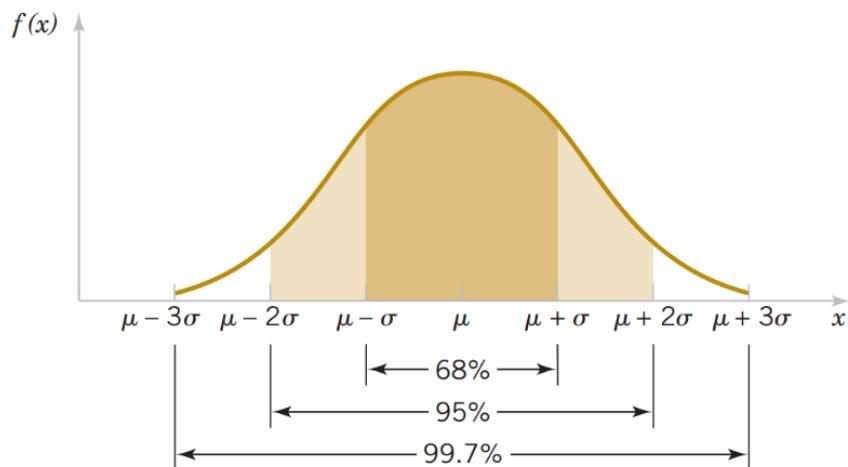
Há vários tipos de amostragem, como as probabilísticas e não probabilísticas e as escolhas delas se baseiam em questões da natureza da pesquisa, por exemplo, em pesquisas que o objetivo é inferir os resultados para a população no geral, a amostragem probabilística é preferível.

4.7.5.1 Determinação do número da amostra

Para determinar o número necessário de amostra, é considerado que a população e a amostra obtém uma distribuição normal, com uma média central, variando para a direita com $(\mu+1\sigma, \mu+2\sigma, \mu+3\sigma)$ e para a esquerda $(\mu-1\sigma, \mu-2\sigma, \mu-3\sigma)$, como mostrado na Figura 4.7.1.

Temos que um desvio-padrão representa 68% da área sobre a curva, 1,96 representa 95% da distribuição, dois desvios-padrão representam 95,5% e três desvios-padrão são 99,7% da distribuição da curva.

Figura 4.7.1: Curva da distribuição normal e porcentagem dos casos



Utilizando a curva da normal, é possível inferir resultados ou proporções da população, utilizando o número de desvio-padrão (Z), o desvio-padrão da média ($\sigma_{\bar{X}}$) e o erro amostral (σ_p).

Pode ser calculado o tamanho amostral com estimadores da porporção populacional (p). Sendo que (\hat{p}) seria a quantidade de elementos favoráveis e (\hat{q}) os desfavoráveis, mas como nem sempre se tem a estimativa de p , considera-se que \hat{p} e \hat{q} como 0,5 (50%) [5].

O tamanho da população ou universo estudado também deve ser lido em conta, onde universos finitos seria aquele menor ou igual a 10 mil e os infinitos maiores que 10 mil.

Assim, temos para universos infinitos:

$$E = \sqrt{\frac{\hat{p} \cdot \hat{q} \cdot z}{n}}$$

onde:

E = erro amostral.

\hat{p} = proporção de favoráveis ao atributo pesquisado;

\hat{q} = proporção de desfavoráveis;

Z = números de desvios-padrão;

N = tamanho da população; e

n = tamanho da amostra.

Como o objetivo é encontrar o tamanho da amostra, isola-se o n :

$$n = \frac{Z^2 \cdot \hat{p} \cdot \hat{q}}{E^2}$$

Quando se trata de um universo finito, é adicionado um fator de correção:

$$E = \sqrt{\frac{\hat{p} \cdot \hat{q} \cdot z}{n}} \sqrt{\frac{N - n}{N - 1}}$$

onde:

E = erro amostral.

\hat{p} = proporção de favoráveis ao atributo pesquisado;

\hat{q} = proporção de desfavoráveis;

Z = números de desvios-padrão;

N = tamanho da população; e

n = tamanho da amostra.

4.7.6 Exemplo de cálculo

1. População finita:

Problema 21. *Uma hamburgueria que serve diariamente 4 mil refeições planeja incluir um novo hambúrguer com carne de cordeiro. Sabe-se que 5% dos clientes já consomem esse tipo de carne, qual o tamanho da amostra que deverá ser pesquisada? Considere o erro como 5% e 95% como segurança.*

De acordo com o problema, se tem:

$$E = 0,05$$

$$\hat{p} = 0,5$$

$$\hat{q} = 0,95$$

$$Z = 1,96$$

$$N = 4000$$

$$n = ?$$

$$E = \sqrt{\frac{\hat{p}\cdot\hat{q}\cdot Z}{n}} \sqrt{\frac{N-n}{N-1}}$$

$$0,05 = \sqrt{\frac{0,5 \cdot 0,95 \cdot 1,96}{n}} \sqrt{\frac{4000-n}{4000-1}}$$

$n = 341$ **Conta errada?**

2. População infinita

Problema 22. *Deseja-se saber quantos brasileiros gostam de brownie. Qual o número de entrevistados com 6% de erro e 95% de confiança?*

$$E = 0,06$$

$$\hat{p} = 0,5$$

$$\hat{q} = 0,5$$

$$Z = 1,96$$

$$N = ?$$

$$n = ?$$

$$n = \frac{Z^2 \cdot \hat{p} \cdot \hat{q}}{E^2}$$

$$n = \frac{1,96^2 \cdot 0,5 \cdot 0,5}{0,06^2}$$

$$n = 267$$

4.8 Análise dos Dados

Para análise dos dados pode ser utilizados programas computacionais, como Excel, SPSS, SAS e RStudio. Em geral, a análise dos dados de uma pesquisa de mercado inicia-se com uma análise exploratória de dados (AED), para que seja possível visualizar o comportamento médio

das amostras. AED é definido como a análise de medidas de centro e medidas de variação, com o suporte de gráficos e ferramentas e estatísticas, como:

(a) **Medidas de Centro**

- (a) Média: medida de centro encontrada com a soma de todos os valores dividido pelo número total de valores.
- (b) Mediana: É o valor do meio de quando os valores estão organizados na forma crescente ou decrescente.
- (c) Moda: O valor que mais aparece.
- (d) Ponto médio: O valor entre o maior valor e o menor valor.

(b) **Variação**

Quantos os valores variam entre sí, podendo ser observado pela variança, desvio-padrão, amplitude dos dados e coeficiente de variação.

(c) **Distribuição**

Irá mostrar como a natureza dos dados se comporta, se é simétrica, assimétrica etc.

(d) **Outliers**

São os valores discrepantes da amostra.

Todos os dados obtidos podem ser organizados em diversos gráficos para uma facil visualização e interpretação.

4.9 Mapa de Preferência

Com o objetivo de analisar os dados obtidos em testes afetivos de como a considerar as respostas individuais de cada um dos consumidores, com relação às características de qualidade de um produto, e não apenas a média do grupo que avaliou o produto, foi criada a técnica do Mapa de Preferência, o qual tem sido muito utilizado no desenvolvimento de novos produtos, facilitando a avaliação de aceitabilidade dele e se o produto se enquadra no segmento de mercado desejado. Ele também é utilizado na otimização de produtos, uma vez que busca encontrar as características sensoriais responsáveis pela direção da preferência, através da realização de análise de dados de análises sensoriais e de pesquisas de mercado [5].

Utilizando o mapa de preferência é possível elaborar uma representação gráfica (Multi-dimensional Scaling - MDS) das preferências individuais dos consumidores, um procedimento estatístico multidimensional, através de procedimentos estatísticos multivariados. Com isso, é

possível visualizar de modo claro se ocorreu a aceitação ou não de um produto pelo consumidor e relacionar essa escolha com as características sensoriais do produto, além de também permitir uma avaliação da aceitabilidade do produto dentro do mercado em função das características de qualidade [6][5].

O Mapa de Preferência pode ser dividido em duas categorias: o Mapa de Preferência Interno (MDPREF), que é a forma mais simples de mapeamento, utilizado apenas quando se utiliza na análise os dados hedônico dos consumidores, ou seja, quando se utiliza dados de aceitação/preferência gerados a partir de testes afetivos. A segunda categoria é o Mapa de Preferência Externo (PREFMAP), onde além dos dados obtidos pelos testes afetivos também utiliza-se dados de uma fonte externa, como medidas descritivas, por exemplo [5] [32]. Observe abaixo algumas diferenças entre os dois tipos de mapas.

Figura 4.9.1: Diferenças básicas entre mapeamento de preferências internas e externas

	Mapa de Preferência Interno	Mapa de Preferência Externo
Enfatiza	Preferência	Percepção sensorial
Posição do produto no mapa	Considera a variação dos dados hedônicos ou de preferência	Considera a variação nos dados sensoriais (geralmente dados descritivos)
Dimensão do mapa	Explica a variabilidade máxima na preferência com diferentes direções entre os produtos	Explica a variabilidade máxima nas direções sensoriais entre os produtos

Fonte: Adaptado de [32]

4.9.1 Análise de componentes principais (PCA)

Para compreender corretamente os mapas de preferência, é necessário abordar primeiro a sua base: a Análise de Componentes Principais (PCA, do inglês Principal Component Analysis) . Essa técnica busca reduzir a dimensionalidade dos dados ao mesmo tempo em que mantém a maior parte da informação contida nela, em outras palavras, deseja-se transformar um conjunto original de variáveis correlacionadas em um novo conjunto de variáveis não correlacionadas (chamadas de componentes principais), os quais são combinações lineares. Logo, cada componente principal é uma combinação linear de todas as variáveis originais [5] [32].

Realizando essa redução da dimensão dos dados, é possível descartar aqueles componentes não tão relevantes para a análise, resultando em uma quantidade mínima de variação nos dados ao fim da pesquisa, além de reduzir sua complexidade, facilitando a visualização e a análise. Esses componentes principais devem ser capazes de extrair as tendências dos dados e/ou ações dos consumidores, entretanto, apesar de ser possível extrair diversos componentes principais,

nem todos podem ser significativos na prática, e portanto o pesquisador deve examinar cada um deles para selecionar as mais relevantes para a análise [4][32].

A análise de componentes principais podem ou não ser normalizada, caso os componentes do conjunto de variáveis originais sejam uma normal variada, então os componentes principais serão não relacionados e possuem uma distribuição normal. A depender do tipo de dados utilizado, a PCA pode ser efetuada através de matrizes de coariância (utiliza os dados originais, e os scores principais são obtidos através das médias das variáveis) ou de correlação (dados obtidos em escalas e/ou com unidade muito diferentes). De modo geral, as variáveis são padronizadas em zero e um [5].

APROFUNDAR MAIS SOBRE A PARTE DAS MATRIZES

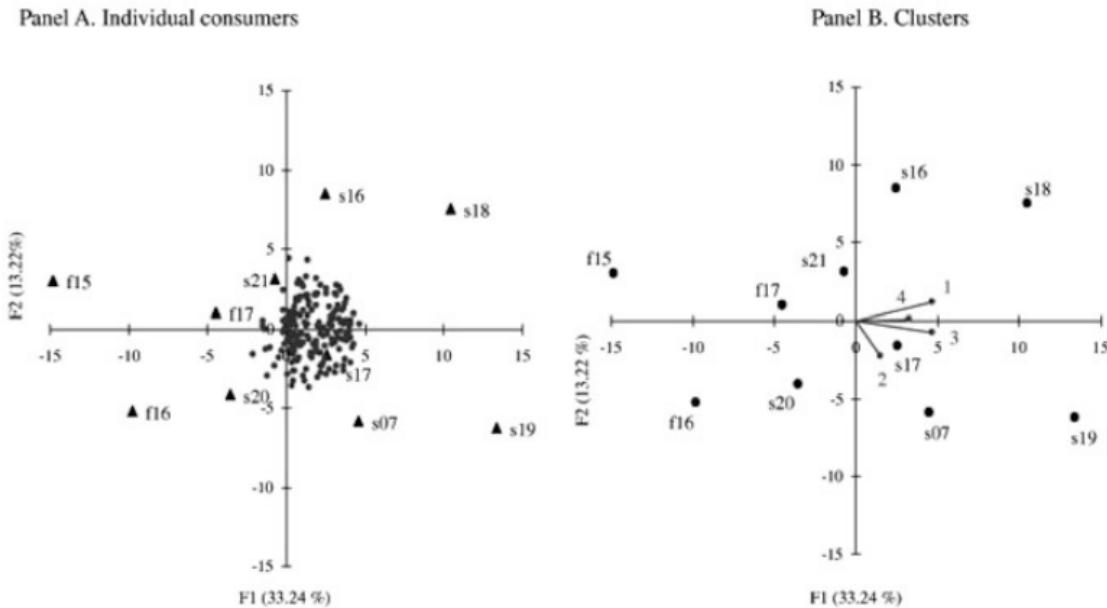
4.9.2 Mapa de preferência interno

O mapeamento interno de preferências utiliza como base o método PCA, onde geralmente as pontuações (scores) do consumidor como variáveis são organizadas em colunas, e as amostras são posicionadas em linhas. O objetivo deste estilo de mapeamento é encontrar um pequeno número de componentes principais, sendo recomendado a utilização de dois ou três no máximo, que possam explicar de modo eficiente a porcentagem de variação das respostas dos scores. Logo, esse modelo de mapa pode ser realizado utilizando como base um modelo vetorial para cada um dos consumidores (modo mais utilizado em estudos sensoriais), os quais constumam ser representados por uma seta apontado para o sentido do aumento de preferência[32].

Através dele é possível concluir o quanto o produto avaliado é aceito e/ou preferido pelos avaliadores, uma vez que ocorre a correlação entre os dados obtidos pela análise descritiva quantitativa, análises físico-químicas ou análises instrumentais, com as dimensões de preferência. Sendo um modelo vetorial, pode ser representado graficamente de diversas maneiras, uma delas é a representação de cada um dos consumidores como um ponto no gráfico com quadrantes[32][5].

Para a formação de um mapa de preferências interno ideal, é necessário fornecer aos consumidores ao menor seis produtos que diferem entre si, existem casos de pesquisas que utilizaram uma quantidade menor de amostras, mas é necessário uma extrema atenção pois o overfitting (sobreajuste de dados) é um problema grave na análise. Além disso, todos os consumidores devem avaliar todas as amostras[32]. Observe a figura abaixo:

Figura 4.9.2: Exemplos de mapas de preferência interno



Fonte: Adaptado de [32]

A figura é um exemplo de mapa de preferência interna utilizando dez amostras de presuntos, onde cada consumidor é representado no mapa por um ponto preto, sendo que é possível desenhar uma linha vetorial que indica quão vem a preferência do indivíduo é explicada pelas simensões que são plotadas. O painel A representa as posições individuais de cada consumidor, e é possível identificar que a grande maioria se localiza no lado direito, já o painel B mostra quatro grupos de consumidores em clusters[32].

4.10 Redes Neurais Artificiais

O cérebro humano é o órgão mais importante do sistema nervoso, capaz de processar uma vasta quantidade de informações em um curto período de tempo, o qual funciona através de uma rede de células complexas: os neurônios. Na tentativa de compreender e simular o funcionamento ideal do cérebro, pesquisadores criaram o modelo de Redes Neurais Artificiais (RNAs), os quais são sistemas computacionais baseados na matemática e em algoritmos denominados lógica de limiar (threshold logic), podendo ser aplicada na criação e formação de inteligências artificiais.

As redes neurais de um cérebro possuem unidades de processamento que são os neurônios, capazes de formar conexões entre si e criar uma rede de processamento de informações, além

disso, pode-se afirmar que quanto mais se utiliza o caminho feito pelo neurônio, mais fortalecido ele será, sendo este um conceito importante para compreender como os humanos são capazes de aprender coisas novas. Logo, as RNAs são capazes de receberem um ou mais sinais como entrada e computam uma ou mais saídas, entretanto, para que o processo ocorra de modo adequado é necessário que as RNAs passam por um estágio de aprendizado, assim como o cérebro, para posteriormente gerarem respostas corretas aos questionamentos, podendo levar a uma generalização da informação aprendida[5].

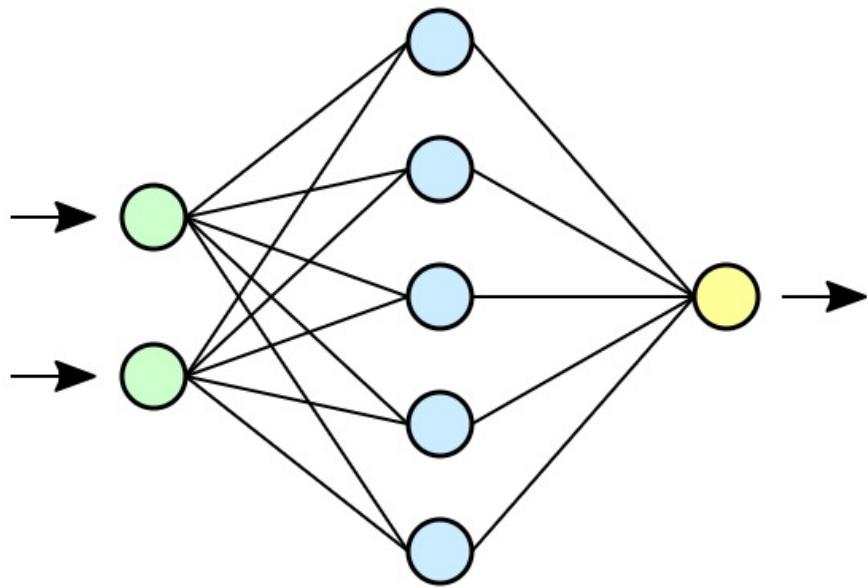
4.10.1 Organização de um RNA

Considere que cada neurônio humano é capaz de receber, combinar e processar sinais provenientes de outros neurônios, caso o sinal recebido seja forte o neurônio é ativado e produz um sinal-resposta de saída. O caminho que o sinal de saída percorre gera efeitos elétricos, que são levados de uma célula a outra devido a suas conexões, essa etapa do processo é denominada sinapse, e sua intensidade ou eficiência pode ser alterada quando o cérebro está na fase de aprendizagem, pois é capaz de formar o mecanismo básico de memorização do cérebro.

De modo parecido, as redes neurais artificiais possuem os nós ou nodos, que são similares aos neurônios, os quais são capazes de receberem sinais de entrada e combiná-los através de somas simples. As RNAs também possuem terminais de saída, como citado anteriormente, podem ser uma única saída ou um conjunto delas que podem se ligar a outros nós, formando um aglomerado de nós chamado de redez[5].

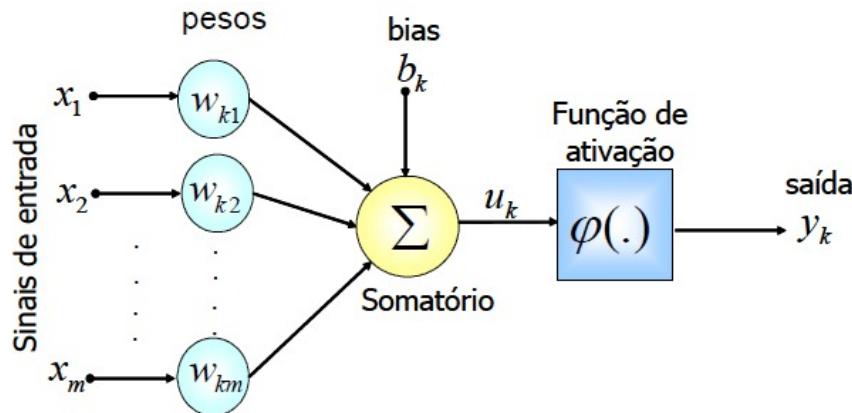
Além disso, as RNAs são organizadas em camadas que formam conexões aleatórias ou não entre si, existindo sempre duas camadas que possuem suas conexões voltadas para fora: a entrada e a saída, já as outras camadas existentes são nomeadas de camadas escondidas ou ocultas, como pode ser observado na imagem 4.10.1, onde a entrada é representada pelo círculo verde, as camadas ocultas são as de tom azul e a saída e o círculo amarelo.

Figura 4.10.1: Exemplo de rede neural artificial



Cada entrada é ligada ao corpo artificial através de pesos, que simulam a força sináptica das conexões neurais, logo, cada sinal de entrada é multiplicado por um peso e posteriormente combinados através de uma função aditiva (soma dos valores), assim, uma função de ativação (ou transferência) determinará se a saída do sinal será ativada ou não, como pode ser observado na figura 4.10.2. Além disso, dependendo do tipo de RNA utilizado, as saídas são realizadas em sistemas binários, capazes de imitar um sistema de portas lógicas que funcionam com um ou mais sinais de entrada e apenas uma saída, podendo ser do tipo e, ou, não. Assim, é possível emular os resultados em um computador digital[5].

Figura 4.10.2: Componentes típicos de um neurônio artificial



Em termos matemáticos, tem-se que a entrada pode ser representada pelo vetor $x^{(k-1)}$ e seus componentes $X_j^{(k-1)}$, com $j=1,n$, os quais resultam em uma saída X_j^k , que pode ou não se tornar uma entrada para outro nodo. Já os fatores de peso são representados pelo vetor W_{j,i^k} , e quando multiplicados pelas entradas podem resultar em resultados positivos ou negativos, os quais tendem a excitar ou inibir, respectivamente, o nó[5].

Com relação às funções de ativação (ou transferência), estas são aplicadas a depender do tipo de rede neural envolvida na resolução do problema. Nela, a saída é disparada ou não dependendo do valor da soma ponderada das entradas. Em termos matemáticos, essas funções podem ser representadas por ϕ , o qual diz a saída do neurônio pelos termos do potencial de ativação I_j^k . Essas funções podem possuir diversas formas, sendo as principais o Threshold (limiar), linear, sigmoidal e piecewise-linear, e abordaremos brevemente cada uma delas[5].

A Função *Threshold*, também conhecida como função limiar possui uma ativação do tipo $\phi(I_j^k)$, que resultam em dois valores finais: igual a um, caso o valor de $I_j^k \geq 0$, ou igual a zero, se $I_j^k < 0$. Portanto, a saída do neurônio j na camada k , utilizando a função de limiar, x_j^k , tem como resultado igual a um se $I_j^k \geq 0$ ou resultado igual a zero se $I_j^k < 0$ [5].

Já a função linear é expressa pela equação $x_j^k = \alpha(I_j^k)$, onde o α representa um número real responsável pela saída linear dos valores de entrada, sendo muito utilizada para se obter valores constantes em um determinado intervalo. A função *piecewise-linear* é derivada da anterior, sendo denominada como linear por partes.

Sendo a função de ativação mais utilizada em RNA's, a função sigmoidal possui como resultados um gráfico de curva muito semelhante com a letra "s". Ela é capaz de realizar um balanço entre um comportamento linear e não linear, além de ser estritamente crescente[5]. Essa função pode ser representada por:

$$\phi(I_j^k) = \frac{1}{1 + \exp(-\alpha * I_j^k)}$$

onde o α simboliza o parâmetro declividade da função, e é através dele que é possível a obtenção de funções sigmoidais de diferentes declividades.

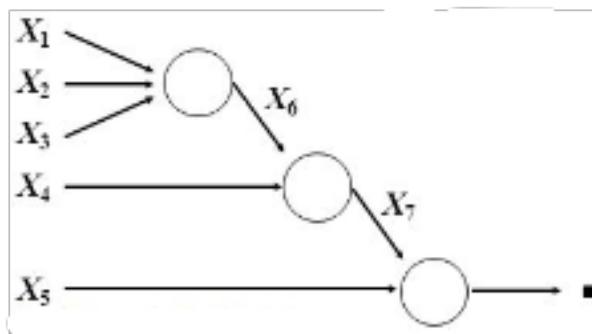
A arquitetura e topologia de uma rede neural artificial são parâmetros muito importantes, pois podem restringir os tipos de problemas que as redes podem resolver. A arquitetura de um RNA representa o número de camadas e nodos em cada camada, já a topologia define o tipo de conexão entre os nodos.

Com relação a arquitetura, elas podem ser de camada única, onde nesses casos só existe um nó entre qualquer entrada e saída, e são ideias para redes que buscam resolver problemas linearmente separáveis. Ou também podem ser de múltiplas camadas (ou redes recorrentes), os quais possuem mais de um nó e uma ou mais camadas ocultas entre a entrada e a saída, onde os nós dessas camadas escondidas são capazes de agirem na rede e extrair mais características da informação. O segundo tipo é ideal para a resolução de problemas com processamento temporal[5].

Tanto o número de camadas quanto o número de nós em redes múltiplas pode variar de acordo com a necessidade e objetivo da RNA, mas pesquisas indicam que entre uma a duas camadas ocultas são suficientes para a realização de tarefas de classificação de dados. Além disso, grande parte dos pesquisadores optam por utilizar poucas camadas (ou o menor tamanho possível da RNA), pois acredita-se que quanto mais simples for a arquitetura melhor será a generalização dos resultados[5].

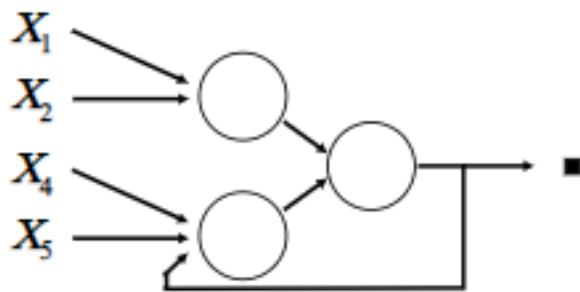
Em relação a topologia, ou seja, às conexões dos nodos, existem 3 tipos principais. O primeiro são as acíclicas ou Feedforward, no qual a saída de um neurônio de uma camada X não pode em hipótese alguma ser utilizada como entrada nos nós em camadas inferiores ou iguais a X. Além disso, nesse tipo de conexão a RNA possui uma não linearidade. Observe abaixo um exemplo de RNA acíclica.

Figura 4.10.3: Exemplo de RNA com conexões acíclicas



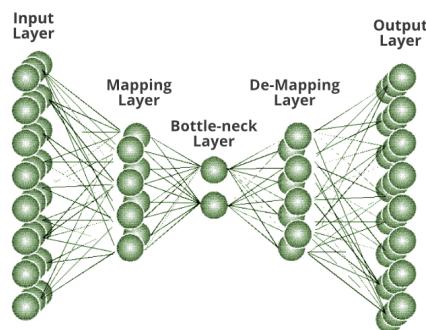
O segundo tipo são as RNA com conexões cíclicas, ou de Feedback, onde ocorre uma realimentação dos dados, pois a saída do neurônio em uma camada X qualquer pode ser usada como entrada de nodos em camas iguais ou inferiores a X, como representado na imagem abaixo.

Figura 4.10.4: Exemplo de RNA com conexões cíclicas



O terceiro e último tipo de conexão nas RNAs são as auto associativas, as quais possuem todas ligações cíclicas, ou seja são capazes de associar um padrão de entrada com ele mesmo e assim gerar uma recuperação ou regeneração de dados.

Figura 4.10.5: Exemplo de RNA com conexões autoassociativas



4.10.2 Etapas para resolução de problemas utilizando RNAs

Para compreender melhor como as RNAs são utilizadas para a resolução de problemas, é preciso conhecer as etapas que ocorrem nesse processo. A primeira etapa, e uma das mais importantes, é a de treinamento ou aprendizagem, no qual a rede passa por processos de adaptações/modificações nos pesos das conexões através da resposta a um estímulo. Nessa etapa, são apresentados à rede um conjunto de dados de entrada (estímulo do ambiente) para que sejam extraídas as informações mais relevantes, em seguida sofre modificações no decorrer da rede devido ao estímulo, ocasionando uma nova resposta do RNA ao ambiente[5].

Esse treinamento pode ocorrer de forma supervisionada, também chamado popularmente de aprendizagem com o professor, pois para cada entrada é feita uma comparação da resposta desejada com a calculada, assim são ajustados os pesos com o intuito de minimizar os erros, resultando em uma saída. Esse método é o mais utilizado em pesquisas, entretanto, suas resoluções são restritas apenas ao que lhe foi ensinado previamente, uma vez que a rede não consegue utilizar outras estratégias de solução além das apresentadas como exemplo[5].

O aprendizado pode ocorrer também de modo não supervisionado, ou seja, sem a presença de um “professor” que direciona o treinamento. Nele são fornecidos à rede apenas os dados de entrada sem as respostas prévias, assim a rede tende a se organizar internamente e desenvolver habilidades para reconhecer os estímulos externos e/ou criar um banco de dados. Esse estilo de aprendizado busca identificar os padrões de entradas sem a presença de um conjunto de alvos de interesse, assim, a rede aprende a realizar um mapeamento de entrada-saída, com propriedades específicas desejadas[5].

A segunda etapa para a solução de problemas em RNAs são os testes de desempenho da rede, os quais têm como principal objetivo avaliar como a rede neural artificial está processando o estímulo apresentado na entrada, na maior parte dos casos essa etapa ocorre em conjunto com o processo de aprendizagem. Nele, são fornecidos à entrada da rede dados que não foram previamente apresentados, assim realiza-se uma comparação entre o resultado final que a RNA chegou e o resultado real previamente calculado (quando existente)[5].

Para compreender melhor a aplicação das RNAs na área de alimentos, veja abaixo um exemplo prático

As cascas foram previamente higienizadas e descascadas, sendo removida a polpa da fruta e separadas em seis grupos: um controle (sem tratamento) e cinco tipos de tratamentos com diferentes concentrações de ácido ascórbico (AA) e cloreto de cálcio ($CaCl_2$), como representado na imagem abaixo.

Figura 4.10.6: Tipos de tratamentos realizados no processamento mínimo das cascas de manga

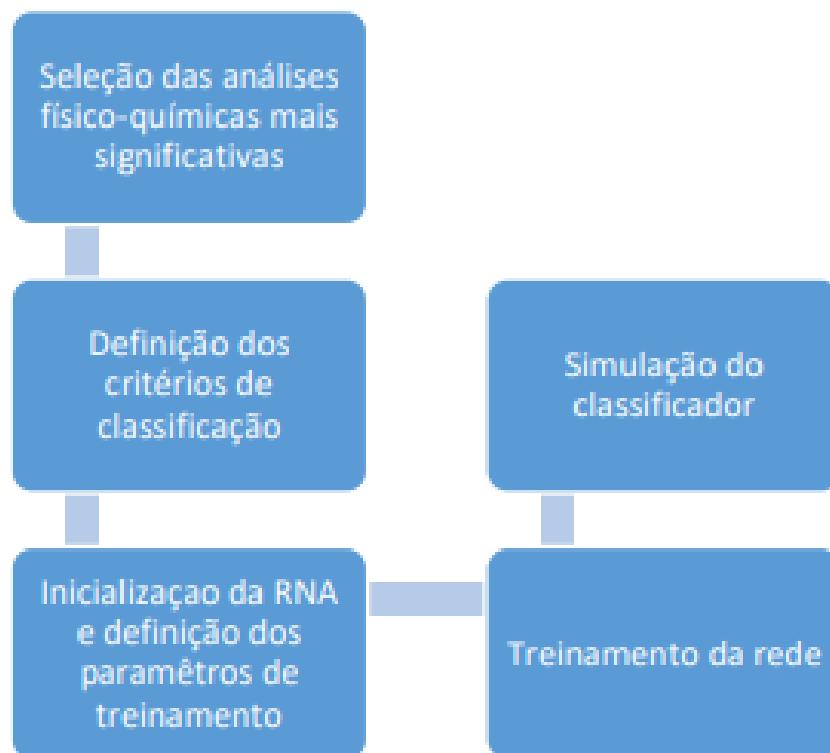
Identificação do tratamento	Tratamento
Controle	sem AA e $CaCl_2$
A	1,0% (AA) + 1,0% ($CaCl_2$)
B	2,0% (AA) + 2,0% ($CaCl_2$)
C	2,0% (AA) + 1,0% ($CaCl_2$)
D	1,5% (AA) + 1,5% ($CaCl_2$)
E*	1,0% (AA) + 2,0% ($CaCl_2$)

*O tratamento E foi utilizado apenas para maior robustez para rede neural.

Fonte: [?]

Para realizar a construção do RNA ideal seguiu-se uma sequência de etapas, como demonstrado no fluxograma abaixo.

Figura 4.10.7: Fluxograma das etapas de construção da Rede Neural Artificial



Fonte: [?]

Efetuou-se então uma análise dos componentes principais (ACP) utilizando uma matriz de correlação de Pearson, assim, com os dados obtidos selecionou-se os mais significantes para a construção da RNA, sendo eles: sólidos solúveis totais, acidez total titulável, teor de umidade, de proteína e de fibra insolúvel em detergente neutro(FDN).

Em seguida, foi definido os critérios para cada componente principal, os quais indicam o valor limitante para a classificação. Esses critérios foram determinados pela literatura, com base nos valores máximos obtidos para as análises mais significativas. Assim, as amostras que apresentarem valores, para cada variável, inferiores ao critério de classificação determinado, serão consideradas aptas para o consumo. Observe os critérios utilizados abaixo.

Figura 4.10.8: Critérios de classificação Rede Neural Artificial

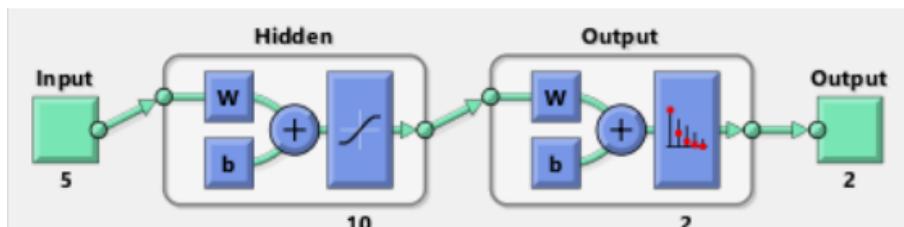
Autor	Análise e resultado
Ajila e Rao, 2013	FDN (%): 40,60
Imra <i>et al.</i> , 2013	Teor de proteínas (%): 4,88
Puligundla <i>et al.</i> , 2014	Sólidos solúveis totais (°BRIX): 25,00
Reinoso, 2017	Acidez total titulável: 0,92 g de ácido cítrico/100 mL de amostra
Alberton, 2014	Teor de umidade (%): 86,80

Fonte: [?]

Após isso, passou-se para a etapa de iniciação do RNA utilizando o software *MATLAB®* que possui uma linguagem de programação técnica baseada em matriz, rico em funções que são frequentemente utilizadas em problemas da área de engenharia, sendo capaz de expressar seus resultados em linguagem matemática. Nesse caso em específico os pesquisadores utilizaram a função patternnet, o qual é capaz de criar uma estrutura neural própria para solucionar problemas de reconhecimento de padrões.

Foram definidos também os parâmetros para o treinamento da rede neural artificial, sendo eles o número de entradas (*input*) igual a 5 (que representam as 5 variáveis avaliadas no estudo), o número de camadas internas igual a 1, o número de neurônios na camada interna igual a 10, e o número de saídas igual a 2 (pois a manga deve ser classificada como apta ou não para o consumo e processamento da farinha). Os valores atribuídos à saída são um para quando a amostra é considerada apta, e zero quando a amostra é inapta.

Figura 4.10.9: Parâmetros para o primeiro treinamento da rede neural

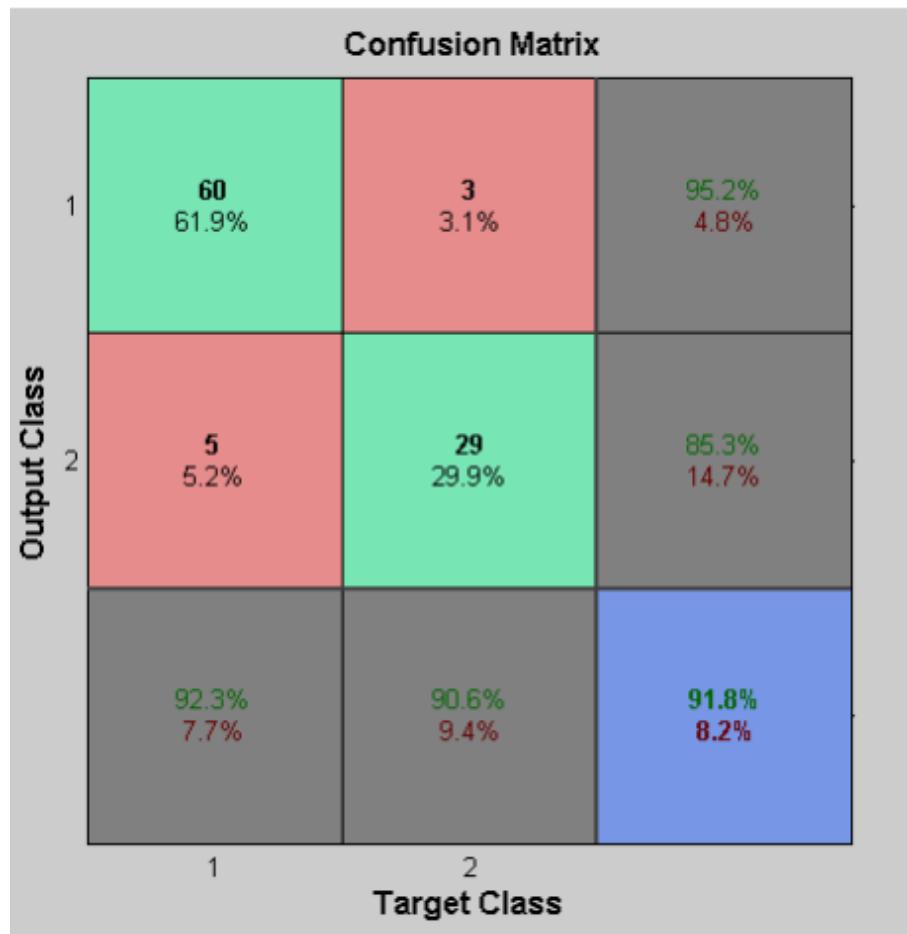


Fonte: [?]

Assim, iniciou-se o treinamento da rede utilizando como base um total de 98 amostras e três etapas internas: um treinamento para a definição dos pesos da rede, utilizando 50% das amostras, uma validação com o intuito de verificar possíveis erros na saída da rede (utilizou-se 25% das amostras), e por último o teste, etapa muito similar a verificação de analisar os

possíveis erros, também utilizando 25% do total de amostras. Após o treinamento, foi realizada a simulação do classificador, utilizando a função net e 97 novas amostras. Observe na figura os resultados da simulação

Figura 4.10.10: Quadro de resultados para Simulação do classificador com 1 camada interna 10 neurônios



Fonte: [?]

É possível observar que ocorreu um erro de 8,2%, sendo que entre as amostras analisadas, 60 delas foram classificadas como aptas para o consumo humano e processamento da farinha, baseado dos parâmetros pré-estabelecidos anteriormente. Entre as amostras, três delas foram classificadas como não aptas para o consumo de modo equivocado, ou seja, ocorreu um erro na classificação. Do mesmo modo, cinco amostras foram classificadas como aptas para o consumo de modo equivocado, quando na realidade eram inaptas.

4.11 Limiares Afetivos

O limiar de detecção (LD), é definido como o mínimo necessário de um determinado estímulo para que seja causado uma reação no indivíduo [5].

É um conceito que pode ser aplicado em diversas situações, como no enriquecimento de um determinado alimento com uma substância que pode causar um sabor amargo em grandes quantidades, então seria de grande importância determinar o ponto no qual essa substância começa a diminuir a aceitação sensorial do produto [5].

Porém, quando o limiar de detecção já está no nível necessário para que o indivíduo perceba o estímulo, não necessariamente acarretou na alteração da aceitação do produto, em diversos casos, é necessário ultrapassar esse limiar de detecção para que haja uma influência na aceitação do produto.

Por conta disso, Prescott et al. (2005) propuseram o primeiro limiar sensorial afetivo: o limiar de rejeição pelo consumidor (*Consumer Rejection Threshold - LRC*), que irá representar justamente em que ponto a preferência do produto é alterada.

Vale ressaltar que Prescott et al. (2005) utilizaram testes de preferência para determinar em que ponto a adição de um componente impactaria em qual amostra é a preferida, porém, nesse tipo de teste, não significa necessariamente que a amostra menos preferida não é aceita sensorialmente.

4.11.1 Limiar de rejeição pelo consumidor (LRC)

O Limiar de rejeição pelo consumidor foi proposto por Prescott et al. (2005), onde consiste em avaliar a preferência dos consumidores utilizando testes pareados de preferência dentro do método de estímulo constante da metodologia de limiar [5].

Para se determinar o LRC, pessoas que consomem o produto habitualmente devem realizar diferentes sessões de testes pareados de preferência, sendo que o número de sessões é igual ao número de amostras estímulo [5].

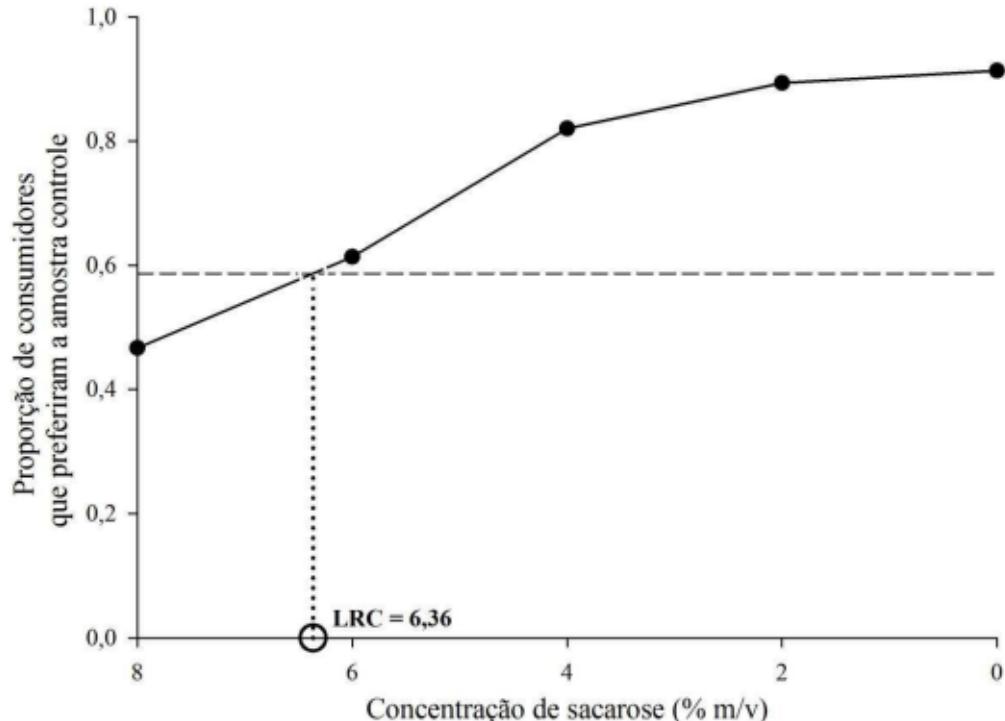
Em cada sessão deve ser indicado na ficha disponibilizada (4.4.1) qual a amostra preferida, e após realizar um branco no paladar, os consumidores recebem um novo par de amostras, sendo que as amostras devem ser apresentadas em ordem crescente ou decrescente de intensidade de estímulo entre as sessões, sendo que a amostra controle deve ser aleatorizada dentro de cada par [5].

4.11.1.1 Cálculo do limiar de rejeição do consumidor

Após coletar os dados, é elaborado um gráfico das proporções de consumidores que preferiram a amostra controle, que é representado pelo eixo Y, em função do estímulo das amostras estímulo, representado pelo eixo X [5], mostrado na Figura 4.11.1. A linha tracejada na

Figura 4.11.1 representa a proporção de consumidores necessária para preferência significativa ($p = 0,05$) de acordo com a distribuição binomial para testes pareados de preferência.

Figura 4.11.1: Proporção de consumidores que preferiram a amostra controle para cada concentração de sacarose da amostra estímulo.



Fonte: [?]

Calcula-se o LRC por interpolação, sendo a intensidade do estímulo correspondente à proporção de consumidores que preferiram a amostra controle necessária para uma significância estatística ($p = 0,05$).

Na Figura 4.11.1 a proporção de consumidores que preferem a amostra controle foi de 6,36% m\% v, indicando que, com a redução da concentração de sacarose para valores a partir de 6,36% m\% v, os consumidores tendem a preferir a amostra controle (9% de sacarose) [?].

4.11.2 Limiares hedônicos

Os limiares hedônicos, sendo o limiar de aceitação comprometida e Limiar de rejeição hedônica, foram propostos por Lima Filho et al. (2015) são determinados por meio da Metodologia dos Limiares Hedônicos (MLH) [5].

Para sua aplicação, alguns pré-requisitos são necessários, como:

1. O estímulo em estudo deve ser quantitativo
2. A variação da intensidade do estímulo deve alterar a aceitação do produto
3. A amostra controle deve ser a de maior aceitação
4. Deve utilizar uma faixa de intensidade do estímulo que inclua o limiar de aceitação comprometida e limiar de rejeição, ou seja, deve existir amostras que possuam um estímulo que tenha aceitação semelhante a amostra controle e também amostras que tenham menor aceitação que a controle.

Deve ser realizado testes preliminares para garantir que os três últimos requisitos estejam sendo atendidos, e também para determinar as intensidades do estímulo das amostras estímulo e do controle [5].

Para que o LAC e o LRH sejam determinados, são realizados os mesmos procedimentos de análise sensorial e coleta de dados, porém, a parte de análise de resultados e cálculo dos limiares são diferentes. Sendo assim, com apenas um experimento de análise sensorial é possível determinar o LAC e o LRH [5].

Na coleta de dados, consumidores habituais do produto a ser analisado irá participar de vários testes de aceitação, sendo que o número de sessões é igual ao número de estímulos, sendo recomendado no mínimo cinco sessões de testes de aceitação.

É apresentado para o consumidor duas amostras, onde uma é a amostra estímulo, em uma das intensidades do estímulo, e a outra a controle. Entre as sessões, as amostras devem ser apresentadas de forma crescente ou decrescente de acordo com a intensidade do estímulo, sendo que a posição da amostra estímulo deve ser aleatorizada.

Para que seja determinado se as amostras seriam apresentadas em ordem decrescente ou crescente, deve-se analisar o objetivo do teste. Por exemplo:

- Caso o objetivo do teste seja reduzir a concentração de sacarose de um alimento, as amostras devem ser servidas em ordem decrescente de intensidade de estímulo, fazendo com que as primeiras sessões tenham as menores concentrações, porém, mais próximas da amostra controle quando comparadas com as últimas amostras estímulo que serão comparadas. Sendo assim, conforme as sessões vão passando a aceitação do produto tenderia a diminuir, já que a concentração de sacarose também teria diminuído.
- Agora para verificar o limite de defeitos que um produto pode ter sem que ocorra alteração da sua aceitação, deve-se fornecer as amostras em ordem crescente, sendo que as primeiras amostras serão as sem defeitos e irão parecer com a amostra controle, mudando isso de acordo com o passar das sessões.

Além de depender do propósito da pesquisa a ordem de apresentação das amostras, a aplicação das amostras dessa forma impede a adaptação e fadiga dos avaliadores.

Para cada sessão do teste de aceitação, é solicitado aos julgadores que provem as amostras da esquerda para a direita, marcando na ficha contendo a escala hedônica de nove pontos, variando de 1 = "desgostei extremamente" a 9 = "gostei extremamente", sendo que após enxaguar a boca com água, será disponibilizado aos avaliadores um novo par de amostras.

4.11.2.1 Cálculo do limiar de aceitação comprometida

Com os dados obtidos em cada sessão do teste de aceitação, é realizado um teste t para amostras pareadas, sendo que de início calculam-se todas as diferenças entre as notas hedônicas do par de cada amostra de cada sessão, sendo subtraido a nota hedônica da amostra controle da nota hedônica da amostra estímulo. As hipóteses testadas são:

$$H_0 : \bar{D} = 0$$

$$H_a : \bar{D} \neq 0$$

ou seja, H_0 seria de que não há diferença significativa entre as amostras quanto a aceitação sensorial, sendo assim, temos a seguinte estatística do teste:

$$t_i = \frac{\bar{d}_i - \bar{D}}{\frac{s(d_i)}{\sqrt{n}}} \quad (4.11.1)$$

onde:

t_i = t calculado da sessão i;

d_i = média das diferenças das notas hedônicas da amostra controle e das notas hedônicas das amostras estímulo da sessão i;

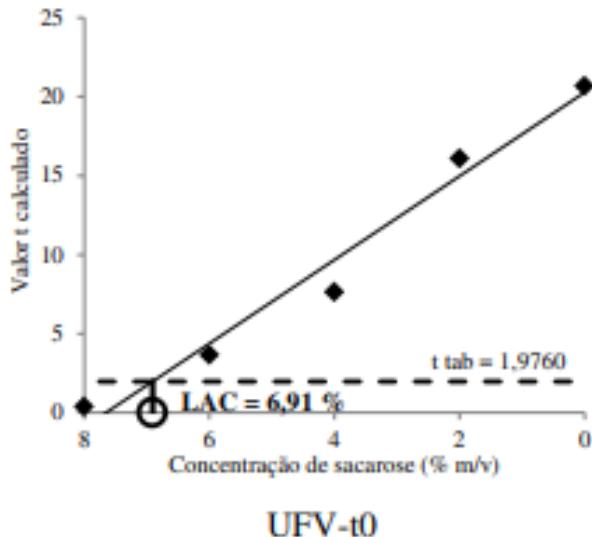
D_i = média das diferenças da população, considera-se 0, sob H_0 ;

$s(d_i)$ = desvio-padrão das diferenças amostrais; e

n = número de consumidores que realizaram o teste.

Quando é calculado o valor de t, é preparado um gráfico do valot calculado, representado pelo eixo (Y1) em função da intensidade do estímulo da amostra estímulo em estudo, representado pelo eixo (X). Com os dados experimentais é ajustado um modelo de regressão e determinado o coeficiente R^2 . Na Figura 4.11.2 está apresentando o t calculado para uma sessão (eixo y) em função da concentração de sacarose (eixo x).

Figura 4.11.2: Valor de t calculado em função das concentrações de sacarose da amostra estímulo desta sessão.



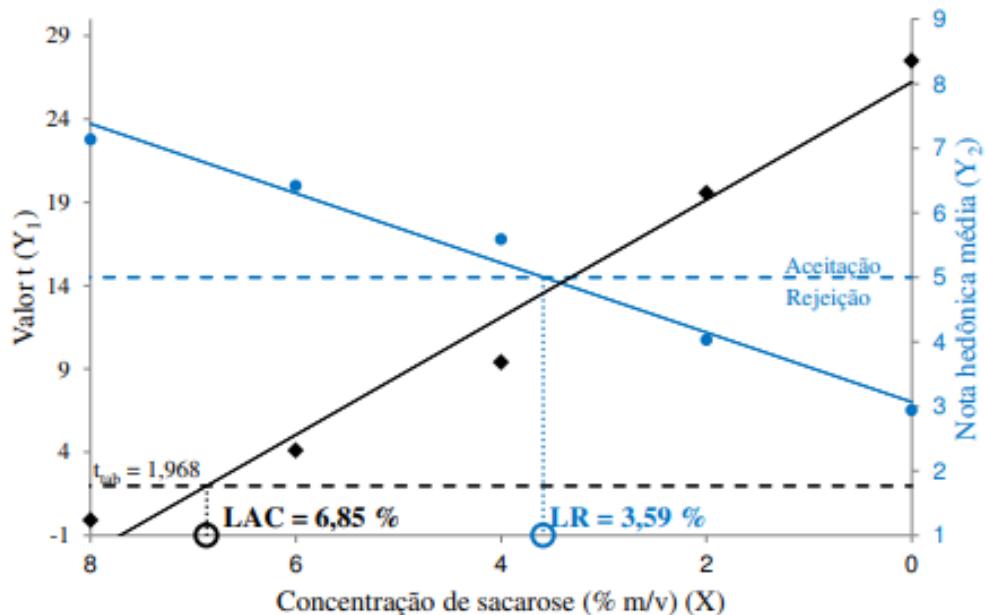
Fonte: [?]

Se o t calculado for maior ou igual ao valor t tabelado para determinado nível de significância α (teste bilateral) e para $n-1$ graus de liberdade, H_0 é rejeitado, ou seja, há diferença significativa entre as amostras quanto à aceitação, sendo que no gráfico é representado por uma linha tracejada, que é referente ao valor t , sendo 1,9670 na Figura 4.11.2.

4.11.2.2 Cálculo do limiar de rejeição hedônica

Para a determinação do LRH, o mesmo gráfico do LAC é utilizado, com a adição de um segundo eixo Y (Y2), que é referente às notas hedônicas médias das amostras estímulo, como mostrado na Figura 4.11.3.

Figura 4.11.3: Valores t calculados e notas hedônicas médias em função da concentração de sacarose, para consumidores com média preocupação com a saúde.



Fonte: [?]

No eixo Y2 o ponto corte é representado no gráfico por uma linha tracejada, referente a nota hedônica 5 (termo hedônica "indiferente"), que indica o momento do início da rejeição sensorial do produto, onde dificilmente o consumidor estará disposto a comprar o produto.

Capítulo 5

Referências

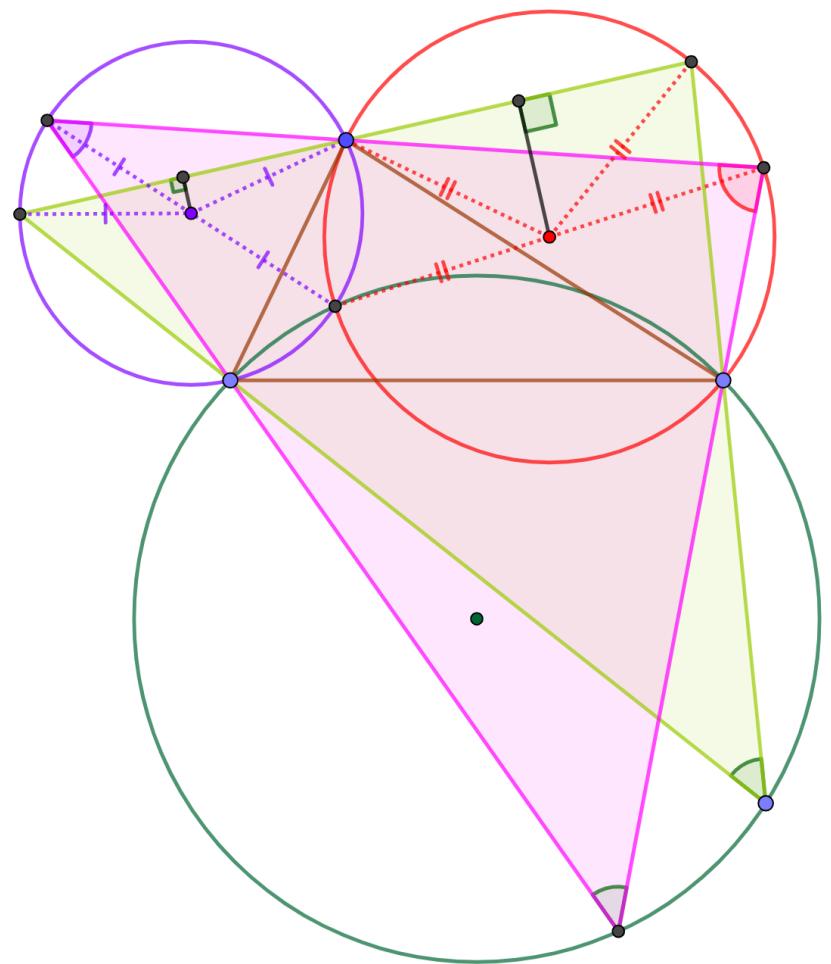
- [1] BANZATTO, DAVID A; KRONKA, SÉRGIO N. Experimentação Agrícola. 4a Edição Jaboticabal, FUNEP, 2006. 237 p. **2.2.1.1**
- [2] DUTCOSKY DEBONI, S. **Análise Sensorial de Alimentos** PUCPRESS, Curitiba, 2019. **1, 1.1, 1.2, 1.2, 1.3, 1.3.1, 1.3.2, 1.3.1, 1.3.1, 1.4, 1.5, 1.6, 2.1, 2.2.1, 2.2.1, 2.2.1.1, 2.2.1.1, 2.2.1.3, 1, 2.2.2, 2.2.4, 2.3.1.1, 2.3.1.4, 2.3.4, 2.3.4, 2.3.4, 2.3.5, 2.3.5, 2.3.6, 2.3.6, 2.4, 2.4, 2.4.1, 2.4.2, 2.5.1, 2.5.2, 3.2, 3.2.1, 3.2.2, 3.2.3, 3.2.4, 3.2.1, 3.2.6, 3.2.1, 3.2.3, 3.2.7, 3.2.8, 3.2.9, 3.2.11, 3.3, 3.4, 3.4.1, 3.4.2, 3.4.3, 3.4.4, 3.4, 3.6, 4.1**
- [3] VAZ DE FARIA, E. **Técnicas de Análise Sensorial** ITAL/LAFISE, Campinas, 2002. 116 p. **1.1, 1.2, 1.4, 2.3.1.1, 2.3.4, 2.3.5, 2.3.5, 3.6**
- [4] HONGYU, Kuang; SANDANIELO, Vera Lúcia Martins; DE OLIVEIRA JUNIOR, Gilmar Jorge. **Análise de componentes principais: resumo teórico, aplicação e interpretação.** E&S Engineering and science, v. 5, n. 1, p. 83-90, 2016. **4.9.1**
- [5] MINIM, V. P. R. Análise sensorial: estudos com consumidores. Vicosa: UFV, 2018. **1, 1.1, 4.1, 4.3.4, 4.3.2, 4.4.2, 4.4.3, 4.4.4, 4.4.5, (iv), 2, 3, 4.7.5.1, 4.9, 4.9.1, 4.9.2, 4.10, 4.10.1, 4.10.1, 4.10.1, 4.10.2, 4.11, 4.11.1, 4.11.1.1, 4.11.2, 4.11.2**
- [6] BEHRENS, Jorge H.; DA SILVA, Maria Aparecida AP; WAKELING, Ian N. Avaliação da aceitação de vinhos brancos varietais brasileiros através de testes sensoriais afetivos e técnica multivariada de mapa de preferência interno. Food Science and Technology, v. 19, p. 214-220, 1999. **4.9**
- [7] TEFF, K. L. Physiological effects of flavour perception. Trends in Food Science and Technology, v. 7, p. 448–453, 1996. **1.2**
- [8] ANZALDUA-MOLARES. La evaluación sensorial de los alimentos en la teoría y la práctica. Zaragoza: Acribia SA, 1994. 198 p. **1.2**

- [9] HUI, Y. H. Sensory evaluation of dairy products. *Dairy scienci and technology handbook*, VCH, v. 1, 1992.
- [10] CHRISTENSEN, R.H.B. **Statistical methodology for sensory discrimination tests and its implementation in sensR** Brock University, Canadá, 2018. [2.3.6.2](#)
- [11] ENNIS, D.M. *Journal of Sensory Studies: The power of sensory discrimination methods* v. 8, p. 353-370, 1993.
- [12] FARIA, E. V. D.; YOTSUYANAGI, K. **Técnicas de análise sensorial** 1 ed. Campinas: ITAL/LAFISE, 2002. [1.1](#), [1.2](#), [1.4](#), [2.3.1.1](#), [2.3.4](#), [2.3.5](#), [2.3.5](#), [3.6](#)
- [13] TEIXEIRA, L. V. Analise sensorial na indústria de alimentos. *Candido Tostes Dairy Institute*, v. 64, p. 12–21, 2009. [1.2](#), [1.3.1](#)
- [14] NARDELI, G.. Proposta de desenvolvimento e validação de um novo software voltado para análise sensorial de alimentos **TCC (Graduação) - Curso de Eng. de Alimentos, Patos de Minas: Universidade Federal de Uberlândia**, 2017.
- [15] ALVEZ, C. J.; OTERO, D. M.; KRUMREICH, F. D. **Testes Discriminativos**. Análise Sensorial Clássica: Fundamentos e Métodos, 2021. 31 p. Disponível em: [.](#) Acesso em: 19 out. 2022.
- [16] GUIMARAES, J. L. **Métodos Discriminativos de Análise Sensorial**. Pirassununga: Faculdade de Zootecnia e Engenharia de Alimentos, 2019. 65 slides. [2.2.1](#), [2.2.1.1](#), [2.2.1.1](#), [2.2.1.2](#)
- [17] ANJOS, Adilson dos. **Métodos Discriminativos**: Teste de ordenação. 2015. 22p. Disponível em: [.](#) Acesso em: 14 out. 2022. [2.3.6](#)
- [18] LÓPEZ LINARES, J.; BRUNO-ALFONSO, A.; BARBOSA, G. F. Bases numéricas na Olimpíada Internacional de Matemática. **Professor de Matemática Online (PMO)**, v. 7, n. 2, p. 195-204, 2019b. Disponível em: <https://doi.org/10.21711/2319023x2019/pmo715>. Acesso em: 14 jun. 2021.
- [19] LÓPEZ LINARES, J. **Soluções detalhadas para 20 problemas da Olimpíada Internacional de Matemática**. Portal de Livros Abertos da USP, Pirassununga: Faculdade de Zootecnia e Engenharia de Alimentos, 2020. 81 p. ISBN 978-65-87023-04-5 (e-book). Disponível em: <https://doi.org/10.11606/9786587023045>. Acesso em: 14 jun. 2021.
- [20] LÓPEZ LINARES, J.; BRUNO-ALFONSO, A.; BARBOSA, G. F. Três problemas sobre série harmônica na Olimpíada Internacional de Matemática. **C.Q.D. – Revista**

- Eletrônica Paulista de Matemática**, v. 17, p. 127-138, fev. 2020. Edição Ermac. DOI: 10.21167/cqdvol17ermac202023169664jllabagfb127138. Disponível em: <https://www.fc.unesp.br/Home/Departamentos/Matematica/revistacqd>. Acesso em: 14 jun. 2021.
- [21] LÓPEZ LINARES, J.; BRUNO-ALFONSO, A.; BARBOSA, G. F. Três problemas sobre desigualdades na Olimpíada Internacional de Matemática. **C.Q.D. – Revista Eletrônica Paulista de Matemática**, v. 18, p. 78-88, jul. 2020. DOI: 10.21167/cqdvol18202023169664jllabagfb7888. Disponível em: <https://www.fc.unesp.br/Home/Departamentos/Matematica/revistacqd>. Acesso em: 14 jun. 2021.
- [22] LÓPEZ LINARES, J. Três problemas sobre partições na Olimpíada Internacional de Matemática. **C.Q.D. – Revista Eletrônica Paulista de Matemática**, v. 19, p. 118-127, dez. 2020. DOI: 10.21167/cqdvol19202023169664jll118127. Disponível em: <https://www.fc.unesp.br/Home/Departamentos/Matematica/revistacqd>. Acesso em: 14 jun. 2021.
- [23] LÓPEZ LINARES, J.; SANTOS, J.P.M.; FIRMIANO, A. Cinco problemas sobre potência de um ponto em relação a uma circunferência e eixo radical em Olimpíadas Internacionais de Matemática. Aceito em **C.Q.D. – Revista Eletrônica Paulista de Matemática**, v. 19, p. -, jul. 2021. DOI:. Disponível em: <https://www.fc.unesp.br/Home/Departamentos/Matematica/revistacqd>. Acesso em: 14 jun. 2021.
- [24] INTERNATIONAL MATHEMATICAL OLYMPIAD. **Timeline**. 2019. Disponível em: <https://www.imo-official.org/organizers.aspx>. Acesso em: 14 jun. 2021.
- [25] DJUKIC, D. *et al. The IMO compendium: a collection of problems suggested for the International Mathematical Olympiads: 1959–2004*. New York: Springer, 2006. Disponível em: <http://web.cs.elte.hu/~nagyzoli/compendium.pdf>. Acesso em: 14 jun. 2021.
- [26] **Problems and Solutions of the 2nd IGO**. 2015. Disponível em: <https://igo-official.ir/events/2/>. Acesso em: 14 jun. 2021.
- [27] NETO, A. C. M. **Geometria**, Coleção ProfMat, SBM, Primeira Edição, ISBN: 9788585818937, 2013.
- [28] Delgado, J. *et al. Geometria Analítica*, Coleção ProfMat, SBM, Segunda Edição, ISBN: 9788583371212, 2017.

- [29] Pinheiro, P. R. O Círculo dos Nove Pontos. **Revista do Professor de Matemática**, v. 14, p. 1-9, 1989. Disponível em:
<https://www.rpm.org.br/cdrpm/14/12.htm>. Acesso em: 14 jun. 2021.
- [30] RODRIGUES, A. R. Napoleão e as Revoluções no Plano Euclidiano. **É Matemática, Oxente!**, v. 1, p. 1-9, 2019. Disponível em:
<http://ematematicaoxente.com.br>. Acesso em: 14 jun. 2021.
- [31] MORGADO, A. C.; CARVALHO, P. C. P. **Matemática discreta**. 2. ed. Rio de Janeiro: SBM, 2015. (Coleção ProfMat).
- [32] LAWLESS, H. T.; HEYMANN, H. **Sensory Evaluation of Food : Principles and Practices**. New York, NY: Springer New York, 2010. **3.3, 3.4, 3.4, 3.4, 3.5, 4.9, 4.9.1, 4.9.1, 4.9.2, 4.9.2**
- [33] CARRA, Thales Andrés. **Do mito ao show de realidade: procusto e o Big Brother** Jornal de Psicanálise, v. 52, n. 96, p. 213-223, 2019. **3.5.1**
- [34] STONE, H.; BLEIBAUM, R.; THOMAS, H. **Sensory Evaluation Practices**. California: Academic Press, 2012. **3.6.2, 3.6.2**
- [35] DOOLEY, L.; LEE, Y.-s.; MEULLENET, J.-F. **The application of check-all-that-apply (cata) consumer profiling to preference mapping of vanilla ice cream and its comparison to classical external preference mapping** Food Quality and Preference, Fayetteville, p. 394–401, 2010 **3.6.2, 3.6.2**
- [36] MEYNERS, M.; CASTURA, J. C. **Novel techniques in sensory characterization and consumer profiling** Check-all-that-apply questions. In: VARELA, P.; ARES, G. (Ed.). Boca Raton, FL: CRC Press, 2014. p. 271–305. **3.6.2, 3.6.2**
- [37] ALCANTARA, M. d.; FREITAS-Sa, D. D. G. C. **Metodologias sensoriais descritivas mais rápidas e versáteis – uma atualidade na ciência sensorial**. Brazilian Journal of Food Technology, Campinas, p. 12, 2018. **3.6.2**
- [38] MEILGAARD, M.; CIVILLE, G. V.; CARR, B. T. **Sensory evaluation, techniques** 4. ed. New York: CRC Press Inc., 2007. **3.6.2**
- [39] VIALI, L. **Testes não paramétricos - KAD (K amostras dependentes)** 2020. 11 p. **3.6.2**
- [40] DELARUE, J.; SIEFFERMANN, J. M. **Sensory mapping using flash profile. Comparison with a conventional descriptive method for the evaluation of the**

flavour of fruit dairy products. *Food Quality and Preference*, Barking, v. 15, n. 4, p. 383-392, 2004. **3.6.3**



ISBN 978-65-87023-14-4 (e-book)

Capítulo 6

ANEXOS

6.0.1 ANEXO A

Tabela 6.0.1: Número de avaliadores necessários para o Teste Triangular

α	pD = 50%	β				
		0,20	0,10	0,05	0,01	0,001
0,20	pD = 50%	7	12	16	25	36
0,10		12	15	20	30	43
0,05		16	20	23	35	48
0,01		25	30	35	47	62
0,001		36	43	48	62	81
0,20	pD = 40%	12	17	25	36	55
0,10		17	25	30	46	67
0,05		23	30	40	57	79
0,01		35	47	56	76	102
0,001		55	68	76	102	130
0,20	pD = 30%	20	28	39	64	97
0,10		30	43	54	81	119
0,05		40	53	66	98	136
0,01		62	82	97	131	181
0,001		93	120	138	181	233
0,20	pD = 20%	39	64	86	140	212
0,10		62	89	119	178	260
0,05		87	117	147	213	305
0,01		136	176	211	292	397
0,001		207	257	302	396	513
0,20	pD = 10%	149	238	325	529	819
0,10		240	348	457	683	1011
0,05		325	447	572	828	1181
0,01		525	680	824	1132	1539
0,001		803	996	1165	1530	1992

Tabela 6.0.2: *

Fonte: ASTM E1885-04.

6.0.2 ANEXO B

Tabela 6.0.3: Número mínimo de respostas corretas necessárias para estabelecer diferença significativa entre as amostras ao nível de erro alfa (α) para o Teste Triangular para o correspondente número de julgadores (n). Rejeite a hipótese de que não há diferença se o número de respostas corretas for maior ou igual ao tabelado

n	α					n	α				
	0,20	0,10	0,05	0,01	0,001		0,20	0,10	0,05	0,01	0,001
6	4	5	5	6		27	12	13	14	16	18
7	4	5	5	6	7	28	12	14	15	16	18
8	5	5	6	7	8	29	13	14	15	17	19
9	5	6	6	7	8	30	13	14	15	17	19
10	6	6	7	8	9	31	14	15	16	18	20
11	6	7	7	8	10	32	14	15	16	18	20
12	6	7	8	9	10	33	14	15	17	18	21
13	7	8	8	9	11	34	15	16	17	19	21
14	7	8	9	10	11	35	15	16	17	19	22
15	8	8	9	10	12	36	15	17	18	20	22
16	8	9	9	11	12	37	16	17	18	20	22
17	8	9	10	11	13	38	16	17	19	21	23
18	9	10	10	12	13	39	16	18	19	21	23
19	9	10	11	12	14	40	17	18	19	21	24
20	9	10	11	13	14	41	17	19	20	22	24
21	10	11	12	13	15	42	18	19	20	22	25
22	10	11	12	14	15	43	18	19	20	23	25
23	11	12	12	14	16	44	18	20	21	23	26
24	11	12	13	15	16	45	19	20	21	24	26
25	11	12	13	15	17	46	19	20	22	24	27
26	12	13	14	15	17	47	19	21	22	24	27
48	20	21	22	25	27	78	30	32	34	37	40
54	22	23	25	27	30	84	33	35	36	39	43
60	24	26	27	30	33	90	35	37	38	42	45
66	26	28	29	32	35	96	37	39	41	44	48
72	28	30	32	34	38	102	39	41	43	46	50

Tabela 6.0.4: *
Fonte: ASTM E1885-04.

H

Tabela 6.0.5: Número de julgadores necessários para Teste Pareado Monocaudal

α		β							
		0,50	0,40	0,30	0,20	0,10	0,05	0,01	0,001
0,5	P max = 75%	2	4	4	4	8	12	20	34
		2	4	4	6	10	14	28	42
		2	6	8	10	14	20	30	48
		6	6	10	12	20	26	40	58
		10	10	14	20	26	34	48	70
		14	16	18	24	34	42	58	82
		22	28	34	40	50	60	80	108
		38	44	52	62	72	84	108	140
0,50	P max = 70%	4	4	4	8	12	18	32	60
		4	4	6	8	14	26	42	70
		6	8	10	14	22	28	50	78
		6	10	12	20	30	40	60	94
		14	20	22	28	40	54	80	114
		18	24	30	38	54	68	94	132
		36	42	52	64	80	96	130	174
		62	72	82	96	118	136	176	228
0,50	P max = 65%	4	4	4	8	18	32	62	102
		4	6	8	14	30	42	76	120
		8	10	14	24	40	54	88	144
		10	18	22	32	50	68	110	166
		22	28	38	54	72	96	146	208
		30	42	54	70	94	120	174	244
		64	78	90	112	144	174	236	320
		108	126	144	172	210	246	318	412
0,50	P max = 60%	4	4	8	18	42	68	134	238
		6	10	24	36	60	94	172	282
		12	22	30	50	84	120	206	328
		22	32	50	78	112	158	254	384
		46	66	86	116	168	214	322	472
		72	94	120	158	214	268	392	554
		142	168	208	252	326	392	536	726
		242	282	328	386	480	556	732	944
0,50	P max = 55%	4	8	28	74	164	272	542	952
		10	36	62	124	238	362	672	1124
		30	72	118	200	334	480	810	1302
		82	130	194	294	452	618	1006	1556
		170	240	338	462	658	862	1310	1906
		282	370	476	620	866	1092	1584	2238
		550	666	820	1008	1302	1582	2170	2928
		962	1126	1310	1552	1908	2248	2938	3812

Tabela 6.0.6: *
 Fonte: ASTM E2164-08.

6.0.3 ANEXO C

6.0.4 ANEXO D

Tabela 6.0.7: Número mínimo de respostas corretas necessárias para estabelecer diferença significativa entre duas amostras para o Teste Pareado Monocaudal

n	Nível de significância (%)						n	Nível de significância (%)					
	0,50	0,20	0,10	0,05	0,01	0,001		0,50	0,20	0,10	0,05	0,01	0,001
4	3	4	4	—	—	—	29	16	18	19	20	22	24
5	4	4	5	5	—	—	30	16	18	20	20	22	24
6	4	5	6	6	—	—	31	16	19	20	21	23	25
7	4	6	6	7	7	—	32	17	19	21	22	24	26
8	5	6	7	7	8	—	33	17	20	21	22	24	26
9	6	7	7	8	9	—	34	18	20	22	23	25	27
10	6	7	8	9	10	10	35	19	21	22	23	25	27
11	6	8	9	9	10	11	36	19	22	23	24	26	28
12	7	8	9	10	11	12	40	21	24	25	26	28	31
13	7	9	10	10	12	13	44	23	26	27	28	31	33
14	8	10	10	11	12	13	48	25	28	29	31	33	36
15	9	10	11	12	13	14	52	27	30	32	33	35	38
16	9	11	12	12	14	15	56	29	32	34	35	38	40
17	9	11	12	13	14	16	60	31	34	36	37	40	43
18	10	12	13	13	15	16	64	33	36	38	40	42	46
19	10	12	13	14	15	17	68	35	38	40	42	45	48
20	11	13	14	15	16	18	72	37	41	42	44	47	50
21	12	13	14	15	17	18	76	39	43	45	46	49	52
22	12	14	15	16	17	19	80	41	45	47	48	51	55
23	12	15	16	16	18	20	84	43	47	49	51	54	57
24	13	15	16	17	19	20	88	45	49	51	53	56	59
25	13	16	17	18	19	21	92	47	51	53	55	58	62
26	14	16	17	18	20	22	96	49	53	55	57	60	64
27	14	17	18	19	20	22	100	51	55	57	59	63	66
28	15	17	18	19	21	23							

Tabela 6.0.8: *
Fonte: ASTM E2164-08.

6.0.5 ANEXO E

Tabela 6.0.9: Valores críticos de F para o teste de Friedman (riscos de 0,05 e 0,01)

	Número de amostras (ou produtos) P										
	3	4	5	6	7		3	4	5	6	7
Nível de significância						Nível de significância					
alpha = 0,05						alpha = 0,01					
7	7,143	7,8	9,11	10,62	12,07	8,857	10,371	11,97	13,69	15,35	
8	6,250	7,65	9,19	10,68	12,14	9,000	10,35	12,14	13,87	15,53	
9	6,222	7,66	9,22	10,73	12,19	9,667	10,44	12,27	14,01	15,68	
10	6,200	7,67	9,25	10,76	12,23	9,600	10,53	12,38	14,12	15,79	
11	6,545	7,68	9,27	10,79	12,27	9,455	10,60	12,46	14,21	15,89	
12	6,167	7,70	9,29	10,81	12,29	9,500	10,68	12,53	14,28	15,96	
13	6,000	7,70	9,30	10,83	12,37	9,385	10,72	12,58	14,34	16,03	
14	6,143	7,71	9,31	10,85	12,34	9,000	10,76	12,64	14,40	16,09	
15	6,400	7,72	9,33	10,87	12,35	8,933	10,80	12,68	14,44	16,14	
16	5,99	7,73	9,34	10,88	12,37	9,790	10,84	12,72	14,48	16,18	
17	5,99	7,73	9,34	10,89	12,38	8,810	10,87	12,74	14,52	16,22	
18	5,99	7,73	9,36	10,90	12,39	8,840	10,90	12,78	14,56	16,25	
19	5,99	7,74	9,36	10,91	12,40	8,86	10,92	12,81	14,58	16,27	
20	5,99	7,74	9,37	10,92	12,41	8,870	10,94	12,83	14,60	16,30	
∞	5,99	7,81	9,49	11,07	12,59	9,210	11,34	13,28	15,09	16,81	

Tabela 6.0.10: **Arrumar Tabela**

6.1 ANEXO F

Tabela 6.1.1: Número mínimo de respostas corretas necessárias para estabelecer diferença significativa entre duas amostras para o Teste Pareado Bicaudal

n	Nível de significância (%)						n	Nível de significância (%)					
	0,50	0,20	0,10	0,05	0,01	0,001		0,50	0,20	0,10	0,05	0,01	0,001
5	4	5	5	—	—	—	29	17	19	20	21	22	24
6	5	6	6	6	—	—	30	18	20	20	21	23	25
7	5	6	7	7	—	—	31	18	20	21	22	24	25
8	6	7	7	8	8	—	32	19	21	22	23	24	26
9	7	7	8	8	9	—	33	19	21	22	23	25	27
10	7	8	9	9	10	—	34	20	22	23	24	25	27
11	8	9	9	10	11	11	35	20	22	23	24	26	27
12	8	9	10	10	11	12	36	21	23	24	25	27	29
13	9	10	10	11	12	13	40	23	25	26	27	29	31
14	9	10	11	12	13	14	44	25	27	28	29	31	34
15	10	11	12	12	13	14	48	27	29	31	32	34	36
16	10	12	12	13	14	15	52	29	32	33	34	36	39
17	11	12	13	13	15	16	56	32	34	35	36	39	41
18	11	13	13	14	15	17	60	34	36	37	39	41	44
19	12	13	14	15	16	17	64	36	38	40	41	43	46
20	13	14	15	15	17	18	68	38	40	42	43	46	48
21	13	14	15	16	17	19	72	40	42	44	45	48	51
22	14	15	16	17	18	19	76	42	45	46	48	50	53
23	14	16	16	17	19	20	80	44	47	48	50	52	56
24	15	16	17	18	19	21	84	46	49	51	52	55	58
25	15	17	18	18	20	21	88	48	51	53	54	57	60
26	16	17	18	19	20	22	92	50	53	55	56	59	63
27	16	18	19	20	21	23	96	52	55	57	59	62	65
28	17	18	19	20	22	23	100	54	57	59	61	64	67

Tabela 6.1.2: *
Fonte: ASTM E2164-08.

Capítulo 7

APÊNDICE PROGRAMAÇÕES

7.1 Problema 3 - Teste de comparação múltipla

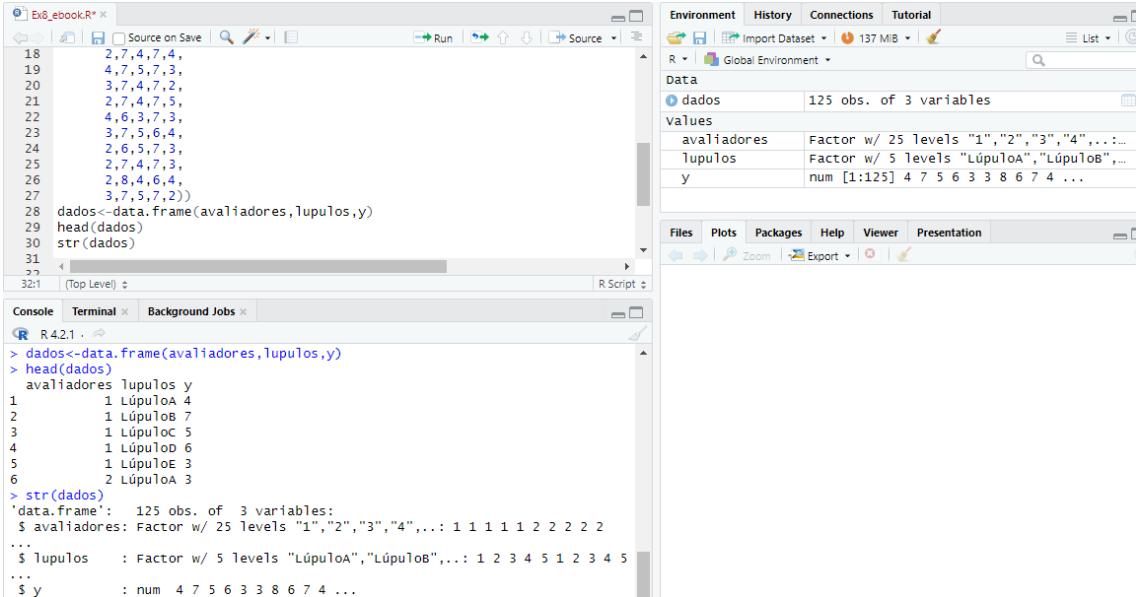
Figura 7.1.1: Importação do banco de dados



A screenshot of an RStudio session titled "Ex6_ebook.R*". The code editor contains the following R script:

```
1 Kavalidores<-as.factor(rep(c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25),each=5))
2 (lupulos<-as.factor(rep(c("LúpuloA", "LúpuloB", "LúpuloC", "LúpuloD", "LúpuloE"), 25)))
3 (y<-c(4,7,5,6,3,
4 3,8,6,7,4,
5 2,7,6,8,2,
6 3,8,5,8,4,
7 2,7,7,7,4,
8 2,8,5,7,3,
9 2,7,4,6,3,
10 4,7,3,6,4,
11 2,7,4,7,3,
12 2,6,3,5,2,
13 3,8,3,6,3,
14 3,7,4,5,3,
15 2,7,4,7,3,
16 2,6,3,6,6,
17 3,8,4,7,5,
18 2,7,4,7,4,
19 4,7,5,7,3,
20 3,7,4,7,2,
21 2,7,4,7,5,
22 4,6,3,7,3,
23 3,7,5,6,4,
24 2,6,5,7,3,
25 2,7,4,7,3,
26 2,8,4,6,4,
27 3,7,5,7,2))
```

Figura 7.1.2: Entendimento do banco de dados



The screenshot shows the RStudio interface. The code editor pane (top left) contains R script code for reading a dataset and printing its head and structure. The environment pane (top right) displays the dataset 'dados' with 125 observations and 3 variables: 'avaliadores' (factor with 25 levels), 'lupulos' (factor with 5 levels), and 'y' (numerical vector). The console pane (bottom left) shows the execution of the script, displaying the head of the dataset and its structure.

```

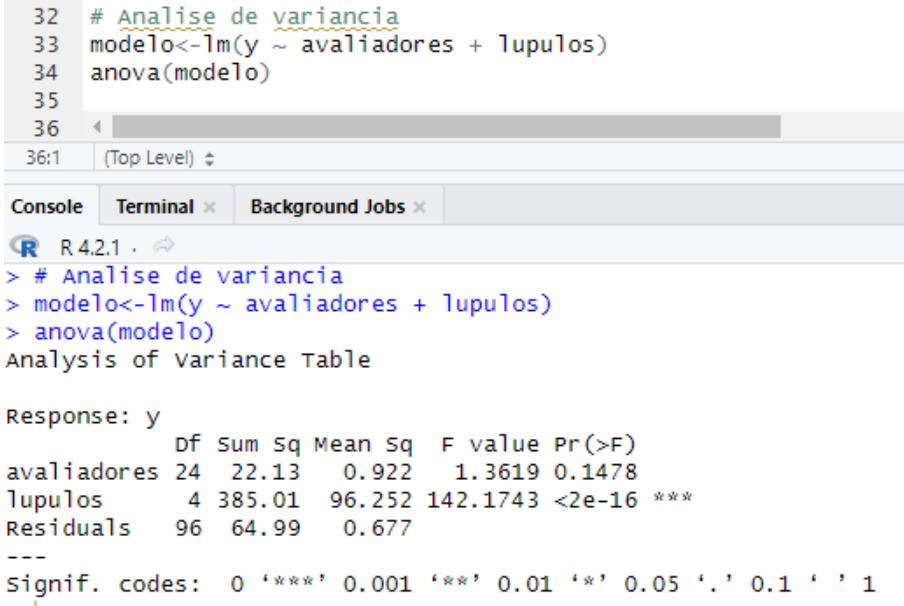
18  2,7,4,7,4,
19  4,7,5,7,3,
20  3,7,4,7,2,
21  2,7,4,7,5,
22  4,6,3,7,3,
23  3,7,5,6,4,
24  2,6,5,7,3,
25  2,7,4,7,3,
26  2,8,4,6,4,
27  3,7,5,7,2))
28 dados<-data.frame(avaliadores,lupulos,y)
29 head(dados)
30 str(dados)
31
32
33
34
35
36
37:1 (Top Level) ⇩ R Script ⇩
```

R 4.2.1 -

```

> dados<-data.frame(avaliadores,lupulos,y)
> head(dados)
  avaliadores lupulos y
1           1 LúpuloA 4
2           1 LúpuloB 7
3           1 LúpuloC 5
4           1 LúpuloD 6
5           1 LúpuloE 3
6           2 LúpuloA 3
> str(dados)
'data.frame': 125 obs. of 3 variables:
$ avaliadores: Factor w/ 25 levels "1","2","3","4",...: 1 1 1 1 1 2 2 2 2 ...
$ lupulos    : Factor w/ 5 levels "LúpuloA","LúpuloB",...: 1 2 3 4 5 1 2 3 4 5 ...
$ y          : num  4 7 5 6 3 3 8 6 7 4 ...
```

Figura 7.1.3: Análise de variância dos dados



The screenshot shows the RStudio interface with the console pane active. The code editor pane contains R script for performing an ANOVA analysis. The console pane displays the results of the ANOVA table, showing the analysis of variance table for the response variable 'y' across factors 'avaliadores' and 'lupulos'. The output includes the F-value, Pr(>F), and significance codes.

```

32 # Analise de variância
33 modelo<-lm(y ~ avaliadores + lupulos)
34 anova(modelo)
35
36
37:1 (Top Level) ⇩
```

R 4.2.1 -

```

> # Analise de variância
> modelo<-lm(y ~ avaliadores + lupulos)
> anova(modelo)
Analysis of Variance Table

Response: y
            Df Sum Sq Mean Sq  F value Pr(>F)
avaliadores 24  22.13   0.922   1.3619 0.1478
lupulos      4 385.01  96.252 142.1743 <2e-16 ***
Residuals   96  64.99   0.677
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figura 7.1.4: Console do teste de Tukey

```
> library(ExpDes.pt)
> drc(lupulos, avaliadores, y, mcomp = "tukey")
-----  
quadro da analise de variancia  
-----  
          GL      SQ      QM      Fc    Pr>Fc  
Tratamento 4 385.01 96.252 142.174 0.00000  
Bloco       24 22.13  0.922   1.362 0.14784  
Residuo     96 64.99  0.677  
Total       124 472.13  
-----  
CV = 17.06 %  
  
-----  
Teste de normalidade dos residuos  
valor-p: 0.01070463  
ATENCAO: a 5% de significancia, os residuos nao podem ser considerados normais!  
-----  
  
-----  
Teste de homogeneidade de variancia  
valor-p: 0.2015048  
De acordo com o teste de oneillmathews a 5% de significancia, as variancias podem ser consideradas homogeneas.  
-----  
  
-----  
Teste de Tukey  
-----  
Grupos Tratamentos Medias  
a      LúpuloB    7.08  
a      LúpuloD    6.64  
b      LúpuloC    4.36  
c      LúpuloE    3.4  
d      LúpuloA    2.64  
-----
```

7.2 Problema 15

Figura 7.2.1: Importação do banco de dados do exercício

```
dados <- read.csv("C:/Users/keyla/Desktop/Plan3 - 2.csv", header = TRUE)
```

Figura 7.2.2: Comparação dos dados pelo método de spearman

```
#avaliamador 1
x <- dados$x.1
y <- dados$x.2
cor.test(x, y, method="spearman")

#avaliamador 2
x <- dados$x.1
y <- dados$x.5
cor.test(x, y, method="spearman")

#avaliamador 3
x <- dados$x.1
y <- dados$x.8
cor.test(x, y, method="spearman")

#avaliamador 4
x <- dados$x.1
y <- dados$x.11
cor.test(x, y, method="spearman")

#avaliamador 5
x <- dados$x.1
y <- dados$x.14
cor.test(x, y, method="spearman")

#avaliamador 6
x <- dados$x.1
y <- dados$x.17
cor.test(x, y, method="spearman")
```

Figura 7.2.3: Console da comparação dos dados pelo método de spearman

```
> #avaliador 1
> x <- dados$x.1
> y <- dados$x.2
> cor.test(x, y, method="spearman")
    Spearman's rank correlation rho

data: x and y
S = 2, p-value = 0.01667
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
0.9428571

> #avaliador 2
> x <- dados$x.1
> y <- dados$x.5
> cor.test(x, y, method="spearman")
    Spearman's rank correlation rho

data: x and y
S = 16, p-value = 0.2972
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
0.5428571

> #avaliador 3
> x <- dados$x.1
> y <- dados$x.8
> cor.test(x, y, method="spearman")
    Spearman's rank correlation rho

data: x and y
S = 0, p-value = 0.002778
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
1

> #avaliador 4
> x <- dados$x.1
> y <- dados$x.11
> cor.test(x, y, method="spearman")
    Spearman's rank correlation rho

data: x and y
S = 4, p-value = 0.03333
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
0.8857143

> #avaliador 5
> x <- dados$x.1
> y <- dados$x.14
> cor.test(x, y, method="spearman")
    Spearman's rank correlation rho

data: x and y
S = 2, p-value = 0.01667
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
0.9428571

> #avaliador 6
> x <- dados$x.1
> y <- dados$x.17
> cor.test(x, y, method="spearman")
    Spearman's rank correlation rho

data: x and y
S = 6, p-value = 0.05833
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
0.8285714
```

7.3 Problema 14

Figura 7.3.1: Inserção dos pacotes necessários para o exercício

```

1 # Passo 1: carregar pacotes
2 library(readr)
3 library(dplyr)
4 library(reshape)
5 library(reshape2)
6 library(PMCMRplus)
7

```

Figura 7.3.2: Importação do banco de dados

```

7
8 # Passo 2: carregar o banco de dados
9 dados <- read_csv("Ex6 - 1 (1).csv")
10 View(dados)

```

Figura 7.3.3: Visualização do banco de dados importado

Avaliadores	Empresa interessada	Empresa Líder A	Empresa Líder B	Empresa Líder C
1	1	2	3	1
2	2	1	3	2
3	3	3	1	2
4	4	1	3	2
5	5	4	3	1
6	6	4	3	2
7	7	3	4	1
8	8	2	3	1
9	9	3	2	1
10	10	3	4	2
11	11	3	2	1
12	12	4	3	2

Showing 1 to 12 of 24 entries, 5 total columns

```

Console Terminal Background Jobs
> R42.1 : C:/Users/key/a/Desktop/
> # Passo 2: Carregar o banco de dados
> dados <- read_csv("Ex6 - 1 (1).csv")
Rows: 24 Columns: 5
-- Column specification --
#> delimitter: ","
#> dbl(5): Avaliadores, Empresa interessada, Empresa Líder A, Empresa Líder B, Empresa Líder C
#> i Use `spec()` to retrieve the full column specification for this data.
#> i Specify the column types or set `show_col_types = FALSE` to quiet this message.
> View(dados)
>

```

Figura 7.3.4: Reestruturação do banco de dados

```
11 # Reestruturando o banco de dados
12 dados1 <- melt(dados,
13   id = "Avaliadores",
14   measured = c("Empresainteressada", "EmpresaLiderA", "EmpresaLiderB", "EmpresaLiderC"))
15
16 # Renomeando as colunas do novo banco
17 colnames(dados1) = c("Avaliador", "Empresa", "Nota")
18 View(dados1)
19
```

Figura 7.3.5: Visualização do banco de dados reestruturado

The screenshot shows a data frame titled 'dados1' in RStudio. The interface includes standard RStudio tools like Source on Save, Run, and Source at the top, and a Filter button above the table. The table has four columns: 'Avaliador' (1 to 15), 'Empresa' (1 to 15), and 'Nota' (2 to 4). The data consists of 15 rows of 15 entries each, all labeled 'Empresa interessada'. A footer message indicates 'Showing 1 to 15 of 96 entries, 3 total columns'.

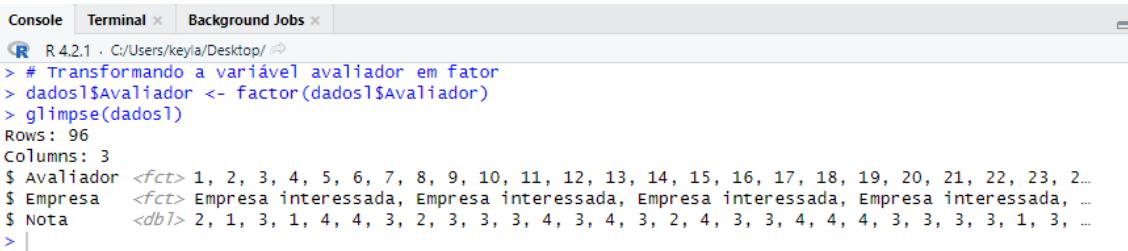
Avaliador	Empresa	Nota
1	1	2
2	2	1
3	3	3
4	4	1
5	5	4
6	6	4
7	7	3
8	8	2
9	9	3
10	10	3
11	11	3
12	12	4
13	13	3
14	14	4
15	15	3

Showing 1 to 15 of 96 entries, 3 total columns

Figura 7.3.6: Comando para transformar variável em fator

```
21 # Transformando a variável avaliador em fator
22 dados1$Avaliador <- factor(dados1$Avaliador)
23 glimpse(dados1)
24
```

Figura 7.3.7: Console do comando para transformar variável em fator

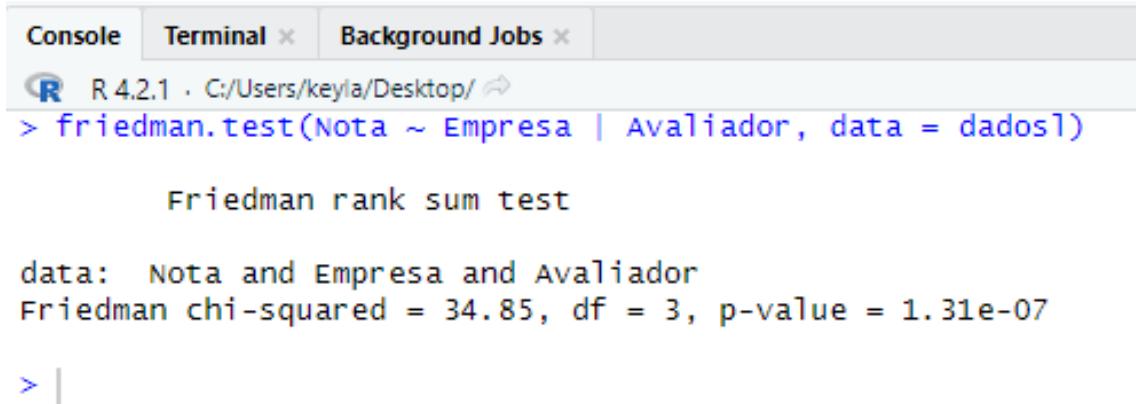


```
Console Terminal × Background Jobs ×
R 4.2.1 · C:/Users/keyla/Desktop/ ↵
> # Transformando a variável avaliador em fator
> dados1$Avaliador <- factor(dados1$Avaliador)
> glimpse(dados1)
Rows: 96
Columns: 3
$ Avaliador <fct> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 2...
$ Empresa   <fct> Empresa interessada, Empresa interessada, Empresa interessada, Empresa interessada, ...
$ Nota      <dbl> 2, 1, 3, 1, 4, 4, 3, 2, 3, 3, 3, 4, 3, 2, 4, 3, 3, 4, 4, 4, 3, 3, 3, 1, 3, ...
> |
```

Figura 7.3.8: Comando para realizar o teste de Friedman

```
26 #Realização do teste de Friedman
27
28 friedman.test(Nota ~ Empresa | Avaliador, data = dados1)
29 #variável dependente ~variável independente | variável que identifica o sujeito (a medida repetida)
```

Figura 7.3.9: Console do comando para realizar o teste de Friedman



```
Console Terminal × Background Jobs ×
R 4.2.1 · C:/Users/keyla/Desktop/ ↵
> friedman.test(Nota ~ Empresa | Avaliador, data = dados1)

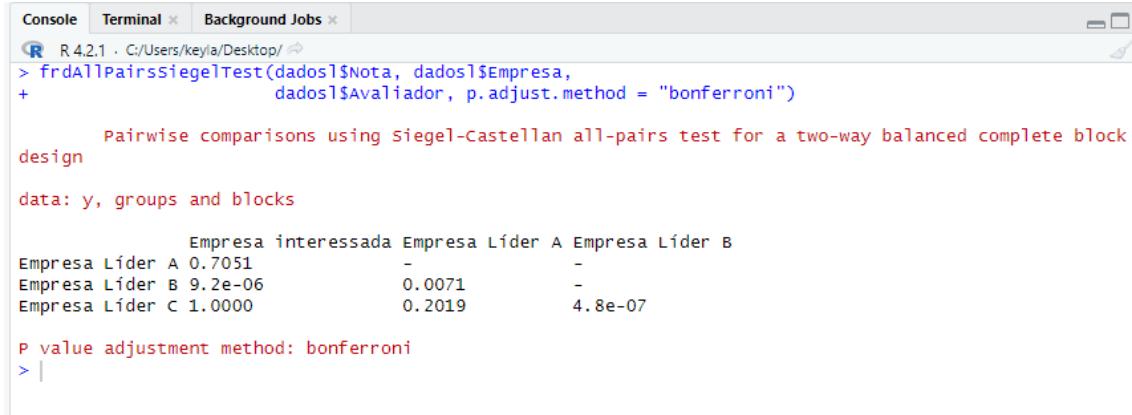
Friedman rank sum test

data: Nota and Empresa and Avaliador
Friedman chi-squared = 34.85, df = 3, p-value = 1.31e-07
> |
```

Figura 7.3.10: Comando bonferroni

```
frdAllPairssSiegelTest(dados1$Nota, dados1$Empresa,
dados1$Avaliador, p.adjust.method = "bonferroni")
```

Figura 7.3.11: Console do comando bonferroni



```

Console Terminal × Background Jobs ×
R 4.2.1 . C:/Users/keyla/Desktop/
> frdAllPairssiegelTest(dados$lNota, dados$Empresa,
+                           dados$Avaliador, p.adjust.method = "bonferroni")
Pairwise comparisons using Siegel-Castellan all-pairs test for a two-way balanced complete block
design

data: y, groups and blocks

    Empresa interessada Empresa Líder A Empresa Líder B
Empresa Líder A 0.7051      -          -
Empresa Líder B 9.2e-06      0.0071     -
Empresa Líder C 1.0000      0.2019    4.8e-07

P value adjustment method: bonferroni
>

```

7.3.1 Análise estatística dos resultados por Excel

Para começar a análise dos dados por excel, deve ser adicionado ao acesso rápido do excel a opção "Análise de dados", que pode ser adicionado pelos seguintes passos após selecionar arquivo → opções:

Figura 7.3.12: Clique em suplementos → Ferramentas de Análise → e Ir

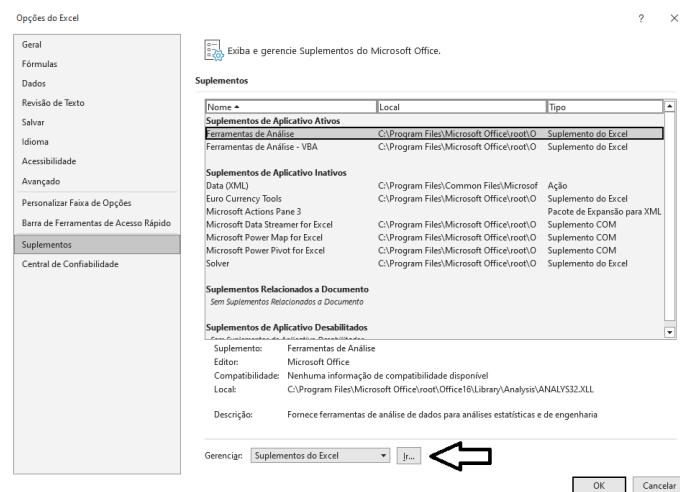
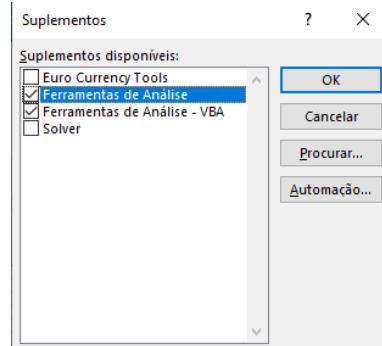


Figura 7.3.13: Selecione Ferramenta de Análise, Ferramenta de Análise - VBA → OK



Com as etapas realizadas, seu excel já está com a opção "Análise de Dados" disponível e pronto para calcular ANOVA e o primeiro passo é inserir os dados da análise final por atributo.

Na figura abaixo os dados foram retirados de (NEVES, WOSIACKI, 2014).

Figura 7.3.14: Dados da análise final por atributos

	Aparência	Amostras	1	2	3
2	Provado				
3			1	2	3
4		8,5	6,45	7,4	
5		8,5	3,8	8,2	
6		8,5	5,9	8,5	
7		2	6,85	6,35	6,9
8		7	6,2	6,85	
9		6,55	6,5	4,75	
10		3	8,4	6,4	6,2
11		8,5	1,1	1,3	
12		8,5	7	7,65	
13		4	9	6,7	4,8
14			9	0,8	6,45
15			9	3,65	2,9
16		5	7,1	2,15	6,4
17			7,75	2,25	7,2
18			7,7	2,65	4,5
19		6	8,2	1,9	7,8
20			8,2	0,65	7,45
21			8,2	1,25	1,6
22		7	9	0,55	8,5
23			9	0,5	8,5
24			9	0	8,45
25		8	6,5	1	2,85
26			7,15	5,65	3,45
27			6,3	3	4,7
28		9	9	4,2	5,15
29			9	4,65	3,9
30			9	5,1	3,6
31		10	6,55	4,5	6,35
32			6,1	2,95	3,5
33			6,7	4	5,8
34		11	8,4	8,4	8,4
35			8,4	0,65	8,3
36			8,5	3,4	8,4
37		12	7,15	3,35	4,25
38			7,1	2,7	7
39			7,4	3,45	6,4
40					
41					
42					
43					
44					

Para realizar a ANOVA desses dados, basta seguir os passos das próximas figuras:

Figura 7.3.15: Vá em Dados → Análise de Dados → Anova: fator duplo com repetição → OK

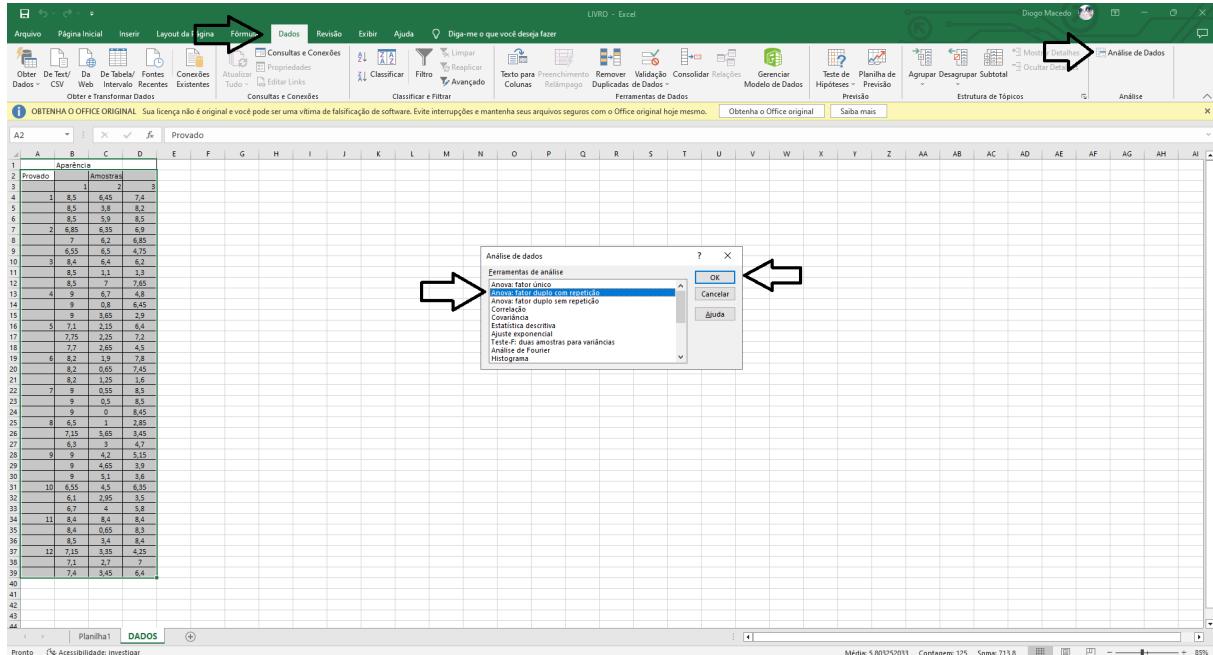
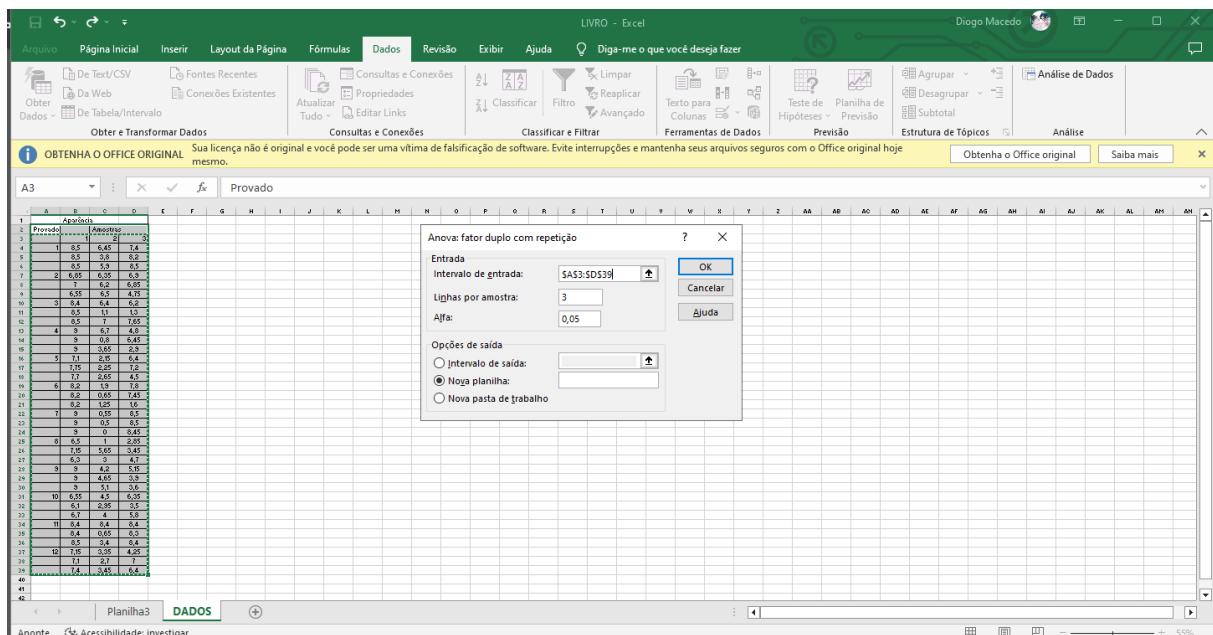


Figura 7.3.16: Selecione os dados no Intervalo de entrada → coloque 3 em Linhas por amostra (3 repetições) → digite o alfa desejado (0,05 no caso)



Após selecionar "OK", a ultima tabela chamada ANOVA é a tabela de interesse, e para

melhor interpretação dela, onde estiver escrito amostra, mudar para PROVADOR, onde estiver coluna, mudar para AMOSTRA e onde estiver dentro mudar para RESÍDUO.

Figura 7.3.17: Tabela de análise de variância com as modificações

ANOVA						
te da varia	SQ	gl	MQ	F	valor-P	F critico
Provador	65,40463	11	5,945875	2,668195	0,006291	1,924308
Amostra	338,5962	2	169,2981	75,97205	1,82E-18	3,123907
Interações	143,5355	22	6,524341	2,92778	0,000322	1,69192
Resíduo	160,4467	72	2,228426			
Total	707,983	107				

O primeiro fator a ser analisado nesta tabela é o coeficiente de variação "Interações", se o p-valor for $\leq 0,05$, significa que pode ter tido problemas com avaliadores que estejam fora do consenso da equipe, tendo uma resposta diferente da maioria dos outros julgadores, podendo ser causado por avaliadores que estejam usando diferentes porções de escala, dentre outros. Quando isso é observado, deve ser construído um gráfico de dispersão para visualização de qual ou quais avaliadores estão fora da curva. Para essa análise os dados da média das repetições por provador por amostra no atributo da aparência, cor caramelo, foi retirado de (NEVES, WOSIACKI, 2014).

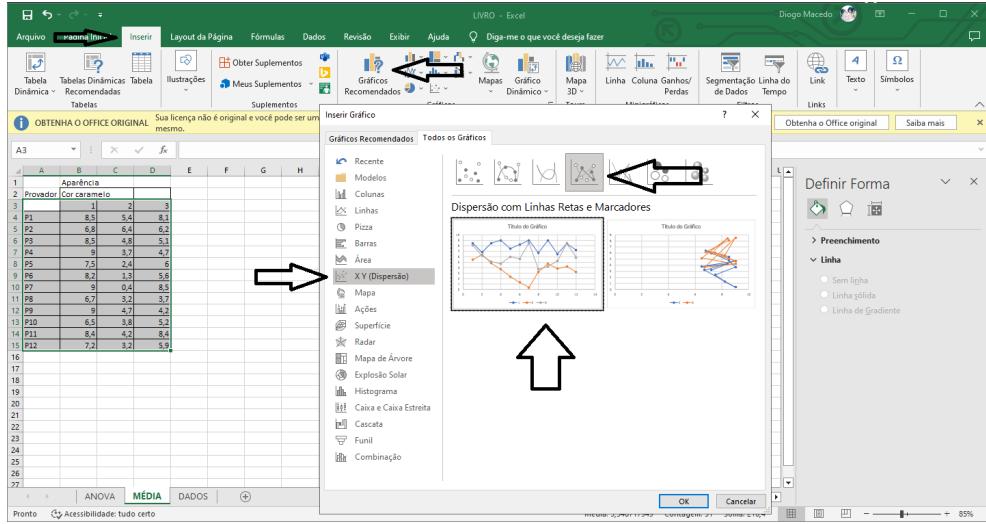
Figura 7.3.18: Média das repetições por provador por amostra no atributo da aparência, cor caramelo

Aparência			
Provador	Cor caramelo	1	2
P1	8,5	5,4	8,1
P2	6,8	6,4	6,2
P3	8,5	4,8	5,1
P4	9	3,7	4,7
P5	7,5	2,4	6
P6	8,2	1,3	5,6
P7	9	0,4	8,5
P8	6,7	3,2	3,7
P9	9	4,7	4,2
P10	6,5	3,8	5,2
P11	8,4	4,2	8,4
P12	7,2	3,2	5,9

Para fazer um gráfico de dispersão com esses dados é simples, basta seguir os seguintes passos:

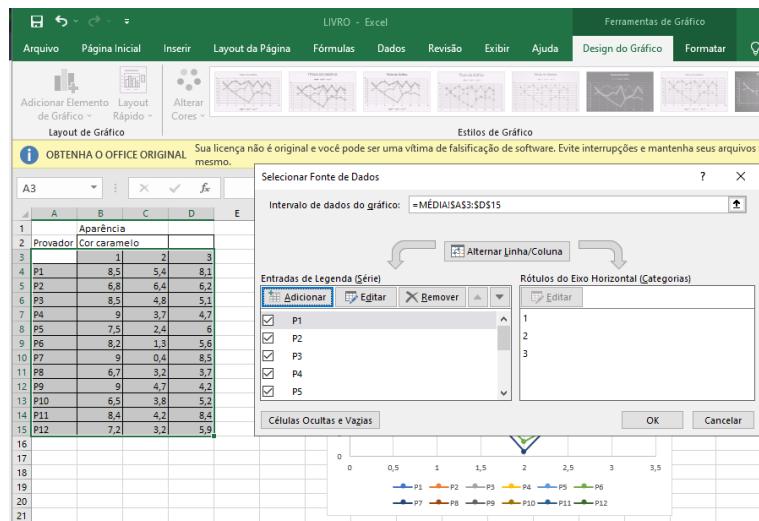
- Inserir;
- Gráficos Recomendados;
- Dispersão;
- Selecionar o gráfico linhas retas + marcadores;

Figura 7.3.19: Montagem do gráfico para avaliação do desempenho por provador



Clique no funil ao lado do gráfico e vá em selecionar dados, em seguida, alterne entre linha/coluna 7.3.20.

Figura 7.3.20: Alinhamento das coordenadas para o gráfico



Em seguida, para ajustar os valores do eixo x do gráfico, clique na linha x do gráfico → e na direita clique no histograma → vá em opções de eixo e coloque em máximo 4 e unidade principal 1:

Figura 7.3.21: Ajuste dos valores do eixo X do gráfico

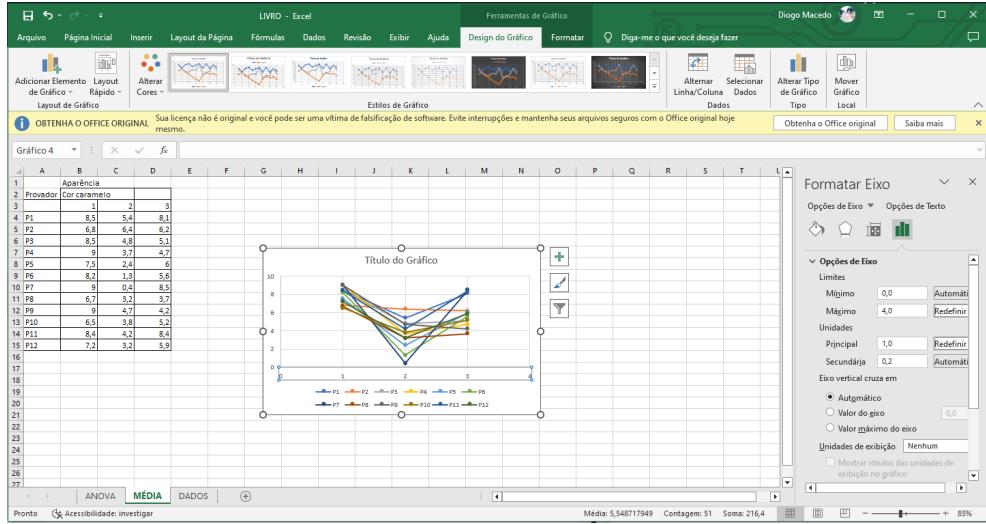
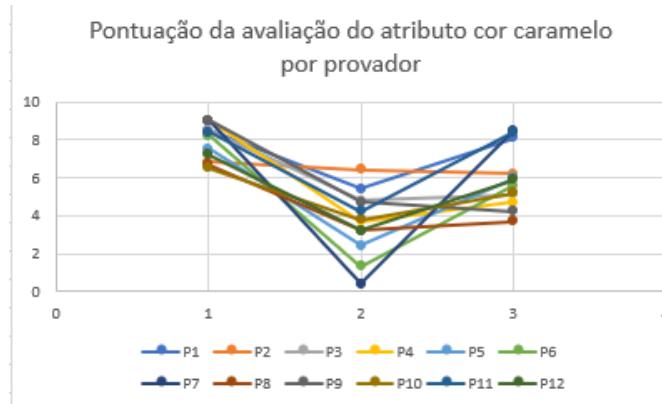


Figura 7.3.22: Pontuação da avaliação do atributo cor caramel por provador



Observando o gráfico final, é possível ver que os avaliadores 2, 7 e 9 estão com um resultado diferente do resto do grupo e para continuar a análise dos dados deve remover esses 3 avaliadores e continuar com outro calculo de ANOVA 7.3.23

Figura 7.3.23: Análise de variância sem os julgadores 2, 7 e 9

ANOVA						
Fonte da variação	SQ	gl	MQ	F	valor-P	F critico
Provador	61,11765	8	7,639707	2,655721	0,015575	2,115223
Amostra	247,9156	2	123,9578	43,09032	6,52E-12	3,168246
Interação	50,34772	16	3,146732	1,09387	0,383612	1,834629
Resíduo	155,3417	54	2,876698			
Total	514,7227	80				

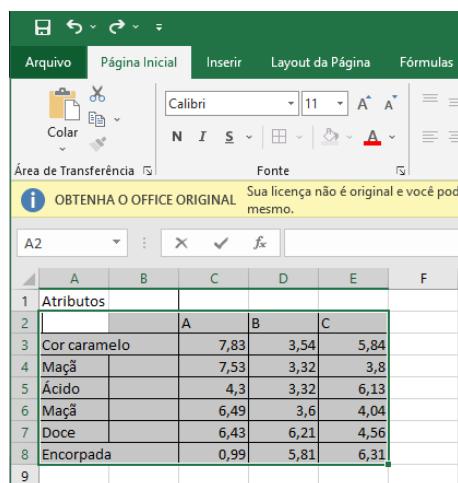
Com essa nova tabela da ANOVA, é possível observar que o p-valor de interações já não é mais significativo, ou seja, agora a observação entre os julgadores é consensual, todos avaliaram da mesma forma, agora podemos prosseguir para os p-valores de provador e amostra.

Na linha da amostra, se o p-valor for menor ou igual a 0,05, é um indicativo de que existem diferenças significativas entre as amostras no atributo estudado e, na linha do provador, quando o p-valor for significativo, caso a análise tenha sido refeita, quer dizer que os provadores utilizaram porções diferentes da escala para analisar a amostra, porém, com o mesmo consenso de respostas, podendo continuar com a análise.

7.3.2 Teste de tukey e análise gráfica

Após verificar que existe diferença entre as amostras, um teste de médias deve ser utilizado, sendo o mais utilizado nesse caso o Teste de Tukey, para verificar entre quais amostras há diferença. Será calculado a diferença mínima significativa (DMS) e comparada com cada diferença de média das amostras e, o primeiro passo a se fazer após realizar o Teste de Tukey, é inserir os dados no Excel 7.3.24, os dados foram retirados de (ELLENDERSEN et al, 2012).

Figura 7.3.24: Média dos atributos sensoriais de três amostras de suco de maçã



Com os dados no excel, selecione os dados a serem analisados, como os em cinza na figura 7.3.24 e faça como na figura 7.3.25:

Figura 7.3.25: Após selecionar os dados, vá em → inserir → gráficos recomendados → todos os gráficos → radar → radar com marcadores

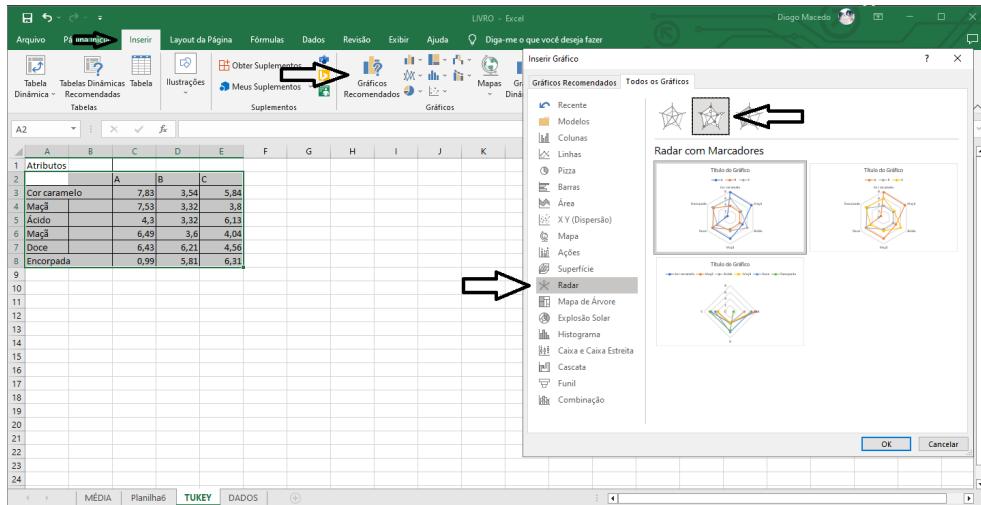
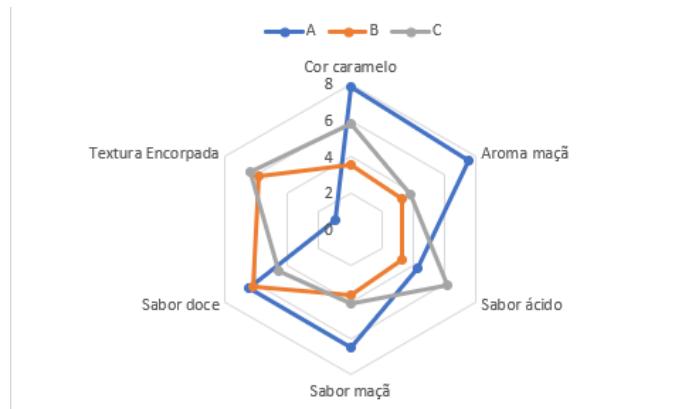


Figura 7.3.26: Gráfico aranha com os atributos avaliados



Com o gráfico aranha feito é possível avaliar e analisar todos os atributos de cada amostra com cuidado, podendo ver em que atributo cada alimento se destaca. É possível observar que quanto mais próximo da extremidade o vetor for, mais intenso é aquele atributo e quanto mais longe, menos intenso é.