

UNIVERSIDADE FEDERAL FLUMINENSE
FACULDADE DE FARMÁCIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS APLICADAS
A PRODUTOS PARA SAÚDE

**DESEMPENHO DO MODELO ANOVA COMPARADO A
TESTES ESTATÍSTICOS NÃO-PARAMÉTRICOS NO
TRATAMENTO DOS RESULTADOS DE TESTES DE ESCALA
HEDÔNICA**

VIVIANE DA SILVA GOMES
Mestranda

Niterói - RJ
2011

UNIVERSIDADE FEDERAL FLUMINENSE
FACULDADE DE FARMÁCIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS APLICADAS
A PRODUTOS PARA SAÚDE

**DESEMPENHO DO MODELO ANOVA COMPARADO A
TESTES ESTATÍSTICOS NÃO-PARAMÉTRICOS NO
TRATAMENTO DOS RESULTADOS DE TESTES DE ESCALA
HEDÔNICA**

VIVIANE DA SILVA GOMES

Mestranda

PROF. LUIS GUILLERMO COCA VELARDE

Orientador

PROF^a CLAUDETE CORRÊA DE JESUS CHIAPPINI

Co-orientadora

Dissertação apresentada ao Programa de Pós-Graduação em Pós-Graduação em Ciências Aplicadas a Produtos para Saúde, da Faculdade de Farmácia, Universidade Federal Fluminense, como requisito parcial para obtenção do grau de Mestre.

Niterói - RJ
2011

G633 Gomes, Viviane da Silva
Desempenho do modelo anova comparado a testes estatísticos não-paramétricos
no tratamento dos resultados de testes de escala hedônica / Viviane da Silva
Gomes; orientador: Luis Guillermo Coca Velarde. – Niterói, 2011.

104f.

Dissertação (Mestrado) – Universidade Federal Fluminense, 2011.

1. Análise de alimentos 2. Análise de variância (Estatística) 3. Alimento;
aspectos econômico 4. Indústria de alimentos I. Velarde, Luis Guillermo Coca
II. Título.

CDD 664

VIVIANE DA SILVA GOMES

**DESEMPENHO DO MODELO ANOVA COMPARADO A
TESTES ESTATÍSTICOS NÃO-PARAMÉTRICOS NO
TRATAMENTO DOS RESULTADOS DE TESTES DE ESCALA
HEDÔNICA**

Dissertação apresentada ao Programa de Pós-
Graduação em Ciências Aplicadas a Produtos
para Saúde, da Faculdade de Farmácia,
Universidade Federal Fluminense.

Aprovada em 1º de julho de 2011

BANCA EXAMINADORA

Dr. Luis Guillermo Coca Velarde
Universidade Federal Fluminense

Dra. Daniela De Grandi Castro Freitas
Empresa Brasileira de Pesquisa Agropecuária

Dra. Maria Cristina Jesus Freitas
Universidade Federal do Rio de Janeiro

Vivian Wahrlich
Universidade Federal Fluminense

Niterói - RJ
2011

DEDICATÓRIA

Dedico este trabalho aos meus familiares e amigos, que me apoiaram e ajudaram de alguma forma na conclusão dessa importante etapa da minha vida. E principalmente a minha fonte inspiradora, minha filha Ana Beatriz, que nasceu durante o curso e me ensinou o verdadeiro significado da palavra AMOR.

AGRADECIMENTOS

A Deus, por ter me dado forças, garra e perseverança para vencer mais esse desafio.

Ao meu Orientador Professor Doutor Luis Guillermo Coca Velarde, que admiro pelo seu profissionalismo, organização, dedicação e pelas horas disponibilizadas para a orientação desse trabalho.

A minha co-orientadora Professora Doutora Claudete Corrêa de Jesus Chiappini, pela dedicação, profissionalismo e por estar sempre disposta a tirar as minhas dúvidas, e também pelas horas dedicadas a esse trabalho.

A todos os professores com quem convivi e de quem tive o privilégio de ser aluna no Programa de Pós-Graduação em Ciências Aplicadas a Produtos para Saúde.

A Coordenação do Curso de Pós-Graduação, que sempre esteve disposta a ajudar durante todo o processo de realização do Mestrado.

A CAPES, por acreditar nesse curso de Pós-Graduação apostando na multidisciplinaridade como caminho para a construção do novo aprendizado acadêmico, na qual se enquadra esse trabalho.

A Universidade Federal Fluminense que, através da Bolsa Reuni, contribuiu para que eu pudesse concluir o Curso de Mestrado.

Aos provadores, que voluntariamente participaram dos testes sensoriais.

RESUMO

Dados de consumidores originados de teste de escala hedônica, apesar de violar o pressuposto básico de normalidade das análises paramétricas, são frequentemente tratados por ANOVA. Este trabalho teve como objetivo avaliar o desempenho da ANOVA comparado a testes estatísticos não-paramétricos no tratamento dos resultados do teste de escala hedônica utilizando dados reais e simulados por meio de programa estatístico. Inicialmente, foi realizado teste de ordenação para selecionar amostras de refresco de maracujá e manga com diferentes concentrações de açúcar (0%, 2,5%, 5%, 7,5% e 10%). Posteriormente, foi realizado teste de escala hedônica em laboratório, com as amostras que apresentaram diferença significativa quanto ao sabor doce no teste de ordenação, com 100 consumidores para cada um dos refrescos. Nesse teste, a normalidade foi avaliada pelo teste de Kolmogov-Smirnov e os dados foram tratados por ANOVA e pelos testes de Kruskal-Wallis e Wilcoxon Mann-Whitney. Foram realizadas também duas simulações utilizando o software estatístico S-Plus, para obter dados simulados de testes de escala hedônica com $n = 5$, $n = 10$, $n = 30$, $n = 50$, $n = 100$ e $n = 1000$ (para cada n foram realizadas 1000 repetições). Os dados das simulações foram tratados por ANOVA e pelo teste de Kruskal-Wallis. No teste de laboratório, o teste de Kruskal-Wallis e o teste de Wilcoxon Mann-Whitney apresentaram os mesmos resultados da ANOVA, mesmo não apresentando distribuição normal ($p < 0,05$). Nas Simulações 1 e 2 as amostras simuladas com $n \geq 50$, número mínimo para testes de escala hedônica em laboratório, tanto a ANOVA quanto o teste de Kruskal-Wallis apresentaram os mesmos resultados ($p < 0,05$). Foi concluído que não é necessário substituir a ANOVA em testes de escala hedônica por testes não-paramétricos, mesmo na ausência de normalidade.

Palavras-chave: Análise de Variância, teste afetivo, Análise sensorial, teste de escala hedônica, teste de Kruskal-Wallis, teste não-paramétrico.

ABSTRACT

Consumers data originated from hedonic scale tests, despite violating the basic assumptions of normality of parametric analysis, are frequently treated with ANOVA. The aim of this study was to evaluate the performance of ANOVA comparing to non-parametric statistical tests on the treatment of hedonic scale test results, using real data that were simulated by statistical software. At first, it was performed ordination test to select passion fruit juice and mango juice samples with different sugar concentration (0%, 2,5% , 5% , 7,5% and 10%). Later, it was performed hedonic scale test in laboratory, with the samples that presented significant difference for sweet flavor on the ordination test, with 100 consumers and for each one of the juices. On this test, the normality was evaluated by Kolmogov – Smirnov test and data were treated with ANOVA and Kruskal – Wallis and Wilcoxon Mann-Whitney tests. Two simulations were also performed using the statistical software S-Plus, to obtain simulated data from hedonic scale test with $n = 5$, $n = 10$, $n = 30$, $n = 50$, $n = 100$ and $n = 1000$ (for each n 1000 repetitions were realized). Data obtained with these simulations were treated with ANOVA and Kruskal – Wallis test. On the laboratory test, the Kruskal – Wallis test and Wilcoxon Mann-Whitney test presented the same results of ANOVA, even not presenting normal distribution ($p < 0,05$). On Simulations 1 and 2 samples simulated with $n \geq 50$, minimum number for hedonic scales tests on laboratory, both ANOVA and Kruskal-Wallis test presented the same results ($p < 0,05$). It was concluded that it is not necessary to substitute ANOVA in hedonic scale tests for non-parametric tests, even with a lack of normality.

Key words: Analysis of Variance, affective tests, hedonic scale test, Kruskal-Wallis test, non-parametric test.

SUMÁRIO

RESUMO.....	11
ABSTRACT.....	12
1. INTRODUÇÃO	13
2. OBJETIVOS	16
2.1. Objetivo geral	16
2.2. Objetivos específicos	16
3. REVISÃO BIBLIOGRÁFICA	17
3.1. Conceito e histórico da análise sensorial	17
3.2. Aplicação da análise sensorial	20
3.3. Percepção sensorial e os sentidos	20
3.3.1. O sentido da visão	22
3.3.2. O sentido do olfato	23
3.3.3. Os sentidos do tato e da audição	24
3.3.4. O sentido do gosto	25
3.4. Fatores que afetam a avaliação sensorial	26
3.5. Provadores em análise sensorial	27
3.6. Métodos sensoriais	28
3.7. Medidas sensoriais e o emprego de escalas	30
3.8. Uso de escalas em teste de aceitabilidade	32
3.9. Tratamento estatístico em análise sensorial	34
3.10. Tipos de variáveis	35
3.11. Variáveis aleatórias	37
3.11.1. Variáveis aleatórias discretas	37
3.11.2. Distribuição uniforme discreta	38
3.11.3. Distribuição de Bernoulli	39
3.11.4. Distribuição binomial	40
3.11.5. Variáveis aleatórias contínuas	40
3.11.6. A distribuição Normal	41
3.12. A escolha do método estatístico	42

3.12.1.	Testes de hipóteses	43
3.13.	Testes de comparação de médias	44
3.14.	Testes paramétricos	46
3.14.1.	Análise de Variância e seus pressupostos.....	46
3.14.2.	Testes para duas amostras.....	49
3.15.	Testes não-paramétricos	50
3.15.1.	Testes não-paramétricos para amostras independentes	53
3.15.1.1.	Teste de Kruskal-Wallis.....	53
3.15.1.2.	Teste da mediana e Teste de Mann-Whitney.....	54
3.15.2.	Teste não-paramétricos para amostras dependentes	55
3.15.2.1.	Teste dos Sinais, de Wilcoxon e de Friedeman.....	55
3.16.	Noções de simulação	57
4.	MATERIAL E MÉTODOS	59
4.1.	Amostras.....	59
4.1.1.	Teste de ordenação.....	59
4.1.2.	Tratamento estatístico do teste de ordenação	60
4.2.	Teste de escala hedônica	61
4.2.1.	Tratamento estatístico do teste de escala hedônica	62
4.3.	Simulação de dados	62
4.3.1.	Simulação 1	63
4.3.2.	Tratamento estatístico do teste de escala hedônica da Simulação 1	67
4.3.3.	Simulação 2	67
4.3.4.	Tratamento estatístico do teste de escala hedônica da Simulação 2	70
5.	RESULTADOS E DISCUSSÃO	71
5.1.	Análise estatística do teste de ordenação.....	71
5.2.	Análise estatística do teste de escala hedônica.....	73
5.2.1.	Teste de Kolmogov-Smirnov	73
5.2.2.	Análise de variância	73
5.2.3.	Teste de Tukey	74
5.2.4.	Teste de Wilcoxon Mann-Whitney	74
5.2.5.	Teste de Kruskal-Wallis	74
5.3.	Simulação de dados	75

5.3.1.	Análise estatística do teste de escala hedônica da Simulação 1	75
5.3.2.	Análise estatística do teste de escala hedônica da Simulação 2	76
6.	CONCLUSÃO	79
7.	REFERÊNCIAS BIBLIOGRÁFICAS	80
8.	APÊNDICE	85
8.1.	Apêndice 1: Rotina do programa estatístico elaborado para a Simulação 1....	86
8.2.	Apêndice 2: Rotina do programa estatístico elaborado para a Simulação 2....	97
9.	ANEXO	103
9.1.	Anexo 1: Tabela de Newell e MacFarlene	104

LISTA DE FIGURAS

Figura 1.	Termos empregados na análise sensorial.....	19
Figura 2.	Exemplo de escala hedônica gráfica facial.....	31
Figura 3.	Exemplo de escala hedônica não estruturada com termos hedônicos como âncora.....	31
Figura 4.	Exemplo de escalas hedônicas estruturadas verbais com 9, 7 e 5 categorias.....	31
Figura 5.	Exemplo de escala ordinal.....	32
Figura 6.	Classificação das variáveis.....	36
Figura 7.	Distribuição Normal Padrão.....	42
Figura 8.	Exemplo de heterocedasticidade.....	48
Figura 9.	Ficha de aplicação do teste de ordenação.....	60
Figura 10.	Ficha de aplicação do teste de escala hedônica.....	62
Figura 11.	Gráficos das probabilidades dos escores dos diferentes tratamentos simulados de teste de escala hedônica de nove pontos do refresco de maracujá de acordo com o gênero e grupo etário da Simulação 1.....	65
Figura 12.	Gráficos das probabilidades dos escores dos diferentes tratamentos simulados de teste de escala hedônica de nove pontos do refresco de manga de acordo com o gênero e grupo etário da Simulação 1.....	66
Figura 13.	Gráficos das probabilidades dos escores dos diferentes tratamentos simulados de teste de escala hedônica de nove pontos das seis diferentes situações, com diferentes distribuições de dados criadas para a Simulação 2.....	69

LISTA DE TABELAS

Tabela 1.	Orientações para utilização de testes paramétricos e não-paramétricos.....	57
Tabela 2.	Ingredientes e quantidades para a preparação dos refrescos de maracujá e manga.....	59
Tabela 3.	Grupos etários utilizados na Simulação 1, sua classificação e proporção aproximada na população brasileira.....	63
Tabela 4.	Médias (μ) e variâncias (σ^2) dos tratamentos da Simulação 2.....	68
Tabela 5.	Soma dos valores de ordenação das amostras dos refrescos de maracujá e manga, quanto ao sabor doce, utilizando escala crescente de ordenação.....	71
Tabela 6.	Comparação da diferença entre os valores de ordenação das amostras do refresco de maracujá com a diferença crítica observada na tabela de Newell e MacFarlane ao nível de 5% de significância.....	72
Tabela 7.	Comparação da diferença entre os valores de ordenação das amostras do refresco de manga com a diferença crítica observada na tabela de Newell e MacFarlane ao nível de 5% de significância.....	72
Tabela 8.	Análise de variância relativa aos dados de aceitabilidade dos refrescos de maracujá e manga.....	73
Tabela 9.	Média dos valores do teste de escala hedônica das amostras dos refrescos de maracujá e manga.....	74
Tabela 10.	Porcentagens de aceitação da hipótese de igualdade dos escores de T1, T2 e T3 (H_0) nos testes Kruskal-Wallis e na análise de variância (ANOVA), na Simulação 1.....	76
Tabela 11.	Porcentagens de aceitação da hipótese de igualdade dos escores de T1, T2 e T3 (H_0) nos testes Kruskal-Wallis e na análise de variância (ANOVA), nas seis situações da Simulação 2.....	78

1. INTRODUÇÃO

A indústria de alimentos utiliza a análise sensorial como uma ferramenta para desenvolvimento, otimização e manutenção da qualidade dos produtos alimentícios e avaliação de seu potencial de comercialização (ELORTONDO; OJEDA; ALBISU; SALMERÓN; ETAYO; MOLINA, 2007; MUÑOZ, 2002; STONE; SIDEL, 1993).

A forma de quantificar as respostas aos estímulos sensoriais é crítica para o bom desempenho da análise sensorial e, dentre as diversas técnicas utilizadas, a mais usual é o emprego de escalas para medir a resposta do provador. Dentre as várias classes de escalas, o analista sensorial deve selecionar a mais apropriada de acordo com o objetivo do teste, o tipo de informação desejada, a qualificação do julgador e o tipo de produto (MEILGAARD; CIVILLE; CARR, 1991). Nos testes de aceitabilidade, a resposta do provador é coletada por meio de escalas hedônicas que podem ser nominais, numéricas, gráficas ou mistas e expressam o gostar ou desgostar do provador em relação a um produto alimentício (STONE; SIDEL, 1993; MEILGAARD; CIVILLE; CARR, 1991).

Análises estatísticas paramétricas são baseadas em suposições básicas de normalidade, independência, e homocedasticidade de respostas experimentais (VILLANUEVA; PETENATE; DA SILVA, 2000; MUNRO, 1997; GAY; MEAD, 1992; GIOVANNI; PANGBORN, 1983). Do mesmo modo, a maioria dos testes estatísticos usados para comparar médias, tais como o teste t, teste de Duncan e teste de Tukey, são testes paramétricos que assumem a normalidade dos resultados (VILLANUEVA; PETENATE; DA SILVA, 2000).

Dados de consumidores obtidos em testes de escala hedônica violam o pressuposto básico de normalidade (VILLANUEVA; PETENATE; DA SILVA, 2000; VIE; GULLI; O'MAHONY, 1991; MCPHERSON; RANDALL, 1985). Um problema observado no tratamento estatístico aplicado aos resultados de teste de escala hedônica é assumir que os dados obtidos apresentam distribuição normal, contradizendo ao encontrado na prática que são frequentemente, informações sensoriais com distribuição assimétrica (LIM; FUJIMARU, 2010; VILLANUEVA; DA SILVA, 2009; VILLANUEVA; PETENATE; DA SILVA, 2005; VILLANUEVA; PETENATE; DA SILVA, 2000; WILKINSON; YUKSEL, 1997). Em testes para consumidores usando escala hedônica estruturada, isto ocorre quando as amostras testadas apresentam consenso, ou seja, uma aceitação muito alta, resultando em altas frequências de marcações nas escalas utilizadas para avaliação sensorial entre os valores 7 e 9, ou uma aceitação muito baixa comparada com as outras amostras testadas, resultando em altas frequências de marcações na escala entre os valores 1 e 3. Giovanni e Pangborn (1983) observaram distribuições bimodais de respostas hedônicas quando usaram escala não-estruturada.

Testes não-paramétricos são boas opções para situações em que ocorrem violações dos pressupostos básicos necessários para a aplicação de testes paramétricos. Podem ser utilizados para tratar os resultados que apresentam distribuição assimétrica ou apresentem comportamento não normal e podem ser utilizados, também, quando a distribuição da variável de interesse não é conhecida (VILLANUEVA; PETENATE; DA SILVA, 2000; MUNRO, 1997). Métodos alternativos têm sido propostos para a análise estatística de dados em análise sensorial com uso de escalas. Entretanto, a literatura mostra que a análise de variância (ANOVA) e outros testes paramétricos ainda são largamente utilizados para o tratamento estatístico de teste de escala hedônica, mesmo que as pressuposições básicas de normalidade, homocedasticidade e independência estejam violadas (VILLANUEVA; PETENATE; DA SILVA, 2000; BROCKHOFF; SKOVGAARD, 1994; GAY; MEAD, 1992; PRITCHETT, 1992; NAES, 1990; MILLER, 1987; WILKINSON; YUKSEL, 1997). Testes não-paramétricos podem ser uma alternativa para tratar os resultados de testes de escala hedônica (VILLANUEVA; PETENATE; DA SILVA, 2000).

Para comparar o desempenho de testes estatísticos pode ser utilizado estudo de probabilidade. Nesse tipo de estudo, se for especificado um espaço amostral e a probabilidade associada aos pontos desse espaço, o modelo probabilístico ficará

completamente determinado e será possível, então, calcular a probabilidade de qualquer evento aleatório de interesse. Entretanto, muitas vezes, mesmo construindo um modelo probabilístico, certas questões não podem ser resolvidas analiticamente e haverá a necessidade de recorrer a estudos de simulação para obter aproximações de quantidades de interesse (BUSSAB; MORETIN, 2005). Utilizando programas estatísticos como, por exemplo, o R, SAS e S-Plus é possível programar funções para gerar dados de acordo com os parâmetros de interesse, para posteriormente serem aplicados testes estatísticos adequados. De modo bastante amplo, estudos de simulação tentam reproduzir num ambiente controlado o que se passa com um problema real. As principais vantagens em se obter dados utilizando simulações por meio de programas estatísticos são: a rapidez com que os resultados são obtidos, a possibilidade de trabalhar com um grande número de amostras e com amostras de diversos tamanhos. Uma outra vantagem, e o baixo custo das análises, pois não existem gastos com amostras.

Este trabalho teve como objetivo comparar o desempenho do modelo ANOVA com testes estatísticos não-paramétricos, no tratamento dos resultados de testes de escala hedônica de nove pontos, utilizando dados reais, obtidos em teste de escala hedônica realizado em laboratório e dados simulados por meio de programa estatístico S-Plus.

2. OBJETIVOS

2.1. Objetivo geral

Este trabalho teve como objetivo avaliar o desempenho do modelo ANOVA comparado a testes estatísticos não-paramétricos no tratamento dos resultados de teste de escala hedônica.

2.2. Objetivos específicos

- Utilizar teste de ordenação como pré-teste para selecionar amostras de refrescos para o teste de escala hedônica;
- Utilizar teste de escala hedônica em experimento realizado em laboratório;
- Simular dados de testes de escala hedônica por meio de programa estatístico;
- Avaliar os resultados dos testes de escala hedônica em laboratório e simulados com diferentes testes estatísticos;
- Verificar a possibilidade de alternativas, dentre os testes estatísticos existentes, que melhor se adaptem a situações em que é necessário o uso de testes de escala hedônica.

3. REVISÃO BIBLIOGRÁFICA

3.1. Conceito e histórico da análise sensorial

A análise sensorial pode ser conceituada como uma disciplina científica usada para evocar (provocar), medir, analisar e interpretar reações das características dos alimentos e bebidas e de outros materiais que são percebidos pelos sentidos da visão, olfato, paladar, tato e audição (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1993; INSTITUTE OF FOOD TECHNOLOGISTS, 1981; AMERINE; PANGBORN; ROESSLER, 1965).

Estudos demonstram que as características sensoriais, em particular o aroma e o sabor, têm efeito sobre a escolha do alimento pelo consumidor. Steiner, em 1979 (CLARK, 1998), realizou um dos primeiros estudos sobre os efeitos do sabor e do *flavor*, observando as expressões faciais de bebês neonatos ao provar soluções doces e amargas. Quando oferecidas soluções com compostos doces, as crianças reagiram com olhos bem abertos e retração dos cantos da boca, evocando uma resposta facial de aceitação, enquanto soluções com gostos amargos tiveram como resposta o fechamento apertado dos olhos, abertura da boca e um repentino giro da cabeça. O estudo de Steiner indica a existência da preferência sensorial inata por sabores adocicados.

Na vida adulta, parte desta reação instintiva dos bebês tende a permanecer e a preferência por sabores doces e aromas prazerosos é enriquecida pelo uso contínuo de alimentos condimentados e bebidas alcoólicas amargas (CLARK, 1998). Esta habilidade, portanto, que leva à capacidade natural de preferir ou rejeitar determinados alimentos ou marcas, vem sendo aproveitada, milenarmente, nas indústrias de cerveja, vinho e destilados da Europa.

A avaliação sensorial foi desenvolvida durante a Segunda Guerra Mundial com o objetivo de estabelecer as razões pelas quais as tropas rejeitavam um grande volume de ração de campanha, embora as dietas estivessem balanceadas nutricionalmente (DELLA MODESTA, 1994).

A metodologia da análise sensorial pode ser dividida em quatro fases de desenvolvimento. A primeira fase, antes de 1940, foi caracterizada pela época artesanal e pré-científica da indústria de alimentos. A qualidade sensorial dos produtos alimentícios era determinada pelo proprietário da empresa (COSTELL; DURAN, 1981).

A segunda fase, no período entre 1940 e 1950, foi época de expansão da indústria de alimentos e da incorporação de pessoal técnico, vindo, na sua maioria, da área química e farmacêutica. Conceitos de controle de processo e de produto final foram introduzidos nessas fases, porém, os métodos de controle eram químicos e instrumentais, tendo a análise sensorial pouca expressão no controle de qualidade dos produtos alimentícios (DUTCOSKY, 1996).

A terceira fase, entre 1950 e 1970, foi caracterizada pelo reconhecimento da utilização do ser humano como instrumento de medida das características sensoriais dos alimentos. Os principais avanços nesse período foram: a definição dos atributos primários que integram a qualidade sensorial dos alimentos e os órgãos sensoriais a eles relacionados (CHIAPPINI; LEITE, 2004).

A quarta fase, após 1970, foi caracterizada pela descoberta de que a qualidade sensorial de um alimento não é só inerente ao alimento, mas sim o resultado da interação entre os estímulos procedentes dos alimentos e as condições fisiológicas, psicológicas e sociológicas do indivíduo ou do grupo que avalia o alimento (DUTCOSKY, 1996).

No Brasil, a prática do uso da análise sensorial foi iniciada em 1954 com degustadores para classificação do café brasileiro. Com o desenvolvimento da análise sensorial na indústria de alimentos, as técnicas de utilização dos sentidos foram sendo aperfeiçoadas e padronizadas e o tratamento estatístico foi introduzido como meio de comparar os resultados (OLIVEIRA, 2010).

Para a utilização adequada das técnicas de avaliação sensorial e a comparação de trabalhos técnicos e científicos na área, foi importante a padronização da linguagem utilizada por seus usuários. No Brasil, a Associação Brasileira de Normas Técnicas (1993) foi responsável pela criação de um glossário com a conceituação dos termos empregados. Na Figura 1 podemos visualizar alguns termos mais utilizados em análise sensorial.

Termo	Significado
Aceitação	Ato de um determinado indivíduo ou população ser favorável ao consumo de um produto alimentício
Aceitabilidade	Grau de aceitação de um produto favoravelmente recebido por um determinado indivíduo ou população, em termo de propriedades sensoriais.
Atributo	Característica perceptível.
Consumidor	Indivíduo que utiliza o produto.
Degustação	Avaliação sensorial de um produto alimentício na cavidade oral.
Degustador	Indivíduo que executa a degustação.
Discriminação	Diferenciação qualitativa e/ou quantitativa entre dois ou mais estímulos.
Equipe selecionada	É o grupo de julgadores selecionados.
Hedônico	Relativo ao gostar e desgostar.
Julgador ou provador	Indivíduo que participa do teste sensorial.
Julgador selecionado	Indivíduo escolhido por sua acuidade em realizar um teste sensorial.
Julgador treinado	Indivíduo selecionado e submetido a treinamento para determinado teste e produto.
Painel (Equipe sensorial)	equipe de julgadores que realizam teste sensorial.
Perito ou Especialista (<i>expert</i>):	Julgador que possui grande experiência com o produto, sendo capaz de realizar individualmente a avaliação sensorial deste.
Preferência	Expressão do estado emocional ou reação afetiva de um indivíduo que o leva à escolha de um produto sobre outro(s).
Sensorial	Relativo ao uso dos órgãos dos sentidos.

Figura 1. Termos empregados na análise sensorial.

Fonte: Adaptado de Associação Brasileira de Normas Técnicas (1993).

Atualmente, a qualidade e a aceitabilidade dos produtos alimentícios podem ser determinadas por meio dos órgãos dos sentidos de forma confiável e reprodutiva, por meio das técnicas de análises sensoriais e do tratamento estatístico dos resultados. A indústria de alimentos e bebidas utiliza-se dessas técnicas no desenvolvimento, no melhoramento e no monitoramento de seus produtos (DUTCOSKY, 1996).

3.2. Aplicação da análise sensorial

A avaliação sensorial fornece suporte técnico para a pesquisa e a produção de alimentos e bebidas. Dentre as muitas aplicações podem ser citadas àquelas inerentes ao processo produtivo como o controle das etapas de desenvolvimento de novos produtos, o melhoramento de produtos, a alteração de processos, a redução de custos, a seleção de novas fontes de matérias-primas e suprimentos, o controle de qualidade na produção, a avaliação da estabilidade no armazenamento e para seleção e treinamento de provedores (ELORTONDO; OJEDA; ALBISU; SALMERÓN; ETAYO; MOLINA, 2007; MUÑOZ, 2002; DUTCOSKY, 1996; STONE; SIDEL, 1993).

Por outro lado, para alcançar o seu objetivo principal, que é o de obter lucros com sua produção, a indústria de alimentos e bebidas necessita de informações acerca da aceitação dos consumidores. A aceitação ou não de um produto alimentício é fundamental para a tomada de decisões dentro da indústria, que só vai empreender esforços em produtos que agradem ao consumidor.

Desta forma, a análise sensorial pode contribuir com técnicas capazes de medir alterações perceptíveis que afetam a aceitabilidade de um produto alimentício (CARDELLO H.; CARDELLO L., 1998) contribuindo com informações para adequá-lo aos padrões de qualidade que satisfaçam as necessidades e expectativas do consumidor (CAPORALE; MONTELEONE, 2004). É nesse contexto que se torna interessante pesquisar as percepções e os anseios do cliente a quem o produto é direcionado (RIBEIRO; DELLA LUCIA; BARBOSA; GALVÃO; MINIM, 2008).

3.3. Percepção sensorial e os sentidos

Os órgãos dos sentidos apresentam receptores sensoriais. Os receptores sensoriais têm como principal propriedade a irritabilidade, que consiste na ação de detectar um estímulo químico ou físico, proveniente do meio ambiente ou do próprio organismo. Os receptores sensoriais consistem em células ou grupos de células que recebem os estímulos, transmitem impulsos nervosos, através dos nervos, até o cérebro. O receptor pode ser a

própria célula nervosa ou uma célula intermediária que recebe e transmite o estímulo a uma célula nervosa (AMERINE, 1965).

A percepção dos atributos de um produto alimentício pelo provador vai variar de acordo com suas experiências pessoais. Quando apresentado pela primeira vez a um produto alimentício a resposta sensorial é resultante da percepção consciente do produto, enquanto para àquele provador que já consumiu o produto em alguma época de sua vida, a resposta sensorial envolve o grau de experiência que ele apresenta com esse produto. Um exemplo disso é o fato de que indivíduos têm dificuldades em identificar e descrever odores e *flavores*, mas recordam facilmente o lugar e quando sentiram esse odor ou *flavor* pela primeira vez (KOSTER M.; PRESCOTT; KOSTER E., 2004).

A avaliação sensorial utiliza, fundamentalmente, mecanismos fisiológicos e psicológicos para perceber atributos que chegam até os órgãos dos sentidos. O uso do olfato, paladar, visão, audição ou tato determina o atributo que se deseja avaliar que podem ser odor, sabor, cor, ruído ou consistência (DUTCOSKY, 1996). Cada atributo sensorial é percebido por meio de mecanismos fisiológicos diferenciados, mas que podem ser resumidos a uma mesma cadeia de sentidos. Nessa cadeia o alimento gera um estímulo que é percebido pelo órgão do sentido; a sensibilização do órgão provoca uma sensação que é mediada até o cérebro; o cérebro percebe que foi estimulado e traduz essa percepção em uma resposta verbal do provador, considerada na análise sensorial, o resultado da avaliação (CHARLES, 1982; DUTCOSKY, 1996).

Desta forma, outros fatores e atributos não sensoriais influenciam na escolha dos alimentos, tais como a identificação do produto, a situação de compra e consumo, a origem, a segurança e as propriedades nutricionais, assim como, a marca e os hábitos pessoais do consumidor (ARAÚJO; SILVA; MINIM, 2003; SIRET; ISSANCHOU, 2000; DRANSFIELD; ZAMORA; BAYLE, 1998; TUORILA; CARDELLO; LESHER, 1994).

3.3.1. O sentido da visão

Os olhos são os órgãos físicos que nos permite o sentido da visão. A receptora dos estímulos visuais é a retina constituída por células especializadas, que são os cones, detectores da cor; e os bastões, que detectam a forma e a luz escura. A visão antecipa a recepção das informações de todos os demais órgãos dos sentidos e fornece informações sobre a aparência do alimento como o estado, o tamanho, a forma, a textura e a cor (MEILGAARD; CIVILLE; CARR, 1991).

Dentre as características sensoriais de aparência, o impacto causado pela cor se sobrepõe ao causado pelos demais componentes da aparência e atributos. A cor frequentemente é relacionada com a qualidade quando, por exemplo, o ovo de casca marrom é considerado pela maioria como de valor nutritivo superior ao de casca branca, embora evidências científicas não comprovem esta suposição. Esta relacionada, também, ao índice de maturação e ao grau de deterioração, influenciando na decisão do consumidor. A indústria de alimentos utiliza a adição de cor nos produtos alimentícios para conferir ou aumentar o desejo do consumidor pelo produto (DELLA MODESTA, 1994; MEILGAARD; CIVILLE; CARR, 1991).

A aparência do produto alimentício influencia a opinião do consumidor em relação aos outros atributos, tendo consequência na decisão de compra. O consumidor espera que o alimento ou bebida apresentem a cor que o caracteriza, relutando em consumir produtos com cores diferentes da tonalidade ou intensidade esperada para aquele produto (OLIVEIRA, 2010).

De um modo geral, o alimento é aceito ou não aceito, em primeiro lugar, pelo julgamento visual de sua aparência e, se a cor não for atraente, apesar de serem agradáveis em outros atributos, dificilmente o alimento ou bebida será consumido (DELLA MODESTA, 1994; MEILGAARD; CIVILLE; CARR, 1991).

3.3.2. O sentido do olfato

A percepção olfativa ocorre no nariz, constituído internamente por epitélio com células receptoras de compostos odoríferos. As células receptoras dotadas de pêlos são terminações dos neurônios olfativos e os compostos odoríferos estimulam as células sensíveis. No que diz respeito à percepção pelo sentido do olfato, o atributo que é percebido quando compostos voláteis são inspirados pelo nariz, voluntariamente ou passivamente, é denominado odor (MANLEY, 1994; MEILGAARD; CIVILLE; CARR, 1991).

Quando um composto odorífero é espalhado no ambiente ou um alimento é mastigado compostos voláteis chegam ao epitélio olfativo. O epitélio olfativo, localizado na parte interna do topo do nariz, é alcançado por pequena fração de ar da inspiração e pelos compostos voláteis. As células receptoras captam os compostos e transmitem aos bulbos olfativos, que realizam uma conexão com o cérebro. A sensibilidade de captar odores varia com o indivíduo e diminui com a idade (MEILGAARD; CIVILLE; CARR, 1991).

O termo aroma é utilizado para denominar os odores dos alimentos, assim como o termo fragrância é utilizado para cosméticos e perfumes. Existem aproximadamente cinco mil aromas diferentes, porém o ser humano detecta entre dois mil a quatro mil, comprovando a alta sensibilidade do sentido do olfato e sua grande capacidade de discriminação (MEILGAARD; CIVILLE; CARR, 1991).

De um modo geral, o sentido do olfato evoca a memória e estimula reações do ser humano, sendo mais sensível e possuindo maior poder de discriminação que o sentido do paladar. O odor pode atrair ou repelir os consumidores, considerando que pode caracterizar um alimento e indicar o seu estado de conservação. Odores caracterizam carne deteriorada, gordura rançosa e alimentos mofados, servindo, assim de sinal de alerta para o consumidor (DELLA MODESTA, 1994; MEILGAARD; CIVILLE; CARR, 1991).

O sentido do olfato é sujeito à adaptação ou fadiga. A adaptação olfativa é um fator importante a ser considerado em análise sensorial. Ocorre quando o provador é exposto continuamente a um determinado odor e sua capacidade de avaliação vai decrescendo à medida que subseqüentes aspirações vão sendo realizadas, impossibilitando, assim, uma avaliação efetiva desse odor. O provador nessa situação pode avaliar odores diferentes

daquele que causou a adaptação. A fadiga, por sua vez, consiste na impossibilidade de avaliar qualquer odor após uma exposição prolongada a um odor, o provador fica incapacitado de avaliar odores por um período de tempo (MANLEY, 1994; MEILGAARD; CIVILLE; CARR, 1991).

3.3.3. O sentido do tato e da audição

A boca e as mãos podem fornecer informações táteis do alimento. A textura é definida como todas as propriedades reológicas (mudanças na forma e no fluxo de um alimento ou bebida), e estruturais (geométrica e de superfície) de um alimento pelos receptores mecânicos, táteis e eventualmente pelos receptores visuais e auditivos. A textura é um importante atributo físico dos alimentos, sendo que as percepções táteis podem influenciar drasticamente o prazer de comer. As células receptoras sensíveis ao tato da boca estão localizadas no epitélio que reveste os lábios, as bochechas, a gengivas, a língua e o palato. As sensações estimuladas pelo manuseio do alimento complementam as informações que chegam pelos outros órgãos sensoriais (MEILGAARD; CIVILLE; CARR, 1991).

A textura é a sensação derivada da degustação de um alimento que estimula os receptores bucais e inclui as sensações na boca, como o calor, o frio e a adstringência, as propriedades mastigatórias, como a dureza, a viscosidade, a elasticidade, a mastigabilidade e a gomosidade, as propriedades residuais e o som. As características sensoriais de textura podem ser captadas, também, pelos dedos e consistem na firmeza, na suavidade e na suculência (DELLA MODESTA, 1994; MEILGAARD; CIVILLE; CARR, 1991).

O som ou ruído, que é produzido durante a mastigação dos alimentos, é apreciado e exigido pelos consumidores em determinados alimentos. Como exemplos, podemos citar o som áspero produzido na mastigação do aipo, da alface e da maçã, o som do borbulhamento das bebidas gasosas e o som da espuma na cerveja. O som também é associado ao preparo de alimentos como o milho-pipoca, que estala e o fritar dos ovos (MEILGAARD; CIVILLE; CARR, 1991).

Portanto, os sentidos do tato e da audição, simultaneamente, permitem a percepção da textura de alimentos e bebidas e fazem parte da satisfação ao comer. São características que serão avaliadas pelo provador na análise sensorial. Durante os testes sensoriais é

recomendado ter atenção à presença de ruídos estranhos que dispersem a atenção do provador e comprometa a avaliação do alimento em estudo.

A importância da textura na aceitabilidade global dos alimentos varia largamente, dependendo do tipo de alimento. Existem alimentos onde essa característica é importante como, por exemplo, na batata chips e no amendoim torrado. A avaliação das propriedades de textura não tem apenas a finalidade de indicar as características do produto final, mas de servir também como uma ferramenta do controle de qualidade da matéria-prima ou do produto em vários estágios do processamento (DELLA MODESTA, 1994; MEILGAARD; CIVILLE; CARR, 1991).

3.3.4. O sentido do gosto

A percepção do gosto ocorre na boca, pelos órgãos gustativos (botões gustativos) quando estimulados por determinadas substâncias solúveis. A gustação envolve a detecção do estímulo dissolvido em água, óleo ou saliva pelos botões gustativos que estão localizados, principalmente, na superfície da língua, assim como na mucosa do palato e em áreas da garganta (DELLA MODESTA, 1994). São percebidos quatro gostos básicos: doce, ácido, amargo e salgado, além do gosto metálico e do umami, gosto característico do glutamato (OLIVEIRA, 2010).

O sabor é a experiência mista e unitária de sensações olfativas, gustativas e táteis percebidas durante a mastigação e a deglutição. O sabor é influenciado pelos efeitos táteis (sólido, mole, crocante, travar da língua/cica e ardência/pimenta), térmicos (quente, frio) e, principalmente, sinestésicos, que é a relação de diferentes atributos sensoriais como, por exemplo, o gosto com o cheiro, ou a visão com o olfato. O termo *flavor* é utilizado para a sensação global da percepção do aroma, do gosto e do tato quando o alimento ou bebida se encontra na boca. O conjunto dessas percepções torna o alimento singular em relação a outro (MEILGAARD; CIVILLE; CARR, 1991).

A sensibilidade máxima e a habilidade sensorial para avaliar um alimento ou bebida ocorrem quando ele apresenta temperatura entre 10°C e 35°C. Com o aumento da temperatura, há um aumento na sensibilidade para o sabor doce e uma diminuição para os

sabores salgado e amargo (MEILGAARD; CIVILLE; CARR, 1991). Para a avaliação sensorial é recomendado que a amostra seja apresentada ao provador na temperatura na qual o produto avaliado é consumido.

3.4. Fatores que afetam a avaliação sensorial

A percepção das características dos produtos alimentícios é afetada por muitos fatores individuais. O julgamento dos atributos sensoriais é influenciado por fatores fisiológicos, comportamentais e cognitivos da experiência do consumidor, acarretando modificações na sua percepção visual, odorífera, tátil, auditiva e gustativa. Sendo assim, o estudo do comportamento do consumidor é tarefa multidisciplinar, que envolve ciência e tecnologia de alimentos, nutrição, fisiologia, psicologia, sociologia e marketing (GUINARD; UOTANI; SCHLICH, 2001; LANGE; ROUSSEAU; ISSANCHOU, 1999).

A expectativa do consumidor em relação a um produto alimentício pode ser gerada por características não sensoriais, tais como informação acerca do produto (SIRET; ISSANCHOU, 2000), informação nutricional (DELIZA; MACFIE, 1994) preço (DI MONACO; CAVELLA; DI MARZO; MASI, 2004), embalagem e rótulo (SMYTHE; BAMFORTH, 2002). Estas expectativas em relação ao produto alimentício em estudo apresentam um papel importante no seu julgamento, uma vez que elas podem melhorar ou piorar a percepção do provador em relação ao produto (DELIZA; MACFIE, 1994). A embalagem pode levar o consumidor a comprar o produto, enquanto as características sensoriais confirmam a apreciação e podem determinar a reincidência na compra (MURRAY; DELAHUNTY, 2000).

Durante o processo de compra, os consumidores buscam informações da memória e do ambiente externo e processam e armazenam os resultados de sua compra em sua memória, para que estes sejam usados em outras compras similares. A presença de uma marca bem estabelecida no mercado, portanto, é uma forte influência na formulação das expectativas sensoriais dos consumidores, assim como em seu comportamento de escolha, compra e de aceitação do produto (DELIZA; MACFIE, 1994; DI MONACO; CAVELLA; DI MARZO; MASI, 2004; JAEGER, 2006).

Os consumidores estão cada vez mais exigentes em relação aos produtos que adquirem, levando a um aumento da competitividade na indústria de alimentos e bebidas, na qualidade dos produtos comercializados e nos investimentos em pesquisa na área de análise sensorial, panorama confirmado pelo grande volume de trabalhos científicos publicados sobre o tema, inclusive sobre o aperfeiçoamento dos métodos em análise sensorial (VILLANUEVA; PETENATE; SILVA, 2005; VILLANUEVA; PETENATE; SILVA, 2000).

3.5. Provadores em análise sensorial

A análise sensorial pode ser realizada por voluntários treinados ou não treinados. Para a formação de uma equipe de provadores treinados, inicialmente os voluntários são selecionados por meio de testes que avaliam a sua habilidade de discriminar para, em seguida, serem treinados de acordo com o objetivo da análise sensorial e o alimento a ser estudado, de modo a se tornar um instrumento de medida acurado (PIGGOTT, 1995).

A maior parte dos estudos deve utilizar provadores com idade adulta, entre dezoito e cinquenta anos, a não ser que o alimento em estudo tenha como público alvo outros grupos etários, como crianças e idosos. A dificuldade da análise sensorial com crianças consiste na falta da capacidade de usar uma terminologia adequada para expressar suas próprias impressões sensoriais, por outro lado, voluntários acima de cinquenta anos podem não apresentar uma boa acuidade sensorial, devido à perda da sensibilidade dos receptores sensoriais. Quando crianças e idosos são os alvos do produto em estudo, é necessário maior tempo de treinamento para obter melhores resultados (OLIVEIRA, 2010).

Uma prática comum é a validação de uma equipe de provadores por meio de diferentes testes ou medidas. A equipe é considerada validada quando os testes ou medidas aplicadas obtêm o mesmo resultado e validam um ao outro. Experimentos realizados com equipes validadas não necessitam de um grande número de provadores na equipe, podendo ter de sete a quinze para testes descritivos ou sete a doze para testes de detecção de diferenças (MEILGAARD; CIVILLE; CARR, 1991; DUTCOSKY, 1996).

Considerando as limitações fisiológicas da equipe de provadores, uma sessão de testes sensoriais pode apresentar o número máximo de seis amostras (BALL, 1997; MEILGAARD; CIVILLE; CARR, 1991). O planejamento de um experimento sensorial deve escolher um desenho experimental eficiente, que necessite o mínimo de recursos e que resulte em boa qualidade de dados. Isto significa usar um pequeno número de provadores treinados, gastar pouco tempo com treinamento, usar métodos sensoriais mais simples, com o mínimo de repetições possíveis. Os resultados devem ser precisos e validados com algum teste interno ou externo ao experimento (PIGGOTT, 1995).

3.6. Métodos sensoriais

Os métodos sensoriais são classificados em afetivos, discriminativos e descritivos. Essa classificação é caracterizada pelo objetivo do teste e pelo critério de seleção dos provadores (CHAVES, 1993).

Segundo Lawless e Claasen (1993), a escolha de um método de análise sensorial é baseada na resposta a três questões fundamentais relacionadas ao produto em estudo. A primeira pergunta é se o produto é aceito pelos consumidores, a segunda é se existe diferença perceptível entre o produto em estudo e algum produto convencional similar e a terceira quais os principais pontos de diferença. As respostas a estas três questões permitem classificar os métodos sensoriais em afetivos ou de aceitabilidade, para resolução da primeira pergunta; discriminativos ou de diferença, para a segunda; e de análise descritiva, para a terceira.

Os métodos afetivos apresentam testes que podem ser classificados nas categorias de testes de preferência e testes de aceitabilidade. O objetivo desses testes é obter a opinião do consumidor em relação a sua preferência ou aceitabilidade para o produto em estudo, seja ele um produto convencional ou um produto novo com potencial de mercado, motivo pelo qual são chamados, também, de testes de consumidor. A preferência pode ser conceituada como uma expressão do grau de gostar ou a escolha de uma amostra em relação à outra. E a aceitabilidade pode ser conceituada como uma experiência caracterizada por uma atitude positiva ou não na utilização do produto em estudo (MEILGAARD; CIVILLE; CARR, 1991).

Os testes afetivos são utilizados, principalmente, pela indústria de produtos de venda direta ao consumidor, embora possa ter importância na estratégia de indústrias de ingredientes como aromas, corantes, aditivos, misturas de pré-preparo de alimentos e organizações prestadoras de serviço. Dentre as principais aplicações dos testes afetivos podemos citar o desenvolvimento de novos produtos, a otimização de produtos e processos e a manutenção da qualidade do produto (OLIVEIRA, 2010).

Nos testes afetivos um grupo de pessoas deve ser selecionado como uma amostragem representativa da população de consumidores ou consumidores potenciais do produto em estudo (CHAVES, 1993). Neste caso, a seleção não se baseia somente nas habilidades sensoriais dos provadores como, também, na probabilidade de que os provadores façam parte da população alvo para o qual o produto é destinado. Em hipótese alguma, devem ser selecionados para testes afetivos, provadores treinados e membros da equipe envolvida no desenvolvimento do produto em estudo. A idade, o gênero, o estado civil, a localização geográfica, a nacionalidade, a religião, a etnia, o grau de instrução e o estado de saúde devem ser levados em consideração na seleção dos provadores. Coletados por meio de questionário, estes parâmetros irão caracterizar e direcionar a seleção do público alvo (DUTCOSKY, 1996).

Os testes afetivos podem ser aplicados em ambiente de laboratório, em cabines individuais, onde as condições de realização dos testes são passíveis de um maior controle, ou em locais centrais, como em supermercados e escolas, onde circulam os consumidores em potencial, e em domicílios. Quando os testes são realizados em locais centrais, são necessários acima de cem julgadores. (CHAVES, 1993).

Em ambiente de laboratório, os testes devem ser realizados com uma equipe composta de vinte e cinco a cinquenta julgadores (CHAVES, 1993). Segundo Stone e Sidel (1993), quando o número de provadores é de no mínimo cinquenta o poder de discriminação da equipe e a validade dos resultados obtidos aumenta consideravelmente, principalmente em testes envolvendo apenas dois produtos.

Os testes afetivos realizados em laboratório possuem a vantagem de oferecer maior controle no preparo e apresentação das amostras, individualidade do provador, luzes apropriadas, maior controle da temperatura de análise, rápido retorno dos resultados e baixo

custo. As desvantagens incluem o preparo do produto fora do ambiente doméstico e o efeito estressante das condições laboratoriais sobre a percepção humana (CHAVES, 1993).

Os testes realizados em locais centrais apresentam a vantagem de acesso a um maior número de consumidores do produto em estudo, resultando em uma amostragem mais representativa do mercado em potencial. As desvantagens incluem o preparo do produto de forma não similar ao preparo doméstico e a menor individualidade dos julgamentos.

Os testes domésticos, realizados na casa do consumidor sob condições normais de preparo e consumo do produto em estudo, em que é obtida a opinião da família inteira, apresentam como vantagens a manutenção do preparo e consumo na forma real, o maior número de informação que se pode obter do produto em estudo e a avaliação de forma continuada pelo uso frequente do produto. As desvantagens dos testes domésticos incluem o alto custo, o longo tempo requerido para a obtenção dos dados e a variação na forma de preparo do produto de um domicílio para outro (CHAVES, 1993).

3.7. Medidas sensoriais e o emprego de escalas

A forma de quantificar as respostas aos estímulos sensoriais é crítica para o bom desempenho da análise sensorial e, dentre as diversas técnicas utilizadas, a mais usual é o emprego de escalas para medir a resposta do provador. Dentre as várias classes de escalas, o analista sensorial deve selecionar a mais apropriada de acordo com o objetivo do teste, o tipo de informação desejada, a qualificação do julgador e o tipo de produto (MEILGAARD; CIVILLE; CARR, 1991).

As escalas são classificadas em escala nominal, ordinal e de intervalos. A escala nominal é utilizada em classificações ou denominações; a escala ordinal é utilizada para uso em ordenação ou classificação em ordens e a escala de intervalos é utilizada para uso em medidas de magnitude ou tamanho, admitindo distâncias iguais entre os pontos da escala (STONE; SIDEL, 2004).

As escalas podem ser classificadas também em verbal ou nominal, numérica ou gráfica, sendo a escala gráfica muito utilizada para testes sensoriais com crianças (Figura 2).

Quanto à forma, as escalas verbais podem ser não estruturadas de 9 e 15 cm, utilizando termos hedônicos como âncoras (Figura 3), e estruturadas (Figura 4). Os termos hedônicos devem ser sempre balanceados, ou seja, o número de categorias positivas deve ser igual ao de categorias negativas, com um ponto zero arbitrário (OLIVEIRA, 2010; STONE; SIDEL, 1993; MEILGAARD; CIVILLE; CARR, 1991).

Em questionários de recrutamento de julgadores treinados, as escalas nominais do tipo estado civil, classe de renda, idade, gênero e comportamento na utilização do produto em estudo são bastante utilizadas. Para essas escalas, as operações matemáticas possíveis são: contagem e distribuição de frequências, moda (categoria que ocorre com maior frequência) e coeficiente de contingência (associação entre variáveis) (MEILGAARD; CIVILLE; CARR, 1991).



Figura 2. Exemplo de escala hedônica gráfica facial.

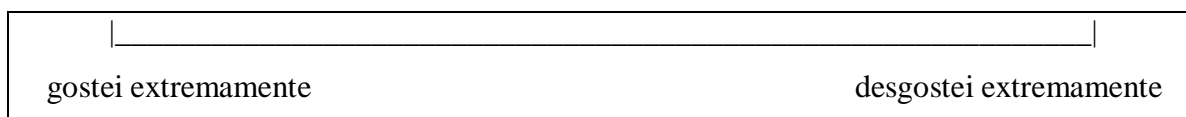


Figura 3. Exemplo de escala hedônica não estruturada com termos hedônicos como âncora.

<input type="checkbox"/> Gostei muitíssimo	<input type="checkbox"/> Gostei muitíssimo	<input type="checkbox"/> Gostei muito
<input type="checkbox"/> Gostei muito	<input type="checkbox"/> Gostei muito	<input type="checkbox"/> Gostei
<input type="checkbox"/> Gostei moderadamente	<input type="checkbox"/> Gostei	<input type="checkbox"/> Não gostei/Não desgostei
<input type="checkbox"/> Gostei ligeiramente	<input type="checkbox"/> Não gostei/Não desgostei	<input type="checkbox"/> Desgostei
<input type="checkbox"/> Não gostei/Não desgostei	desgostei	<input type="checkbox"/> Desgostei muito
<input type="checkbox"/> Desgostei ligeiramente	<input type="checkbox"/> Desgostei	
<input type="checkbox"/> Desgostei moderadamente	<input type="checkbox"/> Desgostei muito	
<input type="checkbox"/> Desgostei muito	<input type="checkbox"/> Desgostei	
<input type="checkbox"/> Desgostei muitíssimo	muitíssimo	

Figura 4. Exemplo de escala hedônica estruturada verbal com 9, 7 e 5 pontos.

A escala de intervalo possui constância do intervalo ou da distância entre seus pontos, a partir de um ponto zero arbitrário. São escalas verdadeiramente qualitativas, cujos dados são considerados matematicamente como variáveis quantitativas, com intervalos iguais, permitindo diversos tratamentos estatísticos paramétricos univariados e multivariados. As escalas hedônicas verbais são exemplos desse tipo de escala.

A escala de ordenação é construída a partir de termos de intensidade, de maior para menor, ou números que representam a ordenação. A forma mais utilizada é solicitar que o provador arrume as amostras em ordem crescente ou decrescente, a partir de algum atributo específico do produto ou da preferência do provador. A escala ordinal fornece a direção da diferença entre as amostras e se existe diferença detectável entre elas. Entretanto, não informa a magnitude da diferença porque os intervalos da escala não podem ser estimados (Figura 5) (MEILGAARD; CIVILLE; CARR, 1991).

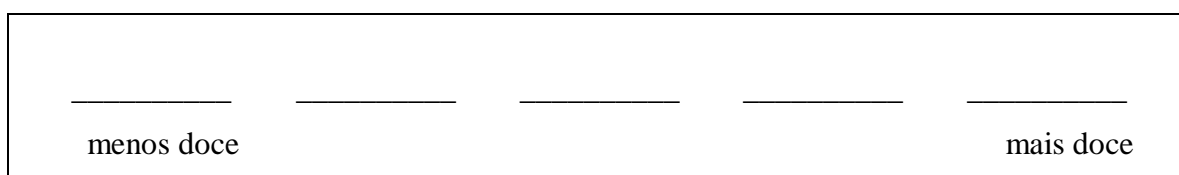


Figura 5. Exemplo de escala de ordenação.

3.8. Uso de escalas em teste de aceitabilidade

A aceitabilidade representa o somatório de todas as percepções sensoriais e expressa o julgamento, por parte do consumidor, sobre a qualidade do produto. A aceitação do consumidor é o objetivo final na ciência e tecnologia de alimentos (DUTCOSKY, 1996).

O teste de aceitabilidade é utilizado quando o objetivo é avaliar o grau com que os consumidores gostam ou desgostam de um produto, consistindo em um julgamento subjetivo. Neste teste o voluntário utiliza uma escala balanceada com igual número de categorias positivas e negativas (DUTCOSKY, 1996). A aplicação deste teste permite avaliar a aceitação global, que é o julgamento de um produto alimentício como um todo, ou avaliar a aceitação de um atributo específico do produto, como a cor, o aroma e o gosto. O recomendado é a apresentação ao provador de uma amostra de cada vez, sequencialmente, uma após a outra. Todos os provadores devem provar todas as amostras, de modo randômico (MEILGAARD; CIVILLE; CARR, 1991).

O uso da escala hedônica estruturada verbal, comparada com a escala numérica, é vantajoso porque os termos hedônicos constituem uma definição de cada ponto da escala que ajudam no julgamento do voluntário. Os termos e as palavras escolhidas são muito importantes e devem expressar bem claro o significado da resposta e dar idéia sucessiva dos intervalos da escala (DUTCOSKY, 1996).

Garrutti, (1985) ao testar nove diferentes escalas, concluiu que não houve superioridade de uma sobre a outra. Baseado nesse resultado, ele sugeriu que escalas de nove categorias são mais sensíveis que escalas menores, que os termos indiferente ou neutro não devem ser utilizados; e que a presença de números positivos e negativos na escala não constitui uma forma ideal para escalas de avaliação.

Além da escolha adequada da escala, as fichas de avaliação apresentadas ao voluntário devem conter instruções que permitam informar o que ele deve saber para dar a resposta que deseja e ajudar o provador a emitir respostas imediatas e sem dificuldades. Em escalas não estruturadas, onde são fixados somente os pontos de cada extremidade da escala, existe uma tendência à distribuição dos julgamentos se posicionarem mais próximos a eles (DUTCOSKY, 1996).

Dentre as vantagens do teste de escala hedônica, podemos apontar a simplicidade da aplicação, a utilização de grande número de voluntários, a facilidade que o voluntário tem para responder sem uma prévia experiência com a amostra, a possibilidade de tratar os resultados estatisticamente; e a obtenção de resultados entre limites amplos, podendo indicar níveis gerais de preferência. No que se refere às desvantagens, podemos citar que a escala hedônica não deve ser utilizada em controle de qualidade devido à ocorrência de variações nem sempre controláveis no teste, exigindo um número maior de provadores do que usualmente utilizado em controle de qualidade (DUTCOSKY, 1996).

3.9. Tratamento estatístico em análise sensorial

A Estatística é a ciência que utiliza um conjunto de métodos para analisar dados. Pode ser aplicada em praticamente todas as áreas do conhecimento humano e em algumas áreas recebe um nome especial, de acordo com a sua aplicação. Em Ciências Biológicas e da Saúde, por exemplo, é chamada de Bioestatística (SOARES; SIQUEIRA, 2002).

Análises estatísticas paramétricas, como a ANOVA, são usadas largamente no tratamento dos resultados em análise sensorial, estas análises são baseadas em suposições estatísticas de independência, normalidade e homocedasticidade de respostas experimentais. Do mesmo modo, a maioria dos testes estatísticos usados para comparar médias, tais como o teste t, teste de Duncan e teste de Tukey, são testes paramétricos que assumem a normalidade dos resultados (VILLANUEVA; PETENATE; SILVA, 2000; VIE; GULLI; O'MAHONY, 1991; McPHERSON; RANDALL, 1985).

Frequentemente, a informação sensorial produz distribuições assimétricas. Em testes realizados com consumidores usando escala hedônica estruturada, isto ocorre quando as amostras apresentam um consenso muito alto, com maior frequência de marcações nas escalas entre os valores 7 (sete) e 9 (nove), ou uma aceitação muito baixa, resultando em maior frequência entre os valores 1 (um) e 3 (três) da escala (WILKINSON; YUKSEL, 1997). É sabido também que intervalos numéricos não refletem, necessariamente, diferenças equivalentes em percepção (GAY; MEAD, 1992; O'MAHONY, 1982). Distribuições bimodais de respostas hedônicas também já foram observadas quando são usadas escalas não-estruturadas, o que viola claramente a suposição de normalidade (GIOVANNI; PANGBORN, 1983).

Dados de consumidores obtidos em testes de escala hedônica violam o pressuposto básico de normalidade (VILLANUEVA; PETENATE; DA SILVA, 2000; VIE; GULLI; O'MAHONY, 1991; MCPHERSON; RANDALL, 1985). Um problema observado no tratamento estatístico aplicado aos resultados de teste de escala hedônica é assumir que os dados obtidos apresentam distribuição normal, contradizendo ao encontrado na prática que são frequentemente, informações sensoriais com distribuição assimétrica (LIM; FUJIMARU, 2010; VILLANUEVA; DA SILVA, 2009; VILLANUEVA; PETENATE; DA

SILVA, 2005; VILLANUEVA; PETENATE; DA SILVA, 2000; WILKINSON; YUKSEL, 1997).

Vários autores (VILLANUEVA; PETENATE; SILVA, 2000; WILKINSON; YUKSEL, 1997; BROCKHOFF; SKOVGAARD, 1994; GAY; MEAD, 1992; PRITCHETT, 1992; NAES, 1990; MILLER, 1987) têm proposto métodos alternativos para a análise estatística de dados em análise sensorial. Entretanto, a literatura mostra que a ANOVA e outros testes paramétricos ainda são largamente utilizados para o tratamento estatístico de teste de escala hedônica, mesmo que as pressuposições básicas de normalidade, homocedasticidade e independência estejam violadas, e indiquem o uso de testes não-paramétricos.

3.10. Tipos de variáveis

Variável é a característica de interesse que é medida em cada elemento da amostra ou população. Como o nome diz, seus valores variam de elemento para elemento. As variáveis podem apresentar valores numéricos ou não numéricos (SOARES; SIQUEIRA, 2002).

As variáveis podem ser classificadas em variáveis quantitativas e qualitativas. As variáveis quantitativas são as características que podem ser medidas em uma escala quantitativa, ou seja, apresentam valores numéricos que fazem sentido. Podem ser discretas ou contínuas (Figura 6) (BUSSAB; MORETTIN, 2005).

As variáveis discretas são as características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores e, assim, somente fazendo sentido um valor inteiro. Geralmente, são resultados de contagens como, por exemplo, o número de filhos, o número de bactérias por litro de leite, o número de cigarros fumados por dia, (SOARES; SIQUEIRA, 2002). Na análise sensorial, é obtido um valor inteiro de respostas como, por exemplo, o número de provadores que escolheram a opção “gostei muito” para uma determinada amostra, e esta informação pode ser considerada uma variável discreta.

As variáveis contínuas são características mensuráveis que assumem valores em uma escala contínua na reta real, para as quais valores fracionários fazem sentido. Usualmente

devem ser medidas por algum instrumento como, por exemplo, o peso aferido por uma balança, a altura medida por uma régua, o tempo verificado em um relógio, a pressão arterial medida no aparelho medidor de pressão arterial e a idade (SOARES; SIQUEIRA, 2002).

Em análise sensorial, podem ser observadas variáveis aleatórias contínuas quando é utilizada uma escala não estruturada. Neste caso, o ponto marcado pelo provador na escala é medido com uma régua para possibilitar, posteriormente, a análise dos resultados. Os dados encontrados em escalas não estruturadas muitas das vezes não são valores inteiros sendo necessário, então, a aproximação destes valores para efeito de cálculo.

Variáveis qualitativas são utilizadas em pesquisas e representam uma qualidade ou atributo do indivíduo pesquisado, sendo exemplos dessas variáveis o gênero, o grau de escolaridade e o estado civil. Essas variáveis podem ser variáveis qualitativas nominais e ordinais (Figura 6). As variáveis qualitativas nominais são aquelas para as quais não existe nenhuma ordenação nas possíveis realizações, e as qualitativas ordinais, são aquelas para as quais existe uma ordem nos seus resultados. Um exemplo de variável nominal obtida em pesquisa é a região de procedência de um indivíduo, enquanto grau de instrução é um exemplo de variável ordinal, considerando que os níveis de escolaridade fundamental, médio e superior correspondem a uma ordenação baseada no número de anos de escolaridade completos (BUSSAB; MORETTIN, 2005). Em análise sensorial, as categorias de gostar e desgostar apontadas na escala hedônica podem ser consideradas variáveis qualitativas ordinais, levando em conta que gostar ou desgostar ligeiramente, moderadamente, muito e muitíssimo expressam um grau lógico de gostar e desgostar.

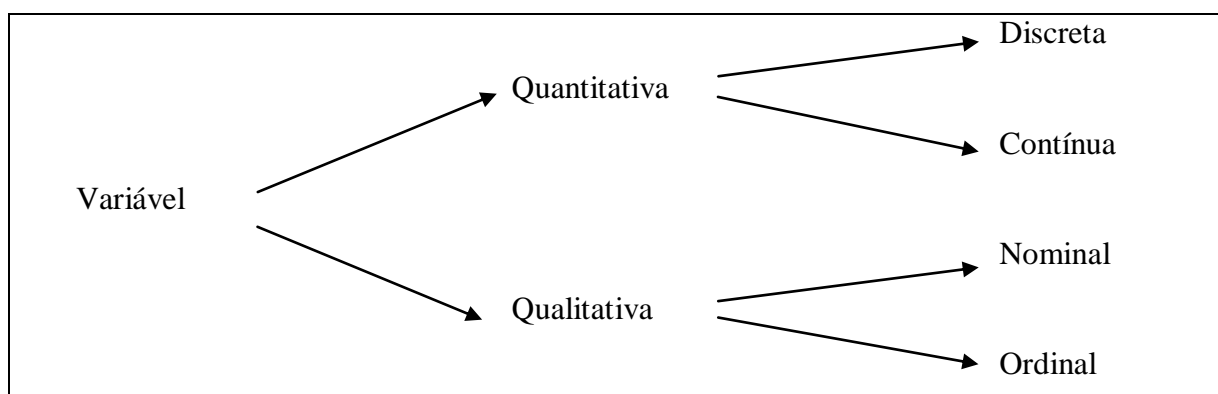


Figura 6. Classificação das variáveis.

Contudo o conhecimento de modelos probabilísticos para variáveis quantitativas é muito importante e as variáveis para as quais são construídos modelos probabilísticos são chamadas de variáveis aleatórias (BUSSAB; MORETTIN, 2005).

3.11. Variáveis aleatórias

Para cada tipo de variável existe uma técnica apropriada para resumir as informações observadas, embora algumas técnicas usadas em um caso possam ser adaptadas para outros. Em algumas situações, são atribuídos valores numéricos às várias categorias de uma variável qualitativa para que possa ser realizado o tratamento dos resultados como se ela fosse uma variável quantitativa, tomando o cuidado para que o procedimento seja passível de interpretação (BUSSAB; MORETTIN, 2005).

Quando se estuda descrição de dados, pode ser observado que os recursos disponíveis para a análise das variáveis quantitativas são muito mais ricos do que para as variáveis qualitativas. Isso sugere o uso de artifícios para transformar as variáveis qualitativas em quantitativas. Como exemplo, podemos citar o caso da proposição, em um questionário, em que as respostas possíveis sejam “sim” ou “não”. Nesse caso, as variáveis podem tomar os valores 1 (um) ou 0 (zero), correspondentes as respostas “sim” ou “não”, respectivamente (BUSSAB; MORETTIN, 2005).

Em testes sensoriais onde são utilizadas escalas hedônicas, as variáveis qualitativas ordinais, que representam o grau de gostar ou desgostar do provador em relação à amostra avaliada, são transformadas em variáveis quantitativas discretas, para que mais recursos estatísticos possam ser utilizados para analisar esses tipos de variáveis. Um exemplo disso, é a aplicação de valores, em uma escala crescente de nove pontos, para os termos desgostei muitíssimo que receberá o valor 1 (um), o desgostei muito o valor 2 (dois) e assim sucessivamente até o termo gostei muitíssimo que receberá o valor 9 (nove).

3.11.1. Variáveis aleatórias discretas

A cada ponto do espaço amostral, que na escala hedônica é denominado categoria, a variável em consideração é associada a um valor numérico, o que corresponde em

Matemática ao conceito de função, mais precisamente, a uma função definida no espaço amostral Ω (conjunto de todos os valores possíveis) e assumindo valores reais. Logo, variável aleatória discreta pode ser definida como uma função X , definida no espaço amostral Ω e com valores num conjunto enumerável de pontos da reta (BUSSAB; MORETTIN, 2005).

Para a compreensão do que é uma variável aleatória discreta, é comum o uso do exemplo do lançamento de uma moeda por duas vezes. O número de “caras” (C) obtidas nos dois lançamentos é definida como variável aleatória Y . Pode ocorrer também “coroa” (R), sendo que pode ser obtido após o lançamento os resultados CC, CR, RC, RR. Em todas as possibilidades nesse lançamento a probabilidade de ocorrer é de $\frac{1}{4}$ (um quarto). Sendo que a probabilidade de não ocorrer “cara” é de $\frac{1}{4}$ (um quarto), a de ocorrer uma “cara” é $\frac{1}{2}$ (um meio) e a de ocorrer duas “caras” é $\frac{1}{4}$ (um quarto) (BUSSAB; MORETTIN, 2005). Em teste de aceitabilidade, no qual é utilizada escala hedônica, as respostas dos avaliadores são consideradas variáveis aleatórias, e a probabilidade de ocorrência de cada uma dessas variáveis será proporcional à sua frequência de ocorrência.

Algumas variáveis aleatórias podem adaptar-se muito bem a uma série de problemas práticos. Portanto, um estudo pormenorizado dessas variáveis é de grande importância para a construção de modelos probabilísticos para situações reais e a consequente estimação de seus parâmetros. Para algumas dessas distribuições existem tabelas que facilitam o cálculo da probabilidade, em função de seus parâmetros que são quantidades desconhecidas na população sobre a qual desejamos obter uma estimativa. E entre os modelos probabilísticos para variáveis aleatórias discretas estão a distribuição uniforme discreta, a distribuição de Bernoulli e a distribuição binomial (BUSSAB; MORETTIN, 2005).

3.11.2. Distribuição uniforme discreta

A distribuição uniforme discreta é o caso mais simples de variável aleatória discreta, em que cada valor possível ocorre com a mesma probabilidade. Um exemplo dessa distribuição ocorre no lançamento de um dado, a probabilidade de ocorrer cada uma das faces é a mesma, ou seja, $\frac{1}{6}$ (um sexto) (BUSSAB; MORETTIN, 2005).

Na análise sensorial de aceitabilidade a distribuição das categorias na escala hedônica pode ser considerada uma distribuição uniforme discreta de valores possíveis, entretanto, muitas vezes os provadores representam grupos da população para os quais os produtos alimentícios testados são direcionados. Produtos voltados para crianças são testados em crianças como provadores. O mesmo acontece com estudos voltados para adolescentes e idosos. Nestes casos, uma tendência dos resultados já é esperada, baseada no conhecimento prévio que o pesquisador possui a respeito dos hábitos alimentares e preferências do grupo. Desta forma, distribuições uniformes discretas dos resultados são dificilmente obtidas nestes testes sensoriais.

3.11.3. Distribuição de Bernoulli

Em muitos experimentos o resultado é a presença ou não de uma determinada característica. Como exemplo pode ser citado o lançamento de uma moeda, que pode apresentar como resultado cara ou coroa; ou a retirada de uma amostra escolhida ao acaso de um lote contendo quinhentas unidades, que pode apresentar defeito ou não. Nesse modelo de experimento o que interessa é a ocorrência de sucesso (cara, peça sem defeito) ou fracasso (coroa, peça defeituosa) e, para cada um desses experimentos, pode ser definida uma variável aleatória X , que assume apenas dois valores, sendo 1 (um) se ocorrer sucesso, e 0 (zero) se ocorrer fracasso. Associado à ocorrência de sucesso existe a probabilidade de obter este resultado (BUSSAB; MORETTIN, 2005).

Na análise sensorial de aceitabilidade a distribuição das categorias na escala hedônica não pode ser considerada uma distribuição de Bernoulli, porque um produto pode ser aceito se a maioria dos provadores optarem pelos pontos da escala referentes às categorias do “gostar” e pode não ser aceito se a maioria dos provadores escolher os pontos da escala relacionados às categorias do “desgostar”. Outra característica da resposta no uso da escala hedônica é a possibilidade do provador ser indiferente em relação ao produto testado, escolhendo o ponto central da escala, onde ele aponta que não gosta e nem desgosta do produto. Existe a possibilidade de dividir a escala em dois grupos de forma a representar “gostar” ou “não gostar” e eliminar a resposta indiferente, viabilizando a utilização da distribuição de Bernoulli, entretanto, ocorre perda de informação ao dicotomizar a escala

hedônica, o que pode resultar em consequências consideráveis para o objetivo do estudo de um produto alimentício.

3.11.4. Distribuição binomial

A distribuição binomial é um exemplo de uma distribuição discreta. Ela lida com situações em que cada resultado de uma série de ensaios independentes resulta em um dentre dois resultados possíveis (SOARES; SIQUEIRA, 2002).

Se um ensaio de Bernoulli é repetido n vezes ou, de maneira alternativa, é obtida uma amostra de tamanho n de uma distribuição de Bernoulli, a variável ou número de sucessos nas n repetições segue uma distribuição binomial. É suposto ainda que as repetições sejam independentes, isto é, que o resultado de um ensaio não tem influência nenhuma no resultado de qualquer outro ensaio. Uma amostra particular será constituída de uma sequência de sucessos e fracassos, ou, alternativamente, de uns e zeros. Como exemplo, podemos citar a repetição de um ensaio de Bernoulli cinco vezes ($n = 5$) onde um resultado pode ser FSSFS (S = Sucesso e F = fracasso) ou a quintupla ordenada (0, 1, 1, 0, 1). O número de sucessos nessa amostra é igual a três, sendo dois o número de fracassos (BUSSAB; MORETTIN, 2005).

Outro exemplo de experimento binomial consiste em pedir a uma pessoa, que se diz um *expert* em vinhos, provar duas taças de vinho e escolher o mais caro. Esta tarefa é repetida 10 vezes, cada vez com um par diferente de vinhos. Ao assumir a princípio, que o apreciador não sabe nada sobre vinhos, que ele na verdade escolhe os vinhos ao acaso e está tentando esconder este fato, e que não existem fatores externos que interferem na decisão do apreciador (tal como uma tendência a escolher o vinho mais escuro), então, em cada ensaio a probabilidade do provador de acertar na escolha é de cinquenta por cento (SOARES; SIQUEIRA, 2002).

3.11.5. Variáveis aleatórias contínuas

As variáveis aleatórias contínuas são variáveis para as quais os possíveis valores pertencem a um intervalo de números reais. De modo geral, podemos dizer que as variáveis

aleatórias cujos valores resultam de algum processo de mensuração são contínuas. Pode ser definida uma função X sobre o espaço amostral Ω que assumem valores num intervalo de números reais, através de uma variável aleatória contínua. A característica principal de uma variável aleatória contínua é que, sendo resultado de uma mensuração, o seu valor pode ser pensado como pertencendo a um intervalo ao redor do valor efetivamente observado (BUSSAB; MORETTIN, 2005).

Na análise sensorial a avaliação do provador com uso de uma escala não estruturada pode ser considerada uma distribuição de variáveis aleatórias contínuas. Como exemplo pode ser citado o experimento onde o provador avalia uma nova formulação de biscoito utilizando uma escala não estruturada, ancorada com os termos “pouco crocante” e “muito crocante” em cada extremidade, e é orientado a marcar nessa escala a sua impressão sensorial quanto ao atributo em questão, que no caso é a crocância do biscoito. Após as avaliações sensoriais, a marcação feita pelos provadores na escala é mensurada com uma régua e os valores obtidos passam a constituir as variáveis contínuas.

3.11.6. A distribuição normal

A distribuição normal é a mais familiar das distribuições de probabilidade e também uma das mais importantes em estatística. É uma distribuição de probabilidade unimodal e simétrica. Em geral, as distribuições de probabilidade para variáveis contínuas não tem um limite para o valor superior e algumas, como a normal, não apresentam limite inferior. A altura da curva de frequência, chamada de densidade de probabilidade, não pode ser considerada como uma probabilidade de valor particular. A área total embaixo da curva é igual a um. Na equação matemática da distribuição normal é importante saber que ela está definida por dois parâmetros que são a média (μ) e o desvio padrão (σ) populacional. A distribuição normal com média zero e desvio padrão igual a um é chamada de distribuição normal padrão (Figura 7) e qualquer distribuição normal pode ser convertida a uma curva normal padrão (SOARES; SIQUEIRA, 2002).

Na distribuição normal a média consiste no centro da distribuição e o desvio padrão no espalhamento (ou achatamento) da curva. Como a distribuição normal é simétrica em

torno da média isso implica que a média, a mediana e a moda são todas coincidentes (SOARES; SIQUEIRA, 2002).

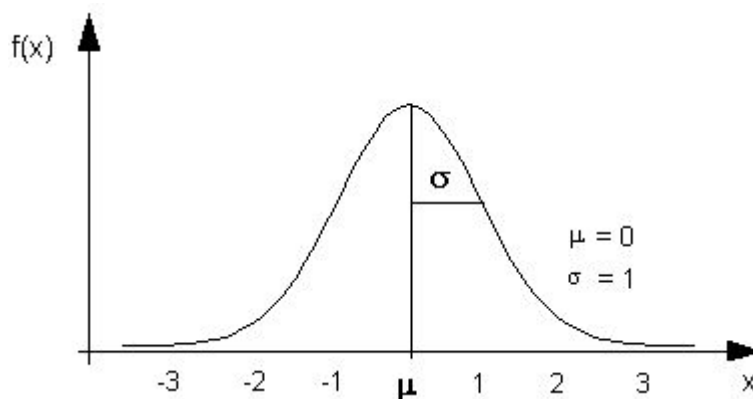


Figura 7. Distribuição Normal Padrão.

3.12. A escolha do método estatístico

Para escolher um método de análise estatística apropriado, diversos fatores devem ser considerados, tais como, o número de grupos de observação, se os dados são ou não independentes, o tipo de dados e a sua distribuição e, principalmente, o objetivo da análise. Os resultados de uma amostra devem ser considerados as melhores estimativas do que é verdade para a população e isso é alcançado com uma amostra representativa da população (SOARES; SIQUEIRA, 2002). Desta forma, os dados obtidos no teste de aceitabilidade a partir de indivíduos de uma amostra podem ser extrapolados para uma população de indivíduos similares.

Segundo Moore e McCabe (2002), os testes de hipóteses estão entre os tipos mais comuns de inferência. Os testes de hipóteses paramétricos são os mais utilizados muitas vezes devido ao desconhecimento dos testes não-paramétricos. A validação dos resultados dos testes paramétricos depende da verificação de suas pressuposições, como por exemplo, a normalidade dos dados, pressuposição básica para a maioria dos testes paramétricos (SIEGEL; CASTELLAN, 2006). Contudo, é importante verificar até que ponto os resultados dos testes paramétricos serão prejudicados quando a pressuposição de normalidade não venha a ser satisfeita.

3.12.1. Testes de hipóteses

Segundo Shimakura (2011), métodos estatísticos são utilizados para o planejamento e a condução de um estudo, a descrição dos dados e para a tomada de decisões, onde têm destaque os testes de hipóteses que se baseiam nos riscos associados às hipóteses formuladas.

A formulação de hipóteses vem sendo muito empregada em pesquisa em diversas áreas do conhecimento e para decidir se uma determinada hipótese é confirmada por um conjunto de dados, é necessário um conjunto de procedimentos objetivos para aceitar ou rejeitar a hipótese (SIEGEL; CASTELLAN, 2006).

A maior parte das análises estatísticas envolve comparações entre tratamentos ou procedimentos ou comparações entre grupos de indivíduos. O valor numérico que corresponde à comparação de interesse é chamado de efeito. O procedimento inicial consiste em estabelecer uma hipótese, denominada hipótese nula (H_0), que estabelece que o efeito seja zero. Adicionalmente, é estabelecida uma hipótese alternativa (H_1), onde o efeito de interesse não seja zero (BUSSAB; MORETTIN, 2005; SOARES; SIQUEIRA, 2002).

Uma vez estabelecidas as hipóteses, a probabilidade de ter obtido dados onde a hipótese nula for verdadeira é avaliado. Esta probabilidade é chamada de valor- p ; quanto menor for o valor- p , menos plausível é a hipótese nula. Quando o valor- p está abaixo de um valor crítico determinado, valor- p menor que 0,05, por exemplo, o resultado é chamado de estatisticamente significativo e a hipótese nula é rejeitada, e se estiver acima desse valor crítico se diz que o resultado é estatisticamente não significativo e a hipótese nula não é rejeitada (BUSSAB; MORETTIN, 2005; SOARES; SIQUEIRA, 2002).

Quando formulamos uma decisão sobre H_0 podem ocorrer dois erros distintos. O primeiro, designado erro tipo I, consiste em rejeitar a hipótese nula quando ela é verdadeira. O segundo, designado erro tipo II, consiste em aceitar H_0 quando ela é falsa. A estes erros estão associados uma probabilidade $P(\text{rejeitar } H_0 \mid H_0 \text{ verdadeira}) = \alpha$ e $P(\text{aceitar } H_0 \mid H_0 \text{ falsa}) = \beta$ (CÂMARA; SILVA, 2001).

Os testes de hipóteses se dividem em paramétricos e não-paramétricos. Os paramétricos são aqueles que utilizam os parâmetros da distribuição, ou uma estimativa destes, para o cálculo de sua estatística. Normalmente, estes são mais rigorosos e possuem mais pressuposições para sua validação. Já os não-paramétricos utilizam, para o cálculo de sua estatística, postos atribuídos aos dados ordenados e são livres da distribuição de probabilidade dos dados estudados (REIS; RIBEIRO, 2007).

No teste de aceitabilidade, se ao calcular as médias dos valores observados com uso de escala hedônica, de três ou mais amostras ou tratamentos, o valor- p assumir um valor menor que 0,05, significa, então, que essas amostras ou tratamentos não são igualmente aceitas. Em seguida, são aplicados testes de comparação de médias, duas a duas, para verificação da diferença entre as amostras.

3.13. Testes de comparação de médias

As médias amostrais apresentam algumas propriedades e, entre elas, podemos citar que a média de uma amostra aleatória pode ser diferente do valor da média da população, devido ao acaso, embora a expectativa seja de que estejam muito próximas. É necessário, portanto, quantificar a incerteza associada a esta estimativa. Outra propriedade das médias amostrais é a sua variabilidade, sendo menor quanto maior o tamanho da amostra e menor que a variabilidade dos indivíduos, aumentando na medida em que aumenta a variabilidade dos indivíduos. É esperado que a média de todas as médias possíveis seja igual à média da população e que o desvio padrão das médias de várias amostras seja σ/\sqrt{n} onde σ é o desvio padrão dos dados originais e n é o número de indivíduos na amostra. Este desvio é chamado erro padrão (BUSSAB; MORETTIN, 2005).

O objetivo de uma pesquisa se depara, em geral, com a necessidade de comparar e testar as médias de amostras ou tratamentos. A hipótese de igualdade de médias de amostras ou tratamentos é testada inicialmente pelo teste F (BORGES; FERREIRA, 2003). Porém, a igualdade de médias também pode ser testada por outros testes paramétricos e por testes não-paramétricos, como o teste Kruskal-Wallis. A significância desses testes em relação a um valor nominal fixado permite inferir que pelo menos um tratamento difere dos demais. Tanto o teste F quanto o teste Kruskal-Wallis não permitem ao pesquisador, no entanto,

descobrir onde estão essas diferenças, necessitando de testes subsequentes de comparação de médias, duas a duas, para descobrir quais amostras ou tratamentos diferem entre si.

A escolha do método adequado para comparar médias deve levar em consideração o nível de significância e o poder do teste de comparação. O nível de significância de um teste é a probabilidade de rejeitar a hipótese de que as médias são iguais (hipótese nula), quando esta hipótese é, na realidade, verdadeira. O poder do teste é a probabilidade de rejeitar a hipótese de que as médias são iguais quando esta hipótese é, na realidade, falsa (VIEIRA, 1999; DAWSON-SANDERS; TRAPP, 1994).

O pesquisador procura mostrar que a probabilidade de rejeitar a hipótese de que as médias são iguais, quando essa hipótese é verdadeira, é baixa e que a probabilidade de rejeitar essa mesma hipótese, quando ela é falsa, é alta. O pesquisador, então, necessita de um teste com baixo nível de significância e poder elevado (VIEIRA, 1999; DAWSON-SANDERS; TRAPP, 1994).

Os testes estatísticos podem ser divididos em dois grandes grupos, conforme fundamentem ou não os seus cálculos na premissa de que a distribuição de frequências dos erros amostrais é normal, as variâncias são homogêneas, e os erros independentes. Se tudo isso ocorrer, é muito provável que a amostra seja simétrica. Nesse caso, a amostra terá com certeza apenas um ponto máximo, centrado no intervalo de classe onde está a média da distribuição e o seu histograma de frequências terá um contorno que seguirá aproximadamente o desenho em forma de sino da curva normal. O cumprimento desses requisitos condiciona a primeira escolha do pesquisador, uma vez que, se forem preenchidos, ele poderá utilizar a estatística paramétrica, cujos testes são em geral mais poderosos do que os da estatística não-paramétrica e, conseqüentemente, devem ter a preferência do investigador, quando o seu emprego for permitido (CAMPOS, 2000).

Os termos paramétrico e não-paramétrico referem-se à média e ao desvio-padrão, que são os parâmetros que definem as populações que apresentam distribuição normal. Entretanto, em muitos artigos científicos e teses acadêmicas que utilizaram testes não-paramétricos, os resultados foram apresentados em termos de média, mais ou menos desvio-padrão da distribuição, ou então em termos de média, mais ou menos erro-padrão da média,

erro este que é também um valor calculado em função do desvio-padrão da amostra (CAMPOS, 2000).

É possível calcular a média de qualquer conjunto de valores numéricos, porém o cálculo do desvio-padrão só é possível em resultados que apresentem distribuição normal, uma vez que, por definição, desvio-padrão é o ponto de inflexão da curva normal. Os desvios-padrão são em número de dois e simétricos em relação à média da distribuição. Portanto, curvas assimétricas jamais podem apresentar desvio-padrão porque, mesmo que tenham pontos de inflexão, como outras curvas matemáticas, esses pontos dificilmente seriam simétricos em relação à média. Portanto, mesmo que distribuições experimentais possam apresentar alguma assimetria, esta deve manter-se dentro de certos limites, aceitáveis em termos estatísticos, e aceitáveis porque atribuídos à variação casual determinada pelos erros não-controlados de amostragem, ou seja, à variação do acaso, típica das variáveis e amostras chamadas aleatórias (SOARES; SIQUEIRA, 2002).

3.14. Testes paramétricos

Em estatística, um teste de hipóteses é um método para verificar se os dados são compatíveis com alguma hipótese, podendo muitas vezes sugerir a não-validade da mesma. O teste de hipóteses é um procedimento estatístico baseado na análise de uma amostra, através da teoria de probabilidades, usado para avaliar determinados parâmetros que são desconhecidos numa população. Um teste de hipóteses pode ser paramétrico ou não-paramétrico. Testes paramétricos são baseados em parâmetros da amostra, por exemplo, média e desvio padrão. O uso tanto dos testes paramétricos como dos não-paramétricos está condicionado à dimensão da amostra e à respectiva distribuição da variável em estudo (FISHER, 1925).

3.14.1. Análise de variância e seus pressupostos

Para realizar uma ANOVA é preciso pressupor que os erros (parte do dado não explicada por fatores fixos e conhecidos) sejam variáveis aleatórias independentes com distribuição normal de média zero, apresentem variância constante e observações independentes, ou seja, os dados devem apresentar distribuição normal, homocedasticidade e

independência. A idéia é comparar a variação devido aos fatores estudados, incluindo tratamentos, com a variação devido ao acaso. Nos resultados destes testes, se o valor do F para tratamento calculado, que é dado pelo quociente entre o quadrado médio de tratamentos e o quadrado médio do resíduo, for maior ou igual ao do F tabelado poderá ser afirmado que existe diferença significativa entre pelo menos duas amostras testadas, o que significa que o valor- p é menor a 0,05. Havendo diferença entre amostras ou tratamentos, é possível determinar quais tratamentos diferem entre si a partir de um teste de comparação de médias dois a dois. Recomenda-se na maioria das vezes o teste de Tukey, em que a diferença entre as médias aritméticas das amostras são comparadas com o valor crítico (d.m.s.) (VIEIRA, 1999).

Para saber se é razoável pressupor que os erros têm distribuição normal de média zero, o pesquisador pode fazer uma análise dos resíduos, que é dado pelo número de repetições (total de amostras multiplicado pelo total de provadores) menos o total de amostras. Para calcular os resíduos podem ser utilizados programas estatísticos, e os resultados devem ser analisados graficamente. Se a pressuposição de que os erros têm distribuição normal de média zero for razoável, os resíduos terão, no gráfico, aparência que lembra uma distribuição normal centrada no zero, ou seja, a configuração típica da distribuição normal. Se a aparência da distribuição dos resíduos for muito diferente da aparência da distribuição normal, é necessário procurar uma explicação para isso. Quando os cálculos resultam em resíduos muito grandes, que indicam valores discrepantes é preciso verificar se esses valores não estão de alguma forma, errados, além de buscar um motivo para rejeitar o valor discrepante, é preciso buscar a causa da discrepância. Contudo, embora seja extremamente útil, esta análise é gráfica, o que significa que não permite associar um nível de probabilidade à conclusão de que a distribuição dos erros não é normal (VIEIRA, 1999).

Testes estatísticos podem ser utilizados para testar a hipótese de que a distribuição dos erros é normal. Os testes mais conhecidos para testar a normalidade dos dados são o teste de χ^2 , o teste de Kolmogorov-Smirnov e o teste de Shapiro-Wilks. O teste F é considerado satisfatório, ou seja, pequenas transgressões à pressuposição de que os erros têm distribuição normal são usuais e não afetam, substancialmente, os resultados. A ANOVA apresenta resultados bons, mesmo que a distribuição dos erros seja apenas aproximadamente normal (VIEIRA, 1999). Adicionalmente, a distribuição da média amostral será

aproximadamente normal, qualquer que seja a distribuição da variável da população, quando o tamanho da amostra é suficientemente grande. Este resultado é conhecido pelo nome de Teorema Central do Limite.

Para realizar uma ANOVA é preciso pressupor que os erros são variáveis aleatórias com variância constante no modelo que estamos considerando, é preciso pressupor que as variâncias de tratamentos são iguais, ou seja, que existe homocedasticidade. Quando se compara dois histogramas, quando um tratamento é capaz de deslocar a curva sem mudar a sua configuração dizemos que existe homocedasticidade, pois o deslocamento está associado à diferença de médias e a configuração à variância. Porém, quando ocorre além do deslocamento, mudança na configuração, implica em que não apenas as médias dos tratamentos são diferentes, mas também as variâncias. Neste caso diz-se que existe heterocedasticidade (VIEIRA, 1999).

Para ilustrar esta propriedade, podemos observar a Figura 8, onde, como exemplo, foi considerado um grupo de alunos que praticam aula de digitação. Com o passar do tempo a quantidade de erros de digitação diminui acarretando em mudança na configuração da curva. A variância nesse caso também não é constante porque, como pode ser observado, ocorreram deslocamento e mudança na configuração da curva com o passar do tempo. Neste caso temos um exemplo de heterocedasticidade.

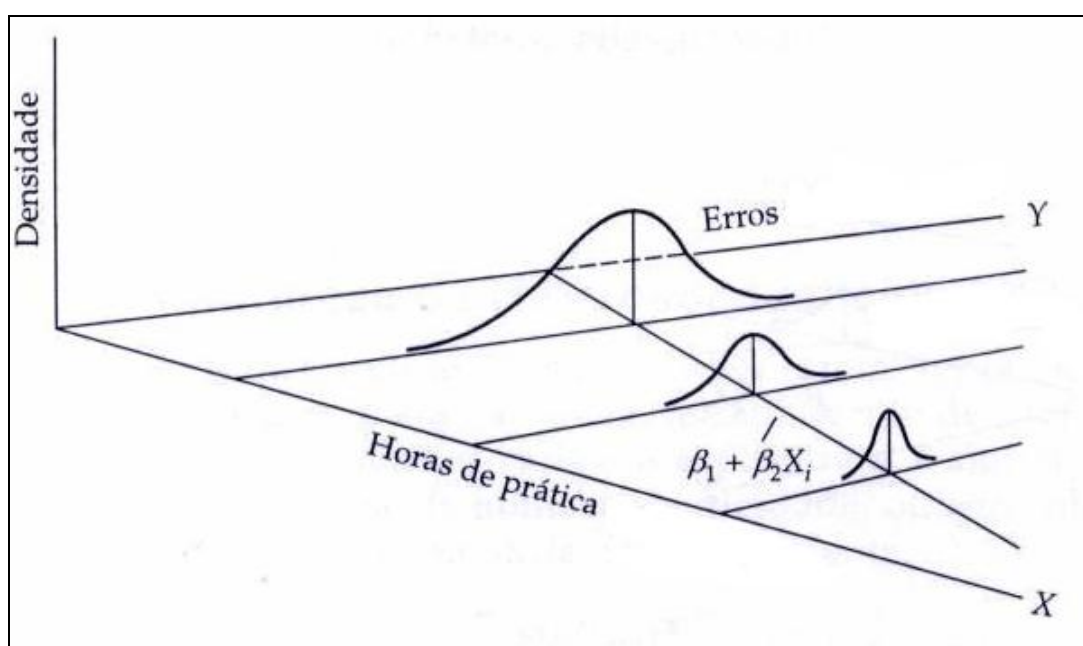


Figura 8. Exemplo de heterocedasticidade.

A realização de ANOVA pressupõe, que os erros são variáveis aleatórias independentes, ou seja, que um resultado obtido não está correlacionado com os outros. A ANOVA só deveria ser aplicada a um conjunto de observações se estiverem satisfeitas as pressuposições de independência, normalidade e homocedasticidade. Contudo, a ANOVA pode ser aplicada quando existe pequeno desvio das pressuposições básicas, mas nunca, porém, quando nenhuma delas está presente (VILLANUEVA; PETENATE; DA SILVA, 2000; MUNRO, 1997; GAY; MEAD, 1992; GIOVANNI; PANGBORN, 1983). Na prática, são poucos os estudos que satisfazem todas as pressuposições, quando é utilizada a ANOVA para tratamento dos resultados (VIEIRA, 1999).

Segundo Vieira (1999), para a utilização da ANOVA é importante saber que:

- a) a não-normalidade tem pouco efeito nas inferências sobre médias, isto é, nas conclusões obtidas da ANOVA e dos testes de comparação múltipla quando o modelo é de efeitos fixos. E tem efeito significativo nas inferências sobre as variâncias, isto é, no modelo de efeitos aleatórios, quando a curtose é diferente de zero;
- b) a heterocedasticidade tem pouco efeito nas inferências sobre médias, isto é, nas conclusões obtidas da ANOVA e dos testes de comparação múltipla desde que o modelo seja de efeitos fixos e o número de repetições seja constante, mas efeito significativo se o número de repetições varia;
- c) a não-independência, isto é, a correlação entre as observações pode ter efeito significativo nas inferências sobre médias.

3.14.2. Teste para duas amostras

O teste t foi proposto por Fisher que também propôs a expressão diferença mínima significativa, identificada pela sigla LSD (least significant difference) em programas de computação que realizam esse teste. Alguns autores referem-se ao teste t como o teste LSD de Fisher. Este teste é o mais utilizado para comparar duas médias. Nesse teste é estabelecido um nível de significância e é realizada a comparação do valor absoluto do “t”

calculado com o valor crítico dado em tabela com os mesmos graus de liberdade. Toda vez que o valor absoluto do “t” calculado for igual ou maior que o da tabela, rejeita-se a hipótese de que as médias são iguais (VIEIRA, 1999).

O teste t para duas amostras é adequado para situações em que as respostas aos dois tratamentos são variáveis quantitativas com distribuição normal com parâmetros (μ_1, σ) e (μ_2, σ), ou seja, é suposto que as variáveis estudadas têm distribuição normal com o mesmo desvio-padrão (SOARES; SIQUEIRA, 2002).

Muitos programas de computação fornecem a probabilidade de o valor de “t” ser, na distribuição teórica, maior que o valor obtido. Essa probabilidade é o valor-*p*, diferente do nível de significância do teste. Então, toda vez que o valor-*p* for menor que o nível de significância estabelecido, rejeita-se a hipótese de que as médias são iguais (SHIMAKURA, 2011).

O teste de Tukey é utilizado em teste de escala hedônica após a ANOVA, para avaliar quais amostras diferem entre si. As amostras testadas são avaliadas duas a duas nesse teste. No teste de Tukey duas médias são estatisticamente diferentes toda vez que o valor absoluto da diferença entre elas for igual ou maior do que a diferença mínima significativa (d. m. s) que é calculada nesse teste. Em programas de computação que realizam esse teste, a diferença mínima significativa é chamada de HDS, diferença honestamente significativa. O teste de Duncan também pode ser utilizado em testes de escala hedônica para comparar médias, porém não é muito utilizado (VIEIRA, 1999).

3.15. Testes não-paramétricos

A Estatística não-paramétrica é considerada, por muitos, uma área de entendimento difícil e, portanto, restrita a um grupo de estudiosos dessa matéria. Os testes não-paramétricos são baseados, fundamentalmente, na ordenação e na realocação de elementos em grupos. Dentre os testes não-paramétricos, podem ser destacados, por sua simplicidade, o teste dos sinais e o teste para duas amostras de Wilcoxon-Mann-Whitney (PONTES A.; PONTES JR.; PONTES L., 2009).

As técnicas dos testes não-paramétricos são, particularmente, adaptáveis aos dados das ciências do comportamento. A aplicação dessas técnicas não exige suposições quanto à distribuição da variável populacional independente dos parâmetros populacionais e de suas respectivas estimativas, sendo sua aplicação interessante nas análises de dados qualitativos (MUNRO, 1997). Assim, em teste de escala hedônica, se a variável populacional analisada não segue uma distribuição normal, deveria ser aplicado um teste não-paramétrico.

Segundo Callegari-Jacques (2003), os métodos não-paramétricos devem ser utilizados:

- 1) Quando não se conhece a distribuição dos dados na população é a mais apropriada. Além disso, são úteis quando essa distribuição é assimétrica e não se deseja realizar uma transformação dos dados, quando existe heterogeneidade nas variâncias ou quando na comparação entre tratamentos, a distribuição é gaussiana em alguns grupos e assimétrica em outros.
- 2) Quando a variável é medida em escala ordinal são os mais indicados.
- 3) Quando as exigências das técnicas clássicas não podem ser satisfeitas, os métodos não paramétricos são mais eficientes do que os testes paramétricos.

E não devem ser utilizados:

- 1) Para tratar dados que satisfazem as exigências dos testes clássicos, os métodos não-paramétricos apresentam uma eficiência menor. Isto é, para se detectar uma diferença real entre duas populações por um teste não-paramétrico, o tamanho amostral deve ser um pouco maior do que seria necessário em um teste clássico.
- 2) Para alguns autores, os testes não-paramétricos extraem menos informação do experimento porque são técnicas empregadas em dados mensurados em escalas não-quantitativas (ou dados quantitativos reduzidos para uma escala qualitativa ordenável).
- 3) Uma análise não-paramétrica pode constituir uma operação cansativa, porém, simples, se a quantidade de dados for grande. Porém se for utilizado programa estatístico, não há problema.

Os testes não-paramétricos têm sido utilizados em substituição aos testes paramétricos usuais quando as pressuposições do modelo não se verificam, ou seja, quando os dados provenientes de um experimento não possuem normalidade ou homogeneidade de variâncias. Esses testes são utilizados também em situações em que a aplicação de testes paramétricos não é possível ou torna-se muito complicada pela falta de informações a respeito da forma da distribuição da população, ou pela dificuldade de obtenção de estimativas confiáveis dos parâmetros populacionais (PONTES; CORRENTES, 2001; VILLANUEVA; PETENATE; DA SILVA, 2000).

Kendall e Stuart (1952) utilizaram os termos distribuição livre (*distribution-free*) e não-paramétrico de formas distintas, considerando que não-paramétrico é a descrição do problema e distribuição livre é o método utilizado para resolver o problema. Entretanto, atualmente, os termos não-paramétricos ou distribuição livre para designar um teste são usados indistintamente.

Os testes não-paramétricos ou testes de distribuição livre constituem uma alternativa para os casos em que não é possível utilizar a estatística paramétrica. O termo distribuição livre é vulgarmente utilizado para indicar que os métodos são aplicáveis independentemente da forma da distribuição, ou que são válidos para um ou mais largo espectro de distribuições (MUNRO, 1997).

Estes métodos são, em geral, fáceis de aplicar, pois podem ser usados quando as hipóteses exigidas por outras técnicas não são satisfeitas (MUNRO, 1997). Apesar de haver certas suposições básicas associadas à maioria dos testes não-paramétricos, essas suposições são em menor número e mais fracas do que as associadas aos testes paramétricos. Servem para pequenas amostras e, além disso, a maior parte dos testes não-paramétricos são aplicados a dados medidos em escala ordinal e, alguns, a dados em escala nominal (VIEIRA, 1999).

Quando um pesquisador utiliza testes não-paramétricos é suposto que a distribuição de seus dados experimentais não seja normal ou que ele não tenha elementos suficientes para poder afirmar que seja. Na dúvida quanto a essa informação é recomendado que o pesquisador faça a opção pelos testes não-paramétricos. Nesse caso, os cálculos de desvio ou

erro padrão não podem ser realizados, embora possa realizar o cálculo de média (CAMPOS, 2000).

Os testes não-paramétricos são boas opções para situações em que ocorrem violações dos pressupostos básicos necessários para a aplicação de um teste paramétrico. Pode ser citado como exemplo, o teste de diferença entre dois grupos quando a distribuição subjacente é assimétrica ou os dados foram coletados em uma escala ordinal, ou quando a distribuição da variável de interesse não é conhecida ou tem comportamento não normal (VILLANUEVA; PETENATE; DA SILVA, 2000; SOARES; SIQUEIRA, 2002). Dados assimétricos como os encontrados em teste de escala hedônica, podem ser analisados com estes tipos de testes.

Os métodos não-paramétricos são relacionados ao desenvolvimento de procedimentos de inferência estatística que não fazem qualquer suposição explícita sobre a forma da distribuição dos dados, tendo, portanto, menores exigências. Além disso, os procedimentos da estatística não-paramétrica em geral são dados sob ótica do desenvolvimento inicial da distribuição exata, o que leva a um maior conhecimento das vantagens e desvantagens do teste que está sendo utilizado. A base para os testes não-paramétricos está na ordenação (*ranks*) dos dados e não no seu valor intrínseco, e na aleatorização, onde se consideram todas as possíveis permutações (rearranjos) dos dados. Se, por um lado, ocorra perda de precisão na troca dos valores da variável por seus respectivos postos, ocorre ganho em eficiência e facilidade no entendimento dos resultados. Como os testes não-paramétricos têm como base as estatísticas de ordem, isso significa que, dada uma amostra aleatória, as estatísticas de ordem dessa amostra aleatória são os valores ordenados e, portanto, dependentes, o que torna as deduções muito mais difíceis (PONTES; CORRETE, 2001).

Um dos grandes problemas ao se trabalhar com métodos não-paramétricos é a obtenção dos níveis ou probabilidade de significância exata para os testes utilizados. A construção de tabelas exatas abrangentes é trabalhosa, exigindo atenção para cada caso em particular. Esta lacuna nas tabelas é sentida, especialmente, quando o número de repetições dos tratamentos não é igual. A realização de comparações múltiplas aumenta o grau de dificuldade uma vez que as tabelas são ainda menos abrangentes (PONTES; CORRENTES, 2001).

3.15.1. Testes não-paramétricos para amostras independentes

3.15.1.1. Teste de Kruskal-Wallis

O teste de Kruskal-Wallis consiste em teste extremamente útil para decidir se k amostras ($k > 2$) independentes provêm de populações com médias iguais. Esse teste só deve ser aplicado se a amostra for pequena e/ou as pressuposições, exigidas para proceder à ANOVA, estiverem seriamente comprometidas. Como o teste de Mann-Whitney, esse teste também condiciona que a variável em análise seja medida em escala ordinal ou numérica (MUNRO, 1997).

O teste não-paramétrico de Kruskal-Wallis foi introduzido por estes autores em 1952, como um competidor ou um substituto do teste F da ANOVA segundo um delineamento inteiramente casualizado (CAMPOS, 2000). Assim como a maioria dos testes não-paramétricos, este também dispõe as respostas dos tratamentos que serão comparados na forma de postos. Quanto maior for a diferença entre a soma dos postos, maior será a evidência de que exista diferença entre os mesmos. Mesmo não precisando da exigência de normalidade ou de outra distribuição qualquer para as populações estudadas, o teste exige que a distribuição dos erros seja a mesma para todos os níveis. Como alternativa não-paramétrica ao teste F da ANOVA, segundo o delineamento em blocos casualizados, existe o teste de Friedman (REIS; RIBEIRO, 2007).

Uma das principais dificuldades na utilização da Estatística Não-Paramétrica é a obtenção de resultados confiáveis. As tabelas disponíveis para o Teste de Kruskal-Wallis e para as comparações múltiplas são pouco abrangentes, em especial nos casos de número diferente de repetições entre tratamentos, fazendo com que o pesquisador seja obrigado a recorrer a aproximações. Estas aproximações diferem dependendo do autor a ser consultado, podendo levar a resultados contraditórios. Além disso, tais tabelas não consideram empates, mesmo no caso de pequenas amostras (PONTES; CORRENTE, 2001).

Nas tabelas para o teste de Kruskal-Wallis os empates não são contemplados e isto leva à utilização de aproximações em delineamentos com um número razoavelmente pequeno de tratamentos e poucas obtenções, podendo resultar em erros nas conclusões (PONTES; CORRENTES, 2001).

3.15.1.2. Teste da Mediana e Teste de Mann-Whitney

O teste da mediana verifica a probabilidade de grupos independentes proverem de populações com a mesma mediana. O teste da mediana é particularmente útil quando existem dados censurados que são aqueles que ficam além dos limites estabelecidos para coleta, embora não se saiba exatamente quais são esses valores. Podemos citar como exemplo os experimentos com animais onde alguma condição específica demora a aparecer ou desaparecer. Se nada acontece a alguns animais até o final do experimento esses dados são censurados (tempo de sobrevivência; limite mínimo em aparelhos de medição). Para esse teste, a variável em análise também deve ser medida em escala ordinal ou numérica (MUNRO, 1997).

O Teste de Wicoxon Mann-Whitney ou Teste de Mann-Whitney é utilizado para testar se duas amostras independentes foram retiradas de populações com médias iguais. Esse teste é, portanto, uma alternativa para o teste “t” para amostras independentes quando a amostra for pequena e/ou as pressuposições, exigidas pelo teste “t”, estiverem seriamente comprometidas. A única exigência do teste de Mann-Whitney é a de que as observações sejam medidas em escala ordinal ou numérica. Nesse teste os dados dos dois grupos são colocados em ordem crescente, e às observações empatadas são atribuídas a média dos postos correspondentes (MUNRO, 1997).

3.15.2. Teste não-paramétricos para amostras dependentes

3.15.2.1. Teste dos Sinais, de Wilcoxon e de Friedman

O teste dos Sinais é utilizado para análise de amostras dependentes, sendo, portanto, uma alternativa para o teste “t”. É aplicado em situações em que o pesquisador deseja determinar se duas condições são diferentes. O teste dos Sinais tem pouco poder, pois usa como informação apenas o sinal das diferenças entre pares. A única pressuposição exigida pelo teste dos Sinais é a de que a distribuição da variável seja contínua. A lógica do teste é que as condições podem ser consideradas iguais quando as quantidades de valores positivos e negativos forem aproximadamente iguais. Este teste é de fácil aplicação e praticamente não exige pressuposições (MUNRO, 1997).

O teste de Wilcoxon consiste em uma extensão do teste dos sinais. Esse teste leva em consideração a magnitude da diferença para cada par e exige que a variável em análise seja medida em escala ordinal ou numérica e a diferença entre duas observações, realizadas no mesmo par, também possa ser ordenada (MUNRO, 1997).

Quando os dados de k amostras dependentes se apresentam pelo menos em escala ordinal, a prova de Friedman é útil para comprovar a hipótese de nulidade, de que as k amostras tenham sido extraídas da mesma população. É, portanto, uma alternativa para a ANOVA. No entanto, o teste de Friedman só deve ser aplicado quando se têm poucos dados e/ou as pressuposições exigidas pela ANOVA, estiverem seriamente comprometidas. Para o teste de Friedman a variável em análise deve ser medida em escala ordinal ou numérica (MUNRO, 1997).

No teste de Friedman, o número de casos é o mesmo em cada amostra. O mesmo grupo de indivíduos é estudado sob cada uma das k condições e o objetivo é comparar grupos de elementos relacionados entre si. Por esta razão, todos os grupos têm o mesmo número de casos (MUNRO, 1997). Este teste é utilizado em testes sensoriais de ordenação, como por exemplo, no teste de ordenação da preferência onde se deseja conhecer a ordem de preferência de um produto.

Campos (2000) elaborou uma tabela com um resumo de orientações sobre a utilização de testes paramétricos e não-paramétricos relacionadas ao número de amostra e aos dados, se são independentes ou não. Estas orientações seguem na Tabela 1 abaixo.

Tabela 1. Orientações para utilização de testes paramétricos e não paramétricos.

Testes Estatísticos			
Paramétricos		Não-Paramétricos	
Independentes	Vinculados	Independentes	Vinculados
2 amostras	2 amostras	2 amostras	2 amostras
Teste <i>t</i> (Student)	Teste <i>t</i> (Student)	Mann-Whitney T. da Mediana χ^2 (2 x 2) Proporções Exato (Fisher)	Wilcoxon T. dos sinais Mac Nemar Binomial
Mais de duas	Mais de duas	Mais de duas	Mais de duas
Análise de variância	Análise de variância	Kruskal-Wallis Mediana (m x n) χ^2 (m x n) Nemenyi	Cochran Friedman

Fonte: CAMPOS, 2000

3.16. Noções de simulação

Uma simulação de dados pode ser entendida como a realização de modelos probabilísticos (Binomial, Normal). Nesse sentido, os valores simulados podem ser considerados como uma amostra. Os números pseudo-aleatórios (NPA) são obtidos por meio de técnicas que usam relações matemáticas recursivas determinísticas. Logo, um NPA gerado numa interação dependerá do número gerado na interação anterior e, portanto, não será realmente aleatório, originando o nome pseudo-aleatório (BUSSAB; MORETTIN, 2005).

Se for especificado um espaço amostral e a probabilidade associada aos pontos desse espaço, o modelo probabilístico ficará completamente determinado e será possível, então, calcular a probabilidade de qualquer evento aleatório de interesse. Entretanto, muitas vezes, mesmo construindo um modelo probabilístico, certas questões não podem ser resolvidas analiticamente e haverá a necessidade de recorrer a estudos de simulação para obter aproximações de quantidades de interesse (BUSSAB; MORETTIN, 2005).

A simulação de variáveis aleatórias deu origem aos chamados métodos Monte Carlo (MMC). Esses métodos apareceram durante a Segunda Guerra Mundial, em pesquisas relacionadas à difusão aleatória de neutrons num material radioativo. Os MMC supõem que o pesquisador dispõe de um gerador de números aleatórios equiprováveis. Um número aleatório (NA) representa o valor de uma variável aleatória uniformemente distribuída no intervalo (0,1). Originalmente, esses números aleatórios eram gerados manualmente ou mecanicamente, usando dados ou roletas. Modernamente, são usados computadores para gerar números que na realidade são pseudo-aleatórios. O nome Monte Carlo está relacionado com a cidade de mesmo nome, no Principado de Mônaco, principalmente devido à roleta, que é um mecanismo simples para gerar números aleatórios (BUSSAB; MORETTIN, 2005).

Nas simulações de dados, as estimativas de significância (α) e de poder do teste (β) podem ser obtidas como por exemplo com a utilização do programa gratuito R, com o programa SAS e com o programa S-Plus. Nesses programas de computação, é possível programar funções para gerar dados de acordo com os parâmetros de interesse, para posteriormente serem aplicados testes estatísticos adequados. De modo bastante amplo, estudos de simulação tentam reproduzir num ambiente controlado o que se passa com um problema real. Sendo que nesse estudo, a solução de um problema real, consistiu na simulação de variáveis aleatórias, que representam dados obtidos em testes de escala hedônica de nove pontos. Sendo que, as principais vantagens em utilizar simulações para obter essas informações são: a rapidez com que os resultados são obtidos, a possibilidade de trabalhar com um grande número de amostras de diversos tamanhos e o baixo custo das análises, pois não existe gasto com amostras.

4. MATERIAL E MÉTODOS

4.1. Amostras

Cinco amostras de refresco de maracujá e cinco de refresco de manga foram preparadas a partir de sucos concentrados engarrafados adquirido no mercado local. Os refrescos foram preparados seguindo as orientações do fabricante quanto às diluições e foram tratados com diferentes concentrações de açúcar que podem ser visualizadas na Tabela 2. Após o preparo, as amostras foram mantidas sob refrigeração até o momento da realização dos testes. As amostras de refrescos de maracujá e manga que apresentaram diferenças significativas, quanto ao sabor doce, no teste de ordenação, foram selecionadas para o teste de escala hedônica e avaliadas quanto ao sabor global.

Tabela 2. Ingredientes e quantidades para a preparação dos refrescos de maracujá e manga.

Amostra	Açúcar		Refresco de maracujá		Refresco de manga	
	%	g	Água (mL)	Suco (mL)	Água (mL)	Suco (mL)
A	0	0	900	100	750	250
B	2,5	25	900	100	750	250
C	5,0	50	900	100	750	250
D	7,5	75	900	100	750	250
E	10	100	900	100	750	250

4.1.1. Teste de ordenação

Preliminarmente, foi realizado teste de ordenação, com a finalidade de determinar diferenças entre a intensidade do sabor doce dos refrescos de maracujá e manga e selecionar as concentrações de açúcar para a realização do teste de escala hedônica. Esse teste foi realizado por trinta provadores, alunos, professores e funcionários da Universidade Federal

Fluminense, segundo a metodologia descrita por Meilgaard; Civille; Carr (1991), em cabines individuais, sob luz branca, no Laboratório de Análise Sensorial da Faculdade de Nutrição Emília de Jesus Ferreiro, da Universidade Federal Fluminense, no Rio de Janeiro – RJ. No teste de ordenação do refresco de maracujá os provadores voluntários (27 mulheres e 3 homens) apresentavam idade entre 18 e 50 anos. No teste de ordenação do refresco de manga os provadores (25 mulheres e 5 homens) também apresentavam idade entre 18 e 50 anos.

As amostras foram servidas em diferentes sessões para não ultrapassar o número de seis amostras analisadas por sessão. Os sucos foram apresentados aos provadores em copos descartáveis brancos, com capacidade para 50mL, contendo aproximadamente 30mL de amostra, codificadas com três dígitos aleatórios. Cada provador recebeu cinco amostras de refresco de maracujá em temperatura de refrigeração, aproximadamente 10°C, acompanhadas da ficha de aplicação do teste de ordenação (Figura 9) e de um copo descartável contendo água em temperatura ambiente para a limpeza da cavidade oral entre as degustações das amostras, com o intuito de anular o *flavor* residual. O provador foi orientado a realizar uma ordenação preliminar, e posteriormente a provar novamente as amostras a fim de verificar se a ordenação estava correta.

Nome:	Data:			
.....				
Idade:	Gênero:			
<p>Você está recebendo 5 amostras codificadas de refresco. Por favor, prove as amostras da esquerda para a direita e ordene em ordem crescente em relação ao sabor doce.</p>				
_____	_____	_____	_____	_____
menos doce				mais doce
Comentários:				

Figura 9. Ficha de aplicação do teste de ordenação.

4.1.2. Tratamento estatístico do teste de ordenação

Para a análise dos resultados foi utilizado o teste de Friedman, que utiliza a Tabela de Newell e MacFarlane (NEWELL E MACFARLANE, 1987) (ANEXO 1). Nesta tabela é observado o valor da diferença crítica entre os totais de ordenação, de acordo com número de amostras, provadores e ao nível de significância observado, que neste caso foi de 5%, para verificar se existem diferenças entre as amostras.

4.2. Teste de escala hedônica

O teste de aceitabilidade foi realizado com três amostras de refresco de maracujá e três de refresco manga utilizando escala hedônica estruturada verbal crescente com nove categorias, segundo a metodologia descrita por Meilgaard; Civille; Carr (1991), em cabines individuais, sob luz branca, no Laboratório de Análise Sensorial da Faculdade de Nutrição Emília de Jesus Ferreiro, da Universidade Federal Fluminense, no Rio de Janeiro – RJ, com 100 consumidores voluntários, alunos, professores e funcionários da Universidade Federal Fluminense. No teste de escala hedônica do refresco de maracujá os provadores (90 mulheres e 10 homens) apresentavam idade entre 18 e 59 anos. No teste de escala hedônica do refresco de manga os provadores (91 mulheres e 9 homens) apresentavam idade entre 18 e 61 anos.

As amostras dos refrescos foram apresentadas aos provadores em copos descartáveis brancos, com capacidade para 50mL, contendo aproximadamente 30mL de amostra, codificadas com três dígitos aleatórios e foram servidas monadicamente aos provadores com base em um delineamento de blocos balanceados com relação à ordem de apresentação das amostras. Inicialmente, foram realizados os testes com as amostras do refresco de maracujá e, em seguida, com as amostras do refresco de manga. O provador recebeu uma amostra de refresco por vez, em temperatura de refrigeração, aproximadamente 10°C, acompanhada da ficha de aplicação do teste de escala hedônica, que pode ser observada na Figura 10 e de um copo descartável contendo água em temperatura ambiente para a limpeza da cavidade oral entre as degustações das amostras, com o intuito de anular o *flavor* residual. Os provadores foram instruídos pelo coordenador do teste a provar as amostras codificadas e assinalar com um X, na ficha de aplicação do teste de escala hedônica, a nota que representasse o quanto ele gostou ou desgostou da mesma quanto ao sabor global.

Nome:	Data:
Idade:	Gênero: F <input type="checkbox"/> M <input type="checkbox"/>
Código da amostra: _____	
<p>Prove a amostra de refresco e assinale com X a nota que descreve o quanto gostou ou desgostou do sabor global.</p>	
<ul style="list-style-type: none"> 1 - desgostei extremamente 2 - desgostei muito 3 - desgostei moderadamente 4 - desgostei ligeiramente 5 - nem gostei/nem desgostei 6 - gostei ligeiramente 7 - gostei moderadamente 8 - gostei muito 9 - gostei extremamente 	
Comentários:.....	

Figura 10. Ficha de aplicação do teste de escala hedônica.

4.2.1. Tratamento estatístico do teste de escala hedônica

Para análise dos resultados foi utilizado o teste de Kolmogorov-Smirnov, que avalia a normalidade dos dados; a ANOVA, tendo como fontes de variação o provador, o gênero, a idade, o tratamento em relação à concentração de açúcar e as variedades dos refrescos; e o teste de Tukey para comparação de médias, duas a duas. Foram utilizados, também, os testes não-paramétricos de Wilcoxon Mann-Whitney e Kruskal-Wallis com a finalidade de comparar os resultados desses testes com os resultados da ANOVA. O programa estatístico utilizado foi o S-Plus versão 8.0 ($p < 0,05$).

4.3. Simulação de dados

Foram realizadas duas simulações de dados segundo a teoria proposta por Bussab; Morettin (2005), utilizando o programa estatístico S-Plus versão 8.0. Nessas simulações foram gerados escores de equipes de provadores de testes de escala hedônica de nove pontos.

4.3.1. Simulação 1

Para caracterizar as amostras da Simulação 1, inicialmente, foram descritos três grupos etários, a saber, 14 a 18 anos, 19 a 30 anos e 31 a 50 anos, utilizando como referência o Institute of Medicine, (2005) e foi pesquisada a proporção aproximada de cada um desses grupos na população brasileira. Foi pesquisada também, a proporção aproximada dos gêneros masculino e feminino, dentro de cada grupo, descrita segundo o Instituto Brasileiro de Geografia e Estatística (IBGE, 2000). Essas proporções definidas podem ser observadas na Tabela 3. Esses grupos etários foram escolhidos porque estão dentro do grupo recomendado como provadores que é de dezoito a cinquenta anos (MEILGAARD; CIVILLE; CARR, 1991).

Tabela 3. Grupos etários utilizados na Simulação 1, sua classificação e proporção aproximada na população brasileira.

Grupo ¹ etário (anos)	Classificação	Proporção na população brasileira (%) ²	Gênero (%) ²	
			Masculino	Feminino
14-18	Adolescente	19	50	50
19-30	Adulto jovem	35	49	51
31-50	Adulto de meia idade	46	48	52

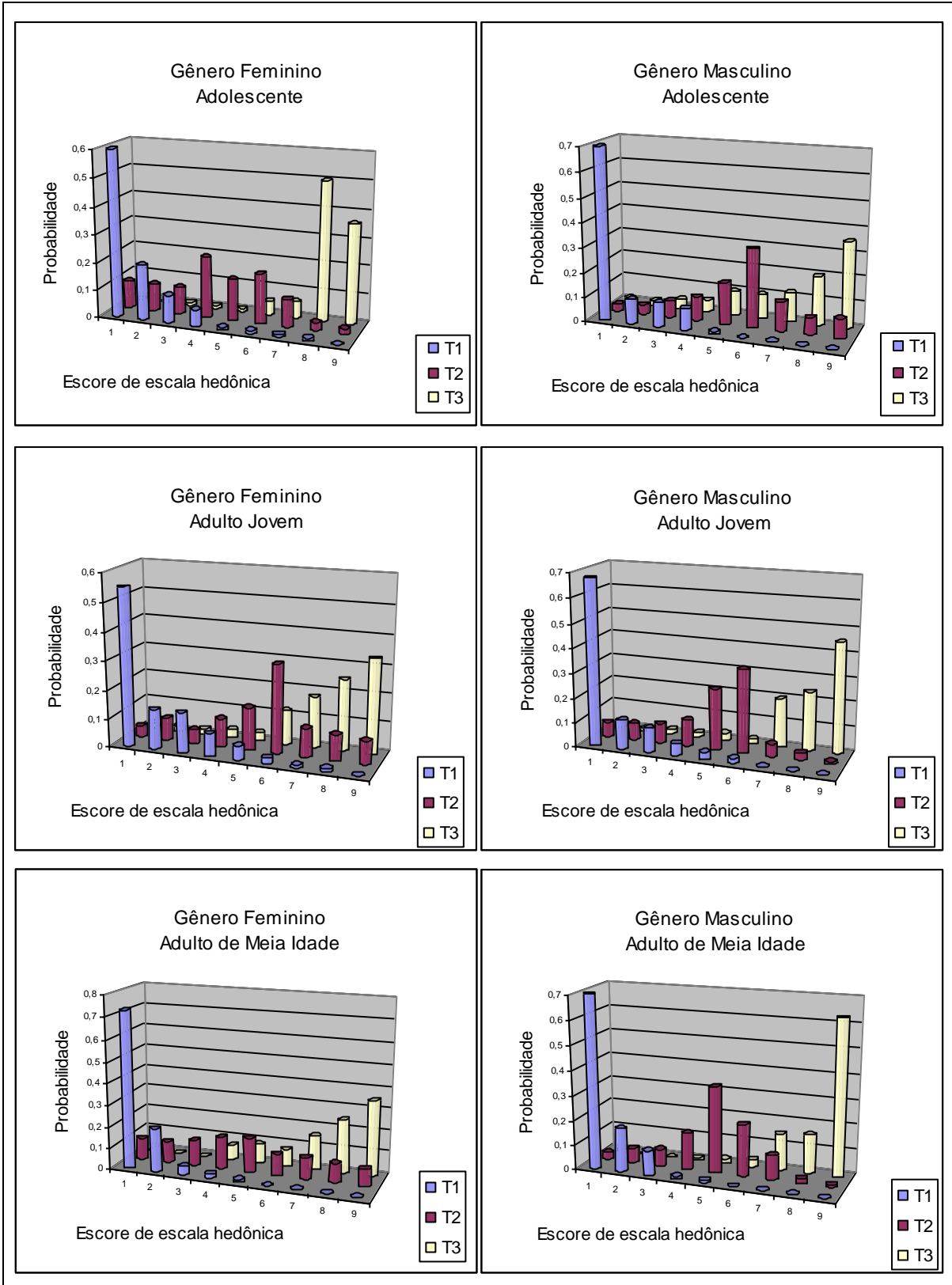
Fonte: ¹IOM, 2005; ²IBGE, 2000

Na segunda etapa foram definidos os tratamentos aplicados, ou seja, foram definidas as concentrações de açúcar que seriam simuladas, tendo sido considerado tratamento 1 (T1) a concentração de 0% de açúcar, tratamento 2 (T2) 5% de açúcar e tratamento 3 (T3) 10% de açúcar. Foram definidas, também, as variedades dos refrescos maracujá e manga. Esses três tratamentos e as variedades dos refrescos foram escolhidos para essa simulação, pois foram os mesmos das análises sensoriais realizadas em laboratório.

Na terceira e última etapa foram definidos os escores dos termos hedônicos, segundo a escala hedônica de nove pontos, resultando em 1 (um) para a categoria desgostei extremamente, 2 (dois) para desgostei muito, 3 (três) para desgostei moderadamente, 4 (quatro) para desgostei ligeiramente, 5 (cinco) para nem gostei/nem desgostei, 6 (seis) para gostei ligeiramente, 7 (sete) para gostei moderadamente, 8 (oito) para gostei muito, e 9

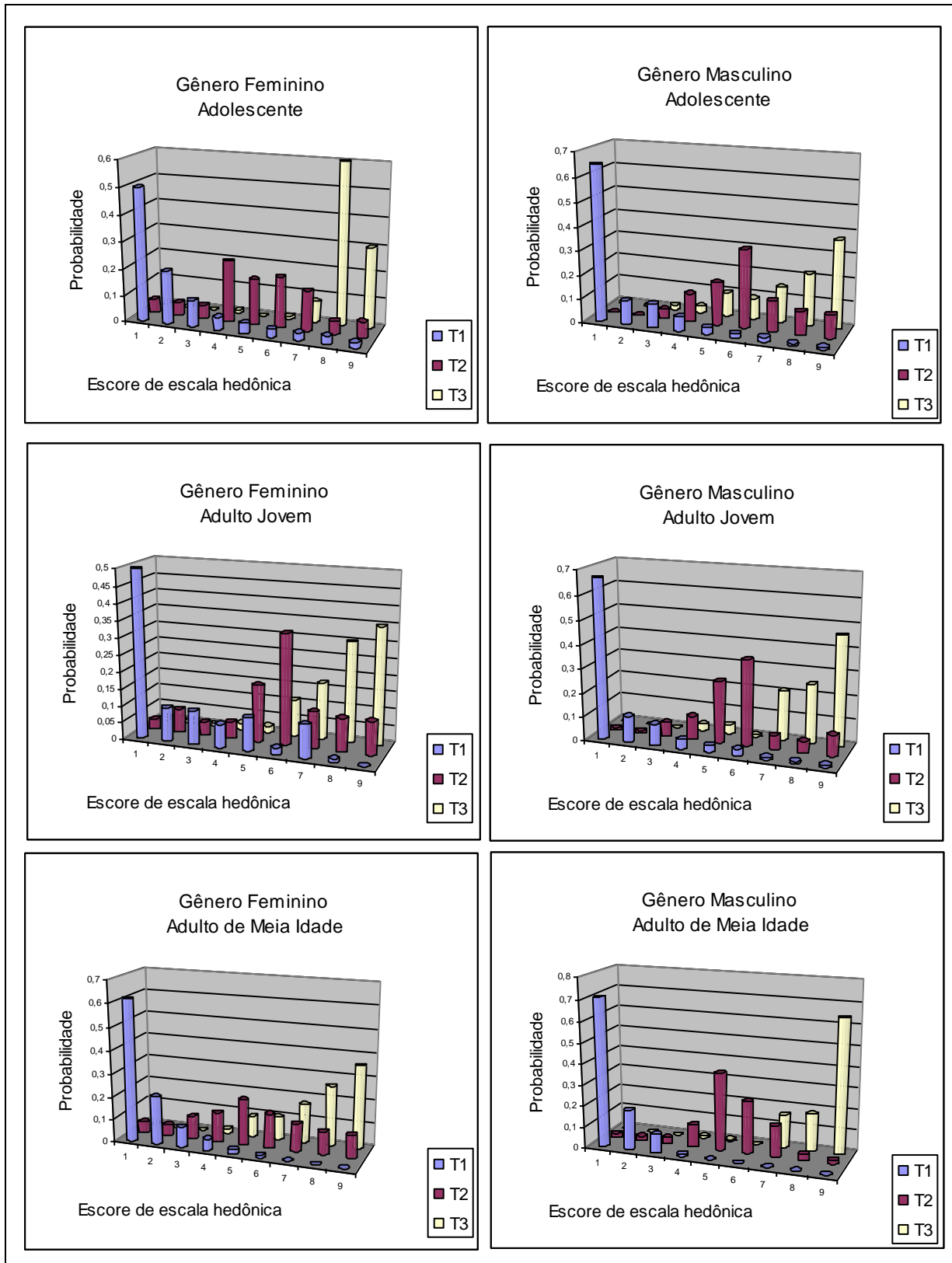
(nove) para gostei extremamente. A probabilidade de ocorrência de cada um dos escores de teste de escala hedônica, foi definida considerando que os refrescos de maracujá e manga com maior concentração de açúcar teriam, em média, maior aceitabilidade independente do gênero e do grupo etário (CLARK, 1998). Foram definidas, aleatoriamente, as probabilidades de ocorrência dos escores de aceitabilidade de refrescos, por indivíduos do gênero masculino e feminino, e para os diferentes grupos etários.

Após definidas as probabilidades de ocorrência dos escores das equipes simuladas considerando os parâmetros gênero, grupo etário, variedade do refresco e tratamento foi elaborada uma rotina para simular dados utilizando o programa estatístico S-Plus versão 8.0 segundo a teoria proposta por Bussab; Morettin (2005) (Apêndice 1). Nessa simulação foram gerados dados de equipes de provadores de testes de escala hedônica de nove pontos com $n = 5$, $n = 10$, $n = 30$, $n = 50$, $n = 100$ e $n = 1000$ (para cada n foram simuladas mil repetições). As distribuições assimétricas das probabilidades dos escores para essa simulação podem ser observadas nas Figuras 11 e 12.



T1, T2 e T3: tratamentos dos refrescos simulados com respectivamente 0%, 5% e 10% de açúcar.

Figura 11. Gráficos das probabilidades dos escores dos diferentes tratamentos simulados de teste de escala hedônica de nove pontos do refresco de maracujá de acordo com o gênero e grupo etário da Simulação 1.



T1, T2 e T3: tratamentos dos refrescos simulados com respectivamente 0%, 5% e 10% de açúcar.
 Figura 12. Gráficos das probabilidades dos escores dos diferentes tratamentos simulados de teste de escala hedônica de nove pontos do refresco de manga de acordo com o gênero e grupo etário da Simulação 1.

4.3.2. Tratamento estatístico do teste de escala hedônica da Simulação 1

Na Simulação 1 os dados simulados foram avaliados por ANOVA e pelo teste não-paramétrico de Kruskal-Wallis, tendo como fonte de variação tratamento em relação à concentração de açúcar e variedade do refresco. Ao final das análises desses testes, tendo como referência a rotina elaborada para essa simulação, o programa estatístico S-Plus versão 8.0 contou o número de vezes em que H_0 (hipótese de igualdade dos escores de T1, T2 e T3) foi aceita no teste ANOVA e no teste Kruskal-Wallis, e calculou a porcentagem de aceitação de H_0 para todas as amostras ($p < 0,05$).

4.3.3. Simulação 2

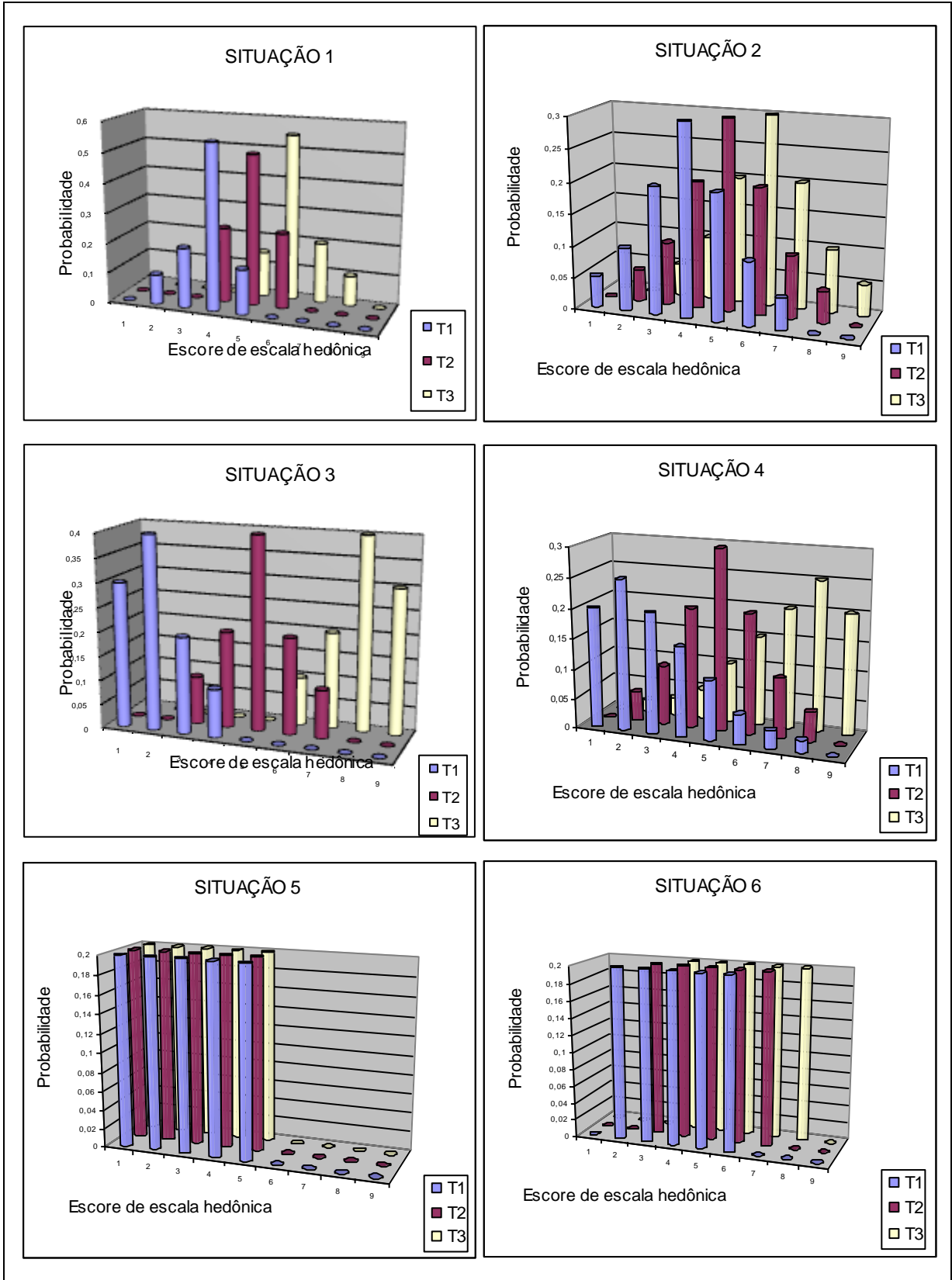
Na Simulação 2 foram gerados dados que simulassem resultados de teste de escala hedônica de nove pontos, com seis diferentes distribuições de dados. Essas seis situações foram definidas considerando que em testes sensoriais comumente são obtidos dados que apresentam distribuições assimétricas, como foi observado por diversos autores (Lim & Fujimaru, 2010; Villanueva & Da Silva, 2009; Villanueva, Petenate, & Da Silva, 2005; Villanueva, Petenate, & Da Silva, 2000; Wilkinson & Yuksel, 1997). Para caracterizar cada uma dessas seis distribuições, foram calculadas as médias e as variâncias de cada situação sendo que, a situação 1, apresentava médias próximas e variâncias pequenas; a situação 2, médias próximas e variâncias iguais e maiores que as da situação 1; a situação 3 apresentava médias mais afastadas e variâncias pequenas; a situação 4 apresentava médias mais afastadas e variâncias grandes, sendo que as variâncias dos tratamentos T1 e T3 eram iguais e um pouco maiores que a variância do tratamento T2; a situação 5 apresenta médias iguais e variâncias grandes e iguais entre os tratamentos; e a situação 6 apresentava médias próximas, e variâncias grandes e iguais entre os tratamentos. Os valores das médias e das variâncias dessas seis situações podem ser observadas na Tabela 4. Em seguida, foi determinada aleatoriamente a probabilidade de ocorrência dos escores, tendo como parâmetro tratamento dos refrescos utilizados no teste de escala hedônica em laboratório (0 %, 5% e 10% de açúcar). As distribuições assimétricas das probabilidades dos escores para essa simulação podem ser observadas na Fig. 13.

Após definidas as probabilidades de ocorrência dos escores das equipes simuladas tendo como parâmetro tratamento, foi elaborada uma rotina para simular dados utilizando o programa estatístico S-Plus versão 8.0, segundo a teoria proposta por Bussab; Morettin (2005) (Appendice 2). Nessa simulação foram gerados dados de equipes de provadores de testes de escala hedônica de nove pontos com $n = 5$, $n = 10$, $n = 30$, $n = 50$, $n = 100$ e $n = 1000$ (para cada n foram simuladas mil repetições).

Tabela 4. Médias (μ) e variâncias (σ^2) dos tratamentos da Simulation 2.

Situation	T1	T2	T3
	μ (σ^2)	μ (σ^2)	μ (σ^2)
1	3,7 (0,7)	5,0 (0,5)	6,25 (0,7)
2	4,0 (2,1)	5,0 (2,1)	6,0 (2,1)
3	2,1 (0,9)	5,0 (1,2)	7,9 (0,9)
4	3,1 (3,0)	5,0 (2,1)	6,93 (3,0)
5	3,0 (2,0)	3,0 (2,0)	3,0 (2,0)
6	4,0 (2,0)	5,0 (2,0)	6,0 (2,0)

Percentual de açúcar : T1 = 0%, T2 = 5%, T3 = 10%



T1, T2 e T3: tratamentos dos refrescos simulados com respectivamente 0%, 5% e 10% de açúcar.
 Figura 13. Gráficos das probabilidades dos escores dos diferentes tratamentos simulados de teste de escala hedônica de nove pontos das seis situações, com diferentes distribuições de dados criadas para a Simulação 2.

4.3.4. Tratamento estatístico do teste de escala hedônica da Simulação 2

Na Simulação 2 os dados simulados foram avaliados por ANOVA e pelo teste não-paramétrico de Kruskal-Wallis. Ao final das análises desses testes, tendo como referência a rotina elaborada para essa simulação, o programa estatístico S-Plus versão 8.0 contou o número de vezes em que H_0 foi aceita no teste ANOVA e no teste Kruskal-Wallis, calculou a porcentagem de aceitação de H_0 para todas as amostras ($p < 0,05$).

5. RESULTADOS E DISCUSSÃO

5.1. Análise estatística do teste de ordenação

Para análise dos dados do teste de ordenação dos refrescos de maracujá e manga, a diferença entre os valores de ordenação (Tabela 5) foi comparada com a diferença crítica observada na tabela de Newell e MacFarlane (NEWELL; MACFARLANE, 1987) (Anexo 1), que para trinta provadores e cinco amostras foi 34. Todas as amostras que diferiram entre si por um valor maior ou igual ao observado na tabela, foram consideradas significativamente diferentes ($p < 0,05$).

Tabela 5. Soma dos valores de ordenação das amostras dos refrescos de maracujá e manga, quanto ao sabor doce, utilizando escala crescente de ordenação.

Provador ($n = 30$)	Amostras				
	A	B	C	D	E
Soma dos valores de ordenação do refresco de maracujá	38 ^a	56 ^a	92 ^b	119 ^b	145 ^b
Soma dos valores de ordenação do refresco de manga	33 ^a	65 ^a	95 ^a	113 ^a	144 ^a

Percentual de açúcar: A = 0%, B = 2,5%, C = 5%, D = 7,5%, E = 10%

Na Tabela 6 pode ser observado que as amostras A e B, C e D, e D e E do refresco de maracujá não apresentaram diferença significativa ($p \geq 0,05$) e, na Tabela 7, podemos observar que o mesmo aconteceu para as amostras A e B, B e C, C e D, e D e E do refresco de manga. Desta forma, foi decidido eliminar as amostras B e D e realizar o teste de escala hedônica subsequente com as amostras A, C, e E, de ambos os refrescos, que apresentavam respectivamente 0%, 5% e 10% de açúcar e que passaram a ser denominadas T1, T2 e T3.

Tabela 6. Comparação da diferença entre os valores de ordenação das amostras do refresco de maracujá com a diferença crítica observada na tabela de New e MacFarlane¹ ao nível de 5% de significância.

Amostras		Diferença	
A	B	38-56	= 18
A	C	38-92	= 54*
A	D	38-119	= 81*
A	E	38-145	= 107*
B	C	56-92	= 36*
B	D	56-119	= 63*
B	E	56-145	= 89*
C	D	92-119	= 27
C	E	92-145	= 53*
D	E	119-145	= 26

Percentual de açúcar: A = 0%, B = 2,5%, C = 5%, D = 7,5% de açúcar, E = 10%

¹Newell; MacFarlane (1987)

Diferença crítica observada na tabela de New e MacFarlane = 34

* diferem significativamente entre si

Tabela 7. Comparação da diferença entre os valores de ordenação das amostras do refresco de manga com a diferença crítica observada na tabela de New e MacFarlane¹ ao nível de 5% de significância.

Amostras		Diferença	
A	B	33-65	= 32
A	C	33-95	= 62*
A	D	33-113	= 80*
A	E	33-144	= 111*
B	C	65-95	= 30
B	D	65-113	= 48*
B	E	65-144	= 79*
C	D	95-113	= 18
C	E	95-144	= 49*
D	E	113-144	= 31

Percentual de açúcar: A = 0%, B = 2,5%, C = 5%, D = 7,5% de açúcar, E = 10%

¹Newell; MacFarlane (1987)

Diferença crítica observada na tabela de New e MacFarlane = 34

* diferem significativamente entre si

5.2. Análise estatística do teste de escala hedônica

5.2.1. Teste de Kolmogov-Smirnov

O teste de Kolmogov-Smirnov, que avalia a normalidade dos dados, apresentou valor- $p = 0$, logo a hipótese nula foi rejeitada, o que significa que os resultados do teste de escala hedônica avaliados não apresentaram distribuição normal ($p < 0,05$). Mesmo com esse resultado a ANOVA foi realizada, com a finalidade de comparar esse teste com outros testes não-paramétricos.

5.2.2. Análise de variância

A ANOVA foi realizada tendo como fontes de variação, provador, gênero, idade, tratamento da amostra, variedade do refresco e resíduo. Na Tabela 8, podem ser observados que apenas os parâmetros provador e tratamento apresentaram diferença significativa ($p < 0,05$). Do ponto de vista do parâmetro provador, este resultado pode ser esperado para equipes não treinadas, desta forma, as análises prosseguiram considerando apenas as diferenças entre tratamentos.

Tabela 8. Análise de variância relativa aos dados de aceitabilidade dos refrescos de maracujá e manga.

Causas de variação	GL	SQ	QM	F	Valor- p
Provador	99	574,285	5,8009	1,6938	0,0001519*
Gênero	1	0,531	0,5307	0,1550	0,6940049
Idade	1	1,558	1,5578	0,4549	0,5003369
Tratamento	2	1196,911	598,4554	174,7487	0,0000000*
Refresco	1	12,292	12,2922	3,5893	0,0587346
Resíduo	495	1695,208	3,4247		

* diferem significativamente entre si

GL : Graus de liberdade, SQ: Soma dos Quadrados, QM: Quadrado Médio, F: fator

5.2.3. Teste de Tukey

Todos os tratamentos das amostras apresentaram diferença significativa entre si e o tratamento T3, com 10% de açúcar, foi o que obteve maior aceitabilidade (Tabela 9). Como o pós-teste não foi o objetivo desse trabalho, o teste de Tukey foi utilizado apenas para ilustrar os resultados dos testes de escala hedônica realizados em laboratório, não sendo considerado, então, na análise dos dados simulados.

Tabela 9. Média dos valores do teste de escala hedônica das amostras dos refrescos de maracujá e manga.

Refresco	T1	T2	T3
Maracujá ¹	3,32	5,78	7,11
Manga ¹	4,06	5,88	7,12

Percentual de açúcar: T1 = 0%, T2 = 5%, T3 = 10% de açúcar

¹ Todos os tratamentos diferiam entre si ao nível de 5% no teste de Tukey

5.2.4. Teste de Wilcoxon Mann-Whitney

Foram avaliados os escores dos testes de escala hedônica com relação ao sabor global, de indivíduos do gênero masculino e feminino, utilizando o teste não-paramétrico de Wilcoxon Mann-Whitney, por meio do programa estatístico S-Plus versão 8.0. O valor obtido foi $p = 0,1764$, ou seja, os escores entre os diferentes gêneros não apresentaram diferença significativa entre si ($p \geq 0,05$). E ao avaliar as variedades dos refrescos, maracujá e manga, utilizando esse mesmo teste, o valor obtido foi $p = 0,2332$, ou seja, os escores para ambos os refrescos também não apresentaram diferença significativa entre si ($p \geq 0,05$). Os resultados desses testes coincidiram com os encontrados na ANOVA, apresentados na Tabela 8.

5.2.5. Teste de Kruskal-Wallis

Foram avaliados os escores dos testes de escala hedônica, com relação ao sabor global, comparando T1, T2 e T3, utilizando o teste não-paramétrico de Kruskal-Wallis, por meio do programa estatístico S-Plus versão 8.0. O valor obtido foi $p = 0$, ou seja, os escores com relação aos diferentes tipos de tratamentos apresentaram diferença significativa entre si

($p < 0,05$). O resultado desse teste também coincidiu com o encontrado na ANOVA, apresentado na Tabela 8.

Nesta etapa, concluímos que para o teste de escala hedônica, realizado no laboratório, o teste de Kruskal-Wallis e o teste de Wilcoxon Mann-Whitney apresentaram os mesmos resultados obtidos na ANOVA. Este fato não justifica, então, a substituição da ANOVA por testes não-paramétricos, para o tratamento dos resultados do teste de escala hedônica, mesmo na ausência de normalidade.

5.3. Simulação de dados

5.3.1. Análise estatística do teste de escala hedônica da Simulação 1

Na Simulação 1, as porcentagens de aceitação de H_0 (hipótese de igualdade dos escores de T1, T2 e T3) nos testes Kruskal-Wallis e na ANOVA, podem ser observadas na Tabela 10. As equipes simuladas com $n = 5$, para ambos os refrescos, apresentaram diferenças entre os resultados, sendo que, o teste de Kruskal-Wallis rejeitou mais vezes H_0 do que o teste ANOVA, o que significa que o teste de Kruskal-Wallis conseguiu detectar as diferenças entre equipes muito pequenas proporcionalmente mais do que a ANOVA. Por outro lado, para as equipes simuladas com $n \geq 10$, de ambos os refrescos, ambos os testes detectaram as diferenças entre os tratamentos, rejeitando H_0 em 100% dos casos.

Mesmo que os testes tenham apresentado diferenças entre as equipes simuladas com $n = 5$, este fato não é suficiente para justificar a substituição da ANOVA pelo seu concorrente não-paramétrico, o teste de Kruskal-Wallis. Pois, teste de escala hedônica em laboratório, segundo Stone e Sidel, (1993), deve ser realizado com o mínimo de 50 provadores e a Simulação 1 mostrou que, para equipes com ≥ 10 , ambos os testes apresentaram os mesmos resultados, mesmo na ausência de normalidade.

Tabela 10. Porcentagens de aceitação da hipótese de igualdade dos escores de T1, T2 e T3

Tamanho da amostra (n)*	Refresco de Maracujá		Refresco de Manga	
	Kruskal-Wallis (Média Ho aceita) ($p < 0,05$)	ANOVA (Média Ho aceita) ($p < 0,05$)	Kruskal-Wallis (Média Ho aceita) ($p < 0,05$)	ANOVA (Média Ho aceita) ($p < 0,05$)
5	2,5	4,5	0,3	1,1
10	0	0	0	0
30	0	0	0	0
50	0	0	0	0
100	0	0	0	0
1000	0	0	0	0

(Ho) nos testes Kruskal-Wallis e na análise de variância (ANOVA), na Simulação 1.

* Para cada n foram realizadas mil repetições

5.3.2. Análise estatística do teste de escala hedônica da Simulação 2

Na situação 1 (Tabela 11), na qual os escores de T1, T2 e T3 apresentam médias próximas e variâncias pequenas, apenas as equipes simuladas com $n = 5$ apresentaram casos em que Ho foi aceita, tendo o teste de Kruskal-Wallis aceitado Ho mais vezes do que a ANOVA. Para as equipes simuladas com $n = 10$, ambos os testes rejeitaram Ho em 100% dos casos.

Na situation 2 (Tabela 11), na qual os escores de T1, T2 e T3 apresentaram médias próximas e variâncias iguais e maiores que as da situação 1, a porcentagem de aceitação de Ho para as equipes simuladas com $n = 5$, $n = 10$ e $n = 30$ foi maior no teste de Kruskal-Wallis do que na ANOVA. Para as equipes simuladas com $n \geq 50$, ambos os testes rejeitaram Ho em 100% dos casos.

Na situação 3 (Tabela 11), na qual os escores de T1, T2 e T3 apresentaram médias mais afastadas e variâncias pequenas, ambos os testes rejeitaram a Ho em 100% dos casos.

Na situação 4 (Tabela 11), na qual os escores de T1, T2 e T3 apresentaram médias afastadas e variâncias grandes, as variâncias dos escores de T1 e T3 foram iguais e um pouco maiores que a variância dos escores de T2, nas equipes simuladas com $n = 5$ e $n = 10$ o teste de Kruskal-Wallis apresentou uma porcentagem maior de aceitação de Ho do que a

ANOVA, apresentando menor poder de detecção de diferença entre os escores do que a ANOVA. Para as equipes simuladas com $n = 10$ as diferenças encontradas foram muito pequenas e para aquelas com $n \geq 30$ ambos os testes rejeitaram a H_0 em 100% dos casos.

Na situation 5 (Tabela 11), onde os escores apresentaram médias iguais e variâncias grandes e iguais, apenas as equipes simuladas com $n = 5$ apresentaram situações onde H_0 foi aceita, sendo que o teste de Kruskal-Wallis apresentou um percentual de aceitação de H_0 maior que a ANOVA. Logo, o teste de Kruskal-Wallis para essa situação apresentou menor poder de detecção de diferença entre os escores de T1, T2 e T3 do que a ANOVA. Para as equipes simuladas com $n \geq 10$, ambos os testes rejeitaram a H_0 em 100% dos casos.

Na situação 6 (Tabela 11), na qual os escores apresentaram médias próximas e variâncias grandes e iguais, as equipes simuladas com $n = 5$, $n = 10$ e $n = 30$ apresentaram situações onde H_0 foi aceita, sendo que o teste de Kruskal-Wallis apresentou uma porcentagem de aceitação de H_0 maior do que a ANOVA. Para essa situação pode ser concluído que o teste de Kruskal-Wallis apresentou menor poder de detecção de diferença entre os escores de T1, T2 e T3 do que a ANOVA. Para as equipes simuladas com $n = 30$ a ANOVA rejeitou H_0 em 100% dos casos e o teste de Kruskal-Wallis apresentou somente um pequeno percentual de aceitação de H_0 . Por outro lado, para as equipes simuladas com $n \geq 50$, número mínimo de provadores para testes de escala hedônica em laboratório (Stone & Sidel, 1993), ambos os testes rejeitaram H_0 em 100% dos casos, detectando diferença entre as amostras.

Tabela 11: Porcentagens de aceitação da hipótese de igualdade dos escores de T1, T2 e T3 (Ho) nos testes de Kruskal-Wallis e na análise de variância (ANOVA), na seis situações da Simulation 2.

	Tamanho da Amostra (n)*	Teste de Kruskal Wallis	ANOVA
Situação 1	5	1,4	0,2
μ (σ^2):	10	0	0
T1 3,7(0,7)	30	0	0
T2 5,0(0,5)	50	0	0
T3 6,2(0,7)	100	0	0
	1000	0	0
Situação 2	5	65,1	47,7
μ (σ^2):	10	31,5	14,8
T1 4,0(2,1)	30	0,3	0
T2 5,0(2,1)	50	0	0
T3 6,0(2,1)	100	0	0
	1000	0	0
Situação 3	5	0	0
μ (σ^2):	10	0	0
T1 2,1(0,9)	30	0	0
T2 5,0(1,2)	50	0	0
T3 7,9(0,9)	100	0	0
	1000	0	0
Situação 4	5	23,4	10,4
μ (σ^2):	10	1,6	0,6
T1 3,1(3,0)	30	0	0
T2 5,0(2,1)	50	0	0
T3 6,9(3,0)	100	0	0
	1000	0	0
Situação 5	5	5,6	0,7
μ (σ^2):	10	0	0
T1 3,0(2,0)	30	0	0
T2 3,0(2,0)	50	0	0
T3 3,0(2,0)	100	0	0
	1000	0	0
Situação 6	5	67,9	48,7
μ (σ^2):	10	31,9	13,6
T1 4,0(2,0)	30	0,4	0
T2 5,0(2,0)	50	0	0
T3 6,0(2,0)	100	0	0
	1000	0	0

* Para cada n foram realizadas mil repetições

6. CONCLUSÃO

Não foi encontrado na literatura consultada estudo semelhante para comparação dos resultados desse estudo. O experimento realizado em laboratório e as simulações de dados permitiram trabalhar com um grande número de provadores e também com diferentes distribuições de dados. A acuracidade dos dados foi estabelecida por critérios estatísticos permitindo aumentar a confiabilidade nos resultados dos testes realizados. Foi concluído que, embora o teste de escala hedônica não produza resultados com distribuição normal, para equipes de provadores com $n \geq 50$, que corresponde ao número mínimo de provadores para testes em laboratório conforme indicado na literatura, o modelo ANOVA apresentou o mesmo desempenho dos testes não-paramétricos utilizados, não justificando a sua substituição por teste não-paramétrico.

7. REFERÊNCIAS BIBLIOGRÁFICAS

AMERINE, M.A.; PANGBORN, R.M.; ROESSLER, E.B. *Principles of sensory evaluation of food*. New York: Academic Press, 1965. 602p.

ARAÚJO, F.B.; SILVA, P. H. A.; MINIM, V. P. R. Perfil sensorial e composição físico-química de cervejas provenientes de dois segmentos do mercado brasileiro. *Ciência e Tecnologia de Alimentos*, 23(2): 121-128, 2003.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. Teste Triangular em Análise Sensorial dos Alimentos e Bebidas - NBR 12995. São Paulo, SP. 1993.

BALL, R.D. Incomplete block designs for minimisation of order and carry-out effects in sensory analysis. *Food Quality and Preference*, 8(2): 111-118, 1997.

BORGES, L.C.; FERREIRA, D.F. Poder e taxas de erro tipo I dos testes Scott-Knott, Tukey e Student-Newman-Keuls sob distribuições normal e não normais dos resíduos. *Revista de Matemática e Estatística*, 21(1): 67-83, 2003.

BROCKHOFF, M.; SKOVGAARD, M. Modeling individual differences between assessors in sensory evaluations. *Food Quality and Preference*, 5(3): 215-224, 1994.

BUSSAB, W.O.; MORETIN, P.A. *Estatística Básica*. 5. ed. São Paulo: Saraiva, 2005. 526 p.

CALLEGARI-JACQUES, Sidia M. Bioestatística: princípios e aplicações. Porto Alegre: Artmed, 2003. p. 255. cap. 18.

CÂMARA, F.G.; SILVA, O. *Estatística não paramétrica: testes de hipóteses e medidas de associação*. Departamento de Matemática - Universidade dos Açores, Ponta Delgada, 2001.

CAMPOS, G.M. *Estatística prática para docentes e pós-graduandos*, 2000. Disponível em: <http://www.forp.usp.br/restauradora/gmc/gmc_livro/gmc_livro_cap19.html>. Acesso em: 9 mar. 2011.

CAPORALE, G.; MONTELEONE, E. Influence of information about manufacturing process on beer acceptability. *Food Quality and Preference*. 15(3): 271-278, 2004.

CARDELLO, H.M. A.B.; CARDELLO, L. Teor de vitamina C, atividade de ascorbato oxidase e perfil sensorial de manga (*Mangífera índica* L.) Var. Haden, durante o amadurecimento. *Ciência e Tecnologia de Alimentos*, 18(2): 211-217, 1998.

CHARLES, H. *Food science*. 2. ed. Rio de Janeiro: Prentice-Hall do Brasil, 1982. 564p.

CHAVES, J.B.P. *Análise Sensorial: Histórico e desenvolvimento*. Viçosa: Imprensa Universitária, 1993. 31p.

CHIAPPINI, C.C.J.; LEITE, S.G.F. Screening of yeasts for the production on the fruit aroma: sensory evaluation by comparative tests. *Alimentaria*, (357): 101-104, 2004.

CLARK, J.E. Taste and flavour: their importance in food choice and acceptance. *Proceedings of the Nutrition Society*, 57(4): 639-643, 1998.

COSTELL, E.; DURAN, L. El analise sensorial en el controle de qualidade de los alimentos. *Revista Agroquímica y Tecnología de Alimentos*, 21(1): 1-10, 1981.

DAWSON-SANDERS B.; TRAPP, R.G. *Basic & Clinical Biostatistics*. 2. ed. Norwalk: Appleton & Lange, 1994. 344 p.

DELIZA, R.; MACFIE, H.J.H. The generation of sensory expectations by external cues and its affects on sensory perception and hedonic ratings: A review. *Journal of Sensory Studies*, 11(2): 103-128, 1994.

DELLA MODESTA, R.C. *Manual de Análise Sensorial de Alimentos e Bebidas: Geral*. Rio de Janeiro: Embrapa-CTAA, 1994. t1.

DI MONACO, R.; CAVELLA, S.; DI MARZO, S.; MASI, P. The effect of expectations generated by brand name on the acceptability of dried semolina pasta. *Food Quality and Preference*, 15(5): 429-437, 2004.

DRANSFIELD, E.; ZAMORA, F.; BAYLE, M.C. Consumer selections of steaks as influenced by information and price index. *Food Quality and Preference*, 9(5): 321-326, 1998.

DUTCOSKY, S.D. *Análise sensorial de alimento*. 20. ed. Curitiba: Champagnat, 1996. 123 p.

ELORTONDO, F.J.P., Ojeda, M., Albisu, M., Salmerón, J., Etayo, I., & Molina, M. Food quality certification: An approach for the development of accredited sensory evaluation methods. *Food Quality and Preference*, 18(2), 425-439, 2007.

FISHER, R. A. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd, 1925. 43 p.

GARRUTI, R.S. *Métodos sensoriais em controle de qualidade na indústria de alimentos*. Campinas: Fundação tropical de pesquisa e tecnologia "André Tosello", 1985.

GAY, C; MEAD, R. A statistical appraisal of the problem of sensory measurement. *Journal of Sensory Studies*, 7(3): 205-228, 1992.

GIOVANNI, M.E.; PANGBORN, R.M. Measurement of taste intensity and degree of liking of beverages by graphic scales and magnitude estimation. *Journal of Food Science*, 48(4): 1175-1182, 1983.

GUINARD, J.X.; UOTANI, B.; SCHLICH, P. Internal and external mapping of preferences for commercial lager beers: comparison of hedonic ratings by consumers blind versus with knowledge of brand and price. *Food Quality and Preference*, 12(4): 243-255, 2001.

INSTITUTE OF FOOD TECHNOLOGISTS. Sensory evaluation guides for testing food and beverage products. *Food Technology*, 35(11): 50-59, 1981.

INSTITUTE OF MEDICINE (IOM). *Food and Nutrition Board. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids (macronutrients)*, Washington, DC: National Academies Press, 2005.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). *Censo Demográfico 2000: Resultados do universo*. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 20 out. 2009.

JAEGER, S. R. Non-sensory factors in sensory science research. *Food Quality and Preference*, 17(1-2): 132-144, 2006.

KENDALL, M.G., STUART, A. *The advanced theory of Statistics*. London: Charles Griffin, 1952. v.2, 690p.

KOSTER, M.A.; PRESCOTT, J.; KOSTER, E.P. Incidental learning and memory for three basic tastes in food. *Chemical Senses*, 29 (5): 441-453, 2004.

LANGE, C.; ROUSSEAU, F.; ISSANCHOU, S. Expectation, linking and purchase behavior under economical constraint. *Food Quality and Preference*, 10(1): 31-39, 1999.

LAWLESS, H.T.; CLAASSEN, Michelle R. Application of the central dogma in sensory evaluation. *Food Technology*, 47(6): 139-223, 1993.

LIM, J.; FUJIMARU, T. Evaluation of the Hedonic Scale under different experimental conditions. *Food Quality and Preference*. 21(5), 521-530, 2010.

MANLEY, C.H. Development and regulation of flavor, fragrance, and color ingredients produced by biotechnology. In: *Bioprocess production of flavor, fragrance and color ingredients*. New York: John Wiley & Sons, 1995. p.19-39.

MEILGAARD, M.C.; CIVILLE, G.V.; CARR, B.T. *Sensory evaluation techniques*. 3. ed. Boca Raton: CRC Press, 1991.

McPHERSON, R.S.; RANDALL, E. Line length measurement as a tool for food preference research. *Ecology of Food and Nutrition*, 17(2): 149-156, 1985.

- MILLER, A. J. Adjusting taste scores for variations in use of scales. *Journal of Sensory Studies*, 2(4): 231-242, 1987.
- MOORE, D.S.; McCABE, G.P. *Introdução à prática da estatística*. Trad. FARIAS, A. A. Rio de Janeiro: LTC, 2002. 536p.
- MUÑOZ, A.M. Sensory evaluation in quality control: an overview, new developments and future opportunities. *Food Quality and Preference*, 13(6), 329-339, 2002.
- MUNRO, B.H. *Statistical methods for health care research*. 3. ed. Philadelphia: Lippincott, 1997. 444 p.
- MURRAY, J.M.; DELAHUNTY, C.M. Mapping consumer preference for the sensory and packaging attributes of cheddar cheese. *Food Quality and Preference*, 11(5): 419-435, 2000.
- NAES, T. Handling individual differences between assessors in sensory profiling. *Food Quality and Preference*, 2(3): 187-199, 1990.
- NEWELL, G.J.; MACFARLANE, J.D. Expanded tables for multiple comparison procedures in the analysis of ranked data. *Journal of Food Science*, 52(6): 1721-1725, 1987.
- OLIVEIRA, A.F. *Apostila de análise sensorial dos alimentos*. Curso de Tecnologia de Alimentos. Universidade Tecnológica Federal do Paraná. Londrina, 2010.
- O'MAHONY, M. Some assumptions and difficulties with common statistics for sensory analysis. *Food Technology*, 36(11): 75-82, 1982.
- PIGGOT, J.R. Design in sensory and consumer science. *Food Quality and Preference*, 6(4): 217-220, 1995.
- PONTES, A.C.F.; CORRENTE J. E. Comparações múltiplas não-paramétricas para o delineamento com um fator de classificação simples. *Revista de Matemática e Estatística*. São Paulo, 19: 179-197, 2001.
- PONTES, A.C.F.; PONTES JR., A. C. F.; PONTES L. O., Testes não-paramétricos: uma ferramenta para o ensaio da Análise Combinatória. In: *Simpósio Nacional de Ensino de Ciência e Tecnologia*, 1., 2009, Paraná. Anais...Paraná: Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Ensino de Ciência e Tecnologia, 2009. 1168-1169 p.
- PRITCHETT, M.P.A. Performance assessment of sensory panelists. *Journal of Sensory Studies*, 7(3): 229-252, 1992.
- REIS, G.M.; RIBEIRO JR., J.I. Comparação de testes paramétricos e não paramétricos aplicados em delineamentos experimentais. In: *Simpósio Acadêmico de Engenharia de Produção*, 3., 2007, Viçosa. Anais... Viçosa: Universidade Federal de Viçosa, 2007.
- RIBEIRO, M.M.; DELLA LUCIA, S.M.; BARBOSA, P.B.F.; GALVÃO, H.L.; MINIM, V.P.R. Influência da embalagem na aceitação de diferentes marcas comerciais de cerveja tipo Pilsen. *Ciência e Tecnologia de Alimentos*, 28(2): 395-399, 2008.

SHIMAKURA, S.E. *A Bioestatística*. Departamento de Estatística, UFPR. Disponível em: <http://leg.ufpr.br/~shimakur/CE005/>. Acesso em: 20 jan. 2011.

SIEGEL, S.; CASTELLAN JR., N. J. *Estatística não-paramétrica para ciências do comportamento*. Trad. Sara Ianda Correa Carmona. 2. ed. Porto Alegre: Artmed, 2006. 448 p.

SIRET, F.; ISSANCHOU, S. Traditional process: influence on sensory properties and on consumer's expectation and liking. Application to paté de campagne. *Food Quality and Preference*, 11(3): 217-228, 2000.

SMYTHE, J.E.; BAMFORTH, C.W. A study of the effect of perceived beer history on reported preferences by sensory panels with different levels of training. *Journal of the Institute of Brewing*, 108(1): 34-36, 2002.

SOARES, J.F.; SIQUEIRA, A.L. *Introdução à Estatística Médica*. 2. ed. Belo Horizonte: Coopmed, 2002. 300 p.

STONE, H.; SIDEL, J.L. *Sensory evaluation practices*. 3. ed. London: Academic Press, 2004. 408p.

STONE, H.; SIDEL, J.L. *Sensory evaluation practices*. 2. ed. San Diego: Academic Press, 1993.

TUORILA, H.; CARDELLO, A.V.; LESHER, L. L. Antecedents and consequences of expectations related to fat-free and regular-fat foods. *Appetite*, 23(3): 247-263, 1994.

VIE, A.; GULLI, O.; O'MAHONY, M. Alternative hedonic measures. *Journal of Food Science*, 56(1): 1-5, 1991.

VIEIRA, S. *Estatística Experimental*. 2. ed. São Paulo: Atlas, 1999. 185 p.

VILLANUEVA, N. D. M.; DA SILVA. Comparative performance of the nine-point hedonic, hybrid and self-adjusting scales in the generation of internal preference maps. *Food Quality and Preference*. 20(1), 1-12, 2009.

VILLANUEVA, N.D.M.; PETENATE, A. J.; DA SILVA, M.A.A. P. Performance of the hybrid hedonic scale as compared to the traditional hedonic, self-adjusting and ranking scales. *Food Quality and Preference*, 16(8): 691-703, 2005.

VILLANUEVA, N.D.M.; PETENATE, A. J.; DA SILVA, M.A.A.P. Performance of three affective methods and diagnosis of the ANOVA model. *Food Quality and Preference*, 11(5): 363-370, 2000.

WILKINSON, C.; YUKSEL, D. Modeling differences between panelists in use of measurement scales. *Journal of Sensory Studies*, 12(1): 55-68, 1997.

8. APÉNDICE

Apêndice 1. Rotina do programa estatístico S-Plus versão 8.0 elaborado para a Simulação 1

```

numaval_5
numitera_10
contadorkruskalmaracuja_0
contadorkruskalmanga_0
contadoranovamaracuja_0
contadoranovamanga_0
for(k in 1:numitera){
print(k)
escores_matrix(0,3*numaval,6)
for(i in 1:numaval){
fetaria_runif(1,0,1)
if(fetaria<=0.19)
{
    indicafaixa_1
    genero_runif(1,0,1)
    if(genero<=0.5){
    indicagenero_0
for(j in 1:3){
escores[3*i-3+j,1]_i
escores[3*i-3+j,2]_indicagenero
escores[3*i-3+j,3]_indicafaixa
escores[3*i-3+j,4]_j
if(j=1){
auxiliar_runif(1,0,1)
if(auxiliar<=0.6){escores[3*i-2,5]_1}
else{if(auxiliar<=0.8){escores[3*i-2,5]_2}
else{if(auxiliar<=0.9){escores[3*i-2,5]_3}
else{if(auxiliar<=0.96){escores[3*i-2,5]_4}
else{if(auxiliar<=0.97){escores[3*i-2,5]_5}
else{if(auxiliar<=0.98){escores[3*i-2,5]_6}
else{if(auxiliar<=0.99){escores[3*i-2,5]_7}
else{if(auxiliar<=1){escores[3*i-2,5]_8}
else{escores[3*i-2,5]_9}
}}}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.5){escores[3*i-2,6]_1}
else{if(auxiliar<=0.7){escores[3*i-2,6]_2}
else{if(auxiliar<=0.8){escores[3*i-2,6]_3}
else{if(auxiliar<=0.85){escores[3*i-2,6]_4}
else{if(auxiliar<=0.89){escores[3*i-2,6]_5}
else{if(auxiliar<=0.92){escores[3*i-2,6]_6}
else{if(auxiliar<=0.95){escores[3*i-2,6]_7}
else{if(auxiliar<=0.98){escores[3*i-2,6]_8}
else{escores[3*i-3+j,6]_9}
}}}}}}}}
}
if(j=2){
auxiliar_runif(1,0,1)

```

```

if(auxiliar<=0.1){escores[3*i-3+j,5]_1}
else{if(auxiliar<=0.2){escores[3*i-1,5]_2}
else{if(auxiliar<=0.3){escores[3*i-1,5]_3}
else{if(auxiliar<=0.52){escores[3*i-1,5]_4}
else{if(auxiliar<=0.67){escores[3*i-1,5]_5}
else{if(auxiliar<=0.85){escores[3*i-1,5]_6}
else{if(auxiliar<=0.95){escores[3*i-1,5]_7}
else{if(auxiliar<=0.98){escores[3*i-1,5]_8}
else{escores[3*i-1,5]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.05){escores[3*i-1,6]_1}
else{if(auxiliar<=0.1){escores[3*i-1,6]_2}
else{if(auxiliar<=0.15){escores[3*i-1,6]_3}
else{if(auxiliar<=0.38){escores[3*i-1,6]_4}
else{if(auxiliar<=0.55){escores[3*i-1,6]_5}
else{if(auxiliar<=0.74){escores[3*i-1,6]_6}
else{if(auxiliar<=0.89){escores[3*i-1,6]_7}
else{if(auxiliar<=0.94){escores[3*i-1,6]_8}
else{escores[3*i-1,6]_9}
}}}}}}
}
if(j=3){
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){escores[3*i,5]_1}
else{if(auxiliar<=0.0){escores[3*i,5]_2}
else{if(auxiliar<=0.01){escores[3*i,5]_3}
else{if(auxiliar<=0.02){escores[3*i,5]_4}
else{if(auxiliar<=0.03){escores[3*i,5]_5}
else{if(auxiliar<=0.08){escores[3*i,5]_6}
else{if(auxiliar<=0.14){escores[3*i,5]_7}
else{if(auxiliar<=0.64){escores[3*i,5]_8}
else{escores[3*i,5]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){escores[3*i,6]_1}
else{if(auxiliar<=0.0){escores[3*i,6]_2}
else{if(auxiliar<=0.0){escores[3*i,6]_3}
else{if(auxiliar<=0.01){escores[3*i,6]_4}
else{if(auxiliar<=0.01){escores[3*i,6]_5}
else{if(auxiliar<=0.02){escores[3*i,6]_6}
else{if(auxiliar<=0.1){escores[3*i,6]_7}
else{if(auxiliar<=0.7){escores[3*i,6]_8}
else{escores[3*i,6]_9}
}}}}}}
}
}
}
else{
indicagenero_1

```



```

for(j in 1:3){
  escores[3*i-3+j,1]_i
  escores[3*i-3+j,2]_indicagenero
  escores[3*i-3+j,3]_indicafaixa
  escores[3*i-3+j,4]_j
  if(j=1){
    auxiliar_runif(1,0,1)
    if(auxiliar<=0.7){ escores[3*i-2,5]_1 }
    else{ if(auxiliar<=0.8){ escores[3*i-2,5]_2 }
    else{ if(auxiliar<=0.9){ escores[3*i-2,5]_3 }
    else{ if(auxiliar<=0.99){ escores[3*i-2,5]_4 }
    else{ if(auxiliar<=1.0){ escores[3*i-2,5]_5 }
    else{ if(auxiliar<=1.0){ escores[3*i-2,5]_6 }
    else{ if(auxiliar<=1.0){ escores[3*i-2,5]_7 }
    else{ if(auxiliar<=1.0){ escores[3*i-2,5]_8 }
    else{ escores[3*i-2,5]_9 }
    } } } } } }
    auxiliar_runif(1,0,1)
    if(auxiliar<=0.65){ escores[3*i-2,6]_1 }
    else{ if(auxiliar<=0.75){ escores[3*i-2,6]_2 }
    else{ if(auxiliar<=0.85){ escores[3*i-2,6]_3 }
    else{ if(auxiliar<=0.91){ escores[3*i-2,6]_4 }
    else{ if(auxiliar<=0.94){ escores[3*i-2,6]_5 }
    else{ if(auxiliar<=0.96){ escores[3*i-2,6]_6 }
    else{ if(auxiliar<=0.98){ escores[3*i-2,6]_7 }
    else{ if(auxiliar<=0.99){ escores[3*i-2,6]_8 }
    else{ escores[3*i-3+j,6]_9 }
    } } } } } }
  }
  if(j=2){
    auxiliar_runif(1,0,1)
    if(auxiliar<=0.03){ escores[3*i-3+j,5]_1 }
    else{ if(auxiliar<=0.07){ escores[3*i-1,5]_2 }
    else{ if(auxiliar<=0.14){ escores[3*i-1,5]_3 }
    else{ if(auxiliar<=0.24){ escores[3*i-1,5]_4 }
    else{ if(auxiliar<=0.41){ escores[3*i-1,5]_5 }
    else{ if(auxiliar<=0.73){ escores[3*i-1,5]_6 }
    else{ if(auxiliar<=0.85){ escores[3*i-1,5]_7 }
    else{ if(auxiliar<=0.92){ escores[3*i-1,5]_8 }
    else{ escores[3*i-1,5]_9 }
    } } } } } }
    auxiliar_runif(1,0,1)
    if(auxiliar<=0.0){ escores[3*i-1,6]_1 }
    else{ if(auxiliar<=0.04){ escores[3*i-1,6]_2 }
    else{ if(auxiliar<=0.04){ escores[3*i-1,6]_3 }
    else{ if(auxiliar<=0.16){ escores[3*i-1,6]_4 }
    else{ if(auxiliar<=0.34){ escores[3*i-1,6]_5 }
    else{ if(auxiliar<=0.67){ escores[3*i-1,6]_6 }
    else{ if(auxiliar<=0.8){ escores[3*i-1,6]_7 }
    else{ if(auxiliar<=0.9){ escores[3*i-1,6]_8 }

```

```

else{escores[3*i-1,6]_9}
}}}}}}
}
if(j=3){
auxiliar_runif(1,0,1)
if(auxiliar<=0.02){escores[3*i,5]_1}
else{if(auxiliar<=0.04){escores[3*i,5]_2}
else{if(auxiliar<=0.08){escores[3*i,5]_3}
else{if(auxiliar<=0.13){escores[3*i,5]_4}
else{if(auxiliar<=0.23){escores[3*i,5]_5}
else{if(auxiliar<=0.33){escores[3*i,5]_6}
else{if(auxiliar<=0.45){escores[3*i,5]_7}
else{if(auxiliar<=0.65){escores[3*i,5]_8}
else{escores[3*i,5]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.01){escores[3*i,6]_1}
else{if(auxiliar<=0.02){escores[3*i,6]_2}
else{if(auxiliar<=0.04){escores[3*i,6]_3}
else{if(auxiliar<=0.07){escores[3*i,6]_4}
else{if(auxiliar<=0.17){escores[3*i,6]_5}
else{if(auxiliar<=0.26){escores[3*i,6]_6}
else{if(auxiliar<=0.41){escores[3*i,6]_7}
else{if(auxiliar<=0.63){escores[3*i,6]_8}
else{escores[3*i,6]_9}
}}}}}}
}
}
}
}
else
{
    if(fetaria<=0.54){
        indicafaixa_2
        genero_runif(1,0,1)
        if(genero<=0.51){
            indicagenero_0
        }
    }
}
for(j in 1:3){
escores[3*i-3+j,1]_i
escores[3*i-3+j,2]_indicagenero
escores[3*i-3+j,3]_indicafaixa
escores[3*i-3+j,4]_j
if(j=1){
auxiliar_runif(1,0,1)
if(auxiliar<=0.55){escores[3*i-2,5]_1}
else{if(auxiliar<=0.69){escores[3*i-2,5]_2}
else{if(auxiliar<=0.83){escores[3*i-2,5]_3}
else{if(auxiliar<=0.91){escores[3*i-2,5]_4}
else{if(auxiliar<=0.96){escores[3*i-2,5]_5}
else{if(auxiliar<=0.98){escores[3*i-2,5]_6}

```

```

else{if(auxiliar<=0.99){escores[3*i-2,5]_7}
else{if(auxiliar<=1.0){escores[3*i-2,5]_8}
else{escores[3*i-2,5]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.5){escores[3*i-2,6]_1}
else{if(auxiliar<=0.6){escores[3*i-2,6]_2}
else{if(auxiliar<=0.7){escores[3*i-2,6]_3}
else{if(auxiliar<=0.77){escores[3*i-2,6]_4}
else{if(auxiliar<=0.87){escores[3*i-2,6]_5}
else{if(auxiliar<=0.89){escores[3*i-2,6]_6}
else{if(auxiliar<=0.99){escores[3*i-2,6]_7}
else{if(auxiliar<=1.0){escores[3*i-2,6]_8}
else{escores[3*i-3+j,6]_9}
}}}}}}}}
}
if(j=2){
auxiliar_runif(1,0,1)
if(auxiliar<=0.04){escores[3*i-3+j,5]_1}
else{if(auxiliar<=0.12){escores[3*i-1,5]_2}
else{if(auxiliar<=0.17){escores[3*i-1,5]_3}
else{if(auxiliar<=0.27){escores[3*i-1,5]_4}
else{if(auxiliar<=0.42){escores[3*i-1,5]_5}
else{if(auxiliar<=0.73){escores[3*i-1,5]_6}
else{if(auxiliar<=0.83){escores[3*i-1,5]_7}
else{if(auxiliar<=0.92){escores[3*i-1,5]_8}
else{escores[3*i-1,5]_9}
}}}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.03){escores[3*i-1,6]_1}
else{if(auxiliar<=0.1){escores[3*i-1,6]_2}
else{if(auxiliar<=0.14){escores[3*i-1,6]_3}
else{if(auxiliar<=0.19){escores[3*i-1,6]_4}
else{if(auxiliar<=0.36){escores[3*i-1,6]_5}
else{if(auxiliar<=0.69){escores[3*i-1,6]_6}
else{if(auxiliar<=0.8){escores[3*i-1,6]_7}
else{if(auxiliar<=0.9){escores[3*i-1,6]_8}
else{escores[3*i-1,6]_9}
}}}}}}}}
}
if(j=3){
auxiliar_runif(1,0,1)
if(auxiliar<=0.02){escores[3*i,5]_1}
else{if(auxiliar<=0.04){escores[3*i,5]_2}
else{if(auxiliar<=0.06){escores[3*i,5]_3}
else{if(auxiliar<=0.09){escores[3*i,5]_4}
else{if(auxiliar<=0.12){escores[3*i,5]_5}
else{if(auxiliar<=0.24){escores[3*i,5]_6}
else{if(auxiliar<=0.42){escores[3*i,5]_7}
else{if(auxiliar<=0.67){escores[3*i,5]_8}

```



```

        genero_runif(1,0,1)
        if(genero<=0.52){
            indicagenero_0
for(j in 1:3){
    escores[3*i-3+j,1]_i
    escores[3*i-3+j,2]_indicagenero
    escores[3*i-3+j,3]_indicafaixa
    escores[3*i-3+j,4]_j
    if(j=1){
        auxiliar_runif(1,0,1)
        if(auxiliar<=0.73){ escores[3*i-2,5]_1 }
        else{ if(auxiliar<=0.93){ escores[3*i-2,5]_2 }
        else{ if(auxiliar<=0.97){ escores[3*i-2,5]_3 }
        else{ if(auxiliar<=0.99){ escores[3*i-2,5]_4 }
        else{ if(auxiliar<=1.0){ escores[3*i-2,5]_5 }
        else{ if(auxiliar<=1.0){ escores[3*i-2,5]_6 }
        else{ if(auxiliar<=1.0){ escores[3*i-2,5]_7 }
        else{ if(auxiliar<=1.0){ escores[3*i-2,5]_8 }
        else{ escores[3*i-2,5]_9 }
        } } } } } }
        auxiliar_runif(1,0,1)
        if(auxiliar<=0.62){ escores[3*i-2,6]_1 }
        else{ if(auxiliar<=0.83){ escores[3*i-2,6]_2 }
        else{ if(auxiliar<=0.92){ escores[3*i-2,6]_3 }
        else{ if(auxiliar<=0.97){ escores[3*i-2,6]_4 }
        else{ if(auxiliar<=0.99){ escores[3*i-2,6]_5 }
        else{ if(auxiliar<=1.0){ escores[3*i-2,6]_6 }
        else{ if(auxiliar<=1.0){ escores[3*i-2,6]_7 }
        else{ if(auxiliar<=1.0){ escores[3*i-2,6]_8 }
        else{ escores[3*i-3+j,6]_9 }
        } } } } } }
    }
    if(j=2){
        auxiliar_runif(1,0,1)
        if(auxiliar<=0.1){ escores[3*i-3+j,5]_1 }
        else{ if(auxiliar<=0.2){ escores[3*i-1,5]_2 }
        else{ if(auxiliar<=0.32){ escores[3*i-1,5]_3 }
        else{ if(auxiliar<=0.47){ escores[3*i-1,5]_4 }
        else{ if(auxiliar<=0.63){ escores[3*i-1,5]_5 }
        else{ if(auxiliar<=0.73){ escores[3*i-1,5]_6 }
        else{ if(auxiliar<=0.83){ escores[3*i-1,5]_7 }
        else{ if(auxiliar<=0.92){ escores[3*i-1,5]_8 }
        else{ escores[3*i-1,5]_9 }
        } } } } } }
        auxiliar_runif(1,0,1)
        if(auxiliar<=0.05){ escores[3*i-1,6]_1 }
        else{ if(auxiliar<=0.1){ escores[3*i-1,6]_2 }
        else{ if(auxiliar<=0.2){ escores[3*i-1,6]_3 }
        else{ if(auxiliar<=0.33){ escores[3*i-1,6]_4 }
        else{ if(auxiliar<=0.53){ escores[3*i-1,6]_5 }

```



```

}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.71){escores[3*i-2,6]_1}
else{if(auxiliar<=0.9){escores[3*i-2,6]_2}
else{if(auxiliar<=0.99){escores[3*i-2,6]_3}
else{if(auxiliar<=1.0){escores[3*i-2,6]_4}
else{if(auxiliar<=1.0){escores[3*i-2,6]_5}
else{if(auxiliar<=1.0){escores[3*i-2,6]_6}
else{if(auxiliar<=1.0){escores[3*i-2,6]_7}
else{if(auxiliar<=1.0){escores[3*i-2,6]_8}
else{escores[3*i-3+j,6]_9}
}}}}}}
}
if(j=2){
auxiliar_runif(1,0,1)
if(auxiliar<=0.03){escores[3*i-3+j,5]_1}
else{if(auxiliar<=0.09){escores[3*i-1,5]_2}
else{if(auxiliar<=0.16){escores[3*i-1,5]_3}
else{if(auxiliar<=0.31){escores[3*i-1,5]_4}
else{if(auxiliar<=0.66){escores[3*i-1,5]_5}
else{if(auxiliar<=0.87){escores[3*i-1,5]_6}
else{if(auxiliar<=0.97){escores[3*i-1,5]_7}
else{if(auxiliar<=0.99){escores[3*i-1,5]_8}
else{escores[3*i-1,5]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.02){escores[3*i-1,6]_1}
else{if(auxiliar<=0.04){escores[3*i-1,6]_2}
else{if(auxiliar<=0.07){escores[3*i-1,6]_3}
else{if(auxiliar<=0.18){escores[3*i-1,6]_4}
else{if(auxiliar<=0.55){escores[3*i-1,6]_5}
else{if(auxiliar<=0.8){escores[3*i-1,6]_6}
else{if(auxiliar<=0.95){escores[3*i-1,6]_7}
else{if(auxiliar<=0.98){escores[3*i-1,6]_8}
else{escores[3*i-1,6]_9}
}}}}}}
}
if(j=3){
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){escores[3*i,5]_1}
else{if(auxiliar<=0.0){escores[3*i,5]_2}
else{if(auxiliar<=0.0){escores[3*i,5]_3}
else{if(auxiliar<=0.01){escores[3*i,5]_4}
else{if(auxiliar<=0.03){escores[3*i,5]_5}
else{if(auxiliar<=0.06){escores[3*i,5]_6}
else{if(auxiliar<=0.21){escores[3*i,5]_7}
else{if(auxiliar<=0.37){escores[3*i,5]_8}
else{escores[3*i,5]_9}
}}}}}}
auxiliar_runif(1,0,1)

```


Apêndice 2. Rotina do programa estatístico S-Plus versão 8.0 elaborado para a Simulação 2

```

numaval_1000
numitera_100
contadorkruskal1_0
contadoranova1_0
contadorkruskal2_0
contadoranova2_0
contadorkruskal3_0
contadoranova3_0
contadorkruskal4_0
contadoranova4_0
contadorkruskal5_0
contadoranova5_0
contadorkruskal6_0
contadoranova6_0
for(k in 1:numitera){
print(k)
escores_matrix(0,3*numaval,7)
for(i in 1:numaval){
for(j in 1:3){
escores[3*i-3+j,1]_j
if(j=1){
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){escores[3*i-2,2]_1}
else{if(auxiliar<=0.1){escores[3*i-2,2]_2}
else{if(auxiliar<=0.3){escores[3*i-2,2]_3}
else{if(auxiliar<=0.85){escores[3*i-2,2]_4}
else{if(auxiliar<=1.0){escores[3*i-2,2]_5}
else{if(auxiliar<=1.0){escores[3*i-2,2]_6}
else{if(auxiliar<=1.0){escores[3*i-2,2]_7}
else{if(auxiliar<=1.0){escores[3*i-2,2]_8}
else{escores[3*i-2,2]_9}
}}}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.05){escores[3*i-2,3]_1}
else{if(auxiliar<=0.15){escores[3*i-2,3]_2}
else{if(auxiliar<=0.35){escores[3*i-2,3]_3}
else{if(auxiliar<=0.65){escores[3*i-2,3]_4}
else{if(auxiliar<=0.85){escores[3*i-2,3]_5}
else{if(auxiliar<=0.95){escores[3*i-2,3]_6}
else{if(auxiliar<=1.0){escores[3*i-2,3]_7}
else{if(auxiliar<=1.0){escores[3*i-2,3]_8}
else{escores[3*i-2,3]_9}
}}}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.3){escores[3*i-2,4]_1}
else{if(auxiliar<=0.7){escores[3*i-2,4]_2}
else{if(auxiliar<=0.9){escores[3*i-2,4]_3}
else{if(auxiliar<=1.0){escores[3*i-2,4]_4}

```

```

else{ if(auxiliar<=1.0){ escores[3*i-2,4]_5}
else{ if(auxiliar<=1.0){ escores[3*i-2,4]_6}
else{ if(auxiliar<=1.0){ escores[3*i-2,4]_7}
else{ if(auxiliar<=1.0){ escores[3*i-2,4]_8}
else{ escores[3*i-2,4]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.2){ escores[3*i-2,5]_1}
else{ if(auxiliar<=0.45){ escores[3*i-2,5]_2}
else{ if(auxiliar<=0.65){ escores[3*i-2,5]_3}
else{ if(auxiliar<=0.8){ escores[3*i-2,5]_4}
else{ if(auxiliar<=0.9){ escores[3*i-2,5]_5}
else{ if(auxiliar<=0.95){ escores[3*i-2,5]_6}
else{ if(auxiliar<=0.98){ escores[3*i-2,5]_7}
else{ if(auxiliar<=1.0){ escores[3*i-2,5]_8}
else{ escores[3*i-2,5]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.2){ escores[3*i-2,6]_1}
else{ if(auxiliar<=0.4){ escores[3*i-2,6]_2}
else{ if(auxiliar<=0.6){ escores[3*i-2,6]_3}
else{ if(auxiliar<=0.8){ escores[3*i-2,6]_4}
else{ if(auxiliar<=1.0){ escores[3*i-2,6]_5}
else{ if(auxiliar<=1.0){ escores[3*i-2,6]_6}
else{ if(auxiliar<=1.0){ escores[3*i-2,6]_7}
else{ if(auxiliar<=1.0){ escores[3*i-2,6]_8}
else{ escores[3*i-2,6]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){ escores[3*i-2,7]_1}
else{ if(auxiliar<=0.2){ escores[3*i-2,7]_2}
else{ if(auxiliar<=0.4){ escores[3*i-2,7]_3}
else{ if(auxiliar<=0.6){ escores[3*i-2,7]_4}
else{ if(auxiliar<=0.8){ escores[3*i-2,7]_5}
else{ if(auxiliar<=1.0){ escores[3*i-2,7]_6}
else{ if(auxiliar<=1.0){ escores[3*i-2,7]_7}
else{ if(auxiliar<=1.0){ escores[3*i-2,7]_8}
else{ escores[3*i-2,7]_9}
}}}}}}
}
if(j=2){
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){ escores[3*i-1,2]_1}
else{ if(auxiliar<=0.0){ escores[3*i-1,2]_2}
else{ if(auxiliar<=0.0){ escores[3*i-1,2]_3}
else{ if(auxiliar<=0.25){ escores[3*i-1,2]_4}
else{ if(auxiliar<=0.75){ escores[3*i-1,2]_5}
else{ if(auxiliar<=1.0){ escores[3*i-1,2]_6}
else{ if(auxiliar<=1.0){ escores[3*i-1,2]_7}
else{ if(auxiliar<=1.0){ escores[3*i-1,2]_8}

```

```

else{escores[3*i-1,2]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){escores[3*i-1,3]_1}
else{if(auxiliar<=0.05){escores[3*i-1,3]_2}
else{if(auxiliar<=0.15){escores[3*i-1,3]_3}
else{if(auxiliar<=0.35){escores[3*i-1,3]_4}
else{if(auxiliar<=0.65){escores[3*i-1,3]_5}
else{if(auxiliar<=0.85){escores[3*i-1,3]_6}
else{if(auxiliar<=0.95){escores[3*i-1,3]_7}
else{if(auxiliar<=1.0){escores[3*i-1,3]_8}
else{escores[3*i-1,3]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){escores[3*i-1,4]_1}
else{if(auxiliar<=0.0){escores[3*i-1,4]_2}
else{if(auxiliar<=0.1){escores[3*i-1,4]_3}
else{if(auxiliar<=0.3){escores[3*i-1,4]_4}
else{if(auxiliar<=0.7){escores[3*i-1,4]_5}
else{if(auxiliar<=0.9){escores[3*i-1,4]_6}
else{if(auxiliar<=1.0){escores[3*i-1,4]_7}
else{if(auxiliar<=1.0){escores[3*i-1,4]_8}
else{escores[3*i-1,4]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){escores[3*i-1,5]_1}
else{if(auxiliar<=0.05){escores[3*i-1,5]_2}
else{if(auxiliar<=0.15){escores[3*i-1,5]_3}
else{if(auxiliar<=0.35){escores[3*i-1,5]_4}
else{if(auxiliar<=0.65){escores[3*i-1,5]_5}
else{if(auxiliar<=0.85){escores[3*i-1,5]_6}
else{if(auxiliar<=0.95){escores[3*i-1,5]_7}
else{if(auxiliar<=1.0){escores[3*i-1,5]_8}
else{escores[3*i-1,5]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){escores[3*i-1,6]_1}
else{if(auxiliar<=0.0){escores[3*i-1,6]_2}
else{if(auxiliar<=0.2){escores[3*i-1,6]_3}
else{if(auxiliar<=0.4){escores[3*i-1,6]_4}
else{if(auxiliar<=0.6){escores[3*i-1,6]_5}
else{if(auxiliar<=0.8){escores[3*i-1,6]_6}
else{if(auxiliar<=1.0){escores[3*i-1,6]_7}
else{if(auxiliar<=1.0){escores[3*i-1,6]_8}
else{escores[3*i-2,6]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){escores[3*i-1,7]_1}
else{if(auxiliar<=0.0){escores[3*i-1,7]_2}
else{if(auxiliar<=0.2){escores[3*i-1,7]_3}

```

```

else{ if(auxiliar<=0.4){ escores[3*i-1,7]_4}
else{ if(auxiliar<=0.6){ escores[3*i-1,7]_5}
else{ if(auxiliar<=0.8){ escores[3*i-1,7]_6}
else{ if(auxiliar<=1.0){ escores[3*i-1,7]_7}
else{ if(auxiliar<=1.0){ escores[3*i-1,7]_8}
else{ escores[3*i-1,7]_9}
}}}}}}
}
if(j=3){
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){ escores[3*i,2]_1}
else{ if(auxiliar<=0.0){ escores[3*i,2]_2}
else{ if(auxiliar<=0.0){ escores[3*i,2]_3}
else{ if(auxiliar<=0.0){ escores[3*i,2]_4}
else{ if(auxiliar<=0.15){ escores[3*i,2]_5}
else{ if(auxiliar<=0.7){ escores[3*i,2]_6}
else{ if(auxiliar<=0.9){ escores[3*i,2]_7}
else{ if(auxiliar<=1.0){ escores[3*i,2]_8}
else{ escores[3*i,2]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){ escores[3*i,3]_1}
else{ if(auxiliar<=0.0){ escores[3*i,3]_2}
else{ if(auxiliar<=0.05){ escores[3*i,3]_3}
else{ if(auxiliar<=0.15){ escores[3*i,3]_4}
else{ if(auxiliar<=0.35){ escores[3*i,3]_5}
else{ if(auxiliar<=0.65){ escores[3*i,3]_6}
else{ if(auxiliar<=0.85){ escores[3*i,3]_7}
else{ if(auxiliar<=0.95){ escores[3*i,3]_8}
else{ escores[3*i,3]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){ escores[3*i,4]_1}
else{ if(auxiliar<=0.0){ escores[3*i,4]_2}
else{ if(auxiliar<=0.0){ escores[3*i,4]_3}
else{ if(auxiliar<=0.0){ escores[3*i,4]_4}
else{ if(auxiliar<=0.0){ escores[3*i,4]_5}
else{ if(auxiliar<=0.1){ escores[3*i,4]_6}
else{ if(auxiliar<=0.3){ escores[3*i,4]_7}
else{ if(auxiliar<=0.7){ escores[3*i,4]_8}
else{ escores[3*i,4]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){ escores[3*i,5]_1}
else{ if(auxiliar<=0.02){ escores[3*i,5]_2}
else{ if(auxiliar<=0.05){ escores[3*i,5]_3}
else{ if(auxiliar<=0.1){ escores[3*i,5]_4}
else{ if(auxiliar<=0.2){ escores[3*i,5]_5}
else{ if(auxiliar<=0.35){ escores[3*i,5]_6}
else{ if(auxiliar<=0.55){ escores[3*i,5]_7}

```

```

else{ if(auxiliar<=0.8){ escores[3*i,5]_8}
else{ escores[3*i,5]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){ escores[3*i,6]_1}
else{ if(auxiliar<=0.0){ escores[3*i,6]_2}
else{ if(auxiliar<=0.0){ escores[3*i,6]_3}
else{ if(auxiliar<=0.0){ escores[3*i,6]_4}
else{ if(auxiliar<=0.2){ escores[3*i,6]_5}
else{ if(auxiliar<=0.4){ escores[3*i,6]_6}
else{ if(auxiliar<=0.6){ escores[3*i,6]_7}
else{ if(auxiliar<=0.8){ escores[3*i,6]_8}
else{ escores[3*i,6]_9}
}}}}}}
auxiliar_runif(1,0,1)
if(auxiliar<=0.0){ escores[3*i,7]_1}
else{ if(auxiliar<=0.0){ escores[3*i,7]_2}
else{ if(auxiliar<=0.0){ escores[3*i,7]_3}
else{ if(auxiliar<=0.2){ escores[3*i,7]_4}
else{ if(auxiliar<=0.4){ escores[3*i,7]_5}
else{ if(auxiliar<=0.6){ escores[3*i,7]_6}
else{ if(auxiliar<=0.8){ escores[3*i,7]_7}
else{ if(auxiliar<=1.0){ escores[3*i,7]_8}
else{ escores[3*i,7]_9}
}}}}}}
}
}
}
valorpkruskal1_kruskal.test(escores[,2],escores[,1])[[3]]
valorpanova1_summary.aov(aov(escores[,2]~escores[,1]))[[5]][1]
if(valorpkruskal1>0.05){ contadorkruskal1_contadorkruskal1+1}
if(valorpanova1>0.05){ contadoranova1_contadoranova1+1}
valorpkruskal2_kruskal.test(escores[,3],escores[,1])[[3]]
valorpanova2_summary.aov(aov(escores[,3]~escores[,1]))[[5]][1]
if(valorpkruskal2>0.05){ contadorkruskal2_contadorkruskal2+1}
if(valorpanova2>0.05){ contadoranova2_contadoranova2+1}
valorpkruskal3_kruskal.test(escores[,4],escores[,1])[[3]]
valorpanova3_summary.aov(aov(escores[,4]~escores[,1]))[[5]][1]
if(valorpkruskal3>0.05){ contadorkruskal3_contadorkruskal3+1}
if(valorpanova3>0.05){ contadoranova3_contadoranova3+1}
valorpkruskal4_kruskal.test(escores[,5],escores[,1])[[3]]
valorpanova4_summary.aov(aov(escores[,5]~escores[,1]))[[5]][1]
if(valorpkruskal4>0.05){ contadorkruskal4_contadorkruskal4+1}
if(valorpanova4>0.05){ contadoranova4_contadoranova4+1}
valorpkruskal5_kruskal.test(escores[,6],escores[,1])[[3]]
valorpanova5_summary.aov(aov(escores[,6]~escores[,1]))[[5]][1]
if(valorpkruskal5>0.05){ contadorkruskal5_contadorkruskal5+1}
if(valorpanova5>0.05){ contadoranova5_contadoranova5+1}
valorpkruskal6_kruskal.test(escores[,7],escores[,1])[[3]]
valorpanova6_summary.aov(aov(escores[,7]~escores[,1]))[[5]][1]

```

```
if(valorpkruskal6>0.05){contadorkruskal6_contadorkruskal6+1}  
if(valorpanova6>0.05){contadoranova6_contadoranova6+1}  
}  
contadorkruskal1_100*contadorkruskal1/numitera  
contadoranova1_100*contadoranova1/numitera  
contadorkruskal2_100*contadorkruskal2/numitera  
contadoranova2_100*contadoranova2/numitera  
contadorkruskal3_100*contadorkruskal3/numitera  
contadoranova3_100*contadoranova3/numitera  
contadorkruskal4_100*contadorkruskal4/numitera  
contadoranova4_100*contadoranova4/numitera  
contadorkruskal5_100*contadorkruskal5/numitera  
contadoranova5_100*contadoranova5/numitera  
contadorkruskal6_100*contadorkruskal6/numitera  
contadoranova6_100*contadoranova6/numitera  
print(contadorkruskal1)  
print(contadoranova1)  
print(contadorkruskal2)  
print(contadoranova2)  
print(contadorkruskal3)  
print(contadoranova3)  
print(contadorkruskal4)  
print(contadoranova4)  
print(contadorkruskal5)  
print(contadoranova5)  
print(contadorkruskal6)  
print(contadoranova6)
```

9. ANEXO

Anexo 1. Tabela de Newell e MacFarlene

Nº.de respostas	DIFERENÇAS CRÍTICAS ENTRE OS TOTAIS DE ORDENAÇÃO ¹																			
	Nº DE AMOSTRAS – Nível de significância 5%										Nº DE AMOSTRAS – Nível de significância 1%									
	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12
8	10	14	18	22	26	30	34	39	43	47										
9	10	15	19	23	27	32	36	41	46	50										
10	11	15	20	24	29	34	38	43	48	53	13	18	23	28	33	38	44	49	54	59
11	11	16	21	26	30	35	40	45	51	56	14	19	24	30	35	40	46	51	57	63
12	12	17	22	27	32	37	42	48	53	58	15	20	26	31	37	42	48	54	60	66
13	12	18	23	28	33	39	44	50	55	61	15	21	27	32	38	44	50	56	62	68
14	13	18	24	29	34	40	46	52	57	63	16	22	28	34	40	46	52	58	65	71
15	13	19	24	30	36	42	47	53	59	66	16	22	28	35	41	48	54	60	67	74
16	13	19	25	31	37	42	49	55	61	67	17	23	30	36	43	49	56	63	70	77
17	14	20	26	32	38	44	50	56	63	69	17	24	31	37	44	51	58	65	72	79
18	15	20	26	32	39	45	51	58	65	71	18	25	31	38	45	52	60	67	74	81
19	15	21	27	33	40	46	53	60	66	73	18	25	32	39	46	54	61	69	76	84
20	15	21	28	34	41	47	54	61	63	75	19	26	33	40	49	55	63	70	78	86
21	16	22	28	35	42	49	56	63	70	77	19	27	34	41	49	56	64	72	80	85
22	16	22	29	36	43	50	57	64	71	79	20	27	35	42	50	58	66	74	82	90
23	16	23	30	37	44	51	58	65	73	80	20	28	35	43	51	59	67	75	84	92
24	17	23	30	37	45	52	59	67	74	82	21	28	36	44	52	60	69	77	85	94
25	17	24	31	38	46	53	61	68	76	84	21	29	37	45	53	62	70	79	87	96
26	17	24	32	39	46	54	62	70	77	85	22	29	38	46	54	63	71	80	89	98
27	18	25	32	40	47	55	63	71	79	87	22	30	38	47	55	64	73	82	91	100
28	18	25	33	40	48	56	64	72	80	89	22	31	39	48	56	65	74	83	92	101
29	18	26	33	41	49	57	65	73	82	90	23	31	40	48	57	66	75	85	94	103
30	19	26	34	42	50	58	66	75	83	92	23	32	40	49	58	67	77	86	95	105
31	19	27	34	42	51	59	67	76	85	93	23	32	41	50	59	69	78	87	97	107
32	19	27	35	43	51	60	68	77	86	95	24	33	42	51	60	70	79	89	99	108
33	20	27	36	44	52	61	70	78	87	96	24	33	42	52	61	71	80	90	100	110
34	20	28	36	44	53	62	71	79	89	98	25	34	43	52	62	72	82	92	102	112
35	20	28	37	45	54	63	72	81	90	99	25	34	44	53	63	73	83	93	103	113
36	20	29	37	46	55	63	73	82	91	100	25	35	44	54	64	74	84	94	105	115
37	21	29	38	46	55	64	74	83	92	102	26	35	45	55	65	75	85	95	106	117
38	21	29	38	47	56	65	75	84	94	103	26	36	45	55	66	76	86	97	107	118
39	21	30	39	48	57	66	76	85	95	105	26	36	46	56	66	77	87	98	109	120
40	21	30	39	48	57	67	76	86	96	106	27	36	47	57	67	78	88	99	110	121
41	22	31	40	49	58	68	77	87	97	107	27	37	47	57	68	79	90	100	112	123
42	22	31	40	49	59	69	78	88	98	109	27	37	48	58	69	80	91	102	113	124
43	22	31	41	50	60	69	79	89	99	110	28	38	48	59	70	81	92	103	114	126
44	22	32	41	51	60	70	80	90	101	111	28	38	49	60	70	82	93	104	115	127
45	23	32	41	51	61	71	81	91	102	112	28	39	49	60	71	82	94	105	117	128
46	23	32	42	52	62	72	82	92	103	114	28	39	50	61	72	83	95	106	118	130
47	23	33	42	52	62	72	83	93	104	115	29	39	50	62	73	84	96	108	119	131
48	23	33	43	53	63	73	84	94	105	116	29	40	51	62	74	85	97	109	121	133
49	24	33	43	53	64	74	85	95	106	117	29	40	51	63	74	86	98	110	122	134
50	24	34	44	54	64	75	85	96	107	118	30	41	52	63	75	87	99	111	123	135
55	25	35	46	56	67	78	90	101	112	124	31	43	54	66	79	91	104	116	129	142
60	26	37	48	59	70	82	94	105	117	130	32	45	57	69	82	95	108	121	135	148
65	27	38	50	61	73	85	97	110	122	135	34	46	59	72	86	99	113	126	140	154
70	28	40	52	64	76	88	101	114	127	140	35	48	61	75	89	103	117	131	146	160
75	29	41	53	66	79	91	105	118	131	145	36	50	64	78	92	106	121	136	151	166
80	30	42	55	68	81	94	108	122	136	150	37	51	66	80	95	110	125	140	156	171
85	31	44	57	70	84	97	111	126	140	154	38	53	68	83	98	113	129	144	160	176
90	32	45	58	72	86	100	114	129	144	159	40	54	70	85	101	116	132	149	165	181
95	33	46	60	74	88	103	118	133	148	163	41	56	71	87	103	120	136	153	169	186
100	34	47	61	76	91	105	121	136	151	167	42	57	73	89	106	123	140	157	174	191

Fonte: NEWELL E MAC FARLANE, 1987.

¹Se a diferença entre os totais de ordenação for maior ou igual ao número tabelado, existe diferença significativa entre as amostras, ao nível de significância observado.