Universidade Estadual da Paraíba
Centro de Ciências e Tecnologia
Departamento de Estatística

Edlaine Martins Barbosa

Teste de Hipótese e Aplicações

Campina Grande
Novembro de 2014

Edlaine Martins Barbosa

Teste de Hipótese e Aplicações

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Orientadora:
Dr^{a}. Divanilda Maia Esteves

Campina Grande
Novembro de 2014

B238t Barbosa, Edlaine Martins.
Testes de Hipóteses e Aplicações [manuscrito] / Edlaine Martins Barbosa. - 2014.
30 p. : il. nao

Digitado.
Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2014.
"Orientação: Profa. Dra. Divanilda Maia Esteves, Departamento de Estatística".

1. Teste de hipótese. 2. Paramétricos. 3. Não paramétricos. I. Título.

21. ed. CDD 519.5

Edlaine Martins Barbosa

Teste de Hipótese e Aplicações

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Aprovado em: 27 / novembro / 2014

Banca Examinador:

DMEsteves

Prof. Drª. Divanilda Maia Esteves
Orientadora

Prof. Ms. Juarez Fernandes de Oliveira
Universidade Estadual da Paraíba

Tiago Almeida de Oliveira

Prof. Dr. Tiago Almeida de Oliveira
Universidade Estadual da Paraíba

Agradecimentos

Primeiramente quero agradecer ao meu anjo protetor, o Senhor Jesus, que em cada momento se fez presente e tão presente que nunca duvidei de sua existência em minha vida.

Quero também agradecer a mainha e painho que me deram a oportunidade de nascer e crescer em um lar digno e fizeram de minhas vitorias as suas, sempre com muita dedicação e trabalho duro me ajudaram a superar as dificuldades que enfrentaria ao longo dessa trajet´oria.

A minha irmã Elaine, minha avó, primas e tios por serem essa família tão normal e única.

A minha UEPB, que guardarei no coração por todos os momentos lá vividos, assim como meus queridos professores Ana Cristina, Ricardo Olinda, Edwirde, Thiago, e em especial ao meu professor Juarez (juju) e minha orientadora Diana Maia que eu escolhi, não por acaso, para a finalizar esse projeto de vida, todos esses citados e muitos outros cada um a sua maneira, deram o sentido real da palavra mestre e me fizeram chegar hoje até aqui.

A todos os meus amigos de curso em especial a Rutineia Gomes, Eder Cabral, Fernanda Matias, Erasnilson Camilo, Fernando Evangelista e André Pereira pelos momentos de alegria proporcionado.

Ao pai dos meus filhos Ubiratan Patrício que teve uma importância fundamental nesta caminhada e serei eternamente grata a dedicação aos meninos enquanto eu tinha que sair para estudar.

Finalmente aos meus príncipes Everton e Daniel, pelo simples fato de existirem, sendo assim a minha maior fonte de inspiração e incentivo desde o primeiro ao último dia desse curso.

Resumo

Teste de hipótese é um método de inferência estatística usado em dados de um estudo científico sobre aspectos desconhecidos da população (que podem ser parâmetros ou mesmo a forma da distribuição). Portanto, é um procedimento que permite decidir se uma dada hipótese é ou não suportada pela informação fornecida pelos dados de uma amostra. Dessa forma, as hipóteses são elaboradas sobre os parâmetros da distribuição de uma ou mais variáveis aleatórias com alternativas que são testadas. Nesse contexto, o objetivo deste trabalho é apresentar inicialmente a teoria do teste de hipótese e posteriormente apresentar um exemplo aos quais foi possível verificar se os dados amostrais possuem evidências que contrariam ou não a hipótese estatística formulada. Nesse estudo, foram utilizados teste de comparação de variância e teste de igualdade de médias aos dados de crescimento dos dentes de porquinhos da Índia disponíveis no Software R, versão 3.0.3. Foi possível verificar que para os testes de comparação de variância não houve evidências para rejeitar a hipótese nula e para o teste de igualdade de médias, concluiu-se que o efeito dos dois tratamentos são estatisticamente diferentes entre si, com superioridade a suplementação via suco de laranja.

Palavras-chave: Teste de Hipótese, Paramétricos, Não Paramétricos.

Abstract

Hypothesis Test is a method of statistical inference data used in a scientific study population unknowns (parameters that can be the same or shape distribution). Therefore, it is a procedure to decide whether a given event is supported or not by the information supplied by data from a sample. Thus, the hypotheses are developed on the parameters of the distribution of one or more random variables with alternatives that are tested. In this context, the objective is to initially introduce the theory of hypothesis testing and then present an example to which it was possible to verify that the sample data have evidence that contradicts or not formulated statistical hypothesis. In this study, we used variance comparison test and for equality of means test to growth data from India's pigs teeth available in the software R, version 3.0.3. It can be seen that the tests for comparison of equal variance there was no evidence to reject the null hypothesis and that test of the comparasion mean, in which it was concluded that the effect of the two treatments are statistically diference to each other, with superiority for orange juice.

Key-words: Hypothesis testing, Parametric, Nonparametric.

Sumário

Lista de Figuras

Lista de Tabelas

1 Introdução

A inferência estatística é um conjunto de métodos e técnicas que permitem induzir, a partir de uma informação empírica proporcionada por uma amostra, qual é o comportamento de uma determinada população com um risco de erro medido em termos de probabilidade. O propósito da inferência estatística é tirar conclusões sobre uma população com base nos dados amostrais de uma população.

Os métodos paramétricos da inferência estatística podem se dividir basicamente em dois: métodos de estimação de parâmetros e testes de hipóteses. Os testes de hipóteses tem como objetivo comparar se determinado aspectos atribuídos a um parâmetro populacional, é compatível com a evidência empírica contida na amostra. Desta forma, teste de hipótese é um processo usado para avaliar a força da evidência de uma amostra em fornecer estrutura para fazer determinações relacionadas a população, em outras palavras, o teste de hipóteses proporciona um método fiável para a compreensão de como se pode extrapolar os resultados observados em uma amostra em estudo para uma população a partir do qual a amostra foi tirada. Cabe ao investigador formular uma hipótese específica, coletar dados com uma amostra, e usar estes dados para decidir se suportam a hipótese especifica do pesquisador (DAVIS et al, 2006).

Neste trabalho, será apresentado a teoria dos testes de hipóteses ressaltando os conceitos fundamentais sob a visão de diversos autores. Inicialmente será apresentado os conceitos de hipótese estatística, teste de hipótese unilateral e bilateral, tipos de erros, poder do teste, propriedades dos testes, nível de significância e valor P. O teste de hipótese é um caso especial do procedimento que usa a estatística amostral a cerca de um parâmetro referente a média, proporção, variância e o desvio padrão, incluindo os testes paramétricos e não paramétricos. Será apresentado ainda de forma específica no trabalho os testes para comparação de duas médias e duas variâncias com uma aplicação para cada método.

Na aplicação, o objetivo do teste de comparação de variância é comparar se existe diferença na dispersão em um tratamento com a vitamina C do suco de laranja e do ácido ascórbico no que diz respeito ao comprimento dos dentes das unidades amostrais utilizadas. Para o teste de igualdade de média, o objetivo é comparar se as médias do comprimento dos dentes dos 60 porquinhos da Índia submetidos a diferentes tratamentos, vitamina C e suco de laranja diferem estatísticamente entre sí.

2 Fundamentação Teórica

2.1 Hipóteses

Hipóteses estatísticas são suposições, especulações ou afirmações referente à distribuição de probabilidade de uma população de interesse, com o objetivo de ser demonstrada ou verificada posteriormente, constituindo uma suposição aceitável. Em geral, as hipóteses são elaboradas sobre os parâmetros da distribuição de uma ou mais variáveis aleatórias com alternativas que são testadas, sendo a distribuição amostral a região de rejeição, onde está incluído todos os valores possíveis que um teste estatístico pode assumir sob a hipótese nula. Logo, a região crítica é constituída por um conjunto de valores assumidos pela variável aleatória ou estatística de teste para os quais a hipótese nula é rejeitada. Assim, a região de rejeição é definida de modo que seja igual a probabilidade α sob a hipótese nula, da ocorrência de um valor da estatística de teste naquele subconjunto. Portanto, teste para a hipótese será sempre com base em resultados amostrais, sendo aceita ou rejeitada. Ela somente será rejeitada se o resultado da amostra for claramente improvável de ocorrer quando a hipótese for verdadeira.

Se o valor da estatística de teste calculado na amostra observada possui um valor pertencente a região de rejeição, rejeita-se a hipótese nula. Além desse critério, pode-se utilizar o p-valor. Quando a probabilidade associada a um valor observado da estatística é igual ou menor de que o valor previamente determinado de α, concluímos que a hipótese nula é falsa. Tal valor observado é chamado significativo. A hipótese de nulidade a ser comprovada é rejeitada sempre que ocorre um valor significativo. Valor significativo é um valor cuja probabilidade de ocorrência, sob a hipótese nula, é igual ou menor do que α (SIEGEL, 1975).

2.2 O que é um teste de hipótese

Teste de hipótese é um método de inferência estatística em que, utiliza dados de um estudo científico. 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. Por exemplo, pode-se estar interessado em determinar se uma moeda é honesta, se certas quantidades são independentes, ou se populações distintas são similares do ponto de vista probabilístico. Cada uma destas afirmações poderão contribuir na formulação de uma hipótese, as quais serão levadas à prova por meio de testes de hipóteses que é um procedimento usual da Inferência Estatística úteis na tomada de decisões (GUIMARÃES, 2008).Naghettini e Pinto (2007), relatam que, testar uma hipótese é recolher evidências nos dados amostrais, que justifiquem a rejeição ou a não rejeição de uma certa alegação sobre um parâmetro populacional ou sobre a forma de um modelo distributivo, tendo-se em conta as probabilidades de serem tomadas decisões incorretas.

De acordo com Scudino (2008), é fundamental considerar alguns conceitos sobre os testes de hipóteses, as quais serão vistas nas seções a seguir os seguintes testes de hipóteses:

- Hipótese nula (H_0) : é a alegação inicialmente assumida como verdadeira para a construção do teste. É o efeito, teoria, alternativa que estamos interessados em testar. A hipótese nula será rejeitada em favor da hipótese alternativa, somente se, a evidência da amostra sugerir que H_0 seja falsa.

- Hipótese alternativa (H_1) : é a afirmação contraditória a H_0, é o que consideramos caso a hipótese nula não tenha evidência estatística que a defenda. As duas conclusões possíveis de uma análise do teste de hipóteses são, então, rejeitar H_0 ou não rejeitar H_0.

Para a realização dos testes de hipóteses, é necessário seguir alguns passos. O procedimento básico de teste de hipóteses relativo ao parâmetro μ de uma população, será decomposto em 4 passos:

1. Definição das hipóteses:
$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

2. Identificação da estatística do teste e caracterização da sua distribuição.

3. Definição da regra de decisão, com a especificação do nível de significância do teste.

4. Cálculo da estatística de teste e tomada de decisão.

2.3 Hipóteses Unilateral e Bilateral

De acordo com Guimarães (2008), a média de uma população é uma das características mais importantes para os testes de hipótese unilateral e bilateral. No qual, é muito comum querer tomar decisões a seu respeito, por exemplo, quando são comparadas duas amostras ou dois tratamentos. Considere as seguintes hipóteses:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$$
$$ou$$
$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$$

A expressão acima em que μ_2 é uma constante conhecida define os chamados testes unilaterais, porque a região de rejeição está somente em uma das caudas da distribuição.

Para o teste bilateral considere a hipótese:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

Esta situação onde μ_2 é uma constante conhecida define os testes bilaterais, no qual a região de rejeição se distribui igualmente em ambas as caudas da distribuição.

Dessa forma, se houver interesse em mostrar que um parâmetro é significativamente superior ou inferior a um determinado valor, deve realizar um teste unilateral com uma única região de rejeição, do tamanho do nível de significância fixado. Porém, se houver interesse em mostrar que um determinado parâmetro é diferente de um determinado valor (sem especificar se inferior ou superior) têm-se que realizar um teste bilateral, no qual a região de rejeição será dividida em duas partes iguais, nas extremidades da curva do teste, em que cada região de rejeição terá metade do nível de significância (GUIMARÃES, 2008).

2.3.1 Região de rejeição ou região crítica

Região de rejeição ou região crítica é o conjunto de valores assumidos pela variável aleatória ou estatística de teste para os quais a hipótese nula é rejeitada. A sua área é igual ao nível de significância, e sua direção é a mesma da hipótese alternativa. A Figura

1 mostra uma representação da região crítica de um teste junto da curva da distribuição da estatística do teste.

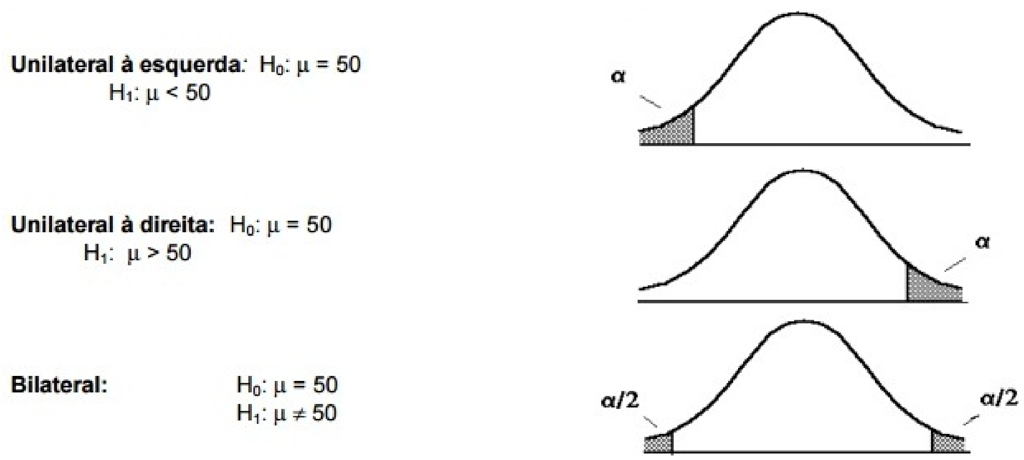

Figura 1: Representação da região crítica de testes uni e bilaterais, conjuntamente com a curva da distribuição da estatística do teste. (Fonte: Kato, 2008)

De acordo com Siegel e Castellan (2006), se o valor da estatística de teste calculado na amostra observada possui um valor que caia na região de rejeição, rejeitamos H_0. Portanto, além do critério da região de rejeição, pode-se utilizar o valor-p, em que observa-se se a probabilidade associada a um valor observado da estatística é igual ou menor ao valor determinado por α, concluímos que H_0 é falsa.

2.4 Tipos de erros

Se uma hipótese for rejeitada quando deveria ser aceita, diz-se que foi cometido o erro do tipo I. Se, por outro lado, for aceita uma hipótese que deveria ser rejeitada, diz-se que foi cometido um erro do tipo II. Em ambos os casos ocorreu uma decisão errada ou um erro de julgamento (SCUDINO, 2008). Sendo assim, há dois possíveis tipos de erros mais comuns quando realizamos um teste estatístico:

- Erro do tipo I : é o erro ao rejeitar H_0 quando, na realidade, H_0 é verdadeira. A probabilidade de cometer este erro do tipo I é designada por α (nível de significância). O erro do tipo I equivale a concluir que o tratamento é eficaz quando na verdade ele não é.

- Erro do tipo II : é o erro ao aceitar H_0 quando, na realidade, H_0 é falsa. A probabilidade de cometer este erro do tipo II é designada por β.

Para um teste de hipóteses, fixa-se a probabilidade máxima de cometer o erro tipo I. Essa probabilidade máxima é chamada de nível de significância, e pode ser elaborada de acordo com o estudo em questão. O valor da probabilidade de se obter o efeito observado, dado que a hipótese nula é verdadeira, é chamado de p-valor.

O p-valor é a probabilidade de que a estatística do teste (como variável aleatória) tenha valor extremo em relação ao valor observado (estatística) quando a hipótese H_0 é verdadeira. Onde também é conhecido como o menor valor do nível de significância para o qual rejeitamos H_0. Desta forma, se o nível de significância (α) sugerido para o teste for menor que o p-valor não rejeitamos a hipótese H_0. De acordo com isto, numa análise de resultados do teste de hipótese o p-valor nos fornece uma idéia de quanto os dados contradizem a hipótese nula. Além disso, ele permite que diferentes experimentadores utilizem seus respectivos níveis de significância nas avaliações.

2.5 Poder do teste

Segundo Filho (2009), o poder do teste é a probabilidade de rejeitar corretamente a hipótese nula, quando esta deve ser rejeitada (evitando o erro tipo II), ou seja, está relacionada com a capacidade do teste em identificar diferenças. Geralmente, o poder do teste é utilizado para interpretar resultados de análises estatísticas inferenciais em que a diferença encontrada não foi estatísticamente significativa.

Considere T um teste estatístico com região crítica C para avaliarmos hipóteses a respeito do parâmetro θ. A função poder do teste é a probabilidade de rejeitarmos H_0 dado o valor de θ. Neste caso, temos que

$$\pi(\theta) = \mathbb{P}[\text{rejeitar } H_0|\theta] = \mathbb{P}[T \in C|\theta],$$

para todo valor de θ.

Fossaluza (2008) relata que, o objetivo do poder do teste é conhecer o quanto o teste de hipóteses controla um erro do tipo II, ou qual a probabilidade de rejeitar a hipótese nula se realmente for falsa. Em geral, é aceito que o poder do teste deva ser de 80% ou mais de encontrar uma diferença estatísticamente significativa quando esta existe mesmo. Embora, a significância estatística seja o parâmetro de controle do teste a ser definido em primeiro lugar. Nesse contexto, busca-se uma equivalência entre as medidas, pois a significância é diretamente proporcional ao poder do teste.

Sabe-se que, o poder de um teste de hipóteses é afetado por três fatores, logo definidos abaixo:

- Tamanho da amostra: Mantendo todos os outros parâmetros iguais, assim quanto maior o tamanho da amostra, maior o poder do teste.
- Nível de Significância: Quanto maior o nível de significância, maior o poder do teste. Dessa forma, ao aumentar o nível de significância, reduz a região de aceitação. Sendo assim, tem-se maior chance de rejeitar a hipótese nula. Isto significa que é menor a chance de cometer um erro do tipo II. Então, o poder do teste aumenta.
- O verdadeiro valor do parâmetro a ser testado: Quanto maior a diferença entre o verdadeiro valor do parâmetro e o valor especificado pela hipótese nula, maior será o poder do teste.

Considere $P(\text{erro tipo I}) = P(\text{rejeitar } H_0|H_0 \text{ é verdadeira}) = \alpha$ e a $P(\text{erro tipo II}) = P(\text{aceitar } H_0|H_0 \text{ é falsa}) = \beta$. Assim, o poder de um teste estatístico como a probabilidade do teste rejeitar H_0 quando H_0 é realmente falsa, ou seja, o poder desse teste será igual a $1 - \beta$.

2.6 Teste mais poderoso

Segundo Kato (2008), o teste de hipótese mais poderoso pode ser definido considerando-se um teste que H_0 pode ser uma hipótese simples ou composta e H_1 é sempre uma hipótese composta. Sendo assim, precisa-se especificar um teste cuja probabilidade máxima de rejeitar H_0 quando ela é verdadeira seja α e que ao mesmo tempo maximize a probabilidade de rejeitar H_0 quando ela é falsa. Em geral, os testes mais poderosos só existem em situações especiais, por exemplo quando a distribuiçao de probabilidade pertence a família exponencial.

Assim, seja $X_1, ..., X_n$ uma amostra aleatória de densidade $f_0(x)$ ou da densidade $f_1(x)$. Logo, $f_0(x) = f(x; \theta_0)$ e $f_1(x) = f(x; \theta_1)$. Então $X_1, ..., X_n$ é uma amostra aleatória de um dos dois membros da família paramétrica $f(x; \theta) : \theta = \theta_0$ ou $\theta = \theta_1$, isto é, $\Theta = \{\theta_0, \theta_1\}$ é um espaço paramétrico com apenas dois pontos, θ_0 e θ_1 conhecidos.

Definição: Um teste com função poder $\pi^*(\theta)$ para testar $H_0 : \theta = \theta_0$ contra $H_1 : \theta = \theta_1$ é dito ser um teste mais poderoso (MP) de tamanho $\alpha \, (0 \leq \alpha \leq 1)$ se e somente se

- $\pi^*(\theta_0) = \alpha$,

- $\pi^*(\theta) \geq \pi(\theta)$ para qualquer outro teste com função poder $\pi(\theta)$ para o qual $\pi(\theta) \leq \alpha$.

Segundo Esteves e Oliveira (2008), um teste é mais poderoso de tamanho α se tem tamanho α e se entre todos os outros testes de tamanho menor ou igual a α, ele tem o maior poder. Ou seja, teste é mais poderoso de tamanho α se ele tem erro de tipo I igual a α e tem o menor tamanho de erro de tipo II, uma vez que $1 - \pi(\theta_1) = \mathbf{P}(\text{aceitar } H_0 | H_0$ é falsa), que é o tamanho do erro de tipo II. Com base nisto, a justificativa para fixar o erro de tipo I em α aparece daquelas situações onde as duas hipóteses são formuladas de tal maneira que erro de tipo I é mais sério do que o erro tipo II. Logo, precisa-se ter certeza de que este erro é realmente pequeno.

2.7 Nível de significância e p-valor

De acordo com Scudino (2008), para testar uma hipótese formulada, a probabilidade máxima com o qual se pode ocorrer o Erro do tipo I é denominada nível de significância do teste. Em geral, o nível de significância é representado por α e também é especificado antes da obtenção das amostras e das hipóteses, de modo que os resultados obtidos na amostra não influenciem a escolha. Geralmente, em um teste de hipóteses usa-se $\alpha = 0{,}05$ ou $\alpha = 0{,}01$, mas nada justifica formalmente a utilização destes valores em particular. Portanto, se escolhido o índice de 0,01, então existe 1 chance em 100, da hipótese ser rejeitada quando ela é verdadeira. Da mesma maneira pode-se dizer que existe uma confiança de 99% de que se tome a decisão correta.

Scudino (2008), salienta que na resposta dos testes de hipóteses, um valor é comparado com o valor do poder do teste, ou seja, p-valor. O p-valor (nível de significância observado) é o menor nível de significância em que H_0 seria rejeitada, quando um procedimento de teste específico é usado em um determinado conjunto de dados. Assim, quando p-valor $\leq \alpha$ implica na rejeição de H_0 no nível α. Ou se p-valor $> \alpha$ implica na não rejeição de H_0 no nível α. Então, em vários estudos as respostas poderão vir referenciando o nível de significância ou p-valor.

Podemos ver que, o p-valor é a probabilidade de observar resultados tão extremos quanto aqueles que foram obtidos se a hipótese nula for verdadeira. A idéia é que se o p-valor for grande ele fornece evidência de que H_0 é verdadeira, enquanto que um p-valor pequeno indica que existe evidência nos dados contra H_0.

2.8 Testes de Hipóteses

Uma hipótese estatística é uma conjectura sobre um parâmetro desconhecido de uma população ou a forma da distribuição de uma característica em estudo na população (CASELLA E BERGER, 2010). Pode ser sobre a forma da distribuição ou sobre o valor de parâmetros desconhecidos da população, conhecida ou não a forma da distribuição. Tal conjectura pode ou não ser verdadeira. A verdade ou falsidade nunca pode ser confirmada, a menos que observássemos toda a população, o que normalmente é impraticável. Então através da informação fornecida por uma amostra que se rejeita ou não a hipótese formulada.

Conforme em Guimarães (2008), existem inúmeros testes estatísticos tanto paramétricos quanto não paramétricos. Os testes paramétricos exigem que seja verificada a pressuposição de que os dados coletados sejam oriundos de alguma distribuição com o objetivo de testar hipóteses acerca de parâmetros, com base em dados amostrais. Porém, os testes não-paramétricos não fazem essa exigência e por isso são considerados menos consistentes, sendo então, uma alternativa a ser usada caso os pressupostos distribucionais não sejam observadas ou, ainda, quando o tamanho da amostra não é suficientemente grande. Neste caso, as hipóteses não são formuladas em termos de parâmetros, já que não há preocupação com a distribuição que os dados seguem. Para cada tipo de situação existem testes específicos a serem utilizados.

Nas próximas seções só será tratado o estudo de testes de hipóteses para igualdade de médias com variâncias desconhecidas e testes para comparação de variâncias, não se esgotando aqui todo o assunto a cerca de testes de hipóteses, para maiores detalhes sobre os diversos testes de hipóteses não abordados neste texto sugere-se consultar Casella e Berger (2010).

2.9 Teste de igualdade de médias

2.9.1 Teste de igualdade de médias para variâncias desconhecidas

Considerando, que as variâncias das populações são iguais, porém, desconhecidas, ou seja, $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Então, para testar a igualdade das médias, considera-se a variável

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Em que a variável tem distribuição t de Student com $n_1 + n_2 - 2$ graus de liberdade. Sendo, S_p o desvio padrão agrupado que é dado por

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

em que,

- S_1^2: variância da amostra proveniente da população 1.
- S_2^2: variância da amostra proveniente da população 2.

Para realizar o teste para igualdade de duas médias com variâncias iguais, porém desconhecidas, devemos realizar os seguintes passos:

- Estabelecer uma das hipóteses, por exemplo:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases} \quad \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases} \quad \text{ou} \quad \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$$

- Fixar o nível de significância α.
- Determinar a região crítica.

2.9.2 Regra de decisão

Se o valor da estatística do teste cair dentro da região crítica, rejeita-se H_0. Ao rejeitar a hipótese nula (H_0) existe uma forte evidência de sua falsidade.

Ao contrário, quando aceitamos, dizemos que não houve evidência amostral significativa no sentido de permitir a rejeição de H_o.

2.10 Teste para comparação de variância

Em Scudino (2008), quando se tem problemas com duas ou mais amostras de distribuições normais é natural que se tenha interesse em comparar as variâncias populacionais. Considerando uma amostra $X_1, ..., X_n$, onde $X_i \sim N(\mu; \sigma^2)$, pode-se testar hipóteses acerca do valor de σ^2. Neste caso, a distribuição F é utilizada para testar as hipóteses associadas. A razão $F = \frac{S_1^2}{S_2^2}$, denominadas de razões de variâncias, são valores de uma variável aleatória com distribuição F. Esta importante distribuição contínua depende de dois parâmetros chamados graus de liberdade do numerador e do denominador. Os valores desses parâmetros são $n_1 - 1$ e $n_2 - 1$ se calcularmos $\frac{S_1^2}{S_2^2}$.

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n retirada de uma população normal $N(\mu, \sigma^2)$. Suponha que desejamos testar uma hipótese sobre a variância σ^2 desta população.

Admitindo-se que a estatística $Q = \frac{(n-1)S^2}{\sigma^2}$ tem distribuição qui-quadrado com $n-1$ graus de liberdade. Denotamos $Q \sim \chi^2_{(n-1)}$. Para executar este tipo de teste, pode-se seguir os passos:

1. Estabelecer uma das hipóteses (bilateral, unilateral é direita ou unilateral é esquerda)
$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases} \quad \begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 > \sigma_0^2 \end{cases} \quad \text{ou} \quad \begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 < \sigma_0^2 \end{cases}$$
2. Fixar o nível de significância α.
3. Determinar a região crítica.
    - Se o teste é bilateral, deve-se determinar os pontos críticos $F_{\alpha/2}$ e $F_{1-\alpha/2}$ da distribuição F com $n_1 - 1$ graus de liberdade no numerador e $n_2 - 1$ graus de liberdade no denominador usando a tabela da distribuição Fisher-Snedecor de modo que $\mathbb{P}[F < F_{\alpha/2}] = \mathbb{P}[F > F_{1-\alpha/2}] = \alpha/2$.
    - Se o teste é unilateral á direita, determina-se o ponto $F_{1-\alpha}$ tal que $\mathbb{P}[F > F_{1-\alpha}] = \alpha$
    - Se o teste é unilateral á esquerda, determina-se o ponto F_α tal que $\mathbb{P}[F < F_\alpha] = \alpha$
4. Calcular, sob a hipótese nula, o valor
$$F_{obs} = \frac{S_1^2}{S_2^2}$$

3 Aplicação

Nesta aplicação, será utilizado teste de comparação de variância e teste de igualdade de médias, aos quais, os testes serão aplicados em um conjunto de dados com duas variáveis. O conjunto de dados em questão é sobre o crescimento dos dentes de porquinhos da Índia, esses dados contém 30 observações, para os testes aplicados têm-se o interesse de saber qual tratamento é mais eficiente no crescimento dos dentes dos porquinhos da Índia, por meio do uso de vitamina C com suco de laranja e com ácido ascórbico em três níveis de dosagem de vitamina C (0,5, 1 e 2 mg). Os dados "ToothGrowth", está disponível no pacote datasets (versão 3.0.3) do Software R versão 3.1.2 (R CORE TEAM, 2013), este banco de dados foi trabalho por BLISS (1925), CRAMPTON (1947), MCNEIL (1977).

As estatísticas de teste devem ter sua distribuição conhecida e ela deve depender do parâmetro que está sendo testado.

3.1 Dados

Deseja-se verificar o crescimento dos dentes de porquinhos da Índia, por meio do uso da vitamina C do suco de laranja e do ácido ascórbico. Será feito o teste de comparação de variância para saber se há diferença no tratamento com a vitamina C do suco de laranja e do ácido ascórbico submetido aos animais, em seguida será feito o teste de igualdade de médias para verificar se há diferença entre as médias do comprimento dos dentes dos animais que foi utilizado o suco de laranja e o comprimento dos dentes dos animais que foi utilizado o ácido ascórbico ao nível $\alpha = 0,05$ de significância. A Tabela 1 apresenta a relação das 30 observações, com as respectivas medidas do comprimento dos dentes dos animais em questão utilizando o suco de laranja e o ácido ascórbico.

Tabela 1: Relação das medidas do comprimento dos dentes de porquinhos da Índia, utilizando o suco de laranja (SL) e o ácido acórbico (AA)

X_1 (SL)	X_2 (AA)
4,2	15,2
11,5	21,5
7,3	17,6
5,8	9,7
6,4	14,5
10,0	10,0
11,2	8,2
11,2	9,4
5,2	16,5
7,0	9,7
16,5	19,7
16,5	23,3
15,2	23,6
17,3	26,4
22,5	20,0
17,3	25,2
13,6	25,8
14,5	21,2
18,8	14,5
15,5	27,3
23,6	25,5
18,5	26,4
33,9	22,4
25,5	24,5
26,4	24,8
32,5	30,9
26,7	26,4
21,5	27,3
23,3	29,4
29,5	23,0

Para iniciar a análise tem-se a tabela com estatísticas descritivas (Tabela 2) que nos permite melhor visualizar os dados. Pode-se observar que os valores mínimos de crescimento dos dentes quando suplementados com vitamina C são menores que os quando suplementados com Suco de laranja, os valores máximos são próximos para os dois tipos de suplementação.

Tabela 2: Estatística descritiva do comprimento dos dentes de porquinhos da Índia, utilizando o suco de laranja (SL) e o ácido acórbico (AA)

Estatísticas	vitamina C	suco de laranja
Minimo	4,20	8,20
1° Qu.	11,20	15,53
Mediana	16,50	22,70
Media	16,96	20,66
3° Qu.	23,10	25,73
Máximo	33,90	30,90

Após uma análise exploratória inicial recorre-se a um diagrama de dispersão para visualizar os dados nas diferentes doses, com gráficos para vitamina C e Suco de laranja em verde e azul, respectivamente.

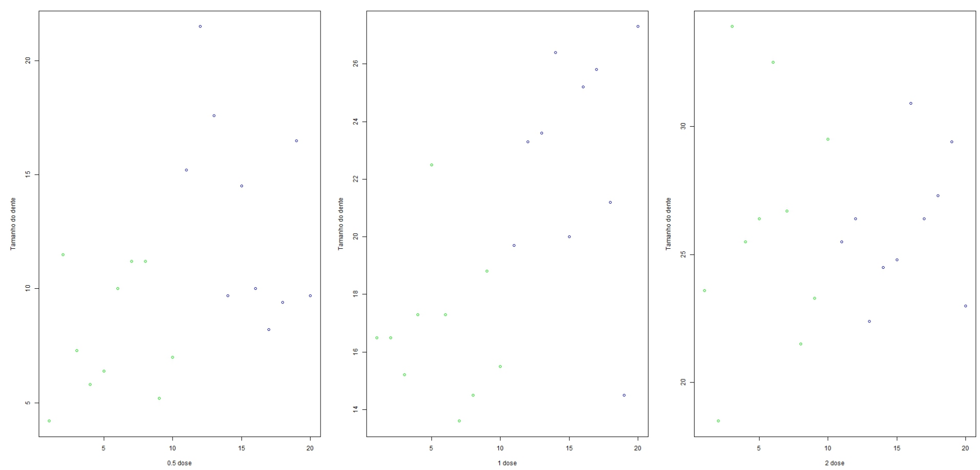

Figura 1: Diagrama de dispersão para o crescimento dos dentes de 60 porquinhos da índia em diferentes doses e dois tipos de suplementação: vitamina C em azul e Suco de Laranja em verde.

Pode-se perceber que na dose 0,5 e 1,0 mg o suco de laranja tem uma média maior que a vitamina C para o crescimento dos dentes, já na dose 2,0 mg as médias são similares.

Na Figura 2, tem-se os tamanhos dos dentes em diferentes suplementações (suco de laranja e vitamina C), com as cores diferenciando as doses (0,5; 1,0 e 2.0) em azul, verde e vermelho, respectivamente.

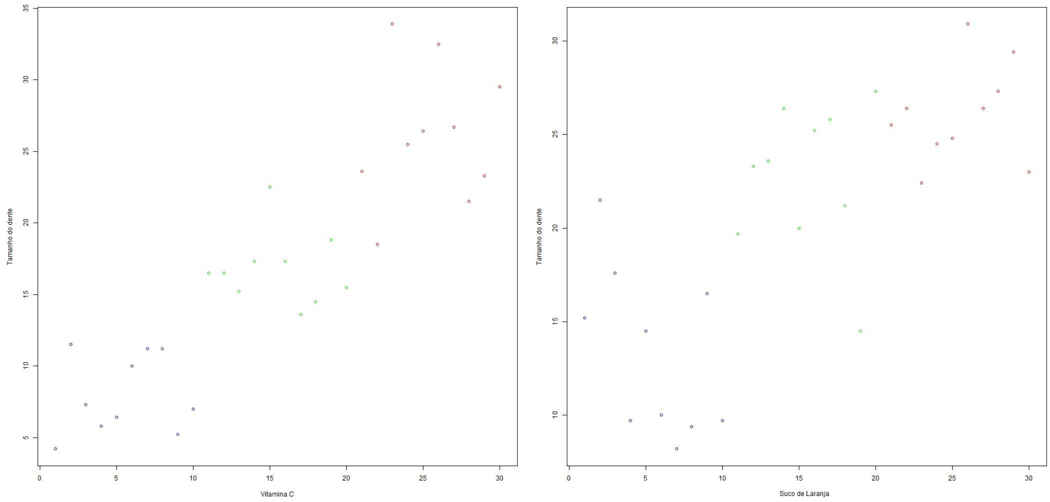

Figura 2: Diagrama de dispersão para o crescimento dos dentes de 60 porquinhos da índia em diferentes tipos de suplementação com as doses (0,5; 1,0 e 2,0) em azul, verde e vermelho, respectivamente.

Pela Figura 3, percebe-se que o suco de laranja tem uma resposta ao crescimento dos dentes dos porquinhos da Índia melhor nas 3 doses analisadas em comparação a vitamina C, porém na dose 2,0 mg os valores são praticamente iguais.

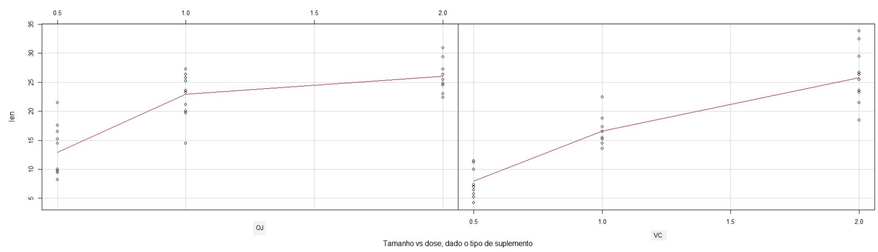

Figura 3: Gráfico com suavização não paramétrica nos dois tipos de suplementação: Suco de laranja (OJ) e vitamina C (VC).

Para que se possa afirmar sobre diferenças entre dos comprimento dos dentes dos porquinhos da Índia deve-se realizar testes de hipóteses.

Teste de comparação de variâncias

Definindo as variáveis,

X_1: comprimento dos dentes que usaram suco de laranja,.

X_2: comprimento dos dentes que usaram ácido ascórbico.

Neste caso, queremos testas as hipóteses:

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

Com $S_1^2 = 66,05$ e $S_2^2 = 42,58$ temos que,

$$F_{obs} = \frac{66,05}{42,58} = 1,55$$

Observando a tabela da distribuição Fisher-Snedecor com 29 graus de liberdade no numerador e 29 no denominador temos que $F_{(29;29;0,05)} = 1,875$.

Conclusão:

Como $F_{obs} = 1,55 < F_{(29;29;0,05)} = 1,875$ logo, existem evidências amostrais suficientes para aceitarmos a hipóteses nula $H_0 : \sigma_1^2 = \sigma_2^2$ e afirmar que as variâncias do comprimento dos dentes dos animais submetidos ao tratamento de vitamina C com suco de laranja e ácido ascórbico são estatísticamente iguais, ou seja, não existe diferença na variabilidade do crescimento utilizando o suco de laranja ou ácido ascórbico.

Teste de igualdade de médias

Assumi-se que variáveis X_1 e X_2 são normalmente distribuídas com variâncias desconhecidas.

Queremos testar a hipótese que μ_1 é o número médio do comprimento dos dentes dos animais que foram submetidos ao suco de laranja e μ_2 o número médio do comprimento dos dentes dos animais que foram submetidos a ácido ascórbico, ou seja, queremos testar as hipóteses:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

Teste t para duas médias com $\alpha = 5\%$

$$S_p = \frac{(30-1)66,05 + (30-1)42,58}{30+30-2} = 22,044$$

$$t_{calc} = \frac{16,96 - 20,7}{\sqrt{22,044\left(\frac{1}{30} + \frac{1}{30}\right)}} = -3,74$$

$$t_{\frac{5\%}{2}(58)} = 2,02$$

Conclusão:

Como $|t_{calc}| > t_{tab}$, logo, há evidências amostrais suficientes para não aceitar a hipótese nula $H_0 : \mu_1 = \mu_2$ ao nível de 5% de significância. Concluímos então que as médias do comprimento dos dentes dos porquinhos da Índia que tomaram suco de laranja e ácido ascórbico são estatísticamente diferentes.

4 Conclusão

Pode-se concluir que testes de hipóteses trata-se de um processo de dedução que tem diversas aplicações em diferentes áreas. Neste trabalho foi abordado a teoria com exemplos baseados em dois tipos de testes de hipóteses que pode se fazer mediante a inúmeras situações do dia a dia. No entanto, é importante ressaltar que o termo hipótese estatística é uma afirmação feita em relação a um parâmetro populacional desconhecido.

Contudo, a ideia básica que se tem é que, a partir de uma amostra da população pode-se estabelecer uma regra de decisão chamada de teste, segundo a qual rejeita-se ou não a hipótese proposta. Normalmente existe uma hipótese que é mais importante para o pesquisador denotada por H_0

Com base na aplicação foi possível verificar que para o teste de comparação de variância, não houve evidência para rejeitar a hipótese nula, ou seja, o resultado do tratamento com a vitamina C do suco de laranja e o do ácido ascórbico referente ao comprimento dos dentes dos porquinhos da Índia são estatísticamente iguais ao nível de 5% de significância. Logo, para o teste de igualdade de médias verificou-se ao nível de 5% de significância que há evidência estatística para rejeitar a hipótese nula, ou seja, aceita-se H_1 concluindo que as médias dos tratamentos da vitamina C do suco de laranja e do ácido ascórbico são estatísticamente diferentes entre si e que os porquinhos submetidos ao suco de laranja apresentam maior crescimento de seus dentes.

5 Bibliografia

BLISS, C. I. The Statistics of Bioassay. Academic Press, 1952.

CASELLA, G., BERGER, R. L. **Statistical inference**. Duxbury Press, 2008.

CRAMPTON, E. W. The growth of the odontoblast of the incisor teeth as a criterion of vitamin C intake of the guinea pig. The Journal of Nutrition 33(5). 1947.

DAVIS, R.B., MUKAMAL, K.J., Hypothesis Testing Means. **Statistical Primer for Cardiovascular Research**. Circulation. 2006;114:1078-1082.

ESTEVES, Gustavo. Henrique; OLIVEIRA, Juarez. Fernandez. **Teste de Hipóteses**. Departamento de Estatística. UEPB. Campina Grande - PB, 2008.

FILHO, G. M. F. **Significância estatística versus poder estatístico**. João Pessoa: Semioblog, 2009.

FOSSALUZA, V. **Teste de hipótese em eleições majoritárias**. São Paulo, 2008.

GUIMARÃES, P. R. B. **Métodos Quantitativos Estatísticos**. 1a ed. Curitiba: Iesde, 2008.

KATO, J. **Estatística Aplicada a Ciências Humanas**. 2a ed. São Paulo: Harbra, 2008.
MCNEIL, D. R. **Interactive Data Analysis**. New York: Wiley. 1977.

NAGHETTINI, M; PINTO, E. J. A. **Hidrologia Estatística**. 1a ed. Belo Horizonte, 2007.

R: A Language and Environment for Statistical Computing. **R Core Team**, R Foundation for Statistical Computing. Vienna. Austria. 2013. ISBN 3-900051-07-0, http://www.R-project.org/

SCUDINO, P. A.; A **Utilização de Alguns Testes Estatísticos para Análise da Variabilidade do Preço do Mel nos Municípios de Angra dos Reis e Mangaratiba, Estado do Rio de Janeiro**. Seropádica - RJ, 2008.

SIEGEL, S.; CASTELLAN JR, N. J. **Estatística não-paramétrica para ciências do comportamento**; [Tradução: CARMONA, S. I. C.], 2a ed. Porto Alegre: Artmed, 2006.

SIEGEL, Sidney. Estatística Não-paramétrica Para as Ciências do Comportamento. São Paulo: McGraw-Hill, 1975