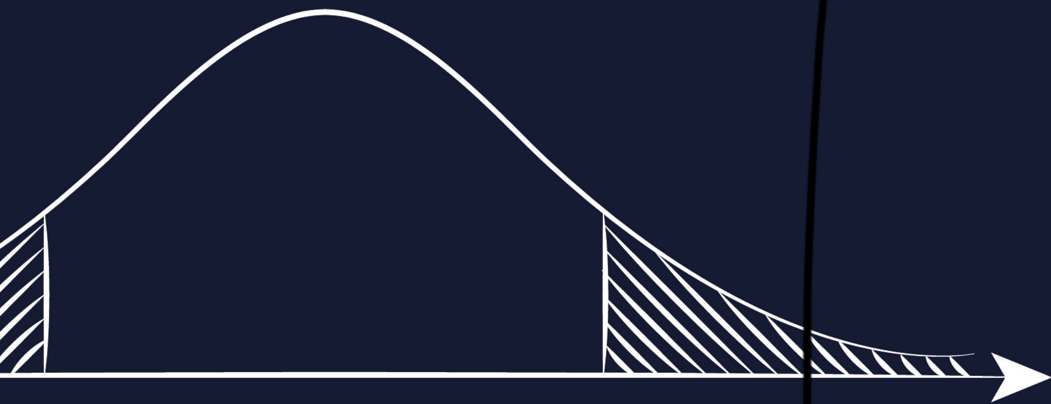
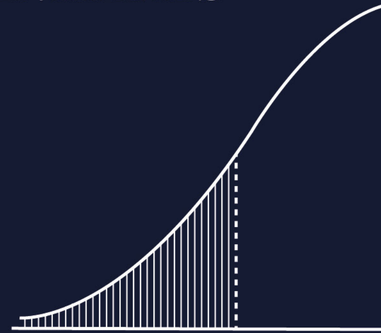
TESTES DE HIPÓTESES ESTATÍSTICAS



JANILSON PINHEIRO DE ASSIS

ROBERTO PEQUENO DE SOUSA

PAULO CÉSAR FERREIRA LINHARES



edufersa

Janilson Pinheiro de Assis
Roberto Pequeno de Sousa
Paulo César Ferreira Linhares

TESTES DE HIPÓTESES ESTATÍSTICAS

10 Anos
edufersa
2020

Dados Internacionais da Catalogação na Publicação (CIP)
Editora Universitária (EdUFERSA)

A848t Assis, Janilson Pinheiro de
Testes de hipóteses estatísticas / Janilson Pinheiro de Assis,Roberto Pequeno de Sousa, Paulo César Ferreira Linhares–Mossoró: EdUFERSA, 2020.
182 p. :Il.

ISBN: 978-65-87108-06-3
E-ISBN: 978-65-87108-05-6

1. Estatística. 2. Teste de hipótese. I. Sousa, Roberto Pequeno de. II. Linhares, Paulo César Ferreira. III. Título.

RN/UFERSA/EdUFERSA CDD: 519.542

Bibliotecário-Documentalista
Mário Gaudêncio (CRB15-476)

Editora filiada:

ABEC BRASIL Associação Brasileira de Editores Científicos

ABEU Associação Brasileira das Editoras Universitárias

Av. Francisco Mota, 572 (Campus Leste, Centro de Convivência) Costa e Silva Mossoró-RN
59.625-900 | +55 (84) 3317-8267 | edufersa.ufersa.edu.br | livraria.edu.br | edufersa@ufersa.edu.br

PREFÁCIO

A teoria da inferência estatística que surge da necessidade do pesquisador fazer tomadas de decisão e realização de generalizações acerca de parâmetros populacionais, bem como sobre a natureza da distribuição de probabilidade desses universos mediante obtenção de conclusões sobre a população partindo-se dos resultados estatísticos através de estimadores ou estatísticas de amostras principalmente aleatórias, faz do teste de hipótese estatístico uma das mais importantes área da teoria estatística. Testes de hipóteses são técnicas ou processos com objetivos de verificação da igualdade ou desigualdade entre duas ou mais medidas, entre valores esperados ou previstos e valores ocorridos, ou entre estatísticas de dois ou mais conjuntos separados no tempo e no espaço, pode-se testar, por exemplo, a eficácia de uma vacina na cura da gripe aviária num lote de aves como galinhas poedeiras, mediante o suporte do cálculo de probabilidades.

O avanço da ciência e a revolução industrial, principalmente a partir do século XVIII e XIX, o surgimento da computação eletrônica na primeira metade do século XX, o aumento da população humana, o avanço da medicina, o surgimento de guerras, de doenças, pragas, etc. a demanda cada vez maior por água, alimentos e energia, fez com que os cientistas se esforçassem em descobrir novas técnicas de análise de dados oriundos de levantamentos e experimentos, com o objetivo de beneficiar a sociedade com suas descobertas, principalmente devido ao desafio de se estudar populações grandes e infinitas, o que muitas vezes torna esta tarefa difícil ou mesmo impossível e cara e demorada. Restou a escolha de tratar as populações mediante as condições de incerteza mediante o uso de amostras oriundas destes universos de interesse, e sendo assim com o desafio das condições de incerteza foi feito o uso da teoria de probabilidades para medir o risco de suas conclusões de forma, econômica, precisa e com elevada exatidão. E assim surgiu a importante ferramenta de inferência estatística denominada de testes de hipóteses ou de significância, paramétrico e não paramétrico, univariado e multivariado, que proporcionou a ciência um grande apoio nas suas descobertas, norteando os pesquisadores nas suas comparações, e atualmente continua se expandindo de forma espetacular principalmente com a utilização de softwares e pacotes estatísticos cada vez mais poderosos e de largo alcance, o que produziu uma imensa gama de artigos científicos sobre este assunto, bem como um grande volume de publicações de excelentes livros com menor ou maior profundidade da teoria .

Este livro teve origem da experiência do primeiro autor, como professor da disciplina de estatística, ministrada mediante o uso de apostilas e diversos materiais didáticos em cursos de graduação ministrados na Escola Superior de Agricultura de Mossoró (ESAM) e atualmente Universidade Federal Rural do Semi-Árido (UFERSA) durante mais de 30 anos. O texto apresenta uma importante ferramenta de inferência estatística denominada de testes de hipótese, paramétrico para a média, diferença entre médias, dados emparelhados, variância, desvio padrão, razão entre variâncias, proporção relativa, diferença entre proporções percentuais quando esta diferença é hipoteticamente nula e também não supostamente nula e o outro grupo dos testes não paramétrico sendo este último grupo representado pelo teste do qui-quadrado nas suas quatro aplicações práticas as quais são: o teste de ajustamento, o de aderência, o de independência e por último o teste de homogeneidade. Vale lembrar que mais recentemente este assunto ou tema passou a ser denominado na literatura estatística como sendo teoria da decisão estatística ou simplesmente teoria da decisão às vezes juntamente com os intervalos de confiança outras vezes sem os intervalos fiduciais.

O texto é apresentado de forma clara e objetiva com diversos exercícios de aplicação e formulários, bem como o uso de tabelas para facilitar o aprendizado e a compreensão deste assunto. O leitor terá facilidade na assimilação do assunto, pois os capítulos seguem uma ordem e sequencia lógica e bem didática.

São mostrados no final do livro diversos exemplos, tabelas auxiliares, o alfabeto grego com todas as letras, formulários dos principais tipos de testes de hipóteses paramétricos numa tabela resumo, bem como referências bibliográficas para o leitor interessado em se aprofundar mais no assunto.

Esta obra tem como principal objetivo atingir o público dos alunos de graduação dos cursos de ciências exatas, engenharias, administração, ciências contábeis, agronomia, zootecnia, engenharia de pesca, biologia, ecologia, medicina veterinária, medicina, bem como os demais cursos que tenham alguma abordagem quantitativa dos seus conteúdos programáticos. No entanto pesquisadores, técnicos, dirigentes de empresas, profissionais que utilizam a estatística em suas atividades profissionais, bem como alunos de pós-graduação podem se nortear e devem se beneficiar com a leitura deste livro.

Mossoró, RN, Brasil, Setembro de 2020.

Professor Dr. Janilson Pinheiro de Assis
Professor Dr. Roberto Pequeno de Sousa
Pesquisador Dr. Paulo César Ferreira Linhares

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

1 INTRODUÇÃO

Em todas as áreas práticas da pesquisa científica ou acadêmica, nas mais diferentes linhas de atuação humana como nas ciências físicas, biológicas, humanas e sociais, nas engenharias, na indústria, no comércio, na geoestatística, etc., o pesquisador ou investigador é frequentemente solicitado a tomar decisões acerca das populações ou universo estatístico (inferência estatística), baseadas nas informações das amostras, ou seja, testar valores de certos parâmetros ou entes desconhecidos desses universos de estudo que são medidas ou grandezas que caracterizam aspectos importantes das populações, tais como localização ou tendência central, variabilidade, dispersão ou escala, assimetria, e achatamento ou curtose, como por exemplo, peso médio de animais recém-nascidos, vida mediana de lâmpadas de led em horas, variância de temperatura ambiente em Graus Celsius, porcentagem de cura de uma vacina contra febre aftosa em animais bovinos adultos, etc., as quais são respostas da população estudada objeto de/da pesquisa, ou mesmo ele, pesquisador, pode ter interesse em testar a natureza da população, sua forma ou natureza (se é Normal ou Gaussiana, Binomial, Gama, etc.). Isto é, quando se pretende tirar conclusões sobre um parâmetro ou sobre a distribuição teórica de probabilidade de uma população. Outras vezes, o investigador pode tomar decisões tirando conclusões sobre a comparação de duas ou mais populações, é o que na estatística se conhece como inferência estatística ou estatística indutiva ou ainda estatística analítica.

Pode-se tomar a decisão, por exemplo, com base nos dados amostrais ou experimentais, se uma marca de fertilizante químico nitrogenado é melhor do que outra marca no que se refere a incrementar o aumento da produtividade em toneladas por hectare de uma cultura agrícola de algodão herbáceo; se um novo medicamento ou se uma nova droga é realmente eficaz na cura de uma doença; se uma ração rica em proteína é mais eficiente na engorda de animais que outra marca de ração, se um processo de enxertia ou propagação vegetativa é mais eficiente que outro; se uma variedade (ou cultivar) de milho é superior (mais produtiva) que outra variedade; se uma vacina é realmente 89% eficiente na imunização de um grupo de ratos ou camundongos submetidos a uma infecção por um determinado tipo de vírus patogênico; se uma raça A de suínos possui ganho médio de peso maior que uma raça B, será que a população dos pesos ao nascer de bezerros machos da raça Nelore tem distribuição normal; se uma nova marca "A" de inseticida é mais potente que uma marca "B", etc., deseja-se, assim, uma resposta do tipo "sim" ou "não" a uma conjectura, suposição, hipótese ou teoria previamente formulada baseada sempre na literatura.

Sendo assim, pode-se afirmar que um dos problemas a serem resolvidos pela inferência estatística é o de testar hipóteses. Isto é, feita determinada afirmação sobre uma população, usualmente sobre um parâmetro θ da distribuição teórica de probabilidade desse universo e, dessa forma, deseja-se saber se os resultados experimentais provenientes de uma amostra, ensaio ou experimento contrariam ou não tal afirmação. Na maioria das pesquisas realizadas, essa afirmação sobre a população é derivada de teorias desenvolvidas no campo substantivo do conhecimento, ou seja, da literatura, ponto inicial de qualquer projeto de pesquisa. A adequação ou não dessa teoria ao universo real ou população objeto de interesse do estudo pode ser verificada ou refutada pela amostra, ou resultados experimentais dentro de determinados níveis de probabilidade de risco, assumindo sempre que toda conclusão obtida pelo pesquisador, deve-se extrapolar as conclusões de tal modo que o processo de repetição da pesquisa seja assumido como se fosse repetido infinitas vezes, daí o nome dessa ciência conhecida como estatística frequentista em contrapartida à estatística Bayesiana.

O objetivo do teste estatístico de Hipótese é, então, fornecer uma metodologia bem objetiva de pesquisa que permita verificar se os dados amostrais ou experimentais trazem evidências que apoiem ou não uma hipótese ou teoria (estatística) formulada previamente, ou seja, antes do início de qualquer experimento, através de uma profunda e exaustiva revisão de literatura.

A ideia central do Teste de Hipótese sobre um parâmetro ou comparação de parâmetros de populações é a de supor verdadeira a hipótese em questão e verificar se a amostra observada é “Verossímil” ou mais provável, ou seja, está refletindo o que há de mais próximo do real estado das condições experimentais dentro do ambiente em que se conduz a pesquisa (BUSSAB; MORETTIN, 2003).

Os testes estatísticos abordados neste livro permitem decidir entre duas hipóteses complementares, a partir de amostras das populações estatísticas estudadas. No transcorrer de todo o do texto que será visto adiante, essas provas baseiam-se em:

i) Formular as hipóteses estatísticas de nulidade H_0 e alternativa H_1 (Ha), que vão ser testadas, como por exemplo, "a variedade de mamoeiro A é equivalente em produtividade a variedade de B" versus "a variedade de mamoeiro A é mais produtiva do que a variedade B", respectivamente.

ii) Usar uma ferramenta estatística inferencial, um método, uma variável teste ou tipo de teste estatístico paramétrico ou não paramétrico para decidir entre as duas hipóteses citadas anteriormente. Isto é, partir de uma amostra ou resultado de um ensaio ou experimento, usar a informação nela contida para tomar a decisão, dentro de determinados níveis de probabilidade de risco ou erro que podem ser de dois tipos, o I e o II, que são, respectivamente, rejeitar a hipótese básica H_0, quando ela for realmente verdadeira ou aceitar a hipótese H_0, quando ela for verdadeiramente falsa. É sempre uma decisão do tipo sim ou não, mutuamente exclusiva, complementar e exaustiva, mas sempre associada a risco ou erros na tomada de decisão, haja vista que o pesquisador, apesar de assumir no inicio da condução do teste de significância que a hipótese H_0 é verdadeira, mas como é uma hipótese ou conjectura, ele jamais saberá a verdadeira situação real da teoria previamente formulada, e assim, com o intuito de aceitá-la ou rejeitá-la conforme o apoio dos resultados amostrais ou experimentais, toda a decisão tomada deve ser incerta, não definitiva e extrapolada para toda a população de interesse de Estudo, mesmo que a decisão tenha sido tomada com base apenas nas informações amostrais, daí o nome do processo de conclusão ser chamado de inferência estatística.

2 TESTE DE HIPÓTESE

É uma regra bem objetiva de decisão, que auxilia o pesquisador a aceitar ou rejeitar uma hipótese estatística, com base nos dados (resultados) amostrais (experimentais).

Podemos ainda afirmar que teste de hipótese ou de significância é o meio pelo qual se obtém a probabilidade de que uma diferença maior ou igual do que a observada tenha sido ocasionada pelo acaso, se realmente não houver diferença, isto é, se a hipótese H_o for verdadeira.

Sendo assim, um teste de hipótese ou teste estatístico ou de significância é um processo estatístico objetivo que auxilia o pesquisador a utilizá-lo para se tirar uma conclusão, tomar uma decisão do tipo sim ou não sobre uma ou mais populações previamente definidas, a partir de uma ou mais amostras oriundas dessas populações. Vale lembrar que todas as amostras devem aleatórias, casuais ou estocásticas escolhidas mediante um mecanismo de sorteio aleatório com elementos independentes e identicamente distribuídos (AAS-i.i.d.) a qual é a única forma de garantir representatividade estatística dos elementos da população interesse de estudo.

3 HIPÓTESE

É uma afirmação, mentalmente elaborada, sobre as possíveis causas ou natureza de um problema, fato observacional ou fenômeno que se deseja investigar. Basicamente consiste em supor conhecida a resposta ou solução para o problema. A hipótese pode ser vista como uma solução provisória para o problema. Se necessário, elaboram-se soluções provisórias complementares. Hipóteses são elaboradas a partir de conhecimentos próprios do pesquisador sobre o problema e de revisão, a mais ampla possível, do que já é sabido sobre o mesmo problema. Formalmente, a hipótese deve ser formulada de tal modo que sua verificação possa ser feita por observação direta, através de procedimentos experimentais ou que deduções nesses fundamentadas e que conduzam a predições verificáveis. Uma hipótese pode consistir de afirmações como: ... numa determinada situação um fenômeno se manifesta, ... possui tais características ou natureza, ... que se relaciona (ou não) com outro fenômeno. Essas afirmações devem ser experimentalmente verificadas, o que equivale a uma confrontação com a realidade empírica. Somente os resultados experimentais autorizam a aceitação ou rejeição de uma hipótese. É conveniente lembrar agora a opinião do cientista Isaac Newton: "... não se inventam hipóteses" (NEWTON, 1934). A hipótese exerce importante papel na definição da metodologia experimental a ser utilizada, composta pelos seguintes elementos: delineamentos experimentais, técnicas de observação e obtenção de dados, variáveis a serem observadas, etc. A hipótese, portanto, tem por função orientar o pesquisador na direção da provável causa do fenômeno ou problema observado e orientar o pesquisador no estabelecimento da metodologia.

A hipótese pode ser definida também como uma explicação provisória a um problema, e ela deve ser submetida à verificação para sua comprovação ou não. No entanto, a comprovação direta de certas hipóteses é empiricamente inviável. Nessas situações, as hipóteses se firmam por meio de comprovações indiretas, isto é, elas são verificadas pelas suas consequências. As hipóteses são as proposições desencadeadoras do processo científico. A investigação não deve produzir mais do que aquilo expresso na formulação antecipada pela hipótese. Deve ficar claro que as hipóteses são sentenças afirmativas ainda que as afirmações sejam prováveis e representem, de alguma forma, relações entre variáveis. Portanto, a hipótese se refere a fatos e fenômenos na tentativa de explicá-los e de verificar o inter-relacionamento das variáveis envolvidas. As investigações executadas para comprovação de hipóteses se acham limitadas pelo próprio conteúdo da hipótese que é, assim, a delimitadora da área de observação e de experimentação.

As hipóteses podem ser explicativas ou preditivas. São explicativas quando representam o resultado de gradativas generalizações de proposições existentes na teoria, sendo formuladas após a observação de um fato. São preditivas quando formuladas, precedendo a observação empírica com base na teoria existente e utilizando uma inferência de tipo dedutiva, ao contrário da hipótese explicativa, que usa a inferência indutiva.

4 HIPÓTESE CIENTÍFICA

É uma proposição preliminar para explicar um fenômeno ou fato observacional cuja verdadeira causa é desconhecida, ou cujas explicações presentes não satisfazem ao corpo do saber. A realização de qualquer pesquisa somente se justifica se tiver por objetivo produzir novos fatos ou evidências que favoreçam a confirmação ou rejeição do saber existente ou que permitam o estabelecimento de novos conceitos sobre a verdadeira causa do fenômeno investigado.

5 HIPÓTESE ESTATÍSTICA

É toda hipótese relativa à caracterização de uma população, podendo dizer respeito tanto à forma como aos parâmetros da respectiva distribuição teórica de probabilidades da variável aleatória de interesse que é a população.

Ou ainda, podemos definir que uma hipótese estatística é uma suposição ou conjectura quanto ao valor de um parâmetro θ ou dos parâmetros θ_i da distribuição de probabilidade de uma variável aleatória ou ainda uma afirmação sobre a comparação de dois ou mais parâmetros que será verificada por um teste paramétrico, ou uma suposição quanto à natureza do formato da distribuição de probabilidade da população que será verificada por um teste não paramétrico como, por exemplo, um teste de qui-quadrado de aderência. E, quanto à distribuição de probabilidade da população, uma hipótese estatística é chamada simples se especifica completamente a distribuição da população, e é chamada composta se não especifica completamente a distribuição. Como exemplos, podemos citar as seguintes situações de conjecturas:

a) $H_0 : X \sim B\left(14; \frac{1}{4}\right)$, isto é, "X" tem distribuição binomial com parâmetros n = 14 e $P = \frac{1}{4}$.

b) $H_a : X \sim B\left(16; \frac{1}{2}\right)$, isto é, "X" tem distribuição binomial com parâmetros n = 16 e $P = \frac{1}{2}$.

c) $H_0 : Y \sim B\left(100; \frac{1}{2}\right)$.

d) $H_a : Y \sim B(100; P)$, onde $P \neq \frac{1}{2}$.

As hipóteses estatísticas 1, 2 e 3 são simples e a hipótese 4 é composta.

A hipótese física correspondente à hipótese estatística H_0 diz que a moeda é equilibrada ou honesta, e a correspondente à hipótese H_a diz que a moeda é não equilibrada ou viciada, pois $P \neq \frac{1}{2}$.

Outros exemplos de hipóteses estatísticas são os seguintes:

e) A média populacional da altura dos brasileiros é 1,65, isto é, $\mu = 1{,}65$.

f) A variância populacional dos salários no Nordeste brasileiro vale R$ 5000^2, isto é, $\sigma^2 = 5000$.

g) A proporção de brasileiros com a doença "X" (dengue) é 40%, ou seja, P= 0,40.

h) A distribuição dos pesos dos alunos em quilogramas da Universidade federal Rural do Semi-Árido (UFERSA) é normal ou gaussiana.

i) A chegada de navios no porto "Y" em Natal, no Estado do Rio Grande do Norte, tem distribuição de Poisson.

j) A distribuição da precipitação média mensal em milímetros de chuva em Mossoró, no Estado do Rio Grande do Norte, durante os seis primeiros meses do ano, é log-Normal.

É bom lembrar que a escolha da hipótese a ser testada deve ser feita antes da coleta dos dados para evitar possíveis interferências subjetivas na interpretação dos resultados.

6 TIPOS DE HIPÓTESES ESTATÍSTICAS A SEREM TESTADAS

6.1 Hipótese de nulidade, nula, básica ou probanda (h_0)

É a hipótese estatística a ser testada, isto é, é a hipótese de que não há diferença entre os dois parâmetros ou processos avaliados, ou ainda, é a hipótese estatística que se enuncia sob a forma de uma ou mais igualdades a zero. Podemos ainda definir que hipótese nula é a hipótese que julgamos inverossímil, geralmente contém o sinal de igual (=).

Para que uma hipótese seja de nulidade, é preciso que:

a) seja unívoca, ou possa reduzir-se a uma unívoca, simples.
b) seja possível comprová-la ou rejeitá-la através de dados experimentais ou de observação.

6.1.1 Exemplo

A verdadeira média populacional da altura dos brasileiros é 1,65 m, H_o: μ = 1,65 m, as duas cultivares de milho Cultive 1 e Cultive 2 possuem, em média, a mesma produtividade média de grãos em quilogramas por hectare, ou seja, H_o: $m_1 = m_2$, isto é, $m_1 - m_2 = 0$ ou

$$H_0 : \mu_1 = \mu_2 \therefore \mu_1 - \mu_2 = 0$$

6.2 Hipótese alternativa (ha ou h_1)

É a hipótese estatística que difere de uma pré-fixada, ou seja, é uma afirmação que oferece uma alternativa à alegação de H_o. Podemos ainda afirmar que a hipótese alternativa é a hipótese que julgamos verossímil e que se pretende verificar, geralmente ela contém o sinal de maior (>), de menor (<) ou diferente ($\neq$).

6.2.1 Exemplos

a) A verdadeira média populacional da altura dos brasileiros é maior que 1,65 m, H_a: μ > 1,65 m; a raça R1 de suínos possui um maior ganho de peso médio em quilogramas que a raça R2, isto é, H_1: $m_1 > m_{2,}$ etc.

b) A verdadeira média populacional da altura dos brasileiros é diferente de 1,65 m. H_1: $\mu \neq 1{,}65$ m.

c) As duas variedades de milho V1 e V2 apresentam produtividades médias em toneladas por hectare, diferentes, H_1: $m_1 \neq m_2$, ou $H_1 : \mu_1 \neq \mu_2$.

6.3 Observações importantes

a) A rejeição da Hipótese de nulidade H_o leva o pesquisador a opção da aceitação da hipótese H_1 com o risco conhecido alfa (α).

b) A Hipótese Alternativa geralmente representa a suposição que o pesquisador quer provar, sendo a Hipótese H_o formulada com o expresso propósito de ser rejeitado.

c) A Hipótese H_o é considerada verdadeira ao longo da realização do Teste de Hipótese até o momento em que haja clara evidência apontando em sentido contrário.

6.4 Estatística teste

Para se optar entre as duas hipóteses estatísticas H_o e H_1 é necessário quantificar a informação contida na amostra, usando, para isso, o que se designa por estatística teste, a qual é uma função das observações amostrais cujo valor vai determinar a conclusão a ser retirada do teste estatístico. No caso de se testar um parâmetro, a estatística teste é, habitualmente, um estimador desse parâmetro.

6.5 Regra de decisão estatística

É o principio que determina a conclusão a ser obtida (rejeitar ou não rejeitar a hipótese H_0) a partir da comparação do valor da estatística teste com um ou mais valores críticos.

Os valores críticos ou tabelados determinam o conjunto de valores da estatística teste que conduz à rejeição da hipótese nula. Este conjunto de valores denomina-se região critica ou região de rejeição da hipótese nula (ou do teste).

7 TIPOS DE TESTES DE HIPÓTESES: NATUREZA DOS TESTES

a) Testes Paramétricos: São aqueles que formulamos Hipóteses com respeito ao(s) valor(res) de um(uns) parâmetro(s) populacional(is). Exemplo: teste referente à média populacional "μ".
b) Testes Não Paramétricos ou de Distribuição Livre: São aqueles nos quais formulamos Hipóteses em que não fazemos nenhuma menção a respeito de parâmetros populacionais ou quando

formulamos Hipóteses com respeito à natureza da distribuição de probabilidades da população. Ou ainda, podemos dizer que é aquele teste cujo modelo não especifica condições sobre os parâmetros da população da qual a amostra foi obtida. Mesmo quando existem certas pressuposições, estas são mais brandas do que aquelas associadas aos testes paramétricos. Além disso, as provas não paramétricas não exigem mensurações tão fortes quanto às provas paramétricas, pois a maior parte das provas não paramétricas se aplica a dados em escala ordinal, e alguns mesmos a dados em escala nominal. Exemplo: Teste do qui-quadrado (χ^2) para ajustamento, Teste do qui-quadrado (χ^2) para aderência, Teste do qui-quadrado (χ^2) para independência, Teste do qui-quadrado (χ^2) para homogeneidade, Teste de Fisher, Teste de Kolmogorov – Smirnov, etc.

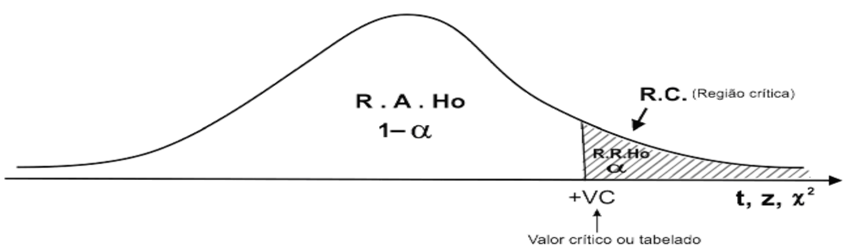
8 TIPOS DE TESTES DE SIGNIFICÂNCIA OU TESTES DE HIPÓTESES, QUANTO À FORMULAÇÃO DAS HIPÓTESES A SEREM TESTADAS

8.1 Teste unilateral ou unicaudal à direita

É aquele teste cuja hipótese alternativa (H_1) contém a desigualdade maior que (>); ou ainda, é aquele cujas hipóteses a serem testadas são do tipo:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}, \text{ por exemplo, } \begin{cases} H_0 : \mu = 150 \\ H_1 : \mu > 150 \end{cases}$$

Figura 1 - Curva do teste de hipótese unilateral à direita, mostrando a região crítica ou de rejeição de H_o



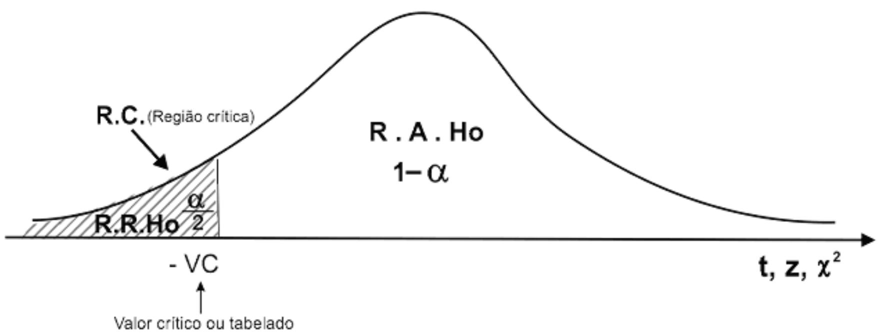
Ou seja, frequentemente, em algumas situações, estamos interessados apenas em valores extremos de um ou de outro lado da média, por exemplo, ou de outro parâmetro de interesse, isto é, em apenas uma cauda da distribuição amostral do estimador, como por exemplo, quando estamos testando a hipótese de um processo (produto, máquina, ração, vacina, sistema, raça, enfim, um tratamento), é melhor do que ou superior a outro processo (o que é diferente de testar se um processo é melhor ou pior do que outro). Em tais casos, a região crítica ou de rejeição da hipótese H_0 situa-se apenas em um lado da cauda da distribuição amostral do estimador envolvido no experimento ou curva do teste (nesse caso, no lado direito do extremo da cauda da curva), com área igual ao nível de significância do teste ou erro do tipo I (alfa, α).

8.2 Teste unilateral ou unicaudal à esquerda

É aquele teste cuja hipótese alternativa (H_1) contém a desigualdade menor que (<), ou ainda, é aquele cujas Hipóteses a serem testadas são do tipo:

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 < \sigma_0^2 \end{cases} \text{, por exemplo, } \begin{cases} H_0 : \sigma^2 = 50 \\ H_1 : \sigma^2 < 50 \end{cases}$$

Figura 2 - Curva do teste de hipótese unilateral à esquerda, mostrando a região crítica ou de rejeição de H_o



Ou seja, frequentemente, em algumas situações, estamos interessados apenas em valores extremos de um ou de outro lado da média, por exemplo, ou de outro parâmetro de interesse, isto é, em apenas no extremo superior direito ou esquerdo de uma cauda da distribuição amostral do estimador, como, por exemplo, quando estamos testando a hipótese de um processo (produto, máquina, ração, enfim, um tratamento), é pior ou inferior a outro processo (o que é diferente de testar se um processo é melhor ou pior do que outro). Em tais casos, a região crítica ou de rejeição da hipótese H_0 situa-se apenas em um lado da cauda da distribuição amostral do estimador envolvido no experimento ou curva do teste (nesse caso, no lado esquerdo do extremo da cauda da curva), com área igual ao nível de significância do teste ou erro do tipo I (alfa, **α**).

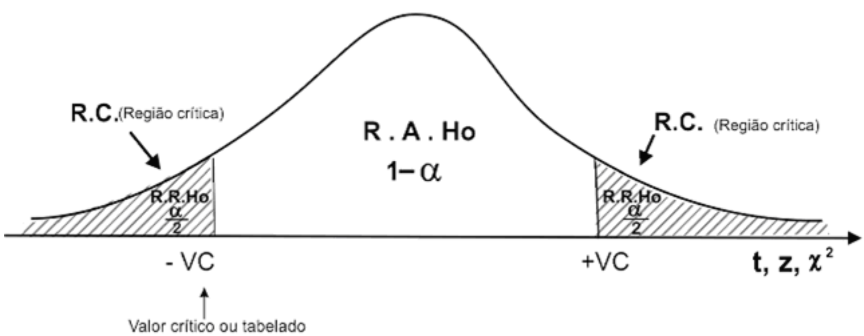
8.3 Teste bilateral ou bicaudal

É aquele teste cuja hipótese alternativa contém a não igualdade ≠ (diferente de) ou ainda, é aquele cujas Hipóteses a serem testadas são do tipo:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$, por exemplo, $$\begin{cases} H_0 : \mu = 100 \\ H_1 : \mu \neq 100 \end{cases}$$, $$\begin{cases} H_0 : P = P_0 \\ H_1 : P \neq P_0 \end{cases}$$, ou

$$\begin{cases} H_0 : P = 0{,}10 = 10\% \\ H_1 : P \neq 0{,}10 = 10\% \end{cases}$$

Figura 3 - Curva do teste de hipótese bilateral, mostrando as regiões críticas ou de rejeição de H_o



ou seja, frequentemente, em algumas situações, estamos interessados nos dois valores extremos superior direito e inferior esquerdo da média, por exemplo, ou de outro parâmetro de interesse, isto é, em duas regiões da cauda da distribuição amostral do estimador, ou seja, em seus escores Z em ambos os lados da média, ou de qualquer outro parâmetro, isto é, em ambas as caudas da distribuição, como, por exemplo, quando estamos testando a hipótese de um processo (produto, máquina, ração, raça, sistema, vacina, enfim um tratamento), é melhor ou pior do que outro (o que é diferente de testar se um processo é melhor do que outro ou se um processo é pior do que outro). Em tais casos, a região crítica ou de rejeição da hipótese H_0 situa-se nos dois lados da cauda da curva da distribuição amostral do estimador envolvido no experimento ou curva do teste (nesse caso, no lado esquerdo e no lado direito dos extremos da cauda da curva), com área igual ao nível de significância do teste ou erro do tipo I, dividido por dois, ou seja, alfa sobre dois, $\left[\frac{\alpha}{2}\right]$.

9 TIPOS DE RISCOS OU ERROS NA TOMADA DE DECISÃO

É sabido, na teoria da estimação, que um estimador do parâmetro θ está sujeito à variabilidade amostral e, assim, $\hat{\theta} \neq \theta$ em geral. Consequentemente, quando $\hat{\theta}$ é usado numa regra de decisão estatística, para se optar entre as duas hipóteses complementares H_0 e H_1, existe a possibilidade de se tirar a conclusão errada, isto é, cometer um erro de inferência.

Dessa forma, um erro de inferência consiste em se obter a conclusão errada em um teste estatístico a partir da informação contida na amostra.

Sendo assim, quando se escolhe entre duas hipóteses complementares, das quais apenas uma é verdadeira, há a possibilidade de se cometer um de dois tipos de erros, os quais são o erro do tipo I e o erro do tipo II, como é mostrado na tabela 1.

A uma decisão está sempre associado um risco, o risco de errar. No caso particular dos testes de hipótese, a tomada de decisão, para a população, é baseada na informação amostral, pelo que se podem cometer erros. Uma das principais características dos testes de hipóteses é permitir controlar ou minimizar o risco associado às decisões erradas.

9.1 Tipos de riscos ou erros na tomada de decisão

a) Probabilidade do Erro do Tipo I (Primeira Espécie), também chamado de nível de significância ou dimensão do teste.
   Constitui-se em se rejeitar a hipótese H_o quando ela é verdadeira, isto é, α é a probabilidade de se cometer o erro do tipo I, ou seja, α = nível de significância do teste, que é a região correspondente à rejeição de H_o quando ela é verdadeira, conhecida também por região crítica ou região de rejeição de H_0.
   De um modo geral, as hipóteses são testadas em três níveis de significância, 1%, 5% ou 10%. Escolhe-se, para apresentar a conclusão, o menor valor de alfa (α) capaz de rejeitar a hipótese H_0. Se nenhum dos níveis conseguir rejeitar H_0, diz-se que o teste é não significativo. Se o menor valor de alfa que rejeita H_0 é, por exemplo, 5%, diz-se que o teste é significativo ao nível de 5% de probabilidade.
   As decisões, associadas a esses níveis de significância alfa, costumam ser assim classificadas: se alfa é menor do que 1%, a conclusão é a de que a diferença é altamente significativa; se

alfa estiver entre 1% e 5%, a diferença é significativa; se alfa estiver entre 5% e 10% (sem incluir o 5% e o 10%), a diferença é provavelmente significativa; e, se alfa for maior do que 10%, a diferença é não significativa.

b) Erro do Tipo II ou de Segunda Espécie:

Constitui-se em se aceitar a hipótese H_o quando ela é falsa. β é a probabilidade de se cometer o erro do tipo II. Então, a probabilidade associada ao erro de Tipo II é β e representa-se por:
$\beta(\theta_1)$ = P[Não rejeitar $H_o \mid H_o$ é falsa] = P [Não rejeitar $H_o \mid H_1$ é verdadeira] = P[Não rejeitar $H_o \mid \theta = \theta_1$]

O valor de β (θ_1) diminui à medida que o verdadeiro valor de θ, θ_1 se afasta de θ_o, dado ser menos provável que não se detecte o verdadeiro valor. Consequentemente, quanto mais próximo estiver θ_1 de θ_o, maior é o valor dessa probabilidade.

c) Poder de um Teste

O poder de um teste é a probabilidade com que ele rejeitará a Hipótese Nula H_o quando esta é falsa. Esta medida é dada por 1- β, onde β é a probabilidade de se cometer um erro tipo II.
Isto é, o poder do teste, π = 1-β, permite medir a capacidade do teste para decidir acertadamente, quando a hipótese H_o é falsa.
Sendo assim, pode-se definir a função poder do teste como a seguir.
Designa-se por função poder de teste, π (θ_1), a probabilidade associada à decisão de rejeição de H_o quando esta é falsa, isto é,
π (θ_1) = P[Rejeitar $H_o \mid H_o$ é falsa] = P[Rejeitar $H_o \mid \theta = \theta_1$]
π (θ_1) = 1 - P [Não rejeitar $H_o \mid \theta = \theta_1$] = 1 - β (θ_1).
Quanto mais perto estiver θ_1 de θ_0, menor é o valor da função poder e menos potente é o teste, uma vez que é menor a capacidade de distinguir os verdadeiros valores dos falsos. Por outro lado, quanto mais afastado estiver o verdadeiro valor de θ, θ_1, em relação a θ_0, maior é a capacidade do teste de tomar decisões corretas.

d) O coeficiente de confiança de um teste é a probabilidade com que ele aceitará a hipótese e nulidade H_o, quando esta for verdadeira. Esta medida é dada por 1-α, na qual α é a probabilidade de se cometer um erro tipo I:
e) P value ou valor P é a probabilidade de significância do teste. O p value, valor p ou p valor, é o menor nível de significância α, a partir do qual se começa a rejeitar a hipótese H_o, isto é, se $\alpha < p$ value, então deve-se rejeitar a hipótese H_o.
Essa abordagem, para expressar a conclusão final de um teste de hipótese, é atualmente mais usada, principalmente depois do advento da computação eletrônica, com o uso intensivo de pacotes estatísticos ou softwares de análise estatística, sendo o interesse principal quantificar a probabilidade do que foi observado ou resultados mais extremos, sob a hipótese de igualdade dos grupos, tratamentos ou fatores experimentais. Assim, essa opção baseia-se na probabilidade de ocorrência de valores iguais ou superiores ao assumido pela estatística do teste, sob a hipótese de nulidade H_0 ser verdadeira.
Esse número é chamado de probabilidade de significância ou valor p ou ainda p value, em inglês, e frequentemente é indicado apenas por p. Como o valor p é calculado, supondo-se que a hipótese H_0 é verdadeira, pode-se fazer duas conjecturas quando se obtém um valor muito pequeno. Um evento que é extremamente raro pode ter ocorrido ou a hipótese H_0 não deve ser verdadeira, isto é, a conjectura inicial e conservadora não parece plausível. Portanto, quanto menor o valor-p, maior a evidência para se rejeitar a hipótese H_0. A Tabela 1 sintetiza o que pode ocorrer num processo de decisão estatística, salientando as probabilidades de decidirmos corretamente ou erroneamente.

Tabela 1 - Resultados Possíveis de um Teste de Hipóteses (Riscos na tomada de decisão ao se realizar um Teste de Hipótese ou de Significância)

SITUAÇÃO REAL	**TOMADA DE DECISÃO**	
	ACEITAR H_o	**REJEITAR H_o**
Se H_o é Verdadeira	Decisão correta Nível de confiança do teste, mensurada pela probabilidade = 1 - α	Decisão errada Erro tipo I ou de primeira espécie Probabilidade = α (nível de significância do teste)
Se H_o é Falsa	Decisão errada Erro tipo II ou de segunda espécie Probabilidade = β	Decisão correta Poder do teste Probabilidade = 1 - β

f) Relação entre o erro tipo I α e o erro tipo II β.

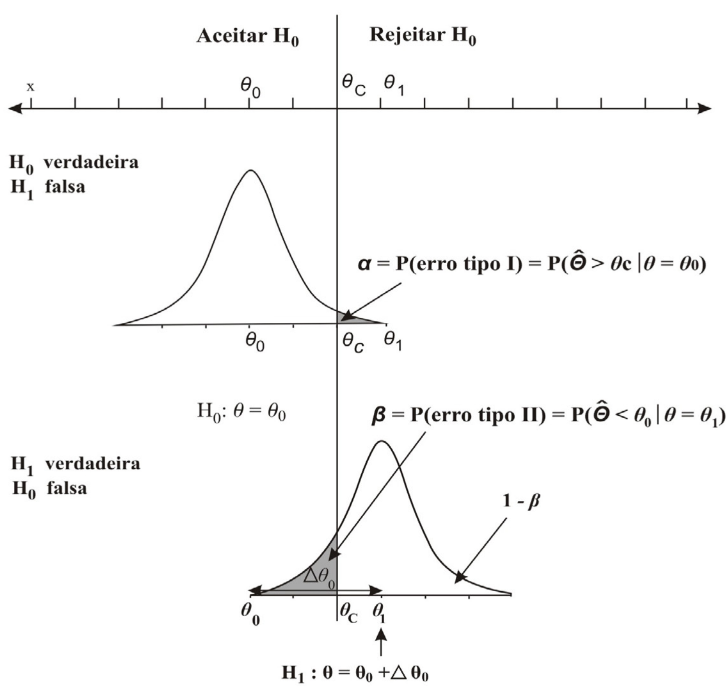
Inicialmente, vamos supor que se pretenda testar a hipótese nula $\boldsymbol{H_0 : \theta = \theta_0}$ contra a hipótese alternativa $\boldsymbol{H_1 : \theta = \theta_1 (a \quad H_1 : \theta = \theta_0 + \Delta\theta_0)}$, conforme se ilustra na Figura 4. Na parte superior dessa figura, exemplifica-se a distribuição da estatística de teste $\hat{\boldsymbol{\Theta}}$, no caso de a hipótese nula ser verdadeira, enquanto que, na parte inferior, exemplifica-se a distribuição da estatística de teste, no caso de a hipótese nula ser falsa. As áreas sombreadas ou hachuradas indicam as probabilidades dos erros de inferência que são determinadas pelo valor crítico $\boldsymbol{\theta_C}$.

Ao se definir o valor crítico $\boldsymbol{\theta_C}$ da regra de decisão estatística, deseja-se controlar os riscos do tipo I o α, e do tipo II que é o β, mantendo-se dentro de limites toleráveis. Como se vê pela Figura 4, se for aumentada a região de aceitação da hipótese $\boldsymbol{H_0}$ diminui-se o risco tipo I que é o α (o que é bom), mas, infelizmente, faz-se aumentar o risco ou erro tipo II que é o β (o que é ruim). Na maioria dos casos práticos, o tamanho da amostra n já está predeterminado e, assim, apenas um dos dois riscos pode ser controlado dentro de um nível pretendido.

No entanto, se o tamanho da amostra n puder ser aumentado, podemos controlar ambos os riscos ou erros tipo I que é o α e erro tipo II que é o β (reduzindo-os simultaneamente). Note-se que aumentar o tamanho da amostra n corresponde a diminuir o valor da estatística teste através da redução da variância do estimador amostral.

Do ponto de vista desse gráfico, para diminuir o erro tipo I o α, é necessário deslocar o valor $\boldsymbol{\theta_C}$ para a direita, fazendo diminuir a região de rejeição da hipótese $\boldsymbol{H_0}$ (isso implica alterar a regra de decisão). Mas esse movimento faz aumentar o erro do tipo II o β.Tendo isso em consideração, na prática costuma fixar-se o risco do erro do tipo I o α, e só depois determinar o valor crítico $\boldsymbol{\theta_C}$.

Figura 4 - Relação entre os erros do tipo I α e do Tipo II β



g) Consistência de um teste.

   Um teste é dito consistente para um determinada alternativa, se o seu poder tende para um (um), quando o tamanho da amostra tende para o infinito (∞).

h) Testes equivalentes.

   Dois testes são equivalentes quando têm o mesmo poder, isto é, um rejeita a hipótese H_0, quando o outro rejeita H_0, e aceita H_0 quando o outro aceita H_0.

É importante destacar as seguintes observações:

a) A hipótese H_o é formulada com o expresso propósito de ser rejeitada.
b) O significado do nível de significância do teste (α) é o seguinte: O valor α % indica que, em 100 experimentos conduzidos, em que a hipótese H_o é verdadeira, rejeitar-se-á a mesma, em média em 100 α % dos experimentos.
c) Ao diminuir-se um dos erros, aumenta-se o valor do outro e vice-versa.
d) O teste de maior poder é aquele que torna mínima a probabilidade do erro tipo II (β)
e) Para reduzir ambos os erros, tem-se que aumentar o tamanho da amostra (n).

A abordagem dos testes de hipóteses para controlar os erros consiste em fixar a probabilidade associada ao erro tipo I, α, e minimizar a probabilidade associada ao tipo II (β).

A razão pela qual se atribui mais importância ao erro tipo I deriva do seguinte ponto de vista: a possibilidade de rejeitar H_o incorretamente é considerada grave, pois essa hipótese corresponde ao que deve ser defendida, a menos que existam evidências convincentes a apontarem em sentido contrário.

Como exemplo, podemos citar o caso do funcionamento de um julgamento num tribunal, onde uma pessoa é inocente até se provar o contrário.

H_o: A pessoa é inocente
H_1: A pessoa é culpada
Erro Tipo I: A pessoa é condenada, mas está inocente.
Erro Tipo II: A pessoa é absolvida, mas é culpada.

O sistema judicial considera que é mais grave culpar um inocente do que absolver um culpado.

10 PROCEDIMENTOS PARA SE EFETUAR UM TESTE DE SIGNIFICÂNCIA

a) Enunciar as hipóteses H_o e H_1

Exemplo:

H_0: $\mu = \mu_0$

H_1: $\mu \neq \mu_o$

H_1: $\mu > \mu_o$

H_1: $\mu < \mu_o$

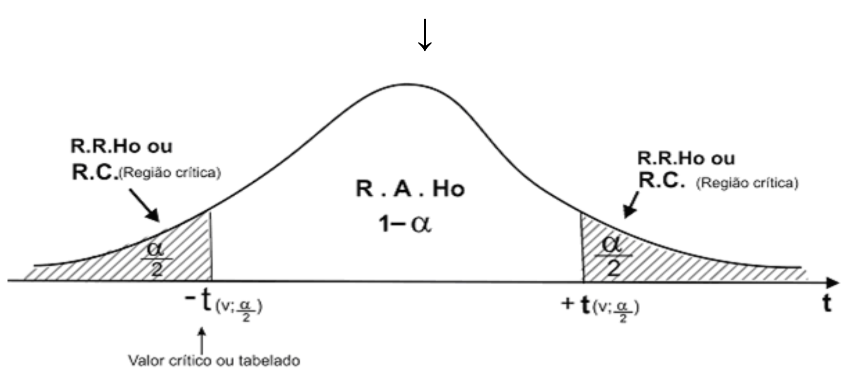
b) Fixar o limite de erro α (nível de significância do teste ou erro tipo 1) e identificar a variável ou estatística do teste.

Por exemplo, admitindo-se que não conhecemos a variância populacional σ^2, a variável do Teste será "t" de "Student", com v = n - 1 graus de liberdade.

c) Determinar a região de aceitação (RA) e a região crítica (RC) ou de rejeição da hipótese Ho[RRH_0] em função do nível α, pelas tabelas estatísticas, sob a curva da distribuição amostral do estimador.

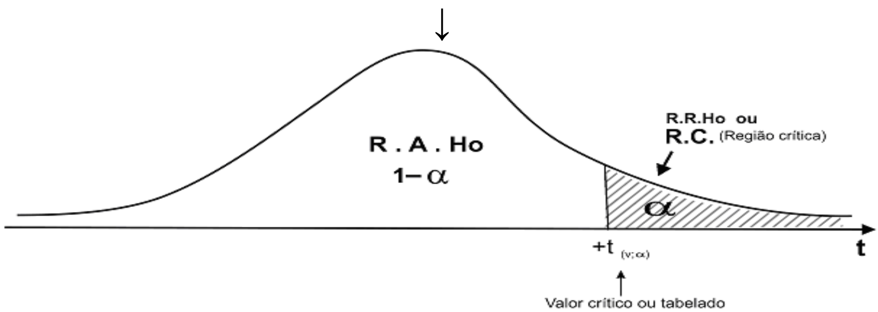
– Distribuição amostral do estimador (curva do teste)

Figura 5 - Curva do teste de hipótese t de "Student" bilateral, mostrando as regiões críticas ou de rejeição de H_o



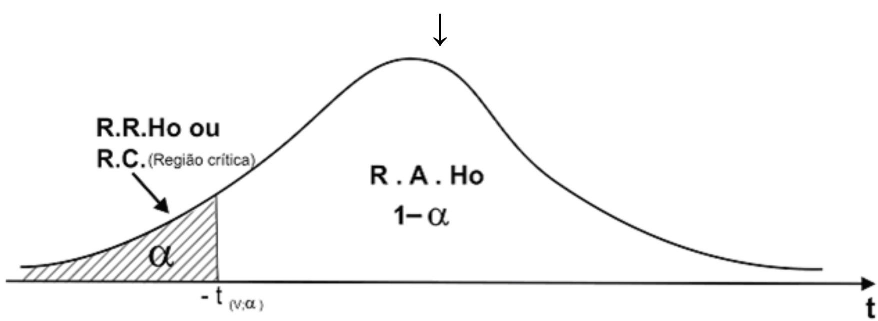
– Distribuição amostral do estimador (curva do teste)

Figura 6 - Curva do teste de hipótese t de "Student" unilateral à direita, mostrando a região crítica ou de rejeição de H_o



– Distribuição amostral do estimador (curva do teste)

Figura 7 - Curva do teste de hipótese t de "Student" unilateral à esquerda, mostrando a região crítica ou de rejeição de H_o



d) Por meio dos elementos ou dados amostrais, calcular o valor da variável ou estatística teste.

Por exemplo:

$$t_{Teste} = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

Onde: $\overline{X}$ = média amostral, $\boldsymbol{\mu_o}$ = valor sugerido da hipótese nula, S = desvio padrão da amostra e n = tamanho da amostra.

É importante lembrar que, se o valor da estatística teste cair dentro da região de rejeição da hipótese $\mathbf{H_0}$, e que, dependendo do valor da área sob a curva a qual é igual ao nível de significância do teste ou erro tipo **I** alfa (α), deve-se colocar no canto superior direito dessa estatística um (1) asterisco (*); se a área for igual a 5% ou 10%, dois asteriscos (**); se a área for igual a 1% e três asteriscos (***); se a área for igual a 0,1%. Por outro lado, se o valor da estatística teste cair na região de aceitação da hipótese $\mathbf{H_0}$, então deve-se colocar o símbolo **n.s.** (não significativo).

e) Concluir pela aceitação ou rejeição de H_o, pela comparação do valor obtido da variável do teste com as regiões críticas e de aceitação fixadas.

H_0: $\mu = \mu_o$
H_1: $\mu < \mu_o$, se $t \leq -t_{(v;\alpha)}$, rejeita-se H_o
H_1: $\mu > \mu_o$, se $t \geq t_{(v;\alpha)}$, rejeita-se H_o
H_1: $\mu \neq \mu_o$, se $t \leq -t_{\left(v;\frac{\alpha}{2}\right)}$, ou $t \geq t_{\left(v;\frac{\alpha}{2}\right)}$, rejeita-se H_o

Se o valor de "t" estiver fora dos intervalos, a decisão será contrária. É conveniente lembrar que todos os testes de médias que serão vistos neste capítulo pressupõem a normalidade da distribuição amostral da variável de teste $\overline{X}$. Como sabemos, essa suposição será rigorosamente válida se a distribuição da população for normal e a amostragem aleatória, e será válida, em geral, com boa aproximação, se a amostra for suficientemente grande, com pelo menos 30 elementos, ou seja, $\boldsymbol{n \geq 30}$.

11 TIPOS DE TESTES PARAMÉTRICOS: TESTES DE HIPÓTESES ESTATÍSTICAS PARAMÉTRICOS: TESTES REFERENTES A MÉDIAS, VARIÂNCIAS E PROPORÇÃO

11.1 Teste de hipótese para a média (μ) de uma população infinita, normal, com variância (σ^2) conhecida [amostra grande (n ≥ 30) ou pequena (n < 30)], ou variância (σ^2) desconhecida [amostra grande (n ≥ 30)].

11.1.1 O teste de hipótese

Tabela 2 (parte 1) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$\mu = \mu_0$	$Z_{TESTE} = \dfrac{\overline{X} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}}$ Com σ conhecido ou não	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$Z \leq -Z_{(\alpha)}$ $Z \geq Z_{(\alpha)}$ $Z \leq -Z_{\left(\frac{\alpha}{2}\right)}$ e $Z \geq Z_{\left(\frac{\alpha}{2}\right)}$

11.1.2 Exercício de aplicação

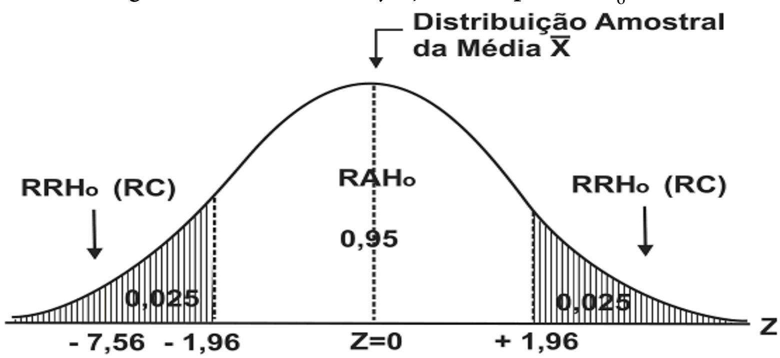
Sabe-se que certa linhagem de camundongo da linhagem C57BL6J, alimentados com uma ração padrão, tem um aumento médio de peso igual a 64 gramas durante os três primeiros meses de vida. Um lote (amostra) de 81 camundongos dessa linhagem foi alimentado com uma nova ração, mantendo-se as condições ambientais padronizadas. O aumento médio de peso observado nos camundongos foi de $\overline{X} = 60,75$ g, e um desvio padrão de S = 3,84 g. A nova ração tem a mesma eficiência alimentar que a padrão? Use α = 0,05. Dados fornecidos: n = 81, $\overline{X} = 60,75$, S = 3,84, $\mu = \mu_0 = 64$

i) $\begin{cases} H_0 : \mu = 64 \\ H_1 : \mu \neq 64 \end{cases}$

ii) α = 0,05; $Z_{\left[\frac{0,05}{2}\right]} = \pm 1,96$

iii)

Figura 8 - Curva do teste de hipótese Z Bilateral, mostrando as regiões críticas ou de rejeição da hipótese H_o



iv) $$Z_{Teste} = \frac{60,75 - 64,00}{\frac{3,84}{\sqrt{81}}} = \frac{-3,25}{0,43} = -7,56^{*}$$

$\left|Z_{Teste}\right| > \left|Z_{\left(\frac{0,05}{2}\right)}\right| \rightarrow$, $\left|-7,56\right| > \left|1,96\right| \rightarrow$ Então rejeita-se a hipótese H_0

v) Conclusões: De acordo com o teste Z ao nível de 5% de significância, pode-se concluir que a nova ração tem menor eficiência alimentar que a ração padrão.

11.2 Teste de hipótese para a média (μ) de uma população infinita, aproximadamente normal com variância (σ^2) desconhecida e amostra pequena (n < 30)

11.2.1 O teste de hipótese

Tabela 2 (parte 2) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$\mu = \mu_0$	$t_{TESTE} = \dfrac{\overline{X} - \mu_0}{\dfrac{S}{\sqrt{n}}}$, $com\ V = n - 1\ G.L.$, e Com σ desconhecido	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$t \leq -t_{(V;\alpha)}$ $t \geq t_{(V;\alpha)}$ $t \leq -t_{\left(V;\frac{\alpha}{2}\right)}$ e $t \geq t_{\left(V;\frac{\alpha}{2}\right)}$

11.2.2 Exercício de aplicação

Em um seringal, localizado no estado do Acre, no qual se utiliza o processo convencional de sangria, a produção média de borracha seca é de 26g / árvore / corte. Tomou-se uma amostra ao acaso, composta de 25 seringueiras, as quais foram sangradas, usando-se um novo processo, tendo produzido 28g em média, com desvio padrão de 4g. O novo processo é mais eficiente que o convencional? Use α = 5%.

11.3 Teste de hipótese para a média das diferenças (μ_D) de duas populações que não são independentes, isto é, as variáveis são emparelhadas (dados emparelhados). E com a amostra de diferenças (n) pequena (n < 30)

11.3.1 Dados emparelhados

Os resultados das duas amostras constituem dados emparelhados, quando estão relacionados dois a dois, segundo algum critério que introduz uma influência marcante entre os diversos pares, que supomos, porém, influir igualmente sobre os valores de cada par.

Ou seja, a característica peculiar das amostras pareadas é que, para cada observação do primeiro grupo, há uma correspondente no segundo grupo. Na técnica conhecida como autoemparelhamento, as medidas são tomadas em um único indivíduo, em dois pontos distintos no tempo. Um exemplo comum de autoemparelhamento é o experimento antes – depois, no qual cada indivíduo é examinado antes que um determinado tratamento seja aplicado e depois que o tratamento for completado. Um segundo tipo de emparelhamento ocorre quando um investigador casa os indivíduos de um grupo com os de um segundo grupo, de modo que os membros de um par sejam tão parecidos quanto possível em relação a características relevantes, tais como idade, sexo, etc.

O emparelhamento é frequentemente empregado na tentativa de controlar as fontes estranhas de variação que poderiam, de outra maneira, influenciar os resultados da comparação. Se as mediadas são feitas no mesmo indivíduo ou elemento e não em dois elementos diferentes, uma certa quantidade de variabilidade biológica por exemplo é eliminada. Não temos de nos preocupar com o fato de um indivíduo ser mais velho do que o outro ou se um é homem e o outro mulher. A intenção do emparelhamento, portanto, é fazer uma comparação mais precisa.

Podemos diferenciar três tipos de pareamento, os quais são: autopareamento, pareamento natural e pareamento artificial.

O autopareamento ocorre quando o indivíduo serve como seu próprio controle, como na situação em que um indivíduo recebe duas drogas administradas em ocasiões diferentes. Outra situação, por exemplo, é aquela em que se mede o nível de colesterol de uma pessoa antes e depois de um programa de dieta ou tratamento. Finalmente, a comparação de dois órgãos no mesmo indivíduo, como braços, pernas, olhos, narinas, segundo alguma característica estudada também constitui um autopareamento.

O pareamento natural consiste em formar pares tão homogêneos quanto possível, controlando os fatores que possam interferir na resposta, sendo que o pareamento aparece de forma natural. Por exemplo, em experimentos de laboratório, podem-se formar pares de cobaias selecionadas da mesma ninhada; em investigações clínicas, gêmeos univitelinos são muito usados.

No pareamento artificial, escolhem-se indivíduos com características semelhantes, tais como, idade, sexo, nível socioeconômico, estado de saúde ou, em geral, fatores que podem influenciar de maneira relevante a variável resposta.

Na prática os pesquisadores, em geral, podem encontrar dificuldades no conhecimento das características que devem ser controladas e, mesmo conhecendo-se, pode ser difícil formar pares homogêneos, como por exemplo, no caso em que o número de fatores é muito grande.

Em muitas situações, embora desejável, torna-se difícil ou mesmo impossível a implementação do planejamento com amostras pareadas ou emparelhadas.

Ora, se os dados das duas amostras estão emparelhados, tem sentido calcularmos as diferenças (d_i), correspondentes a cada par de valores, reduzindo, assim, os dados a uma única amostra de "$\underline{n}$" diferenças. Por outro lado, testar a hipótese de que a diferença entre as médias das duas populações emparelhadas seja igual a um certo valor d_0, equivale a testar a hipótese de que a média de todas as diferenças (referentes às populações) seja igual a d_0, o que decorre diretamente das propriedades da média, ou seja, vamos testar simplesmente a hipótese $H_0 : \mu_d = d_0$, contra uma hipótese alternativa H_1, que poderá corresponder a um teste unilateral ou bilateral.

Os exemplos 1 e 2 ilustram situações em que os dados obtidos de duas amostras são correlacionados (dados emparelhados).

a) Exemplo 1: Quando certo caráter é medido no mesmo indivíduo, em épocas diferentes, os valores obtidos nas duas mensurações tendem a ser mais parecidos entre si do que se houvessem sido obtidos de indivíduos diferentes. Como por exemplo, a medição de taxas de crescimento de plantas de uma determinada cultura, antes e depois de se aplicar uma substância inibidora da fotossíntese.
b) Exemplo 2: As eficiências de duas rações podem ser comparadas, utilizando-se vários pares de animais irmãos de uma mesma leitegada (suínos, por exemplo), ou de uma mesma ninhada (camundongos, por exemplo).

O experimento consiste em alimentar cada membro, de cada par, com uma das rações alocadas ao acaso. Indivíduos desses tipos (irmãos-germanos), pelas suas semelhanças genéticas, tendem a apresentar respostas correlacionadas aos estímulos a que são submetidos.

11.3.2 O teste de hipótese

Tabela 2 (parte 3) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_0	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_0]
$\mu = \mu_D$	$t_{TESTE} = \dfrac{\overline{X} - \mu_D}{\dfrac{S}{\sqrt{n}}}$, $com\ V = n - 1\ G.L.$ Com σ desconhecido	$\mu < \mu_D$ $\mu > \mu_D$ $\mu \neq \mu_D$	$t \leq -t_{(V;\alpha)}$ $t \geq t_{(V;\alpha)}$ $t \leq -t_{\left(V;\frac{\alpha}{2}\right)}$ $e\quad t \geq t_{\left(V;\frac{\alpha}{2}\right)}$

11.3.3 Exercício de aplicação

Num ensaio com gado de leite da raça Holandês Preto e Branco, estudou-se o efeito de um novo componente de ração na produção de leite em quilogramas por dia. Para isso, foi tomada uma amostra aleatória de 10 vacas (de mesmo porte e homogêneas quanto ao peso, raça, ordem de parição, etc.), que foram alimentadas com uma ração básica durante um certo período de tempo, após o qual foram obtidas as produções diárias de leite em cada animal. Em seguida, adicionou-se à ração básica um "Novo componente H", e as vacas foram alimentadas com essa mistura durante um certo período, ao final do qual mediram-se novamente as produções diárias de leite em cada animal. Os dados desse experimento constam na tabela seguinte:

Tabela 3 - Produções diárias de leite de 10 vacas após aplicação de uma ração básica e após a aplicação de uma ração básica + um novo componente H

VACA	Ração Básica (X_i)	Ração Básica + H (X_2)	Ganho de Produção (Kg/dia) $di = X_{i2} - X_{i1}$	d_i^2
1	9,0	10,5	1,5	2,25
2	9,5	9,5	0,0	0,00
3	10,0	9,5	-0,5	0,25
4	11,0	13,0	2,0	4,00
5	9,8	12,8	3,0	9,00
6	8,9	9,9	1,0	1,00
7	9,2	11,7	2,5	6,25
8	9,7	10,5	0,8	0,64
9	11,0	11,5	0,5	0,25
10	9,5	12,0	2,5	6,25
SOMA	---------	---------	13,3	29,89

Supondo-se que a variável aleatória "D" relativa a ganho de produção leiteira tenha distribuição aproximadamente normal com média "μ" e variância σ^2, pede-se:

a) Obter as estimativas por ponto para $\mu_{\overline{D}}$ e para $\sigma^2_{\overline{D}}$
b) Fazer um teste de significância para: H_0: $\mu_{\overline{D}} = 0$, com $\alpha = 0,01$.

11.4 Teste de hipótese para a diferença entre as médias $(\mu_1 - \mu_2)$ de duas populações infinitas, com variâncias $(\sigma_1^2 \; e \; \sigma_2^2)$ conhecidas [amostras grandes ($n_1 \geq 30$ e $n_2 \geq 30$) ou pequenas ($n_1 < 30$ e $n_2 < 30$)] ou com variâncias desconhecidas [amostras grandes ($n_1 \geq 30$ e $n_2 \geq 30$)].

11.4.1 O teste de hipótese

Tabela 2 (parte 4) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$\mu_1 = \mu_2$	$Z_{TESTE} = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ $Z_{TESTE} = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$, σ_1 ***e*** σ_2 ***conhecidos.*** Ou não	$\mu_1 < \mu_2$ $\mu_1 > \mu_2$ $\mu_1 \neq \mu_2$	$Z \leq -Z_{(\alpha)}$ $Z \geq Z_{(\alpha)}$ $Z \leq -Z_{\left(\frac{\alpha}{2}\right)}$ *e* $Z \geq Z_{\left(\frac{\alpha}{2}\right)}$

11.4.2 Exercício de aplicação

Um experimento aleatório conduzido No Estado de São Paulo consistiu em comparar os pesos ao nascer de animais bovinos (bezerros) machos das raças Gir e Guzerá da estação experimental de zootecnia de Sertãozinho, do instituto de Zootecnia da secretária da Agricultura de São Paulo. Ao nível de 5% de Probabilidade.

-Pesos ao nascer (Kg) de Bezerros Machos Zebu. Onde:

$$n_1 = 40;\ \overline{X}_1 = 28{,}3\ Kg;\ S_1^2 = 108{,}37\ Kg^2;\ S_1 = 10{,}41\ Kg$$

$$n_2 = 50;\ \overline{X}_2 = 23{,}5\ Kg;\ S_2^2 = 49{,}70\ Kg^2;\ S_2 = 7{,}05\ Kg$$

11.5 Este de hipótese para a diferença entre as médias $(\mu_1 - \mu_2)$ de duas populações infinitas, aproximadamente normais, com variâncias $(\sigma_1^2 \; e \; \sigma_2^2)$ desconhecidas e estatisticamente iguais $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$, ou seja, populações homocedásticas e amostras pequenas ($n_1 < 30$ e $n_2 < 30$)

11.5.1 O teste de hipótese

Tabela 2 (parte 5) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_0	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_0]
$\mu_1 = \mu_2$	$t_{TESTE} = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_P \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$, $com\ V = n_1 + n_2 - 2,\ e$ $\sigma_1 = \sigma_2$ *mas desconhecidos, sendo* $S_P^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$, $e\ S_P = \sqrt{S_P^2}$	$\mu_1 < \mu_2$ $\mu_1 > \mu_2$ $\mu_1 \neq \mu_2$	$t \leq -t_{(V;\alpha)}$ $t \geq t_{(V;\alpha)}$ $t \leq -t_{\left(V;\frac{\alpha}{2}\right)}$ $e \quad t \geq t_{\left(V;\frac{\alpha}{2}\right)}$

11.5.2 Exercício de aplicação

Realizou-se um experimento aleatório em uma granja localizada em Mossoró, RN, com 340 aves de pintos machos de um dia de idade, da linhagem "Ross", com o objetivo de estudar o efeito da fonte de fósforo no ganho de peso das aves, em gramas. Foram utilizadas 10 repetições por tratamento e 17 pintos por parcela, e os resultados para ganho de peso médio por parcela, em gramas, estão mostrados na tabela abaixo,

Tabela 4 - Dados de ganhos de peso de pintos machos de um dia de idade, submetidos a tratamentos diferentes

Termofosfato Magnesiano (X_1)		**Fosfato Bicálcico (X_2)**	
63,0	60,9	90,0	91,8
62,5	62,5	90,6	92,0
61,9	63,8	91,7	92,3
60,7	62,0	90,9	90,6
63,2	61,8	90,3	91,2

onde temos os seguintes dados:

$$\boldsymbol{n_1 = 10; \overline{X}_1 = 62{,}23\ g; S_1^2 = 0{,}96\ g^2; S_1 = 0{,}98\ g}$$

$$\boldsymbol{n_2 = 10; \overline{X}_2 = 91{,}14\ g; S_2^2 = 0{,}61\ g^2; S_2 = 0{,}78\ g}$$

a) verificar se existe efeito significativo da fonte de fosfato, considerando $\alpha = 0{,}05$.

11.6 Teste de hipótese para a diferença entre as médias $(\mu_1 - \mu_2)$ de duas populações infinitas, aproximadamente normais, com variâncias $(\sigma_1^2\ e\ \sigma_2^2)$ desconhecidas e estatisticamente diferentes $(\sigma_1^2 \neq \sigma_2^2)$, ou seja, populações heterocedásticas e amostras pequenas ($n_1 < 30$ e $n_2 < 30$)

11.6.1 O teste de hipótese

Tabela 2 (parte 6) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$\mu_1 = \mu_2$	$t_{TESTE} = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$, *com* $V = \dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$ $\sigma_1 \neq \sigma_2$ *e desconhecidos.*	$\mu_1 < \mu_2$ $\mu_1 > \mu_2$ $\mu_1 \neq \mu_2$	$t \leq -t_{(V;\alpha)}$ $t \geq t_{(V;\alpha)}$ $t \leq -t_{\left(V;\frac{\alpha}{2}\right)}$ *e* $t \geq t_{\left(V;\frac{\alpha}{2}\right)}$

Segundo Welch (1947), o número de graus de liberdade pode ser calculado através da seguinte fórmula:

$$V = \frac{\left[\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right]^2}{\left\{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1 + 1} + \dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2 + 1}\right\}} - 2$$

11.6.2 Exercício de aplicação

Considere os ganhos de peso diários em filhos de 2 (dois) reprodutores Suínos, da raça Duroc, acasalados cada um com matrizes da mesma raça, obtidos em um ensaio conduzido numa granja produtora de suínos, localizada em Mossoró, RN. Ao nível de 5% de significância, concluir se os ganhos médios de peso diários são estatisticamente iguais,

Tabela 5 - Ganhos de peso diário em quilogramas em filhos de dois reprodutores suínos da raça Duroc

Reprodutores	**matrizes**	**Ganhos de peso (Kg)**				
R_1	M	0,95	0,86	0,92	0,93	0,94
R_2	M	0,80	0,82	0,78	0,80	0,79

onde são fornecidos os seguintes dados referentes as análises desse experimento:

$$n_1 = 5; \overline{X}_1 = 0,900\ Kg; S_1^2 = 0,00125\ Kg^2; S_1 = 0,035\ Kg; W_1 = 0,000250$$

$$n_2 = 5; \overline{X}_2 = 0,798\ Kg; S_2^2 = 0,00022\ Kg^2; S_2 = 0,015\ Kg; W_2 = 0,000044$$

11.7 Teste de hipótese para a variância de uma população normal

Tabela 2 (parte 7) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$\sigma^2 = \sigma_0^2$	$\chi^2_{TESTE} = \frac{(n-1)S^2}{\sigma_0^2}$ Com V = n - 1	$\sigma^2 < \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 \leq \chi^2_{(1-\alpha;V)}$ $\chi^2 \geq \chi^2_{(\alpha;V)}$ $\chi^2 \leq \chi^2_{\left(1-\frac{\alpha}{2};V\right)}$ $e \quad \chi^2 \geq \chi^2_{\left(\frac{\alpha}{2};V\right)}$

11.8 Teste de hipótese para a razão $\left(\frac{\sigma_1^2}{\sigma_2^2}\right)$ entre as variâncias $\left(\sigma_1^2 \ e \ \sigma_2^2\right)$, de duas populações normais

Tabela 2 (parte 8) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$\sigma_1^2 = \sigma_2^2$	$f_{TESTE} = \frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{S_1^2}{S_2^2}$ Com $V_1 = n_1 - 1$ e $V_2 = n_2 - 1$	$\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$f \leq f_{(1-\alpha;V_1,V_2)}$ $f \geq f_{(\alpha;V_1,V_2)}$ $f \leq f_{\left(1-\frac{\alpha}{2};V_1,V_2\right)}$ $e\ f \geq f_{\left(\frac{\alpha}{2};V_1,V_2\right)}$

11.9 Teste de hipótese para a proporção (p) de uma população infinita, normal, e com amostra grande $(n \geq 30)$

11.9.1 O teste de hipótese

Tabela 2 (parte 9) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_0	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_0]
$P = P_0$	$Z_{TESTE} = \dfrac{\hat{P} - P_0}{\sqrt{\dfrac{P_0 q_0}{n}}}$ **Com população infinita,** e $n\hat{P} > 5$, $n\hat{q} > 5$, $n \geq 30$, $\hat{P} = \dfrac{X}{n}$ e $q_0 = 1 - P_0$	$P < P_0$ $P > P_0$ $P \neq P_0$	$Z \leq -Z_{(\alpha)}$ $Z \geq Z_{(\alpha)}$ $Z \leq -Z_{\left(\frac{\alpha}{2}\right)}$ *e* $Z \geq Z_{\left(\frac{\alpha}{2}\right)}$

11.9.2 Exercício de aplicação

Em uma granja localizada no município de Mossoró, RN, uma amostra aleatória de 800 coelhos apresentou 480 machos. Ao nível de 1% de significância, pode-se concluir que há prevalência (predominância) de coelhos machos nessa granja?

11.10 Teste de hipótese para a diferença entre as proporções $(P_1 - P_2)$, de duas populações infinitas, normais, com amostras grandes $(n_1 \geq 30 \; e \; n_2 \geq 30)$, quando a hipótese de nulidade (h_o) da diferença entre as proporções se refere a um valor diferente de zero

Tabela 2 (parte 10) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$P_1 - P_2 = P_0$ $P_0 \neq 0$	$Z_{TESTE} = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - \left(P_1 - P_2\right)}{\sqrt{\frac{\hat{P}_1 \hat{q}_1}{n_1} + \frac{\hat{P}_2 \hat{q}_2}{n_2}}}$ **Com populações infinitas e** $n_1 \hat{P}_1 > 5; n_2 \hat{P}_2 > 5$; $n_1 \hat{q}_1 > 5; n_2 \hat{q}_2 > 5$; $n_1 \geq 30; n_2 \geq 30$	$P_1 < P_2$ $P_1 > P_2$ $P_1 \neq P_2$	$Z \leq -Z_{(\alpha)}$ $Z \geq Z_{(\alpha)}$ $Z \leq -Z_{\left(\frac{\alpha}{2}\right)}$ *e* $Z \geq Z_{\left(\frac{\alpha}{2}\right)}$

11.11 Teste de hipótese para a diferença entre as proporções $(P_1 - P_2)$, de duas populações infinitas, normais, com amostras grandes $\left(n_1 \geq 30 \; e \; n_2 \geq 30\right)$, quando a hipótese de nulidade (h_o) da diferença entre as proporções se refere a um valor igual a zero

Tabela 2 (parte 11) - Testes de hipóteses estatísticas para médias, diferença entre médias, variância, razão entre variâncias, desvio padrão, proporção, diferença entre proporções

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$P_1 - P_2 = P_0$ $P_0 = 0$	$Z_{TESTE} = \dfrac{\left(\hat{P}_1 - \hat{P}_2\right) - \left(P_1 - P_2\right)}{\sqrt{\dfrac{P' q'}{n_1} + \dfrac{P' q'}{n_2}}}$ $P' = \dfrac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2}$ e $q' = 1 - P'$	$P_1 < P_2$ $P_1 > P_2$ $P_1 \neq P_2$	$Z \leq -Z_{(\alpha)}$ $Z \geq Z_{(\alpha)}$ $Z \leq -Z_{\left(\frac{\alpha}{2}\right)}$ e $Z \geq Z_{\left(\frac{\alpha}{2}\right)}$

Vale salientar que, se a população for finita $\left(n \geq 5\% N\right)$ e a amostragem for sem reposição (ASR), é necessário e deve-se utilizar o fator de correção para população finita [FCPF], empregando a seguinte fórmula: FCPF = $\dfrac{N - n}{N - 1}$, o qual deve ser multiplicado pelo erro padrão do estimador usado no teste. Por exemplo, se o estimador for a média da amostra $\overline{X}$, cujo erro padrão é dado por $S_{(\overline{x})} = \dfrac{S}{\sqrt{n}}$, então temos que usar a multiplicação $S_{(\overline{x})} = \dfrac{S}{\sqrt{n}} \cdot \sqrt{\dfrac{N - n}{N - 1}}$.

12 TIPOS DE TESTES NÃO PARAMÉTRICOS

12.1 Teste de qui-quadrado (χ^2)

12.1.1 Introdução

Este tipo de teste não paramétrico ou de distribuição livre tem por objetivo auxiliar o pesquisador na resolução de problemas de pesquisa científica que ocorrem frequentemente na ciência, nos quais o experimentador ou pesquisador deseja testar uma hipótese baseada em alguma Lei Natural, distribuição teórica de probabilidade, homogenidade de amostras ou elaborada por raciocínio indutivo, bem como na verificação ou ajustamento de frequencias ou proporções genéticas em ensaios biológicos ou não. Nesse caso, os dados das variáveis respostas experimentais fornecem uma amostra e o pesquisador verifica se o material observado está ou não em concordância com a hipótese de nulidade estabelecida. Raramente os valores observados ou coletados no experimento concordarão exatamente com os valores teóricos, hipotéticos ou esperados. Entretanto, o que interessa para o pesquisador é saber quais são os limites que os desvios entre os valores observados e esperados não devam exceder para que essas discrepâncias possam ser consideradas como erros de amostragem, erros casuais ou aleatórios, e não como uma discordância definitiva entre a hipótese e os fenômenos experimentais observados. Sendo assim, nestas circunstâncias, o investigador obtém frequências absolutas ou contagem do número de ocorrências nas diferentes categorias, classes ou atributos estudados, as quais correspondem àquilo que denominamos de frequência observada.

12.1.2 Estatística teste

O teste não paramétrico usado como ferramenta nas condições expostas aqui deve-se ao matemático e estatístico inglês Karl Pearson (1857-1936), e é conhecido por Teste de qui-quadrado (χ^2). Uma medida das diferenças entre as frequências observadas e esperadas é medida pela estatística qui-quadrado (χ^2), definida por

$$\chi^2_{Teste} = \frac{\sum_{i=1}^{k}\left(f_o - f_e\right)^2}{f_e}$$

onde F_o é a frequência observada e f_e é a frequência esperada ou teórica.

As frequências observadas são frutos da experiência, ou seja, são obtidas na amostra ou experimento, e as frequências esperadas, teóricas ou calculadas são resultantes da hipótese H_0 formulada.

Considere a tabela 6, a seguir, como sendo uma estrutura em que o pesquisador deve descrever as frequências observadas e esperadas após a condução de sua pesquisa, e que possa, assim, comparar, através do teste de qui-quadrado, as possíveis diferenças existentes entre as frequências práticas e as frequências teóricas.

Tabela 6 - Frequências observadas e esperadas do teste qui-quadrado

Evento	**E_1**	**E_2**	**E_3**	**...**	**E_K**
Frequência observada	f_{O1}	f_{O2}	f_{O3}	...	f_{OK}
Frequência esperada	fe_1	fe_2	fe_3	...	fe_K

Dessa forma, a estatística teste de qui-quadrado pode ser obtida conforme a seguinte equação:

$$\chi^2_{Teste} = \frac{\left(f_{o_1} - f_{e_1}\right)^2}{f_{e_1}} + \frac{\left(f_{o_2} - f_{e_2}\right)^2}{f_{e_2}} + \frac{\left(f_{o_3} - f_{e_3}\right)^2}{f_{e_3}} + \cdots + \frac{\left(f_{o_k} - f_{e_k}\right)^2}{f_{e_k}}$$

$$\chi^2_{Teste} = \frac{\sum_{j=1}^{k}\left(f_{o_j} - f_{e_j}\right)^2}{f_{e_j}}$$

Sendo a frequência total igual a N, então a soma das frequências observadas é sempre igual à soma das frequências esperadas e são iguais à frequências total N, isto é,

$$\sum_{j=1}^{k} f_{o_j} = \sum_{j=1}^{k} f_{e_j} = N\,.$$

Quando $\chi^2 = 0$, então as frequências teóricas e observadas concordam exatamente, ao passo que, quando $\chi^2 > 0$, isso não ocorre, e mais: quanto maior for o valor de χ^2, maior será a discrepância entre as frequências observadas e esperadas.

O número de graus de Liberdade V da distribuição de qui-quadrado (χ^2) é dado por:

a) V = K-1, quando as frequências esperadas puderam ser calculadas sem que se façam estimativas dos parâmetros populacionais a partir das distribuições amostrais. É importante lembrar que "K" é o número de eventos ou categorias em que foi dividida a amostra.
b) V = K-P-1, quando para a determinação das frequências esperadas P parâmetros populacionais tiveram suas estimativas calculadas a partir das distribuições amostrais, como, por exemplo, para a distribuição normal temos que estimar dois parâmetros: a média (μ) e o desvio padrão (σ).

12.1.3 Exemplo de aplicação

Lançar um dado (hexaedro) honesto 60 vezes e observar os seguintes resultados descritos na seguinte tabela:

Tabela 7 - Frequências observadas e esperadas, obtidas de um experimento aleatório que consistiu no lançamento de um dado honesto

Face do Dado	F_o	F_e	(F_o-F_e)	$(F_o-F_e)^2$	$(F_o-F_e)^2\ F_e$
1	8	10	-2	4	0,4
2	10	10	0	0	0,0
3	9	10	-1	1	0,1
4	12	10	2	4	0,4
5	10	10	0	0	0,0
6	11	10	1	1	0,1
Total	60	60	-	-	1,0

Onde, F_o = número de ocorrências verificadas na realização da experiência, e F_e = frequências esperadas ou calculadas, baseadas na hipótese de que cada face tem 1/6 de sucesso em cada jogada.

H_o: O dado é honesto

P (Face i) $= \frac{1}{6}$, i = 1,2,3,4,5,6

H_1: O dado não é honesto

P (Face i) $\neq \frac{1}{6}$, para algum i.

V = k-1= 6-1=5

$$\chi^2_{[5;0,05]} = 11,070$$

$$\chi^2_{Teste} = 1,00^{n.s.}$$

Portanto, de acordo com o teste do qui-quadrado, com cinco graus de liberdade e 5% de significância, aceita-se a hipótese H_o, ou seja, o dado é considerado não viciado ou honesto.

12.1.4 Procedimentos para se efetuar o teste do qui-quadrado (χ^2)

a) Enunciar as hipóteses de nulidade e alternativa, dadas por:
   H_o: $F_{oi} = F_{ei}$ ($\chi^2 = 0$),
   H_1: $F_{oi} \neq F_{ei}$ ($\chi^2 > 0$)

b) Escolher o nível de significância α

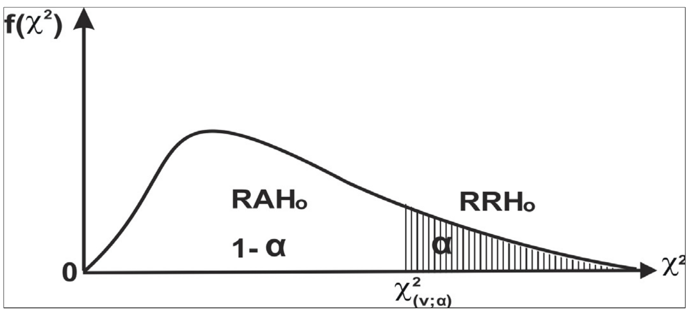
c) Calcular a estatística teste dada por:

$$\chi^2_{Teste} = \chi^2_{Calculado} = \frac{\sum_{i=1}^{k}\left(f_{o_i} - f_{e_i}\right)^2}{f_{e_i}}$$

Esta quantidade segue aproximadamente a distribuição de χ^2 com V = K-1 graus de liberdade.

d) Construir a região de rejeição da hipótese H_0, a qual é dada por $\left[\chi^2_{Teste} = \chi^2_{Calculado}\right] \geq \left[\chi^2_{Tabelado} = \chi^2_{Crítico}\right]$, sob a curva do teste que é a curva da distribuição teórica de probabilidade de qui-quadrado. Note-se que, nesse teste, a região de rejeição está sempre na cauda superior da curva, ou seja, a prova é um teste unilateral, pois o que se quer provar é a hipótese de que as frequências observadas sejam diferentes das frequências esperadas ou calculadas, ou seja, que $f_o \neq f_e$.

Figura 9 - Curva do teste de hipótese de qui-quadrado, mostrando as regiões de aceitação e de reijeição de Ho



e) Obter as conclusões do teste. Se $\left[\chi^2_{Calculado}\right] \geq \left[\chi^2_{Tabelado}\right]$, então rejeita-se a hipótese H_o; se $\left[\chi^2_{Calculado}\right] < \left[\chi^2_{Tabelado}\right]$, então aceita-se a hipótese H_o.

É importante lembrar que, nesse Teste, não há uma hipótese alternativa claramente definida; portanto o erro do tipo II e, consequentemente, o poder do teste, não pode ser computado ou determinado exatamente.

12.1.5 Exemplo

Suponhamos que um jogador deseja saber se um dado é viciado. Para isso, ele contou o número de vezes que cada face apareceu voltada para cima em 60 jogadas. Admitindo-se que o dado é honesto, a probabilidade de aparecer cada face, numa jogada é $\frac{1}{6}$. A previsão teórica para as frequências esperadas é $F_{ei} = \frac{1}{6} \times 60 = 10$ (i = 1, 2, 3, 4, 5, 6). O resultado das jogadas foi o seguinte:

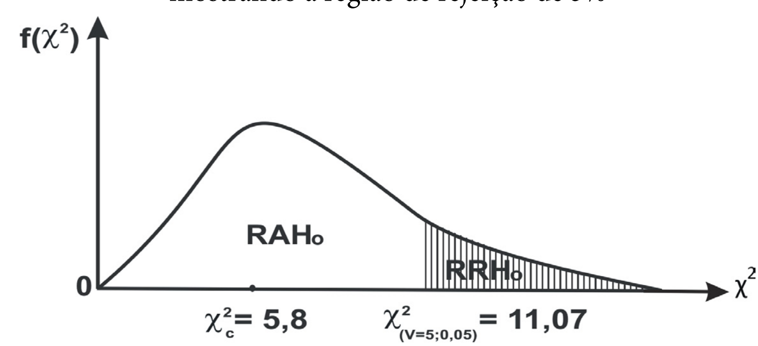
Tabela 8 - Frequências observadas e esperadas, obtidas de um experimento referente ao lançamento de um dado honesto

Faces	1	2	3	4	5	6
F_o	4	12	9	11	10	14
F_e	10	10	10	10	10	10

O nível de significância é α = 0,05

A região de rejeição da hipótese H_0 é a seguinte.

Figura 10 - Curva do teste de hipótese de qui-quadrado, mostrando a região de rejeição de 5%



O número de graus de liberdade é dado por: V = 6-1 = 5, portanto o valor crítico ou tabelado de qui-quadrado é: $\chi^2_{[5;0,05]} = 11,07$.

O valor da estatística teste é o seguinte:

$$\chi^2_{Teste} = \frac{(4-10)^2}{10} + \frac{(12-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(11-10)^2}{10} + \frac{(10-10)^2}{10} + \frac{(14-10)^2}{10}$$

$$\chi^2_{Teste} = 5,80^{n.s.}$$

Como o valor do qui-quadrado teste calculado é menor que o valor do qui-quadrado tabelado, então aceita-se a hipótese H_0.

As conclusões são de que, de acordo com o teste do qui-quadrado com cinco graus de liberdade, e um nível de significância de 5% de probabilidade, pode-se concluir, portanto, que a hipótese de nulidade será aceita, isto é, as frequências observadas não diferem das frequências calculadas mais do que seria esperado na variação acidental do processo de amostragem. A decisão é considerar o dado como honesto ou não viciado.

12.1.6 O teste de qui-quadrado (χ^2) nas tabelas de contingência 2x2

Alguns pesquisadores preferem calcular a estatística de qui-quadrado (χ^2) nas tabelas de contingência 2x2, de uma maneira diferente, usando apenas as frequências observadas. Se um atributo ou categoria for qualificado em A e B e o outro em C e D, e as frequências observadas de cada célula forem classificadas em a, b, c e d, conforme mostra a seguinte disposição tabular, então temos que:

Tabela 9 - Frequências observadas numa tabela 2x2

	C	D	TOTAL
A	a	b	a+b
B	c	d	c+d
Total	a+c	b+d	a+b+c+d = N

A estatística de qui-quadrado (χ^2) terá para o seu valor a seguinte equação:

$$\chi^2_{teste} = \frac{N\left[a.d - b.c\right]^2}{\left[a+b\right]\left[c+d\right]\left[a+c\right]\left[b+d\right]}$$

Se for aplicada no Teste χ^2 a correção de Yates, temos a seguinte estatística teste resultante:

$$\chi^2_{testeCorrigido} = \frac{\left[\left|a.d - b.c\right| - \frac{N}{2}\right]^2 .N}{[a+b][a+c][c+d][b+d]}$$

Ou a seguinte estatística

$$\chi^2_{Corrigido} = \frac{\sum_{i=1}^{k}\left(\left|f_{o_{ij}} - f_{e_{ij}}\right| - 0{,}5\right)^2}{f_{e_{ij}}}$$

Essa correção é justificada devido ao fato de a distribuição da estatística teste qui-quadrado

$$\chi^2_{Teste} = \chi^2_{Calculado} = \frac{\sum_{i=1}^{k}\left(f_{o_i} - f_{e_i}\right)^2}{f_{e_i}}$$

ser discreta e representar uma aproximação da distribuição de qui-quadrado, que é continua. A aproximação será mais eficiente à medida que n aumentar. Para v =1 graus de liberdade (caso da tabela 2x2), a aproximação é menos eficiente. Por isso é recomendado o uso da chamada correção de continuidade, devida a Yates. Esse tipo de correção é similar ao que foi efetuada no uso da aproximação da distribuição normal para distribuição binomial.

Esse teste, corrigido quando utilizado para o teste de associação (independência), é mais conservativo que aquele em que se utiliza a estatística sem a correção, diminuindo a chance de se rejeitar H_0.

É importante observar que, em tabelas 2x2, a raiz quadrada do qui-quadrado calculado ou qui-quadrado teste é igual ao Z calculado ou Z teste. Por exemplo, se o valor do qui-quadrado teste se for igual a 3,84, então o Z calculado é igual a 1,96, pois $(3{,}84)^{\frac{1}{2}} = 1{,}96 = Z_{calculado}$.

Portanto, em tabelas 2x2, para verificar se existe efeito de tratamento, no caso de duas amostras independentes, o pesquisador pode optar: (i) testar a igualdade das proporções com o calculo da estatística normal padrão $Z \sim N(0;1)$; ou, (ii) testar a associação entre tratamento e resposta com o cálculo da estatística teste qui-quadrado (χ^2) com 1 grau de liberdade.

12.1.7 Limitações do teste de qui-quadrado (χ^2)

a) A experiência e a teoria indicam que podemos empregar a prova de qui-quadrado (χ^2), nas tabelas de 1 x k ou h x 1, quando as frequências esperadas em todas as células ou classes sejam pelo menos iguais a 5. Se a frequência esperada numa célula ou classe for menor do que 5, esta célula deve ser combinada com uma ou mais células adjacentes, até que a condição seja satisfeita.
b) A distribuição teórica de probabilidade de qui-quadrado é uma distribuição contínua de valores quando o número de graus de liberdade é igual a um (V = G.L. = 1), porém a distribuição empírica do qui-quadrado teste não se aproxima suficientemente da distribuição teórica de probabilidade do qui-quadrado para permitir aplicar-se testes adequados. Para aproximar melhor essas duas distribuições, usa-se a correção para continuidade proposta pelo matemático e estatístico Inglês Frank Yates (1902 – 1994), que será apresentada no próximo item.
c) Segundo Cochran (1954), o Teste de χ^2 não deve ser usado nas tabelas de contingência tipo 2x2, nos seguintes casos:
   – Se a frequência total **N** for menor do que **20**.
   – Se a frequência total **N** estiver entre **20** e **40** e a frequência esperada mínima for menor do que 5.
   – Em todos os casos em que a frequência esperada for inferior a **1**.

– Mesmo que a frequência total **N** seja maior do que **40**, mas se tenha $\mathbf{F_e \leq 5}$, deve-se fazer a correção de Yates, a qual foi proposta pelo matemático e estatístico Inglês Frank Yates (1902 –1994), que será mostrada no próximo item.

d) Nas tabelas de contingência 2x2 que apresentarem frequência esperada mínima inferior a **20**, recomenda-se empregar a correção de continuidade ou correção de Yates, que consiste em subtrair 0,5 da diferença em valor absoluto, entre as frequências observadas e esperadas, ou seja,

$$\chi^2_{testeCorrigido} = \frac{\left[\left|f_o - f_e\right| - 0,50\right]^2}{f_e}$$

Como regra empírica, o fator de correção tem pouco efeito e pode ser omitido quando a frequência total $N \geq 50$. Outra regra é a de que a correção não deve ser aplicada a nenhuma casela, classe ou célula para a qual a diferença entre as frequências observadas e as frequências esperadas seja menor do que 0,50 (F_o - F_e < 0,50). Outra regra é a de que não se faz correção de continuidade (correção de Yates) em tabelas com dois ou mais graus de liberdade. Outra regra também observada é a de que, mesmo em tabelas com um grau de liberdade, só tem sentido aplicar tal correção se a hipótese nula tiver sido rejeitada. E, finalmente uma última regra, é a de que essa correção deve também ser usada se a frequência total N estiver entre 20 e 40, ou seja, $20 \leq N \leq 40$, e todas as frequências esperadas são iguais ou superiores a 5. Esse teste de qui-quadrado (χ^2), assim corrigido, não goza da propriedade de aditividade.

Correções de continuidade são rotineiramente recomendadas nos testes de qui-quadrado com i grau de liberdade (V = G.L = 1), e somente nesses casos. O número de graus de liberdade é 1 em tabelas de entrada simples com duas categorias e em tabelas 2x2. A correção para continuidade mais conhecida para este teste é a correção de Yates, embora sejam encontradas várias outras propostas na literatura como em Zar (1999), o qual mostra como empregar a correção de Haber, melhor que a de Yates, mas de cálculo menos simples.

e) O teste de qui-quadrado (χ^2) é válido apenas para as frequências absolutas, se o pesquisador tiver os dados em porcentagens que poderão ser transformadas em frequências absolutas, desde que se conheça a frequência total, e a transformação se dá por $f_i = \frac{N.A}{100}$, em que "N" é a frequência absoluta total e A é a porcentagem. Ou então modificando-se a fórmula de cálculo do qui-quadrado teste.
f) Só se deve aplicar o teste de qui-quadrado (χ^2) em tabelas de 2xK, com K>2, quando
   a) a frequência esperada mínima não for inferior a **1** (um);
   b) que só em poucos casos seja a frequência esperada menor do que 5 (menos de 20% das frequências esperadas forem abaixo de 5).
g) Frequências esperadas mínimas e correção de continuidade.

Os valores calculados da estatística do teste qui-quadrado são baseados em contagem discreta, enquanto a distribuição qui-quadrado é uma distribuição contínua. Quando as frequências F_e. Para as caselas, classes ou células da tabela não são pequenas, tal fato não é importante em termos da extensão na qual a distribuição da Estatística de Teste é aproximada pela distribuição qui-quadrado. Uma regra frequentemente utilizada é a de que a frequência esperada F_e, para cada casela ou categoria, deve ser pelo menos igual a cinco (5). As caselas que não se ajustam a esse critério devem ser combinadas com as categorias adjacentes, quando for possível, a fim de atender o requisito. O número reduzido de categorias passa a ser a base para determinar os graus de liberdade (G.L.) aplicáveis a essa situação.

Demonstra-se que, quando há apenas um grau de liberdade associado com o teste qui-quadrado (Tabela 2x2 por exemplo), a menos que a amostra seja muito grande, o valor calculado de χ^2 é sistematicamente supervalorizado devido ao caráter discreto dos dados. O estatístico Yates demonstrou que a seguinte fórmula, que inclui uma correção de continuidade, é apropriada quando o grau de liberdade for igual a 1 (G.L=1). Como regra empírica, o fator de correção tem pouco efeito e pode ser omitido quando N ≥ 50. Além disso, não deve ser aplicada a nenhuma casela para a qual a diferença entre a frequência observada e a frequência esperada (F_o - F_e) seja menor do que 0,5. Essa correção, já mostrada anteriormente, é dada pela seguinte equação:

$$\chi^2_{testeCorrigido} = \frac{\left[\left|f_o - f_e\right| - 0,50\right]^2}{f_e}$$

Fonte: Pimentel e Gomes (1985).

Segundo Sidia (2003), estudos recentes, como Everitt (1992) e Zar (1999), sugerem que as exigências quanto às frequências esperadas são rigorosas demais, sendo que muitos valores das frequências esperadas (fe) podem ser iguais a 1 sem afetar, de modo significativo, o resultado do teste. Sendo assim, uma abordagem mais moderna deve levar em conta as seguintes condições:

– Em teste de ajustamento (tabelas de entrada única) Tabelas com apenas duas categorias (k=2). A frequência esperada (fe) deve ser 5 ou mais em cada categoria, e usa-se a correção de continuidade de Yates para o cálculo do qui-quadrado calculado ou qui-quadrado teste, como será explicado a seguir.
  Se alguma fe < 5 é preferível obter diretamente o valor – p associado ao teste de hipóteses pela distribuição binomial;

Tabelas com K > 2 e todos os valores esperados iguais: para testes, usando α = 0,05, os valores da fe devem ser iguais ou maiores do que 1,0; para α = 0,01, a fe deve ser igual ou maior do que 2,0;
Tabelas com K > 2 e valores esperados diferentes: aplica-se o teste de qui-quadrado se forem satisfeitas três exigências: $n \geq 10$ e $\frac{n^2}{K} \geq 10$ e $\frac{n}{K} \geq 2$ para testes com $\alpha = 0{,}05$, se $\alpha = 0{,}01$ a última exigência fica assim $\frac{n}{K} \geq 4$.

– Em tabelas de contingência (dupla entrada)

=> Tabelas 2x2 (com duas linhas e duas colunas): nenhuma fe pode ser menor do que 5. Além disso, deve-se utilizar a correção de Yates no cálculo da estatística teste. Se o valor esperado mínimo não for alcançado, usar o teste exato de Fisher;
=> Tabelas 2xk (com duas linhas e mais de duas colunas): o teste de qui-quadrado pode ser aplicado se todos os valores das frequências esperadas forem menores ou igual a um;
=> Tabelas hxk (com mais de duas linhas e mais de duas colunas): o teste de qui-quadrado é um procedimento seguro se o número esperado médio for 6,0 ou maior para testes com α = 0,05, e 10,0 ou maior para testes com α = 0,01.O esperado médio pode ser obtido, dividindo-se o total de indivíduos estudados pelo número de caselas ou classes.

Vale lembrar que o moderno desenvolvimento da computação eletrônica e programas estatísticos não é difícil, obter atualmente, o nível crítico amostral exato do qui-quadrado calculado para amostras que apresentam valores das frequencias esperadas ou calculadas (fe) pequenos demais.

h) Para as tabelas com h linhas e k colunas, ou seja, hxk em geral (de 3x3, 3x4, 5x7, 4x8, etc.). O Teste de qui-quadrado (χ^2) pode ser aplicado de maneira semelhante e com as restrições mostradas nos itens anteriores.
i) Sobre os níveis de significância do Teste de qui-quadrado (χ^2). Uma particularidade importante do Teste de qui-quadrado (χ^2), é a de que, sendo feito com quadrados de desvios, não se distingue o sinal desses desvios, ou seja, ele corresponde a um teste bilateral para desvios. Se o pesquisador tiver através de conhecimento prévio, uma hipótese ou por trabalhos antigos no que se refere aos desvios, se para mais ou para menos, as tabelas originais de qui-quadrado (χ^2) podem ser utilizadas, mas tomando a metade da probabilidade ou área indicada, isto é, os limites de 10% correspondem ao nível de significância de 5%, e assim por diante.
j) O Teste de qui-quadrado é, específico para dados agrupados, em que as classes que possuírem valores menores que três ou cinco devem ser agrupadas em outras classes, sendo um fator limitante para uso em dados com poucas classes. Esse teste é baseado na soma dos erros absolutos das frequências, que é comparada com um valor tabelado de acordo com o nível de significância desejado e os graus de liberdade da distribuição. Isso favorece o aspecto cumulativo dos erros pelo somatório.
k) O teste de qui-quadrado é considerado mais rigoroso do que o teste de Kolmogorov – Smirnov (que tem o mesmo objetivo do teste de qui-quadrado), o que se deve principalmente, por exemplo, quando aplicado no ajuste em séries históricas de precipitação pluviométrica e, em outros tipos de séries de dados, aos seguintes aspectos:

- Considerando que uma distribuição sob teste tenha duas ou mais classes coprobabilidades observadas diferentes das estimadas e, consequentemente, frequências da mesma forma, quando se aplicam esses valores à equação de definição do teste de qui-quadrado, ou seja, a equação

$$\chi^2_{Teste} = \chi^2_{Calculado} = \frac{\sum_{i=1}^{k}\left(f_{o_i} - f_{e_i}\right)^2}{f_{e_i}},$$

tem-se um somatório dos erros absolutos; já quando se aplica o teste de Kolmogorov – Smirnov por exemplo, tem-se um único valor, que é o módulo da diferença.

Isso mostra que os erros, no teste de qui-quadrado, são considerados de forma cumulativa e em todas as classes e que, no teste de Kolmogorov – Smirnov, por exemplo, eles são considerados somente na classe em que foi maior;

- O quadro de frequências de uma série histórica de precipitação pluviométrica, por exemplo, apresenta maiores valores nas classes iniciais e menores valores nas classes finais, e a definição do teste de qui-quadrado determina que devem ser reunidas em uma única classe e as classes com frequências estimadas inferiores a três ou cinco. Os modelos matemáticos ajustados e testados geralmente superestimam as classes iniciais e subestimam as classes finais, com algumas exceções.

Para atender as definições do teste de qui-quadrado, as classes estimadas com frequência inferior a três ou cinco devem ser somadas a outra classe mais próxima. Decorre daí que o somatório ocorrerá também nas classes de frequência observadas, gerando um erro absoluto grande, que, somado aos anteriores, resulta em valores de qui-quadrado maiores que os valores tabelados, não aprovando, assim, a distribuição sob teste, quando a estimação não for boa;

- Quando a situação anterior ocorre, esse fato não afeta, por exemplo, a aplicação do teste de Kolmogorov – Smirnov, ou seja, não surgirá problema algum, pois, independentemente da distribuição de classes, o que interessa é o módulo da maior diferença, permitindo que o teste aprove a maioria das distribuições, com muitos erros, mas de pequena proporção.
- Os valores de qui-quadrado calculado ou qui-quadrado teste, como já foi mostrado, são comparados com valores críticos tabelados.

Esses valores críticos são obtidos de tabelas referenciadas pelo nível de significância e pelo grau de liberdade, no caso do qui-quadrado, e pelo nível de significância e pelo número de observações, no teste de Kolmogorov – Smirnov. Observa-se que, nesse teste, independentemente da capacidade da distribuição em estimar as frequências observadas e do número de classes, o valor crítico tabelado depende unicamente do número de observações, algo que não varia de distribuição para distribuição, dependendo apenas da série sob teste. Considerando agora o teste de qui-quadrado, vê-se que o grau de liberdade depende dos parâmetros da distribuição, em torno de dois ou três, e do número de classes (características dos dados). Analisando a observação feita no item "xi.1", este número reduz quando a distribuição subestima as classes finais, devido ao agrupamento de algumas classes em outras, e o número de gruas de liberdade fica menor, reduzindo o valor crítico tabelado; isso mostra que o valor crítico tabelado para o teste de qui-quadrado depende da capacidade da distribuição em estimar as frequências observadas, o que não ocorre, por exemplo, quando se aplica o teste de Kolmogorov – Smuirnov.

12.1.8 Aplicações práticas do teste de qui-quadrado (χ^2)

12.1.8.1 Teste de ajustamento

12.1.8.1.1 Objetivo

São usadas para verificar se resultados experimentais estão coerentes com resultados teóricos, conhecidos a "priori", ou são usados na comprovação de uma lei natural (lei da segregação dos sexos da genética), ou verificar hipóteses elaboradas por raciocínio indutivo verificar se um dado ou hexaedro é honesto, ou seja, P(Face i) = 1/6.

A hipótese H_o diz que os resultados experimentais obedecem à Lei Natural, e a hipótese alternativa H_1 diz o contrário da hipótese H_o.

A estatística teste é o qui-quadrado de Pearson dado por

$$\chi^2_{Teste} = \chi^2_{Calculado} = \frac{\sum_{i=1}^{k}\left(f_{o_i} - f_{e_i}\right)^2}{f_{e_i}}$$

o qual possui distribuição teórica de probabilidade de qui-quadrado com V = K – 1 graus de liberdade, onde "K" é o numero de categorias, atributos, classes, casas, caselas ou células na amostra em estudo.

12.1.8.1.2 Exercício de aplicação

Em um experimento aleatório foram postos 1000 ovos de galinhas poedeiras da linhagem Shaver White para chocar. No dia da eclosão, nasceram 950 pintos, sendo 580 pintos machos e 370 pintos fêmeas. Sendo a relação esperada de pintos machos para pintos fêmeas de 1:1 pretende-se saber se as observações estão de acordo com essa hipótese ou lei natural da segregação dos sexos. Utilize-se um nível de significância α = 1%.

A resolução do teste de ajustamento é a seguinte:

1) H_o: Os dados da proporção de pintos machos e pintos fêmeas seguem a proporção 1:1, ou seja, $P_{[1]} = P_{[Macho]} = \frac{1}{2}, P_{[2]} = P_{[Fêmea]} = \frac{1}{2}$ ou obedecem a essa lei natural.

   H_1: Os dados da proporção de pintos machos e pintos fêmeas não seguem a proporção 1:1, ou não obedecem a essa lei natural $P_{[1]} = P_{[Macho]} = \frac{1}{2}, P_{[2]} = P_{[Fêmea]} = \frac{1}{2}$.
2) α = 1%, e o número de graus de liberdade é dado por: V = 2 – 1 = 1
3) Construção da curva do teste

Figura 11 - Curva do teste de qui-quadrado, mostrando a região de rejeição de H_o correspondente a 1% de probabilidade

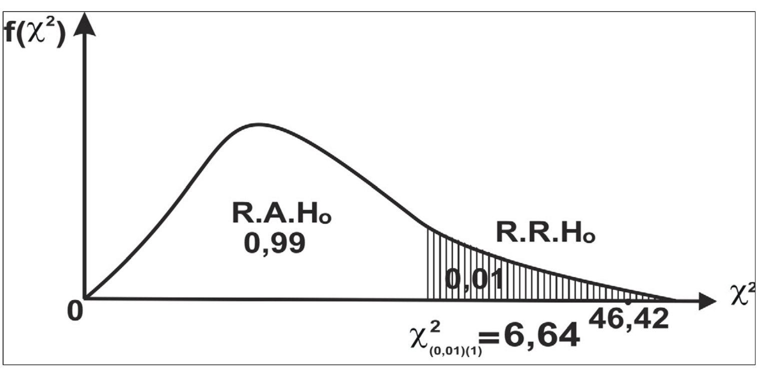


Tabela 10 - Frequências observadas e esperadas do número de pintos machos e fêmeas após a eclosão dos ovos

Frequências	Sexos		Total
	Machos	Fêmeas	
Fo	580	370	950
Fe	475	475	950

As frequências esperadas são obtidas assim:

Para pintos machos $f_{e_1} = N.P_1 = 950.\frac{1}{2} = 475$ e

para pintos fêmeas $f_{e_2} = N.P_2 = 950.\frac{1}{2} = 475$

4) $\chi^2[Observado] = \frac{(580-475)^2}{475} + \frac{(370-475)^2}{475} = 46,42^{**}$

$\chi^2[Observado] > \chi^2[Crítico]$, então rejeita-se a hipótese H_{0}.

Como já comentado, sendo o valor do qui-quadrado observado ou qui-quadrado teste é maior que o valor do qui-quadrado crítico ou tabelado, então rejeita-se a hipótese H_{o}.

5) Conclusões: Baseado no Teste do Qui-Quadrado com 1 grau de liberdade e ao nível de 1% de probabilidade, pode-se concluir que a relação entre pintos machos e pintos fêmeas é diferente da relação esperada ou da lei natural 1:1.

12.1.8.2 Teste de aderência

12.1.8.2.1 Objetivo

É usado para verificar se um conjunto de dados amostrais ou uma distribuição de frequências, denominada de como distribuição empírica, segue ou pode ser representada por uma distribuição teórica ou especial de probabilidades conhecida, como a normal, binomial, Poisson, etc., ou ainda, quando é usado para determinar quão aproximadamente as distribuições teóricas, como a normal, a binomial, a Poisson, etc., se ajustam às distribuições empíricas, isto é, as obtidas por meio dos dados amostrais ou experimentais.

A hipótese nula H_0 estabelece qual é o tipo ou natureza da distribuição teórica de probabilidade da população da qual a amostra, ou as amostras foram obtidas, isto é, se a amostra n pertence a população. A hipótese nula seria H_0 : a população é o do tipo estabelecida versus a hipótese alternativa H_1: a população não é aquela estabelecida na hipótese de nulidade H_0.

Se chamarmos de F_{oi} os valores observados na Amostra para cada classe de frequência e de E_i, os valores que seriam esperados, se a variável aleatória x_n se distribuísse segundo a função ∅. A prova poderia ser realizada por meio da distribuição de qui-quadrado (χ^2) para nk -2 graus de liberdade (nk = número de classes de frequência). Nesse caso, o teste de qui-quadrado (χ^2) seria definido como

$$\chi^2 = \sum_{i=1}^{nk} \frac{\left[f_{o_i} - f_{e_i} \right]^2}{f_{e_i}} = \left[\sum_{i=1}^{nk} \frac{f_{o_i}^2}{f_{e_i}} \right] - n$$

Ou seja, a estatística teste é o qui-quadrado de Pearson dado por

$$\chi^2_{Teste} = \chi^2_{Calculado} = \frac{\sum_{i=1}^{k} \left(f_{o_i} - f_{e_i} \right)^2}{f_{e_i}}$$

o qual sob a hipótese H_0 possui distribuição teórica de probabilidade de qui-quadrado com V = K – 1 graus de liberdade, ou V = K – P – 1, quando, ao se utilizar no cálculo das frequências esperadas (f_e), a estimativa de “P” parâmetros populacionais baseada nas distribuições amostrais do estimador sob estudo, onde “K” é o numero de categorias, atributos, classes, casas, caselas ou células na amostra em estudo.

12.1.8.2.2 Exercício de aplicação

Avaliaram-se 30 parcelas (área de terra de 20 m2) de 27 plantas de mandioca (Manihot esculenta, Crantz). Cada uma, de acordo com o número de plantas doentes (atacadas por bacteriose) colhidas. Os dados estão apresentados na tabela abaixo (Distribuição de frequências ou Distribuição empírica).

Tabela 11 - Distribuição do número de plantas doentes (atacadas por bacteriose), encontradas em N=30 parcelas com plantas de mandioca (<u>Manihot</u> <u>esculenta</u>, Crantz)

	NÚMERO DE PLANTAS DOENTES COLHIDAS (X_i)	NÚMERO DE PARCELAS COM X PLANTAS DOENTES ($f_i = f_o$)	$X_{i.}\ f_i$	PROBABILIDADES $P(X_i)$	fe=N.P(X_i)	$f_{ecorrigida}$
	0	14	0	0,38	11,4	11,4
	1	8	8	0,37	11,1	11,1
	2	4	8	0,18	$\left.\begin{matrix}\mathbf{5,4}\\\mathbf{1,8}\\\mathbf{0,3}\end{matrix}\right\}\mathbf{7,5}$	7,5
	3	3	9	0,06		
	4	1	4	0,01		
SOMA	----	30	29	1,00	30	30

Com $P(X = x_i) = \dfrac{e^{-0,97}.0,97^{X_i}}{X_i!}, X = 0,1,2,3,4$

Onde $\overline{X} = \mu = \mu_X = \lambda = \dfrac{\sum_{i=1}^{K} X_i.f_i}{N} = \dfrac{29}{30} \cong 0,97\ PLANTAS$, sendo,

$P(X = 0) = \dfrac{e^{-0,97}.0,97^0}{0!} = 0,379083038 = 0,38$, V= K – P – 1 = 3 – 1 – 1 = 3 – 2= 1,

a) Testar, a 5% de significância, a hipótese de que os dados referentes à distribuição empírica do número de plantas de mandioca

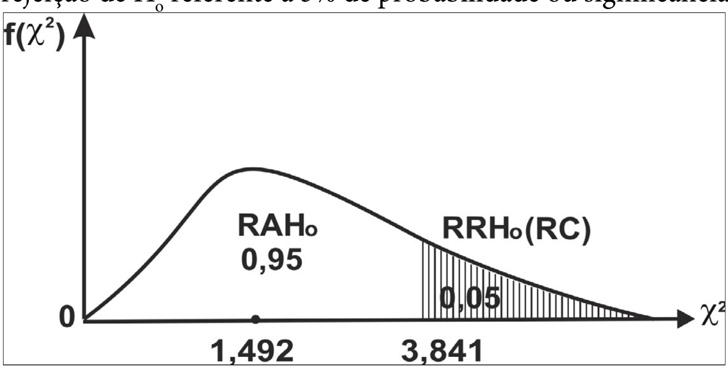
(<u>*Manihot esculenta*</u>, Crantz,) doentes, encontradas nas 30 parcelas, aproxima-se ou pode ser representada por uma distribuição teórica de probabilidades de "POISSON".

H_o: A distribuição empírica do número de plantas de mandioca doentes (atacadas por bacteriose) colhidas segue uma distribuição teórica de probabilidade de Poisson.

$H_{1:}$ A distribuição empírica do número de plantas de mandioca doentes (atacadas por bacteriose) colhidas não segue uma distribuição teórica de probabilidade de Poisson.

$$\alpha = 0{,}05, \quad V = K - P - 1 = 3 - 1 - 1 = 3 - 2 = 1, \chi^2_{(1;0{,}05)} = 3{,}841$$

Figura 12 - Curva do teste de qui-quadrado, mostrando a região de rejeição de H_o referente a 5% de probabilidade ou significância



$$\chi^2_{[Teste]} = \frac{\sum_{i=1}^{3}\sum_{j=1}^{3}\left(f_{0_{i,j}} - f_{e_{i,j}}\right)^2}{f_{e_{i,j}}}$$

$$\chi^2_{[Teste]} = \frac{(14-11{,}4)^2}{11{,}4} + \frac{(8-11{,}1)^2}{11{,}1} + \frac{(8-7{,}5)^2}{7{,}5} =$$

$$\chi^2_{[Teste]} = 0{,}593 + 0{,}866 + 0{,}033 = 1{,}492^{n.s}$$

Tabela 12 - Frequências observadas e esperadas oriundas de um experimento referente à contagem do número de plantas doentes colhidas de mandioca (Manihot esculenta, Crantz)

	NÚMERO DE PLANTAS DOENTES COLHIDAS (X_i)	NÚMERO DE PARCELAS COM X PLANTAS DOENTES ($f_i = f_o$)	$X_{i.}\ f_i$	PROBABILIDADES $P(X_i)$	$fe=N.P(X_i)$	$f_{ecorrigida}$	$[f_o - f_e]$	$[f_o - f_e]^2$	$[f_o - f_e]^2 / f_e$
	0	14	0	0,38	11,4	11,4	2,6	6,76	0,593
	1	8	8	0,37	11,1	11,1	-3,1	9,61	0,866
	2	4	8	0,18	**5,4**	7,5	0,5	0,25	0,033
	3	3	9	0,06	**1,8** } **7,5**				
	4	1	4	0,01	**0,3**				
SOMA	----	30	29	1,00	30	30	-	-	1,492 n.s

Veja como foi calculado o valor das estimativas da média da distribuição, bem como os valores das probabilidades dos diversos eventos aleatórios.

$$\overline{X} = \mu = \mu_{[X]} = \mu_X = \lambda = \frac{\sum_{i=1}^{5} X_i f_i}{N} = \frac{29}{30} \cong 0,97$$

$$P(X = x_i) = \frac{e^{-0,97} . 0,97^{X_i}}{X_i!} \quad \forall\ X_i = 0,1,2,3,4$$

$$P\ (X = 0) = \frac{0,97^0 . e^{-0,97}}{0!}, = 0,379053038 \cong 0,38$$

Como o número de graus de liberdade é dado por

V = 3 - 1 - 1 = 3 - 2 = 1,

Então o valor tabelado ou crítico sob a curva da distribuição de probabilidade de qui-quadrado é 3,841.

$\chi^2_{[1;0,05]} = 3,841$, sendo o valor da estatística teste $\chi^2_{Teste} = 1,492^{n.s}$.

Como $\chi^2_{[1;0,05]} = 3,841 > \chi^2_{[Teste]} = 1,492^{ns}$, então aceita-se a hipótese H_0.

– Conclusões: De acordo com o teste do qui-quadrado com um grau de liberdade e 5% de significância, pode-se concluir que a distribuição empírica do número de plantas de mandioca doentes (atacadas por bacteriose), colhidas nas parcelas, aproxima-se ou segue, ou ainda pode ser representada por uma distribuição teórica de probabilidade de Poisson.

12.1.8.3 Teste de independência

12.1.8.3.1 Objetivos

O teste é usado, nessa situação, quando duas variáveis categóricas estão classificadas segundo atributos, categorias, eventos ou variáveis qualitativas que necessariamente não identifiquem distintas populações. Nesse caso, o pesquisador está preocupado em medir o grau de associação entre as variáveis e colocará à prova ou teste as seguintes hipóteses:

H_o = As variáveis são independentes (o que é o mesmo que testar a hipótese de que duas probabilidades são iguais), contra.

H_1 = As variáveis não são independentes, ou seja, elas apresentam algum grau de associação entre si.

Ou ainda podemos escrever as hipóteses assim:

H_o = Não existe associação entre tratamento e resposta.

H_1 = Existe associação entre tratamento e resposta.

As frequências esperadas encontram-se sempre sujeitas a uma hipótese particular H_o, e uma hipótese comumente admitida é que as duas classificações são independentes uma da outra.

Para aplicação desse teste, os dados devem estar dispostos em uma tabela de contingência constituída de h linhas e k colunas. Vale lembrar que contingência significa dependência, de modo que uma tabela de contingência nada mais é do que uma tabela que mostra como duas características ou atributos dependem uma da outra.

Da mesma forma que nos casos anteriores, a estatística teste de qui-quadrado (χ^2) é calculada pela fórmula de Karl Pearson, mostrada abaixo, que tem uma distribuição amostral que muito se aproxima da distribuição teórica de probabilidade, ou melhor ainda, da função de densidade de qui-quadrado (χ^2), desde que as frequências esperadas não sejam muito pequenas.

$$\chi^2_{[Teste]} = \frac{\sum_{i=1}^{3}\sum_{j=1}^{3}\left(f_{0_{i,j}} - f_{e_{i,j}}\right)^2}{f_{e_{i,j}}}$$

Ou seja, a estatística teste é o qui-quadrado de Pearson dado por

$$\boldsymbol{\chi^2_{Teste} = \chi^2_{Calculado} = \frac{\sum_{i=1}^{k}\left(f_{o_i} - f_{e_i}\right)^2}{f_{e_i}}}$$

o qual possui distribuição teórica de probabilidade de qui-quadrado com o número de graus de liberdade "V", obtido conforme é mostrado no próximo item, sendo que "K" é o numero de categorias, atributos, classes, casas, caselas ou células em que a amostra em estudo é dividida.

Uma tabela de contingência pode ser visualizada conforme a tabela 13 abaixo, onde "A" e "B" são as variáveis qualitativas, eventos ou categorias em estudo, $f_{o_{11}}$, é a frequência observada das categorias A_1 e B_1, simultaneamente, e "N" é o valor da soma total das frequências observadas, além de que $M_{1.}$ representa o total marginal da primeira linha, e $N._1$ o total marginal da primeira coluna.

Tabela 13 - Tabela de contingência de duas variáveis qualitativas, categorias ou atributos sob a hipótese de independência entre elas

CATEGORIA B	CATEGORIA A				TOTAL
	A_1	A_2	…	A_k	
B_1	$f_{o_{11}}$	$f_{o_{12}}$	…	$f_{o_{1K}}$	$M_1.$
B_2	$f_{o_{21}}$	$f_{o_{22}}$	…	$f_{o_{2K}}$	$M_2.$
⋮	⋮	⋮		⋮	⋮
B_L	$f_{o_{L1}}$	$f_{o_{L1}}$	…	$f_{o_{LK}}$	$M_L.$
TOTAL	$N._1$	$N._2$	…	$N._k$	N

As hipóteses a serem testadas são as seguintes.

$$H_0: P(A_1 \cap B_1) = P(A_1).P(B_1); P(A_1 \cap B_2) = P(A_1).P(B_2); P(A_1 \cap B_L) = P(A_1).P(B_L); \cdots;$$
$$; \cdots; P(A_K \cap B_1) = P(A_K).P(B_1); P(A_K \cap B_2) = P(A_K).P(B_2); P(A_K \cap B_L) = P(A_K).P(B_L)$$

Contra uma hipótese alternativa, dada por:

$$H_1: P\left(A_i \mid B_j\right) \neq P(A_i).P\left(B_j\right) \forall \; par\,(i,j)$$

As frequências teóricas, esperadas ou calculadas fe, são obtidas sob a hipótese Ho, a qual sugere independência entre as variáveis ou características e, para exemplificar como se obtém tais frequências, mostraremos como é calculada apenas a primeira fe, e as demais são obtidas por subtração das outras frequências, haja vista que a soma das frequências observadas deve ser igual à soma das frequências esperadas. Sendo assim, temos que

$$\frac{f_{o_{11}}}{N} = \frac{M_{1.}}{N}\frac{N_{.1}}{N} \therefore \frac{f_{o_1}}{N}N = \frac{M_{1.}}{N}\frac{N_{.1}}{N}N \therefore f_{o_1} = \frac{M_{1.}N_{.1}}{N}.$$

Ou seja, a primeira frequência esperada é obtida através da razão entre o produto dos totais marginais da primeira linha pelo total marginal da primeira coluna pelo total geral N, e as demais frequências esperadas são obtidas, como já mencionado anteriormente, através da subtração do total marginal de uma linha ou de uma coluna menos essa frequência esperada já obtida.

12.1.8.3.2 O número de graus de liberdade

O número de graus de liberdade "V" para h > 1 e K > 1, onde "h" é o número de linhas e "K" é o número de colunas da tabela de contingência, é determinado por:

a) V = K - 1, nas tabelas de simples entrada, onde k é o número de classes, casas ou células.

b) V = K - P - 1, nas tabelas de simples entrada, quando as frequências esperadas são calculadas através da estimativa de "P" parâmetros populacionais, usando-se as estatísticas Amostrais.

c) V = (h-1) (k-1), nas tabelas h x k, quando as frequências esperadas podem ser calculadas sem que se tenha que estimar os parâmetros populacionais por meio das estatísticas amostrais do estimador em questão ou sob estudo.

d) V = (h - 1) (k - 1) - P, nas tabelas h x k, quando as frequências esperadas somente podem ser calculadas mediante à estimativa de "P" Parâmetros Populacionais, através das estatísticas amostrais do estimador sob estudo.

12.1.8.3.3 Coeficiente de contingência

a) Introdução

Muitos pesquisadores consideram que a análise estatística está pronta logo que terminam de aplicar o teste de qui-quadrado. Não deveriam, porque é importante estimar o grau de associação entre as duas variáveis. Isso porque o teste de qui-quadrado serve para verificar a significância da associação, mas não para medir o grau da associação entre duas variáveis.

O fato acima ocorre devido à significância de todo teste estatístico depender muito do tamanho da mostra "n". No caso do teste de qui-quadrado não é diferente, pois a significância depende não só das diferenças entre as proporções, mas também do tamanho da amostra. O grau de associação, no entanto, independe do tamanho da amostra, pois ele é função das proporções observadas.

Além das considerações anteriores, verifica-se que o valor do qui-quadrado (χ^2) é função do tamanho amostral, o que torna essa estatística imprópria para medir o grau da associação entre duas características. Sendo assim, o qui-quadrado teste (χ^2) deve ser usado para verificar a significância e não o grau de associação.

Sendo assim, veja como medir o grau de associação entre duas variáveis.

b) Coeficientes de associação

É fácil verificar que o valor do qui-quadrado teste é zero

$$\left\{\chi^2_{[teste]} = 0\right\}$$

somente quando as frequências observadas são iguais às frequências esperadas, indicando não-associação entre as duas variáveis qualitativas. Por outro lado, quando

$$\left\{\chi^2_{[teste]} > 0\right\}$$

temos a indicação de associação entre as duas variáveis ou atributos. O valor do teste de qui-quadrado pode variar de 0 (zero) a $+\infty$ (mais infinito). Como ele não tem limite superior, fica difícil para o pesquisador utilizar o seu valor para interpretar a associação entre as duas vaiáveis. Para solucionar esse problema, medidas de correlação ou associação baseadas nessa estatística foram criadas pelos estatísticos.

O Matemático e Estatístico Inglês Karl Pearson (1857 – 1936) propôs o chamado coeficiente de contingência, representado pela letra “C”, o qual é uma medida do grau de relacionamento, associação, dependência ou correlação das classificações em uma tabela de contingência, e é definido conforme a seguinte equação:

$$C = \sqrt{\frac{\chi^2_{[Calculado]}}{\chi^2_{[Calculado]} + n}}$$

Onde “n” é o número de observações, sendo que alguns autores utilizam “N” para representar esse número total ou frequência observada total, que é o mesmo que o número de elementos de uma amostra.

Quanto maior o valor de “C”, maior é o grau de associação entre as variáveis. O máximo valor de “C”, que será exemplificado a seguir, dependerá do número de linhas e de colunas da tabela e nunca excederá o valor 1 (um). Se, por exemplo, o número de linhas e de colunas da tabela for igual a “M”, então o valor máximo de “C” é dado pela seguinte equação:

$$C = C_{[Máximo]} = \sqrt{\frac{[M-1]}{M}}$$

Se considerarmos, por exemplo, que um pesquisador que estudar a relação ou associação entre duas variáveis qualitativas, dadas pela cor da corola e forma da folha em plantas de algodão herbáceo, para tanto ele construiu uma tabela de contingência 2x2, que mostra nesse experimento aleatório um valor de qui-quadrado teste obtido igual a 10, 67, e um número de observação "n" igual a 100, e se estudarmos essa correlação, temos que:

$$C = \sqrt{\frac{[10,67]}{[10,67+100]}} = 0,3105$$

indicando, assim, que existe associação; porém, ainda não podemos dizer se a associação é fraca ou não. Para tanto, precisamos conhecer qual é o maior valor que esse coeficiente pode assumir. É possível mostrar que esse valor é dado por

$$C = C_{[Máximo]} = \sqrt{\frac{[M-1]}{M}}$$

"M" é o mínimo entre o número de colunas e o número de linhas da tabela de contingência. Em tabelas cujo número de linhas ou colunas for igual a 2 (tabela 2x2), o maior valor que esse coeficiente pode assumir é

$$C_{[Máximo]} = \sqrt{\frac{[2-1]}{2}} = 0,71$$

Esse valor só é atingido quando temos associação perfeita. Assim, podemos concluir que existe uma associação moderada entre a cor da corola e a forma da folha de plantas de algodão herbáceo.

Como o limite superior do coeficiente de contingência de Pearson depende das dimensões da tabela, foi proposto um outro coeficiente, com valor variando entre 0 (zero) e 1 (um), dado pela seguinte equação,

$$C^* = \frac{C}{\sqrt{\frac{M-1}{M}}}$$

Para o exemplo, anterior temos que:

$$C^* = \frac{0,3105}{\sqrt{\frac{2-1}{2}}} = \frac{0,3105}{0,7071} = 0,4391$$

Sendo assim, o que se verifica é uma confirmação de uma associação moderada entre as variáveis. Vale a pena ressaltar que, mesmo sendo uma associação moderada, ela deve ser levada em consideração na interpretação dos resultados experimentais.

Existe ainda um outro coeficiente denominado de correlação de atributos para tabelas K x K – leia-se M x M, cujo campo de variação vai de 0 (zero) a 1 (um), ou seja, $0 \leq r \leq 1$, cujo valor é obtido através da seguinte equação,

$$r = \sqrt{\frac{\chi^2_{[Teste]}}{N[M-1]}}$$

c) Coeficientes de associação em tabelas 2x2

O Coeficiente Fi - φ

Conforme Vieira (2003), o coeficiente é uma medida da associação bastante conhecida e muito usada pelos pesquisadores das áreas de Psicologia e Sociologia, e é definido conforme a seguinte equação:

$$\varphi = \sqrt{\frac{\chi^2}{n}}$$

Em que χ^2 é o valor não corrigido do teste de qui-quadrado e "n" é o tamanho da amostra.

A interpretação do coeficiente φ é feita da seguinte maneira:

- Se o valor for igual a 1, o que, no caso de φ, só acontece quando as amostras são de mesmo tamanho, a associação é perfeita;
- Se o valor for igual a 0 (zero), a associação é nula;

- Quanto mais próximo estiver de 1, maior será o grau de associação entre as variáveis e, quanto mais próximo de zero, menor é a associação. Como regra prática, valores de φ menores do que 0,30 ou 0,35 podem ser tomados como indicadores de pequena associação.

  O coeficiente φ funciona bastante bem quando o estudo é transversal, mas não é tão bom no caso de estudos prospectivos e retrospectivos.

  O coeficiente φ, que varia entre 0 e 1, tem uma significância que depende da significância do teste qui-quadrado (χ^2).

  O Coeficiente - Gama γ

  Esse coeficiente conhecido como coeficiente gama, em referência à letra grega gama γ, mede o grau de associação com que duas categorias ordenadas de variáveis tendem a crescer e, portanto, decrescer juntas.

  O coeficiente γ é definido conforme a equação seguinte:

  $$\gamma = \frac{(ad - bc)}{(ad + bc)}$$

  em que a, b, c e d são as frequências observadas nas caselas da tabela 2x2, referentes às duas categorias estudadas.

  Como o coeficiente γ varia entre -1 e +1, a interpretação do resultado do valor obtido é feita da mesma forma que a interpretação do coeficiente de correlação.
- Se for igual a 1 ou a -1, a associação é perfeita, positiva ou negativa;
- Se o valor for igual a zero (0), a associação é nula;
- Quanto mais próximo estiver de 1, maior será o grau de associação positiva entre as variáveis, e quanto mais próximo de -1, maior será o grau de associação negativa entre as variáveis.

d) Exercício de aplicação 1

Considere o seguinte quadro de resultados do ensaio clínico de uma dada vacina, conduzido por um médico veterinário da UFERZAM, em Mossoró, RN, contra uma doença conhecida como BRUCELOSE BOVINA, aplicado num rebanho de 400 cabeças de gado bovino da raça Gir.

Tabela 14 - Frequência do número de animais da raça Gir, sadios e doentes, submetidos aos tratamentos Vacinados e Não Vacinados, através de um ensaio clínico de uma vacina conduzido por um médico veterinário

GRUPO	DIAGNÓSTICO		TOTAL
	SADIO	DOENTE	
Vacinados	250 (180,5)	6 (75,5)	256
Não vacinados	32 (101,5)	112 (42,5)	144
TOTAL	282	118	400

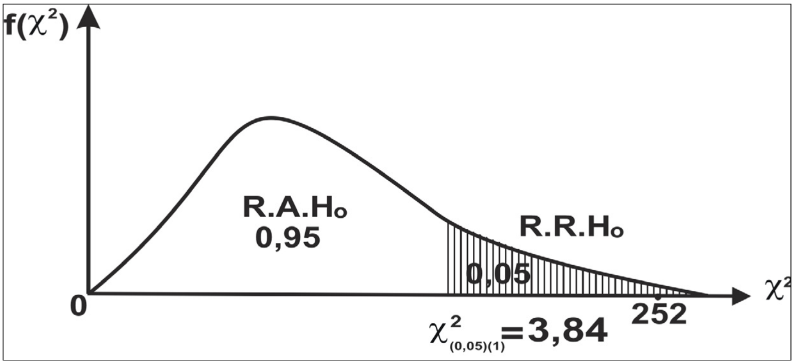
– Concluir sobre a associação (independência) entre a vacinação e a enfermidade, considerando um nível de significância α = 0,05 = 5%.

=> H_o: Não existe associação entre a vacinação e a enfermidade, isto é, essas variáveis são independentes.

H_1: Existe associação entre a vacinação e a enfermidade, isto é, essas variáveis são dependentes.

=> α = 0,05, como o número de graus de liberdade é dado por: V = (2-1) (2-1) = 1, então o valor crítico ou tabelado sob a curva da distribuição de probabilidades de qui-quadrado é: $\chi^2_{[1;0,05]} = 3,84$.

Figura 13 - Curva do teste de qui-quadrado, mostrando a região crítica ou de rejeição de H_o igual a 5% de probabilidade



$$\chi^2_{[teste]} = \frac{(250-180,50)^2}{180,50} + \frac{(32-101,50)^2}{101,50} + \frac{(6-75,50)^2}{75,50} + \frac{(112-42,50)^2}{42,50}$$

$$\chi^2_{[teste]} = 251,98 \cong 252^{*}$$

Como o valor do $\chi^2_{[teste]} > \chi^2_{[1;0,05]}$, então rejeita-se a hipótese H_0, pois 252 > 3,84.

Conclusão: De acordo com o teste do qui-quadrado com 1 grau de liberdade e um nível de significância de 5% de probabilidade, pode-se concluir que existe associação entre a vacinação e a enfermidade, ou seja, a doença depende da vacinação.

Vale a pena mostrar a seguinte observação sobre como foram obtidos os cálculos para a determinação das frequências esperadas.

Considere as seguintes variáveis qualitativas ou atributos as quais podem ser vistas ou consideradas também como eventos aleatórios.

A = Grupo	B = Diagnóstico
X = Vacinados	Z = Sadio
Y - Não vacinados	W = Doente

Como as frequências esperadas estão sempre sujeitas a uma particular hipótese H_0, que é aquela de que as variáveis são independentes, então, se usarmos o teorema dos eventos independentes para dois eventos, o qual é dado por $P(A \cap B) = P(A).P(B)$, temos que:

$$P(X \cap Z) = \frac{f_{e_1}}{400} \quad P(X) = \frac{256}{400}, \quad P(Z) = \frac{282}{400}$$

$$P(X \cap Z) = P(X).P(Z) \quad \therefore \quad \frac{f_{e_1}}{400}.400 = \frac{256}{400}.\frac{282}{400}.400$$

$$f_{e_1} = \frac{256.282}{400} = 180{,}48 = 180{,}50.$$

As outras frequências esperadas são determinadas através das diferenças entre os totais marginais de colunas ou de linhas, senão vejamos a seguir.

Como já foi dito anteriormente que em cada categoria a soma total das frequências observadas deve ser igual à soma total das frequências esperadas, e sabendo dessa propriedade, podemos determinar as outras frequências esperadas através da diferença entre o total marginal da primeira coluna menos o valor da primeira frequência esperada, localizada na primeira casela a qual neste caso é: 282 – 180,5 = 101,5, e as outras também, por, diferença são as seguintes: 256 – 180,5 = 75,5 e 144-101,5 = 42,5. Sendo assim, só precisamos calcular uma frequência esperada, pois as outras já estão previamente determinadas, usando esse conhecimento, como mostrado anteriormente. Nesse caso, dizemos que temos apenas um grau de liberdade ou que necessitamos calcular apenas uma só frequência esperada ou teórica.

e) Exercício de aplicação 2

Os dados seguintes mostram o número de sementes de ervas daninhas contidas em 98 subamostras de *Phleum Praetense* (Gramínea americana), em que cada subamostra pesou 250 gramas e continha muitas sementes, das quais somente uma pequena porcentagem era nociva.

Tabela 15 - Distribuição do número de sementes de ervas daninhas encontradas em N = 98 subamostras de Phleum Praetense

Número de sementes X	Frequência Observada Fo	(X) x (Fo)	Probabilidade P(X)	Frequência Esperada Fe = N x P(X)	F_e corrigida
0	3	(0). (3) = 0	0,048801200	4,7825176	5
1	17	17	0,147379600	14,4432000	14
2	26	52	0,222543200	21,8092330	22
3	16	48	0,224026800	21,9546260	22
4	18	72	0,169140200	16,5757390	17
5	9	45	0,102160700	10,0117480	10
6	3	18	0,51420800	5,0392384	8
7	5	35	0,022184400	2,1740712	-
8	0	0	0,008374620	0,8207127	-
9	1	9	0,002810150	0,2753947	-
10	0	0	0,000848666	0,831692	-
11 ou mais	0	0	0,0003096634	0,0303470	-
TOTAL	98	296	1,0000000000	97,9999960	98

Com $P(X) = P(X = x_i) = \dfrac{3,02^{x_i}.e^{-3,02}}{x_i\ !}$, onde,

$E(X) = \mu_{[X]} = \mu_X = \mu = \lambda = \dfrac{298}{98} = 3,02$ Sementes,

Sendo, por exemplo, $P(X = 0) = \dfrac{3,02^0.e^{-3,02}}{0!} = 0,048801200$.

– Testar a hipótese a 5% de significância, de que os dados referentes à distribuição empírica do número de sementes de ervas daninhas encontradas nas 98 subamostras de Phleum Praetense, se aproxima, se ajusta ou ainda pode ser representada por uma distribuição teórica de probabilidades de Poisson.

12.1.8.4 Teste de homogeneidade

12.1.8.4.1 Objetivos

Nesse tipo de prova, temos a suspeita de que nossa amostra tem origem de mais de uma população. A prova objetiva determina se essas populações são distintas ou não.

Se o pesquisador dispõe de um conjunto de dados que ele supõe tenha sido obtido de uma mesma população, o teste de χ^2 pode ser usado para verificar a homogeneidade das amostras. Esse teste é feito, comparando-se a variância estimada com a variância prevista teoricamente. Note-se que não há interesse em se ajustar uma distribuição teórica, mas apenas saber se as diversas amostras provêm de uma mesma população. (Distribuição de probabilidades).

Coloca-se assim em prova as hipóteses:

H_o: As amostras representadas nas colunas foram tiradas de uma mesma população (sobre a qual nada sabemos), de tal maneira que os dados sejam homogêneos a esse respeito.

H_1: As amostras representadas nas colunas são provenientes de diferentes populações.

A estatística teste é o qui-quadrado de Pearson dado por

$$\chi^2_{Teste} = \chi^2_{Calculado} = \frac{\sum_{i=1}^{k}\left(f_{o_i} - f_{e_i}\right)^2}{f_{e_i}}$$

o qual possui distribuição teórica de probabilidade de qui-quadrado com o número de graus de liberdade "V" determinado conforme é mostrado no próximo item, onde "K" é o numero de categorias, atributos, classes, casas, caselas ou células na amostra em estudo.

12.1.8.4.2 O número de graus de liberdade

O número de graus de liberdade "V" para h > 1 e K > 1, onde "h" é o número de linhas e "K" é o número de colunas da tabela de contingência, são determinados por:

- V = K - 1, nas tabelas de simples entrada, onde k é o número de classes, casas ou células.
- V = K - P - 1, nas tabelas de simples entrada, quando as frequências esperadas são calculadas através da estimativa de "P" parâmetros populacionais, usando-se as estatísticas Amostrais.
- V = (h-1) (k-1), nas tabelas h x k, quando as frequências esperadas podem ser calculadas sem que se tenha que estimar os parâmetros populacionais por meio das estatísticas amostrais do estimador em questão ou sob estudo.
- V = (h - 1) (k - 1) - P, nas tabelas h x k, quando as frequências esperadas somente podem ser calculadas mediante à estimativa de "P" parâmetros populacionais, através das estatísticas amostrais do estimador sob estudo.

12.1.8.4.3 Exercício de aplicação

Os dados da Tabela 16 referem-se à distribuição nas 4 classes de sangue: O, A, B e AB, de residentes numa Zona Rural conhecida como Alagoinhas e na zona urbana do município de Mossoró-RN. Os resultados foram obtidos de duas amostras de indivíduos dos quais deu o seguinte resultado: Aplique um teste de qui-quadrado ao nível de 5% de probabilidade, para verificar se as duas amostras são provenientes de uma mesma população.

Tabela 16 - Frequências observadas referentes à contagem de pessoas classificadas de acordo com o tipo sanguíneo numa zona rural, e na zona urbana do município de Mossoró-RN, e pessoas residentes na zona urbana

	Frequências das classes de sangue				Totais
	O	A	B	AB	
Residentes da zona rural	56	60	18	6	140
Residentes da zona urbana	120	122	42	11	295
Totais	176	182	60	17	435

– H_o: As Amostras de pessoas residentes na zona rural de Alagoinhas e na zona urbana no município de Mossoró-RN, representadas nas colunas, foram selecionadas de uma mesma população, ou seja, que os dados sejam homogêneos a esse respeito.
  H_1: As Amostras de pessoas residentes na zona rural de Alagoinhas e na zona urbana no município de Mossoró-RN, representadas nas colunas, foram selecionadas de populações diferentes, ou seja, que os dados sejam heterogêneos a esse respeito.
– $\alpha = 5\ \%$

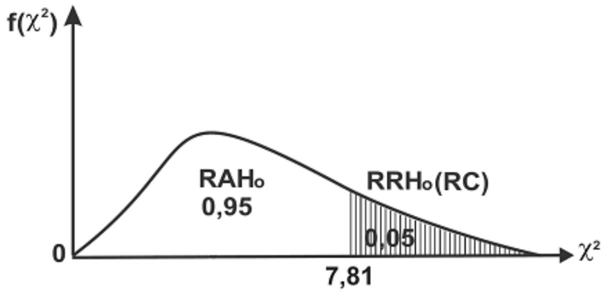
Com o número de graus de liberdade é dado por:

V = (h - 1) (K - 1), V = (2 - 1) (4 - 1) = 1 x 3 = 3 G.L.

Temos que o valor tabelado é:

$$\chi^2_{[3;0,05]} = 7,81$$

Figura 14 - Curva do teste de qui-quadrado, mostrando a área ou região crítica ou de rejeição de H_o correspondente a 5% de probabilidade



iv) A determinação das frequências esperadas é feita como mostramos abaixo:

$$\begin{array}{l} 435 \rightarrow \; 176 \quad \therefore \quad f_{e_1} = \dfrac{140.176}{435} = 56,44 \\ 140 \rightarrow \; f_{e_1} \end{array}$$

$$\begin{array}{l} 435 \rightarrow \; 182 \quad \therefore \quad f_{e_2} = \dfrac{140.182}{435} = 58,57 \\ 140 \rightarrow \; f_{e_2} \end{array}$$

$$\begin{array}{l} 435 \rightarrow \; 60 \quad \therefore \quad f_{e_3} = \dfrac{140.60}{435} = 19,31 \\ 140 \rightarrow \; f_{e_3} \end{array}$$

Podemos agora obter as outras frequências restantes por simples subtrações, assim:

$$f_{e_4} = 140,00 - \left(f_{e_1} + f_{e_2} + f_{e_3} \right) = 140,00 - 134,52 = 5,48$$

$$f_{e_5} = 176,00 - \left(f_{e_1} \right) = 176,00 - 55,64 = 119,36$$

$$f_{e_6} = 182,00 - \left(f_{e_2} \right) = 182,00 - 58,57 = 123,43$$

$$f_{e_7} = 60,00 - \left(f_{e_3}\right) = 60,00 - 19,31 = 40,69$$

$$f_{e_8} = 17,00 - \left(f_{e_4}\right) = 17,00 - 5,48 = 11,52$$

Sendo assim, temos que o valor do teste de qui-quadrado é dado por:

$$\chi^2_{[Teste]} = \chi^2_{[Observado]} = \frac{(56,00-56,64)^2}{56,64} + \frac{(60,00-58,57)^2}{58,57} + \frac{(18,00-19,31)^2}{19,31} + \frac{(6,00-5,48)^2}{5,48} + \frac{(120,00-119,36)^2}{119,36} + \frac{(12,00-123,43)^2}{123,43} + \frac{(42,00-40,69)^2}{40,69} + \frac{(11,00-11,52)^2}{11,52}$$

$$\chi^2_{[Teste]} = \chi^2_{[Observado]} = 0,0072 + 0,0034 + 0,0349 + 0,0166 + 0,0889 + 0,0422 + 0,0513 + 0,0244$$

$$\chi^2_{[Teste]} = \chi^2_{[Observado]} = 0,2689^{n.s.}$$

Como o valor do qui-quadrado teste é menor que o valor do qui-quadrado tabelado, então aceita-se a hipótese H_0.

$\chi^2_{[Teste]} = \chi^2_{[Observado]} = 0,2689^{n.s.} < \chi^2_{[3;0,05]} = 7,81$, aceita-se H_0.

– Conclusões: Concluí-se que, de acordo com o teste de qui-quadrado, com 3 graus de liberdade e 5% de significância, que as duas amostras de pessoas oriundas das zonas rural e urbana estudadas em Mossoró-RN podem ser consideradas como pertencentes a uma mesma população.

Note-se que esse teste só deve ser recomendado quando nada se souber acerca da distribuição de frequência da população, caso contrário o teste sofre modificações.

12.1.9 Exercício de teste não paramétrico [teste do qui-quadrado (χ^2)]

a) Em um experimento, foram postos 2000 ovos de EMA (Rhea americana) para chocar. No dia da eclosão, nasceram 1800 filhotes, sendo 980 Machos e 820 Fêmeas. Sendo a relação esperada de 1:1, pretende-se saber se as observações estão de acordo com essa hipótese. Utilize um nível de significância $\alpha = 1\%$.

b) Um experimentador conduziu um ensaio com o objetivo de obter flores de cores variadas da espécie de planta *Abutilon darwinii*, usada em paisagismo e, ao conduzir a pesquisa para alcançar o objetivo, ele obteve 120 flores de cor magenta com estigma verde, 48 de cor magenta com estigma vermelho, 36 flores vermelhas com estigma vermelho. A teoria mendeliana prevê que flores desses tipos devem ser obtidas nas razões de 9:3:3:1. Esses resultados experimentais são compatíveis com a teoria de Mendel.

c) Um experimento consistiu no cruzamento de tomates (*Lycopersicon esculentum*), da variedade Santa Cruz, com hipocótilo roxo e folha recortada (AACC), com a variedade folha de batata de hipocótilo verde e folha normal (aacc), no qual obteve-se a geração F_1 de hipocótilo roxo e folha recortada (AaCc), que, autofecundadas, deram origem à geração F_2. Os resultados estão mostrados a seguir. Verifique se os dados estão concordantes com a teoria ou Lei de Mendel, a qual prevê razões teóricas de 9:3:3:1.

Tabela 17 - Dados das frequências observadas de plantas de tomates (Lycopersicon esculentum) da variedade Santa Cruz

FENÓTIPO	**CLASSES**	f_o	**fe**	$\frac{(f_o - f_e)^2}{f_e}$
HIPOCÓTILO ROXO E FOLHA RECORTADA	1	105		
HIPOCÓTILO ROXO E FOLHA NORMAL	2	37		
HIPOCÓTILO VERDE E FOLHA RECORTADA	3	22		
HIPOCÓTILO VERDE E FOLHA RECORTADA	4	8		
TOTAL	-----------	172		

d) Os dados abaixo foram obtidos de um experimento conduzido por pesquisadores da empresa de pesquisa agropecuária Embrasil, da classificação de 1282, plantas de algodão herbáceo (*Gossypium hirsutum L.*), de geração F_2, de acordo com a cor da corola e a forma das folhas. Deseja-se saber se as duas classificações são independentes. Utilize um nível de significância de 1 de probabilidade [α = 0,01].

Tabela 18 - Frequências observadas de plantas de algodão herbáceo (Gossypium hirsutum L), de geração F_2

FORMA DA FOLHA	**COR DA COROLA**		**TOTAL**
	AMARELA	**BRANCA**	
ESTREITA	717 ( )	249 ( )	966
LARGA	236 ( )	80 ( )	316
TOTAL	953	329	1282

e) Os dados abaixo foram obtidos do livro de Viana et al. (1978), os quais sugerem que o gado charolês seja menos fecundo do que os demais gados que constam na tabela. Para verificar se há diferenças significativas de fecundidade entre tipos de gados, pode-se aplicar o teste de qui-quadrado para independência. Sendo assim, verifique se esses dados garantem, com evidência suficiente, o que há independência entre o tipo de acasalamento e a natureza das vacas estudadas. Use um nível de significância de 5% de probabilidade.

Tabela 19 - Frequências observadas referentes a quatro tipos raças e mestiços de vacas quanto ao tipo de acasalamento

NATUREZA DAS VACAS	TIPOS DE ACASALAMENTOS		TOTAL
	FECUNDOS	NÃO FECUNDOS	
CHAROLESA	515 ( )	1287 ( )	1802
INDUBRASIL	506 ( )	665 ( )	1171
NELORE	58 ( )	70 ( )	128
$\frac{1}{2}$ CHAROLÊS - ZEBU	205 ( )	93 ( )	298
TOTAL	1284	2115	N = 3399

f) Pesquisadores da empresa de pesquisas zootécnicas ZOOPESQBRASIL sabem que a doença COCCIDIOSE, a qual ataca galinhas da raça Plymouth Rock Barrado (Carijó), é não contagiosa. Para testar essa teoria, 30000 galinhas foram subdivididas em três categorias de 10000 animais, numa divisão aleatória, obtida através de sorteio. Um desses grupos não manteve qualquer contato com as galinhas atacadas pela doença; outro grupo teve contato moderado; e outro teve pleno contato com as galinhas infectadas. Os dados mostrados a seguir foram obtidos após um período de 6 meses, anotando-se o número de galinhas com a doença em cada um dos 3 grupos. Esses dados garantem, com evidência suficiente, que há uma dependência entre o contato entre galinhas sadias e doentes e a incidência da doença. Utilize um nível de significância de 5 % de probabilidade.

Tabela 20 - Frequências observadas referentes de dois grupos de galinhas da raça Plymouth Rock Barrado (Carijó), submetidos a três tipos de contato

CONTATO	**TIPOS DE ACASALAMENTOS**			**TOTAL**
	NENHUM CONTATO	**CONTATO MODERADO**	**CONTATO PLENO**	
NÚMERO DE GALINHAS DOENTES	87 ( )	89 ( )	124 ( )	300
NÚMERO DE GALINHAS SADIAS	9913 ( )	9911 ( )	9876 ( )	29700
TOTAL	10000	10000	10000	30000

13 ASPECTOS GERAIS DAS PROVAS DE SIGNIFICÂNCIA

Em toda pesquisa científica deve se atentar para condições essenciais na aplicação correta dos testes de hipóteses. Com o uso intensivo da computação eletrônica na pesquisa científica, principalmente a partir da década de 90 mediante o uso de pacotes estatísticos tais como o R, o SPSS, etc. nas análises desses dados, vários aspectos aqui apresentados tem um grau de importância maior que os conceitos mostrados no capítulo da teoria básica dos testes estatísticos de significância.

13.1 P-value, valor-p ou nível descritivo do teste

Para decidir-se aceitar ou rejeitar a hipótese de nulidade compara-se o valor p com um nível de significância usado normalmente como 5%, o qual é a probabilidade de risco, erro do tipo I ou de primeira espécie que o pesquisador arrisca ao conduzir ou aplicar o teste de hipótese, em rejeitar a hipótese de nulidade quando na realidade ela for verdadeira. Se o valor p é inferior a este nível de significância, o resultado é significativo. Se o valor p for superior então o resultado é não significativo.

Os valores p são estabelecidos sem nenhum pressuposto daí não se deve atribuir grande importância a estes valores. Por exemplo, em duas pesquisas em que os níveis descritivos p forem 0,047 e 0,058 e sejam interpretados de forma diferente para um nível de significância de 0,05 devem levar a conclusões semelhantes e não a conclusões totalmente opostas.

13.2 Comparação entre hipóteses bilaterais e hipóteses unilaterais

Hipótese alternativa para o teste de comparação de médias, por exemplo, é dada porH1: $\mu 1 \neq \mu 2$, entretanto pode ser desdobrada como nas seguintes hipóteses H1: $\mu 1 > \mu 2$ ou H1: $\mu 1 < \mu 2$.

Essa hipótese estabelece que uma média de tratamento pode ser superior ou inferior a outra. Neste caso a hipótese é denominada bilateral. Por outro lado o valor p bilateral é a probabilidade de se obter, uma diferença igual ou superiora observada.

Nos testes unilaterais as comparações entre tratamentos são estabelecidas em uma determinado sentido. Por exemplo, ao se comparar uma nova ração com uma ração padrão, a pesquisa se propõe a avaliar se a nova ração deve ser indicada.

Ao se testar hipóteses unilaterais, baseadas na distribuição teórica de probabilidade t de Student ou normal, utiliza-se apenas uma das extremidades da distribuição. Por exemplo, fixando-se um nível de significância de 5%, o valor adotado para testes baseados na distribuição normal é 1,96. Convém salientar que o nível descritivo ou valor p para o teste bilateral é o dobro do valor p correspondente ao teste unilateral.

Existem casos em que hipótese unilateral é a melhor com relação ao objetivo da pesquisa a ser realizada. Em estudos comparando o desempenho de uma variedade tradicional de milho por exemplo, a hipótese alternativa mais interessante é que a uma nova variedade é mais produtiva, dai a tendência de se estabelecer à hipótese unilateral. Lógico que neste caso não há nenhum interesse de que a nova variedade seja menos produtiva, o que justifica a escolha da hipótese unilateral.

O uso de um teste unilateral também pode ter como argumento a afirmação de que um efeito de tratamento em direção a uma das caudas da curva de probabilidade não tenha sentido.

A oposição de pesquisadores ao uso de hipóteses unilaterais é que, mesmo que se tenha grande informação de o efeito de um tratamento seja maior que outro fator experimental, jamais o indivíduo terá absoluta certeza do que realmente pode ocorrer. Por exemplo, mesmo que se tenha grande expectativa de que um novo tipo de inseticida seja mais eficiente que um produto testemunha, pode existir uma probabilidade ainda que rara de que os seus efeitos sejam inferiores.

Quando o pesquisador for implementar um teste de hipótese de qualquer tipo, ele deve levar em conta as seguintes considerações: i) a hipótese a ser testada deve ser escolhida antes ou a priori e não pelos resultados do experimento; ii) o teste que verifica hipótese bilateral é mais rigoroso do que um teste que adota hipótese unilateral; iii) na literatura científica argumenta-se que a diferença entre testes unilaterais e bilaterais não é condição única para interpretação dos resultados, pois na metodologia o autor deve citar o tipo de teste adotado; iv)existem autores que insistem no uso de teste unilateral mas muitas revistas científicas rejeitam a publicação de tais resultados.

Muitos autores recomendam que o nível descritivo do teste ou valorp unilateral nunca deve ser usado. Isto de deve segundo esse argumento é a criação da padronização dos resultados apresentados, pois um determinado nível de probabilidade terá um mesmo significado em todos os trabalhos publicados; eles também argumentam que pesquisas em que o uso do teste unilateral é aplicado são difíceis de encontrar na prática; e finalmente o último argumento é de que em pesquisas que são de extrema importância tais como a criação e de uma nova vacina, o nível descritivo é apenas um nível de probabilidade que auxilia o pesquisador na hora de aceitar ou rejeitar a hipótese de nulidade. Sendo assim a eficiência de um produto não pode ter como base apenas um valor, por exemplo, menos que 1%, nestas circunstancias o teste de hipótese bilateral é mais rigoroso.

13.3 Tomada de decisão nos testes de significância

Quando através do resultado de um teste bilateral rejeita-se a hipótese de nulidade, conclui-se que existem diferenças estatísticas entre os tratamentos. Uma das formas de se mensurar qual é o tratamento mais eficiente, é através da qualificação dos efeitos dos tratamentos aplicando-se testes de comparação de médias, análise de regressão, etc.

Um nível de significância grande como, por exemplo, 15% mostra apenas que probabilisticamente os tratamentos são semelhantes. Neste caso pode ser que esses fatores sejam diferentes, pois as evidências experimentais ou amostrais não permitem ao pesquisador rejeitar tal hipótese de igualdade de efeitos de tratamentos. Em uma outra situação se o valor p for pequeno como por exemplo 0,5%, então os efeitos dos tratamentos serem estatisticamente iguais não parece ser aceitável , sendo assim um dos tratamentos deve ser superior ao outro. O nível de significância de 5 % é muito usado na maioria das pesquisas científicas como um valor padrão, isso mostra que o pesquisador acerta 19 vezes em cada vinte, assumindo que a hipótese de nulidade seja verdadeira. É importante salientar que nunca se sabe que uma hipótese de nulidade é realmente verdadeira, o que se pode afirmar é que assumindo que ela é realmente verdadeira as evidências experimentais ou amostrais não permitem a sua rejeição, mas no futuro com a repetição da pesquisa no tempo e no espaço pode ser que ela venha a ser rejeitada.

13.4 Comparação entre os conceitos de diferença estatística e diferença prática

Na teoria da inferência estatística principalmente nos testes de hipóteses o termo significativo é interpretado de forma frequentemente errada, isso se deve ao fato de interpretar-se certo este termo na aplicação do teste, mas na aplicação prática usá-lo de forma errada, por isso que alguns pesquisadores defendem que todos devem diferenciar a significância estatística da significância prática. Por exemplo, uma diferença estatística de sete dias no tempo de recuperação de dois tipos de cirurgias de hérnia umbilical pode na realidade não ter importância prática, no entanto o resultado de um teste de hipótese pode ser não significativo do ponto de vista estatística, mas as conclusões apontam para um resultado importante do ponto de vista prático. Isso de deve muito ao refinamento da técnica experimental incluindo ai o tamanho da amostra dentre outros fatores importantes. Deve ser lembrado que um resultado é significativo se os limites do intervalo de confiança não englobar o valor do parâmetro estabelecido na hipótese de nulidade por um valor tão ou mais importante do ponto de vista prático e aplicado

13.5 Efeitos que causam interpretações errôneas nas diferenças entre tratamentos

Nos resultados dos testes de hipótese t de Student, Z normal ou qui-quadrado, uma diferença detectada entre os efeitos dos tratamentos estudados pode muitas vezes ser atribuído não aos tratamentos propriamente ditos e comparados na pesquisa, mas sim a um terceiro fator experimental ou não experimental denominado na linguagem científica de fator de confundimento. Por exemplo, na experimentação animal o estudo referente a avaliação e composição de carcaça de gado de corte, deve ser influenciada pelo sexo, raças e manejos. Uma maneira de garantir eficiência e precisão nas análises é o adotar o uso de técnicas de análise multivariada, o uso da técnica de covariância, isso garante a estimativa da variação individual evitando outros fatores que possam infiltrar durante a condução do experimento. Deve-se atentar para o uso correto dos princípios básicos da experimentação tais como o princípio da repetição, o princípio da casualização, o princípio do controle local, bloqueamento ou estratificação, o princípio da uniformidade dos elementos experimentais, refinamento da técnica experimental como o princípio da uniformidade na aplicação dos tratamentos e o princípio da uniformidade do meio. A aplicação e observação desses princípios básicos definirá um experimento mais eficiente e sensível para detectar pequenas diferenças e por resultar em valor mais realístico da variação individual.

13.6 Interpretação dos resultados do experimento

Quando o pesquisador for realizar as análises de testes de significância, é recomendável que ele exponha todos os resultados de grandezas estatísticas obtidas, para que assim os cálculos possam ser refeitos e o valor da probabilidade do nível descritivo ou de significância do teste de hipótese possa ser redefinido, este último não deva ser apresentado em forma de intervalo, mas sim expondo o valor aproximado com pelo menos três casas decimais, pois é o que é mais comum de se observar nas saídas ou output dos programas computacionais, pacotes estatísticos tais como o R, SPSS, S-PLUS, etc.. Sendo assim este deve apresentar os valores das médias, desvios padrões, erro padrão da média, tamanho de amostra, etc. para cada grupo, fator ou tratamento estudado.

13.7 Pressuposições na análise de dados estatísticos

Na apresentação anterior dos testes de hipóteses, foi colocado no início as pressuposições para a validade das análises mediante o correto uso de tais ferramentas de inferência estatística. Esses pressupostos são a população dos dados é infinita ou finita, ela tem distribuição normal ou não tem, qual é o tamanho da amostra, se a amostragem foi realizada com reposição ou sem reposição, as variâncias são conhecidas ou desconhecidas, as proporções são do tamanho adequado, as frequências esperadas são de no mínimo o valor exigido, etc. Sendo assim, após a escolha da prova de significância a ser utilizado, essas exigências podem e devem ser questionadas, pois se elas não forem atendidas em sua totalidade, o que deve acontecer com as análises, podem ser relevadas ou a falta de sua observação provoca danos as análises, ou ainda ajustes e transformações nos dados serão necessárias, se sim quais modificações. Muitas vezes é necessário antes realizar uma análise exploratória dos dados, uma análise gráfica e de resíduos, transformações dos dados, aplicação de teste de hipótese não paramétrico de aderência e ajustamento, etc.

13.7.1 Características gerais do teste não paramétrico do qui-quadrado

O teste não paramétrico do qui-quadrado criado pelo estatístico inglês Karl Pearson é proposto para ser aplicado em situações em que são válidas as seguintes conjecturas: i)a probabilidade de ocorrer um evento ou sucesso não varia de unidade ou elemento para outro elemento dentro dos grupos de tratamento testemunha e tratamento propriamente dito. Se esta condição não for verificada deve ser usada uma análise alternativa com o uso de tabelas de contingência 2x2; ii) a resposta que ocorre para um elemento, de quaisquer dos grupos, não influencia, ou ocorre de forma independente quando comparado ao resultado de outras unidades. Se esta condição não for verificada deve ser usada uma análise alternativa de dados dependentes. Ou seja, o teste é utilizado para verificar se a frequência absoluta com que um determinado acontecimento observado em uma amostra se desvia significativamente ou não da frequência teórica com que ele é esperado segundo a hipótese de nulidade estabelecida a priori. Também é utilizado para comparar a distribuição de diversos acontecimentos em diferentes amostras, a fim de avaliar se as proporções observadas destes eventos mostram ou não discrepâncias estatisticamente significativas ou se as amostras diferem significativamente quanto às proporções desses acontecimentos. As seguintes condições devem ser estabelecidas antes de se conduzir o teste: i) os grupos devem ser independentes, ii) os itens de cada grupo são selecionados aleatoriamente, iii) as observações devem ser frequências absolutas ou contagens, iv) cada observação pertence a uma e somente uma categoria ou atributo e v) a amostra deve ser relativamente grande com pelo menos cinco observações em cada célula, casa ou casela, e, no caso de poucos grupos, pelo menos 10 devem ser adotados. Um exemplo é a aplicação em testes usando tabelas de contingência h por k com duas linhas e duas colunas, ou seja, em tabelas 2x 2.

Para se aplicar o teste de aderência do qui-quadrado, deve-se levar em consideração o seguinte: seja pi a probabilidade associada à categoria ou atributo i, para i = 1,..., k. O objetivo do teste de aderência de qui-quadrado é testar as hipóteses H0 : p1 = po1 , ... , pk = pokHa: existe pelo menos uma diferença sendo poi a probabilidade especificada para a categoria i, para i = 1, ..., k, fixada através do modelo probabilístico de interesse. Se fei ou Ei é o total de indivíduos esperados na categoria i, quando a hipótese H0 é verdadeira, então: Ei = n poi , para i = 1, ...,k.

13.7.2 Considerações sobre o teste paramétrico t de student

O teste paramétrico de t de Student, quando é usado para testar apenas média, comparação entre médias, deve ser aplicado corretamente quando as exigências ou pressuposições forem verificadas, tais como se o estudo envolve apenas uma população ou tratamento, a população deve ser aproximadamente normal ou gaussiana e infinita ou finita e o uso de amostragem com reposição, caso contrário deve ser usado o fator de correção para população finita e amostragem sem reposição, a variância desta população deve ser desconhecida, e o tamanho da amostra aleatória simples selecionada mediante sorteio e utilizada na pesquisa tem ou deve ter menos de 30 elementos, numa segunda situação onde se tem comparações de duas médias de tratamentos usando duas populações, existe um subcaso onde as amostras não são independentes, mas sim pareadas ou os dados são denominados emparelhados ou dependentes, e o pesquisador usa média de diferenças entre valores, mas neste subcaso as pressuposições continuam válidas tais como: o estudo envolve apenas uma população ou tratamento, a população deve ser aproximadamente normal ou gaussiana e infinita ou finita e o uso de amostragem com reposição, caso contrário deve ser usado o fator de correção para população finita e amostragem sem reposição, a variância desta população deve ser desconhecida, e o tamanho da amostra aleatória simples selecionada mediante sorteio e utilizada na pesquisa tem ou deve ter menos de 30 elementos. No caso de comparação de médias quando as amostras são independentes, têm-se duas populações infinitas ou com tratamentos estudados, e esta tem distribuição normal ou gaussiana pelo menos aproximadamente, cujas variâncias são supostamente iguais. Isso permite ao pesquisador testar a hipótese de igualdade das médias o que é equivalente à hipótese de que os dois tratamentos sejam em médias comparáveis. E o outro subcaso é aquele em que as variâncias são diferentes ou as populações são heterocedásticas,tendo que se fazer um ajuste no número de graus de liberdade para que o teste se torne de uso correto. Trabalhos publicados na área de estatística mostram que o teste t de Student é robusto no que se refere ao pressuposto da exigência de que os dados tenham que possuir distribuição normal, ou seja, o teste t de Student pode e deve ser aplicado mesmo para variáveis que sejam moderadamente assimétricas ou que efetivamente não tenham distribuição gaussiana. Por outro lado a pressuposição de que as variâncias tem que ser estatisticamente semelhantes, ou seja, as populações devem ser homocedásticas, é um exigência que deve ser avaliada com bastante rigor, pois sua violação pode resultar em tomadas de decisões incorretas. O uso de testes não paramétricos tais como os testes de Bartlett e de Hartley podem ser usados para mensurar a homocedasticidade das populações estudadas, por outro lado se os dados apresentam heterocedasticidade então é recomendável o uso de transformações de dados tais como as transformações log x, arcsenx , raiz quadrada, etc., ou então aplicar o uso de testes não paramétricos tais como; o teste do sinal, o teste de Wilcoxon, teste de Mann-Whitney, etc.

Muitos autores da área estatística afirmam que se o tamanho das amostras de dois tratamentos estudados é igual e as amostras são grandes com pelo menos 30 elementos e ainda, se a distribuição dos dados é supostamente normal e as variâncias são estaticamente semelhantes então o valor do nível de significância calculado é válido, mas se os tamanhos das amostras e os desvios padrão forem estatisticamente diferentes, o nível de significância pode sofrer alteração de um valor que é função de quantas vezes um valor de um desvio padrão é maior que o outro desvio padrão da análise. No teste Z apoiado na distribuição normal para comparação da diferença entre duas médias, exige-se que as populações sejam também infinitas ou finitas e o uso de amostragem com reposição, porém que as duas variâncias dessas populações sejam conhecidas, neste caso os tamanhos das amostras aleatórias simples podem ser com menos de 30 elementos ou mais de 30 elementos, pois conforme o teorema do limite central isto garante que a distribuição teórica de amostragem da diferença entre as médias amostrais seja normal ou gaussiana. No entanto a teoria mostra que se as variâncias das duas populações forem desconhecidas devem ser usadas duas amostras aleatórias simples com pelo menos trinta componentes. O mesmo deve ser observado para o teste referente a proporção e diferença entre duas proporções de uma ou duas população infinita ou finita e o uso de amostragem com reposição respectivamente, pois nesta situação a distribuição por amostragem da proporção relativa amostral só converge para a distribuição normal quando o tamanho da amostra aleatória simples adotada for de pelo mesmo 30 componentes.

Em todos esses testes paramétricos para média, diferença entre duas médias e proporção, se a população for finita e adotar-se o plano amostral de amostragem sem reposição o erro padrão do estimador amostral deve ser corrigido ou multiplicado pelo fator de correção de população finita, pois este é responsável pela suavização, minimização ou para frear o aumento da variabilidade entre os estimadores amostrais medida pelo erro padrão desse estimador ou grandeza amostral adotada quando se faz seleção sem reposição em população finita. Já os testes para a variância e para o desvio padrão de uma população normal devem ser apoiados pela distribuição teórica de probabilidade de qui-quadrado e o teste para razão entre duas variâncias de populações normais deve ter como distribuição teórica de probabilidade para tomada de decisão a distribuição de Snedecor-Fisher ou distribuição F de probabilidades.

Finalmente vale destacar o efeito que a heterocedasticidade ou variâncias diferentes nas populações estudadas influi no significado das hipóteses elaboradas e postas em prova, pois nessas situações, por exemplo, um teste de comparação de médias pode não ter muito interesse prático, devido ao fato de que se dois fatores experimentais testados possuírem a mesma média aritmética, estes podem ter efeito completamente opostos e, portanto, não serem considerados estatisticamente semelhantes.

14 EXERCÍCIOS PROPOSTOS SOBRE TESTES DE HIPÓTESES (OU SIGNIFICÂNCIA) PARAMÉTRICOS E NÃO PARAMÉTRICOS [TESTE DO QUI – QUADRADO (χ^2)]

14.1 Responda as seguintes questões.

a) Qual é o significado de testar uma hipótese? Defina hipótese de nulidade (H_0) e hipótese alternativa (H_1). Qual é o procedimento geral? O que é teste paramétrico? E não-paramétrico?

b) Qual é o significado de erro de tipo I, de erro tipo II? E de valor-p?

c) Qual é o significado de nível de significância? E nível de confiança?

d) O que é poder de teste?

14.2 Um fornecedor de mancais comprometeu-se a enviar para uma firma lotes que não continham mais de 2% de defeituosos. O comprador extrai amostras, ao receber a remessa, para verificar a qualidade.

a) Indique H_0 e H_1.

b) O fornecedor não deseja remeter lotes com elevado risco de devolução em razão de número excessivo de unidades defeituosas, mas também não deseja remeter lotes com percentagem de defeituosos muito menos que a estabelecida de modo que ele também, fornecedor, faz seu teste antes de proceder à remessa. Indique H_0 e H_1.

14.3 A tensão de ruptura dos cabos produzidos por um fabricante apresenta média de 1800 kg e desvio padrão de 100 kg. Mediante nova técnica de fabricação, proclamou-se que a tensão de ruptura foi aumentada. Testando uma amostra de 50 cabos produzidos pelo 2º método obteve-se X = 1850 kg. Teste a declaração ao nível de 0,01 de significância.

14.4 Uma amostra de 80 fios de aço produzidos por uma fábrica A deu para a ruptura os valores $\overline{X}_A$ = 1230 e S_A = 120. Uma amostra de 100 fios de aço do mesmo tipo de uma fábrica B deu $\overline{X}_B$ = 1190 e S_B = 90. Há realmente diferenças significativas entre os valores médios? α = 1% e α = 15%.

14.5 Um fabricante de cigarros afirma que o teor médio de nicotina, para determinada marca por ele fabricada, não excede a 26,2 mg. Uma amostra foi retirada e analisada, fornecendo os valores 27, 26, 25, 31, 29, 28, 34, 30, 28, 29. Os resultados amostrais confirmam a afirmativa do fabricante? α = 1%.

14.6 São dadas duas amostras aleatórias de tamanho n_1 = 11 e n_2 = 14, retiradas de duas populações normais independentes, com $\overline{X}_1$ = 75 e $\overline{X}_2$ = 60, S_1 = 6,1 e S_2 = 5,3. Teste a hipótese de que $\mu_1 = \mu_2$ contra a alternativa de que $\mu_1 \neq \mu_2$, ao nível de 0,05.

14.7 Um centro de recrutamento do exército sabe, a partir da experiência passada, que o peso dos recrutas do exército é normalmente distribuído com uma média μ de 80 kg e um desvio padrão σ de 10 kg. O centro de recrutamento deseja testar ao nível de significância de 1% se o peso médio dos recrutas deste ano é maior do que 80 kg. Para fazer isso, ele toma uma amostra aleatória de 25 recrutas e encontra um peso médio para essa amostra de 85 kg. Como pode esse teste ser realizado?

14.8 Os 64 alunos que prestaram exames para admissão em um curso de Doutorado em Agronomia, com área de concentração em Fitotecnia em 1999 na UFERSAM, tiveram média de 640 pontos, com um desvio padrão de 20. Em 1982, os 81 alunos que prestaram exames tiveram média de 650 pontos com desvio padrão de 40.

São os candidatos de 1981 menos qualificados que os de 1982, ao nível de significância de 1%?

14.9 Suponhamos que, de uma população de alturas de pés de sorgo com média de 238 cm, foi retirada uma amostra aleatória de 25 indivíduos, com um $\overline{X} = 240$ e um desvio padrão de 15 cm, após um período de adubação. Deseja-se verificar se as plantas adubadas são estatisticamente maiores que as da população. Use $\alpha = 5\%$.

14.10 O fabricante de uma droga medicinal reivindicou que ela era 90% eficaz, em curar uma alergia, em um período de 8 horas. Em uma amostra de 200 pessoas que tinham alergia, a droga curou 160 pessoas. Determine se a pretensão do fabricante é legitima. Use $\alpha = 1\%$ e $\alpha = 5\%$.

14.11 Para avaliação do QI de 1725 crianças na escola primária, através de um teste apropriado, obteve-se os seguintes resultados:

Tabela 21 - Frequências observadas da avaliação do quociente de inteligência de 1725 crianças de uma escola primária

Nível Econômico	**Rude**	**Normal**	**Superior**
Alto	81( )	322( )	233( )
Médio	141( )	457( )	153( )
Baixo	127( )	163( )	45( )

Há independência entre a inteligência e o nível econômico com base nestes dados? Use $\alpha = 5\%$.

14.12 A tabela a seguir mostra a distribuição de determinado tipo de acidente.

Tabela 22 - Frequências observadas do número de acidentes

Nº de acidentes (X)	Nº de dias observados(fo)	P(X acidentes)	Frequências esperadas (fe)
0	21	0,4066	
1	18	0,3659	
2	7	0,1657	
3	3	0,0494	
4	1	0,0124	

Calcule as frequências esperadas e verifique se esse tipo de acidente pode ser considerado como regido pela lei de POISSON; sabe-se que o parâmetro "p" foi estimado. Use $\alpha = 5\%$.

14.13 Num estudo de hereditariedade comercial, Lindstrom encontrou 98 plântulas com a cor verde e 24 com a cor amarela, numa descendência de 122. Suponho que o verde domina o amarelo na razão 3:1, pretende-se saber se as observações estão de acordo com esta hipótese. Use $\alpha = 5\%$.

14.14 Em seus experimentos com ervilhas, Mendel observou simultaneamente a forma e a cor de suas ervilhas. Num total de n = 556 ervilhas, obteve os seguintes resultados: redondas e amarelas 315, redondas e verdes 106, angulares e amarelas 101, angulares e verde 32. De acordo com sua teoria da hereditariedade, os números deveriam apresentar-se na proporção de 9:3:3:1. Há evidência para duvidar dessa teoria, ao nível de (a) 0,01 e (b) 0,05.

14.15 Uma Universidade implantou novo sistema de avaliação de aprendizagem. Uma enquete realizada com alunos de três faculdades apresentou os seguintes resultados, quanto à opinião dos alunos relativa ao no novo sistema:

Tabela 23 - Frequências observadas do número de alunos de três faculdades, favoráveis e contrários ao um novo sistema de avaliação de aprendizagem

	Favoráveis	**Contrários**	**Indiferentes**	**TOTAL**
Faculdade A	50 ( )	80 ( )	10 ( )	
Faculdade B	60 ( )	40 ( )	20 ( )	
Faculdade C	80 ( )	50 ( )	10 ( )	
TOTAL				

Testar a hipótese de que não há dependência significativa, no nível 0,05, de opinião entre os estudantes das três faculdades, em relação ao novo sistema de avaliação.

14.16 A tabela seguinte apresenta as médias das vendas, em cada dia da semana, de uma casa comercial. Testar, no nível 0,05, a hipótese de que o volume de vendas não depende do dia da semana.

Tabela 24 - Frequência observada de médias de vendas, por dia da semana, em uma casa comercial

	2ª feira	3 ª feira	4 ª feira	5 ª feira	6 ª feira	Sábado
Vendas (R$ 1000)	200	150	180	210	190	150

14.17 Lançando-se 18 vezes um dado, os resultados obtidos foram:

Tabela 25 - Frequência observada do número de faces de um dado honesto lançado 18 vezes

Ponto	1	2	3	4	5	6
Frequência	1	4	2	3	5	3

Testar, no nível de 0,05, a compatibilidade entre as frequências observadas e esperadas.

14.18 Sabe-se que certa linhagem de camundongo, alimentados com uma ração padrão, tem um aumento médio de peso igual a 64g, durante os três primeiros meses. Um lote de 81 camundongos dessa linhagem foi alimentado com uma nova ração, mantendo-se as condições ambientais padronizadas. O aumento médio de peso observados nos camundongos foi de $\overline{X}$ = 60,75g, e um desvio padrão de S= 3,84g. A nova ração tem a mesma eficiência alimentar que a padrão? Use α =0,05.

14.19 Em um seringal, no qual se utiliza o processo convencional de sangria, a produção média de borracha seca é de 26g/árvore/corte. Tomou-se uma amostra ao acaso, composta de 25 seringueiras, as quais foram sangradas, usando-se um novo processo, tendo produzido 28g em média, com desvio padrão de 4g. O novo processo é mais eficiente que o convencional? Use α =5%.

14.20 Num ensaio com gado de leite, estudou-se o efeito de um novo componente de ração na produção de leite. Para isso, foi tomada uma amostra aleatória de 10 vacas (de mesmo porte e homogêneas quanto ao peso, raça, ordem de parição, etc.), que foram alimentadas com uma ração básica durante um período de tempo após o qual foram obtidas as produções diárias de leite, em cada animal. Em seguida, adicionou-se à ração básica um “Novo componente H”, e as vacas foram alimentadas com essa mistura durante um determinado período, ao final do qual mediram-se novamente as produções diárias de leite, em cada animal. Os dados desse experimento constam na tabela a seguir:

Tabela 26 - Produções diárias de leite de 10 vacas, após aplicação de uma ração básica e após aplicação de uma ração básica + um Novo Componente H

VACA	Ração Básica (X_1)	Ração Básica + H (X_2)	Ganho de produção (kg/dia) $d_i = X_{i2} - X_{i1}$	d_i^2
1	9,0	10,5	1,5	2,25
2	9,5	9,5	0,0	0,00
3	10,00	9,5	-0,5	0,25
4	11,0	13,0	2,0	4,00
5	9,8	12,8	3,0	9,00
6	8,9	9,9	1,0	1,00
7	9,2	11,7	2,5	6,25
8	9,7	10,5	0,8	0,64
9	11,0	11,5	0,5	0,25
10	9,5	12,0	2,5	6,25
SOMA	-	-	**13,3**	**29,89**

Supondo-se que a variável aleatória D relativa a ganho de produção tenha distribuição aproximadamente normal com média μ e variância σ^2, pede-se:

a) Obter as estimativas por ponto para $\mu_{\overline{D}}$ e para $\sigma^2_{\overline{D}}$.

b) Fazer um teste de significância para: $\mu_D = 0$ com $\alpha = 0{,}01$

14.21 Um experimento aleatório consiste em comparar os pesos ao nascer de animais machos das raças Gir e Guzerá da Estação Experimental de Zootecnia de Sertãozinho, do Instituto de Zootecnia da Secretaria da Agricultura de São Paulo. Ao nível de 5% de probabilidade.

Pesos ao nascer (kg) de bezerros Zebu. Onde,

$n_1 = 40$; $\overline{X}_1 = 28{,}3$ kg; $S_1^2 = 108{,}37$ kg^2; $S_1 = 10{,}41$ kg

$n_2 = 50$; $\overline{X}_2 = 23{,}5$ kg; $S_2^2 = 49{,}70$ kg^2; $S_2 = 7{,}05$ kg

14.22 Realizou-se um experimento com 340 pintos machos de um dia de idade, da linhagem "Ross", com o objetivo de estudar o efeito da fonte de fósforo no ganho de peso. Foram utilizadas 10 repetições por tratamento e 17 pintos por parcela e os resultados para ganho de peso médio por parcela (em g), estão no quadro abaixo.

Tabela 27 - Dados de ganhos de peso de pintos machos de um dia de idade, submetidos a tratamentos diferentes

Termosfosfato magnesiano (X_1)		**Fosfato bicálcico (X_2)**	
63,0	60,9	90,0	91,8
62,5	62,5	90,6	92,0
61,9	63,8	91,7	92,3
60,7	62,0	90,9	90,6
63,2	61,8	90,3	91,2

Onde: $n_1 = 10$; $\overline{X}_1 = 62{,}23g$; $S_1^2 = 0{,}96\ g^2$; $S_1 = 0{,}98g$

$n_2 = 10$; $\overline{X}_2 = 91{,}14g$; $S_2^2 = 0{,}61\ g^2$; $S_2 = 0{,}78g$

pede-se:

a) Verificar se existe efeito significativo da fonte de fosfato, considerando, $\alpha = 0{,}05$.

14.23 Considere os ganhos de peso diários em filhos de 2 (dois) reprodutores suínos da Raça Duroc, acasalados cada um com matrizes da mesma raça. Ao nível de 5% de significância, concluir se os ganhos médios de peso são diários são estatisticamente iguais.

Tabela 28 - Ganhos de peso diário em kg de filhos de dois reprodutores suínos da raça Duroc

Reprodutores	**Matrizes**	**Ganhos de peso (kg)**				
R_1	M	0,95	0,86	0,92	0,93	0,94
R_2	M	0,80	0,82	0,78	0,80	0,79

Onde: $n_1 = 5$; $\overline{X}_1 = 0{,}920$kg; $S_1^2 = 0{,}00125$kg^2; $S_1 = 0{,}035$kg; $W_1 = 0{,}000250$

$n_2 = 5$; $\overline{X}_2 = 0{,}798$kg; $S_2^2 = 0{,}00022$kg^2; $S_2 = 0{,}01$kg; $W_2 = 0{,}000044$

14.24 Em uma granja, uma amostra aleatória de 800 coelhos da raça Norfolk, apresentou 480 machos. Ao nível de 1% de significância, pode-se concluir que há prevalência ou predominância de coelhos machos nessa granja?

REFERÊNCIAS

ANDRADE, D. F.; OGLIARI, P. J. **Estatística para as ciências agrárias e biológicas**: com noções de experimentação. Florianópolis: UFSC, 2007.

BARBETTA, P. A. **Estatística aplicada às ciências sociais**. 5. ed. Florianópolis: UFSC, 2005.

BERQUÓ, E. S. ; SOUZA, J. M. P.; GOTLIEB, S. L. D.; SOUZA, J. F. P. **Bioestatística**. São Paulo: EDUSP, 1980. 325 p.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. 5. ed. São Paulo: Saraiva, 2003.

CAMPOS, H. **Estatística experimental não paramétrica.** 4. ed. Piracicaba, SP: Edusp, 1983.

COCHRAN, W. G. Some methods for strengthening the common χ^2 test. **Biometrics,** v. 10, p. 417-451, 1954.

COSTA NETO, P.L.O. **Estatística**. São Paulo: Edgard Blücher, 2000. 264 p.

CURI, P. R. **Metodologia e análise da pesquisa em ciências biológicas.** 2. ed. Botucatu, SP: [s.n.], 1998.

EVERITT, B. S. **The analysis of contingency tables**. 2. ed. London: Chapman & Hall, 1992.168 p.

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

FISHER, R. A. On the mathematical foundations of theoretical statistics. **Philosophical Transactions of the Royal Society of London a: mathematical, physical and engineering sciences**. v. 222, i. 594-604, p. 368-368, jan. 1922. Series A. Disponível: https://doi.org/10.1098/rsta.1922.0009. Acesso em: 09 abr. 2020.

MEMÓRIA, J. M. **Curso de estatística aplicada à pesquisa científica**. Viçosa: Imprensa Universitária, 1973. 304p. (apostila).

MURTEIRA, B.; RIBEIRO, C. S.; SILVA, J. A.; PIMENTA, C. **Introdução à estatística.** [*S.l.*]: McGraw-Hill, 2001.

NEWTON, I. **Mathematical principles of natural philosophy**. Los Angeles: University of California Press, 1934.

NUNES, C.; AFONSO, A. **Apontamentos de introdução às probabilidades e à estatística** Évora, PT: Universidade de Évora, 2005. (Manuais da Universidade de Evora, v. 2).

NUNES, R. P. **Métodos para a pesquisa agronômica.** Fortaleza: UFC, 1998.

OLIVEIRA, L. M. **Distribuições estatísticas**. Vicosa: UFV, 1970. (Apostila).

PAGANO, M.; GAUVREAU, K. **Principios de bioestatistica.** São Paulo: Thomson Learning, 2006.

PEDROSA, A. C.; GAMA, S. M. A. **Introdução computacional à probabilidade e estatística.** Porto, PT: Porto, 2004.

PIMENTEL GOMES, F. **Iniciação à estatística**.
São Paulo: Nobel, 1976.

PIMENTEL GOMES, F. **Curso de estatística experimental**. 15. ed. Piracicaba: Nobel, 2009.

SAMPAIO, I. B. M**. Estatística aplicada à experimentação animal.** Belo horizonte: Fundação de Ensino e Pesquisa em Medicina Veterinária e Zootecnia, 1998.

SIDIA, M. C-J. **Bioestatística**: princípios e aplicações. Porto Alegre: Artmed, 2003.

SOARES, J. F.; SIQUEIRA, A. L. Introducao a estatística médica. *In*: SIDIA, M. C-J. **Bioestatistica**: princípios e aplicações. Porto Alegre: Artmed, 2003.

SPIEGEL, M. R. **Estatística**. 4. ed. São Paulo: Makron Books, 2000.

STEVENSON, W.J. **Estatística aplicada à administração**. São Paulo: Harper & Row do Brasil, 1986. 498 p.

VIANA, A.T.; PIMENTEL GOMES, F.; SANTIAGO, M. **Formação do gado canchim pelo cruzamento charolês – zebu.** 2. ed. São Paulo: Nobel: 1978.

VIEIRA, S. **Bioestatística: tópicos avançados.** Rio de Janeiro: Campus, 2003.

WELCH, B. L. The generalization of Student's problem when several different population variances are involved. **Biometrika,** v. 34, p. 28-35, 1947.

ZAR, J. H. **Biostatistical analysis**. 4. ed. Upper Saddle River, USA: Prentice Hall, 1999.

APÊNDICE A

ALFABETO GREGO

	SÍMBOLOS	
NOME DA LETRA	**MAIÚSCULA**	**MINÚSCULA**
Alfa	Α	α
Beta	Β	β
Gama	Γ	g
Delta	Δ	d
Épsilon	Ε	ε
Zeta	Ζ	ζ
Eta	Η	h
Téta	Θ	θ
Iota	Ι	ι
Capa	Κ	κ
Lambda	Λ	λ
Mu(mi)	Μ	μ
Nu(ni)	Ν	ν
Csi	Χ	ξ
Omicron	Ο	ο
Pi	P	p
Ró	Ρ	ρ
Sigma	S	s
Tau	Τ	t
Upsilon(ipsilon)	Υ	u
Fi	F	j
Chi(qui)	Χ	χ
Psi	Ψ	ψ
Omega	W	ω

APÊNDICE B

Tabela 29 - Áreas ou probabilidades sob a curva normal padrão ente z = 0,00 e um valor positivo de Z Para os valores das probabilidades entre os valores negativos de Z e Z = 0,00, as áreas são obtidas por simetria

(continua)

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916

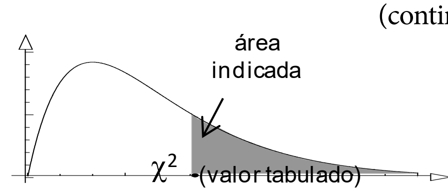
Tabela 29 - Áreas ou probabilidades sob a curva normal padrão ente z = 0,00 e um valor positivo de Z Para os valores das probabilidades entre os valores negativos de Z e Z = 0,00, as áreas são obtidas por simetria

(conclusão)

z	**0,00**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4965	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4983	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,49	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500

Tabela 30 - Distribuição de qui-quadrado para diversos níveis de significância

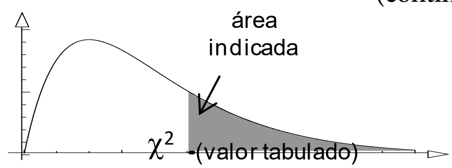
(continua)



	Área na cauda superior							
gl	0,999	0,9975	0,995	0,99	0,975	0,95	0,9	0,75
1	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,10
2	0,00	0,01	0,01	0,02	0,05	0,10	0,21	0,58

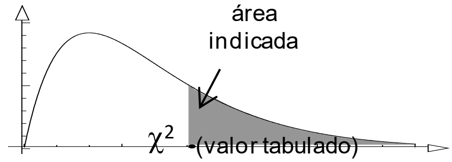
Tabela 30 - Distribuição de qui-quadrado para diversos níveis de significância

(continua)



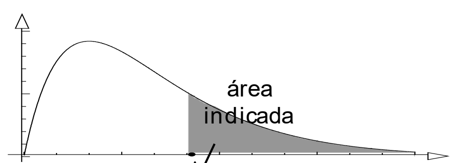
gl	Área na cauda superior							
	0,999	0,9975	0,995	0,99	0,975	0,95	0,9	0,75
3	0,02	0,04	0,07	0,11	0,22	0,35	0,58	1,21
4	0,09	0,14	0,21	0,30	0,48	0,71	1,06	1,92
5	0,21	0,31	0,41	0,55	0,83	1,15	1,61	2,67
6	0,38	0,53	0,68	0,87	1,24	1,64	2,20	3,45
7	0,60	0,79	0,99	1,24	1,69	2,17	2,83	4,25
8	0,86	1,10	1,34	1,65	2,18	2,73	3,49	5,07
9	1,15	1,45	1,73	2,09	2,70	3,33	4,17	5,90
10	1,48	1,83	2,16	2,56	3,25	3,94	4,87	6,74
11	1,83	2,23	2,60	3,05	3,82	4,57	5,58	7,58
12	2,21	2,66	3,07	3,57	4,40	5,23	6,30	8,44
13	2,62	3,11	3,57	4,11	5,01	5,89	7,04	9,30
14	3,04	3,58	4,07	4,66	5,63	6,57	7,79	10,17
15	3,48	4,07	4,60	5,23	6,26	7,26	8,55	11,04
16	3,94	4,57	5,14	5,81	6,91	7,96	9,31	11,91
17	4,42	5,09	5,70	6,41	7,56	8,67	10,09	12,79
18	4,90	5,62	6,26	7,01	8,23	9,39	10,86	13,68
19	5,41	6,17	6,84	7,63	8,91	10,12	11,65	14,56
20	5,92	6,72	7,43	8,26	9,59	10,85	12,44	15,45
21	6,45	7,29	8,03	8,90	10,28	11,59	13,24	16,34
22	6,98	7,86	8,64	9,54	10,98	12,34	14,04	17,24
23	7,53	8,45	9,26	10,20	11,69	13,09	14,85	18,14
24	8,08	9,04	9,89	10,86	12,40	13,85	15,66	19,04
25	8,65	9,65	10,52	11,52	13,12	14,61	16,47	19,94
26	9,22	10,26	11,16	12,20	13,84	15,38	17,29	20,84
27	9,80	10,87	11,81	12,88	14,57	16,15	18,11	21,75

Tabela 30 - Distribuição de qui-quadrado para diversos níveis de significância (conclusão)



gl	Área na cauda superior							
	0,999	0,9975	0,995	0,99	0,975	0,95	0,9	0,75
28	10,39	11,50	12,46	13,56	15,31	16,93	18,94	22,66
29	10,99	12,13	13,12	14,26	16,05	17,71	19,77	23,57
30	11,59	12,76	13,79	14,95	16,79	18,49	20,60	24,48
35	14,69	16,03	17,19	18,51	20,57	22,47	24,80	29,05
40	17,92	19,42	20,71	22,16	24,43	26,51	29,05	33,66
45	21,25	22,90	24,31	25,90	28,37	30,61	33,35	38,29
50	24,67	26,46	27,99	29,71	32,36	34,76	37,69	42,94
100	61,92	64,86	67,33	70,06	74,22	77,93	82,36	90,13

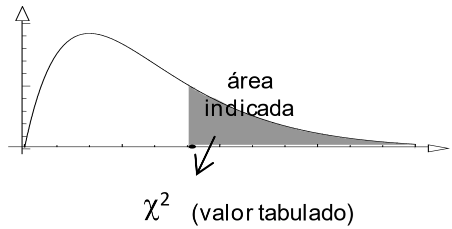
Tabela 31 - Distribuição de qui-quadrado para diversos níveis de significância (continua)



gl	Área na cauda superior							
	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001
1	1,32	2,71	3,84	5,02	6,63	7,88	9,14	10,83
2	2,77	4,61	5,99	7,38	9,21	10,60	11,98	13,82
3	4,11	6,25	7,81	9,35	11,34	12,84	14,32	16,27
4	5,39	7,78	9,49	11,14	13,28	14,86	16,42	18,47
5	6,63	9,24	11,07	12,83	15,09	16,75	18,39	20,51

Tabela 31 - Distribuição de qui-quadrado para diversos níveis de significância

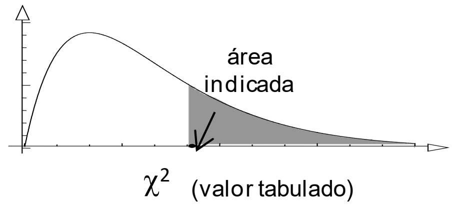
(continua)



gl	Área na cauda superior							
	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001
6	7,84	10,64	12,59	14,45	16,81	18,55	20,25	22,46
7	9,04	12,02	14,07	16,01	18,48	20,28	22,04	24,32
8	10,22	13,36	15,51	17,53	20,09	21,95	23,77	26,12
9	11,39	14,68	16,92	19,02	21,67	23,59	25,46	27,88
10	12,55	15,99	18,31	20,48	23,21	25,19	27,11	29,59
11	13,70	17,28	19,68	21,92	24,73	26,76	28,73	31,26
12	14,85	18,55	21,03	23,34	26,22	28,30	30,32	32,91
13	15,98	19,81	22,36	24,74	27,69	29,82	31,88	34,53
14	17,12	21,06	23,68	26,12	29,14	31,32	33,43	36,12
15	18,25	22,31	25,00	27,49	30,58	32,80	34,95	37,70
16	19,37	23,54	26,30	28,85	32,00	34,27	36,46	39,25
17	20,49	24,77	27,59	30,19	33,41	35,72	37,95	40,79
18	21,60	25,99	28,87	31,53	34,81	37,16	39,42	42,31
19	22,72	27,20	30,14	32,85	36,19	38,58	40,88	43,82
20	23,83	28,41	31,41	34,17	37,57	40,00	42,34	45,31
21	24,93	29,62	32,67	35,48	38,93	41,40	43,77	46,80
22	26,04	30,81	33,92	36,78	40,29	42,80	45,20	48,27
23	27,14	32,01	35,17	38,08	41,64	44,18	46,62	49,73
24	28,24	33,20	36,42	39,36	42,98	45,56	48,03	51,18
25	29,34	34,38	37,65	40,65	44,31	46,93	49,44	52,62
26	30,43	35,56	38,89	41,92	45,64	48,29	50,83	54,05
27	31,53	36,74	40,11	43,19	46,96	49,65	52,22	55,48
28	32,62	37,92	41,34	44,46	48,28	50,99	53,59	56,89
29	33,71	39,09	42,56	45,72	49,59	52,34	54,97	58,30
30	34,80	40,26	43,77	46,98	50,89	53,67	56,33	59,70

Tabela 31 - Distribuição de qui-quadrado para diversos níveis de significância

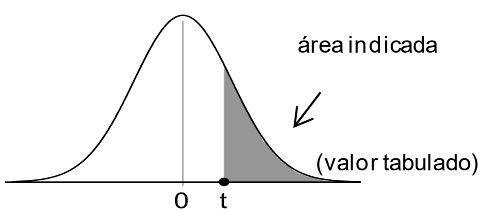
(conclusão)



gl	Área na cauda superior							
	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001
35	40,22	46,06	49,80	53,20	57,34	60,27	63,08	66,62
40	45,62	51,81	55,76	59,34	63,69	66,77	69,70	73,40
45	50,98	57,51	61,66	65,41	69,96	73,17	76,22	80,08
50	56,33	63,17	67,50	71,42	76,15	79,49	82,66	86,66
100	109,1	118,5	124,3	129,6	135,8	140,2	144,3	149,4

Tabela 32 - Tabela da distribuição t de "Student" com valores críticos ou tabelados unilaterais para diversos níveis de significância

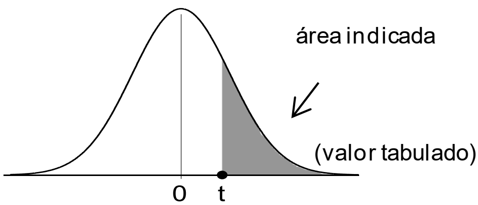
(continua)



gl	Área na cauda superior								
	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
1	1,000	3,078	6,314	12,71	31,82	63,66	127,3	318,3	636,6
2	0,816	1,886	2,920	4,303	6,965	9,925	14,09	22,33	31,60
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,21	12,92
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,894	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408

Tabela 32 - Tabela da distribuição t de “Student” com valores críticos ou tabelados unilaterais para diversos níveis de significância

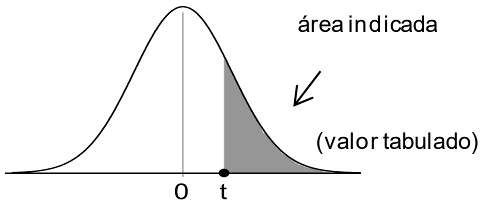
(continua)



gl	Área na cauda superior								
	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	0,685	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768
24	0,685	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	0,684	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	0,684	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	0,684	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,689
28	0,683	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	0,683	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,660
30	0,683	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646
35	0,682	1,306	1,690	2,030	2,438	2,724	2,996	3,340	3,591
40	0,681	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551

Tabela 32 - Tabela da distribuição t de “Student” com valores críticos ou tabelados unilaterais para diversos níveis de significância

(conclusão)



gl	Área na cauda superior								
	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
45	0,680	1,301	1,679	2,014	2,412	2,690	2,952	3,281	3,520
50	0,679	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496
z	0,674	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,291

Tabela 33 - Distribuição de t de “Student” bilateral

(continua)

gl\P	**0,90**	**0,80**	**0,70**	**0,60**	**0,50**	**0,40**	**0,30**	**0,20**	**0,10**	**0,05**	**0,02**	**0,01**	**0,001**
01	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
02	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
03	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,541	12,924
04	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
05	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
06	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
07	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,365	3,499	5,408
08	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
09	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883

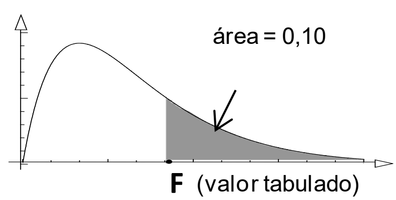
Tabela 33 - Distribuição de t de “Student” bilateral

(conclusão)

gl\P	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01	0,001
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,726
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,856	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,856	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

Tabela 34 - Distribuição “F” de Snedecor-Fisher, mostrando os valores críticos ou tabelados para o nível de significância de 0,10

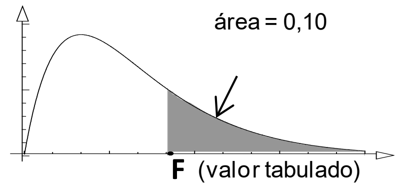
(continua)



g.l. do denominador	graus de liberdade no numerador									
	1	2	3	4	5	6	7	8	9	10
1	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86	60,19
2	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39
3	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23
4	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92
5	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30
6	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94

Tabela 34 - Distribuição “F” de Snedecor-Fisher, mostrando os valores críticos ou tabelados para o nível de significância de 0,10

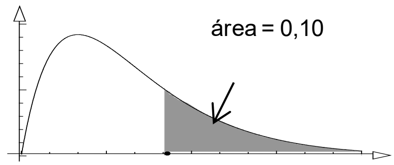
(continua)



g.l. do denominador	graus de liberdade no numerador									
	1	2	3	4	5	6	7	8	9	10
7	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32
11	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25
12	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24	2,21	2,19
13	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20	2,16	2,14
14	3,10	2,73	2,52	2,39	2,31	2,24	2,19	2,15	2,12	2,10
15	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09	2,06
16	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09	2,06	2,03
17	3,03	2,64	2,44	2,31	2,22	2,15	2,10	2,06	2,03	2,00
18	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,00	1,98
19	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02	1,98	1,96
20	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96	1,94
21	2,96	2,57	2,36	2,23	2,14	2,08	2,02	1,98	1,95	1,92
22	2,95	2,56	2,35	2,22	2,13	2,06	2,01	1,97	1,93	1,90
23	2,94	2,55	2,34	2,21	2,11	2,05	1,99	1,95	1,92	1,89
24	2,93	2,54	2,33	2,19	2,10	2,04	1,98	1,94	1,91	1,88
25	2,92	2,53	2,32	2,18	2,09	2,02	1,97	1,93	1,89	1,87
26	2,91	2,52	2,31	2,17	2,08	2,01	1,96	1,92	1,88	1,86
27	2,90	2,51	2,30	2,17	2,07	2,00	1,95	1,91	1,87	1,85
28	2,89	2,50	2,29	2,16	2,06	2,00	1,94	1,90	1,87	1,84
29	2,89	2,50	2,28	2,15	2,06	1,99	1,93	1,89	1,86	1,83
30	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85	1,82
35	2,85	2,46	2,25	2,11	2,02	1,95	1,90	1,85	1,82	1,79

Tabela 34 - Distribuição “F” de Snedecor-Fisher, mostrando os valores críticos ou tabelados para o nível de significância de 0,10

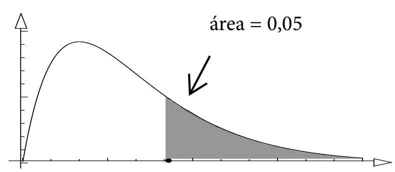
(conclusão)



g.l. do denominador	graus de liberdade no numerador									
	1	2	3	4	5	6	7	8	9	10
40	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79	1,76
45	2,82	2,42	2,21	2,07	1,98	1,91	1,85	1,81	1,77	1,74
50	2,81	2,41	2,20	2,06	1,97	1,90	1,84	1,80	1,76	1,73
100	2,76	2,36	2,14	2,00	1,91	1,83	1,78	1,73	1,69	1,66

Tabela 35 - Distribuição “F” de Snedecor-Fisher, mostrando valores os críticos ou tabelados para o nível de significância de 0,05

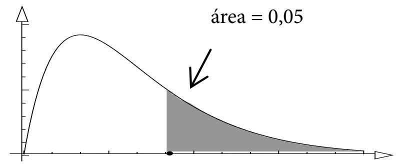
(continua)



gl do denominador	graus de liberdade no numerador									
	1	2	3	4	5	6	7	8	9	10
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14

Tabela 35 - Distribuição “F” de Snedecor-Fisher, mostrando valores os críticos ou tabelados para o nível de significância de 0,05

(conclusão)



gl do denominador	graus de liberdade no numerador									
	1	2	3	4	5	6	7	8	9	10
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16
35	4,12	3,27	2,87	2,64	2,49	2,37	2,29	2,22	2,16	2,11
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
45	4,06	3,20	2,81	2,58	2,42	2,31	2,22	2,15	2,10	2,05
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93

Tabela 36 - Valores críticos de "Z" para alguns níveis de significância (α) mais comuns na construção de intervalos de confiança e aplicação de testes de hipóteses

A tabela seguinte, que fornece os valores críticos ou tabelados de "Z" para ambos os testes unilateral e bilateral e também para intervalos de confiança em vários níveis de significância α, pode revelar-se útil como referência. Os valores críticos de "Z", para outros níveis de significância, são determinados mediante o uso das tabelas de áreas sob a curva normal.

NÍVEL DE SIGNIFICÂNCIA α	**0,10**	**0,05**	**0,01**	**0,005**	**0,002**
VALORES CRÍTICOS DE Z PARA TESTES DE HIPÓTESES E INTERVALOS DE CONFIANÇA UNILATERAIS	-1,280 OU 1,280	-1,645 OU 1,645	-2,330 OU 2,330	-2,580 OU 2,580	-2,880 OU 2,880
VALORES CRÍTICOS DE Z PARA TESTES DE HIPÓTESES E INTERVALOS DE CONFIANÇA BILATERAIS	-1,645 E 1,645	-1,960 E 1,960	-2,580 E 2,580	-2,810 E 2,810	-3,080 E 3,080

APÊNDICE C

Tabela 37 - Resumo teórico dos testes de hipóteses paramétricos, para médias, variâncias e proporções

H_o	σ^2 CONHECIDO?	TIPO DE POPULAÇÃO	ESTATÍSTICA TESTE	H_1	REJEITAR H_o SE
$\mu = \mu_0$	SIM	NORMAL	$Z = \dfrac{\overline{X} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}} \cap N(0,1)$	$\mu < \mu_0$	$Z_{obs} \leq -Z_{(\alpha)}$
				$\mu > \mu_0$	$Z_{obs} \geq Z_{(\alpha)}$
				$\mu \neq \mu_0$	$\lvert Z_{obs} \rvert \geq Z_{\left(\frac{\alpha}{2}\right)}$
	NÃO	NORMAL	$t = \dfrac{\overline{X} - \mu_0}{\dfrac{S}{\sqrt{n}}} \cap t_{(v=n-1)}$	$\mu < \mu_0$	$t_{obs} \leq -t_{(V;\alpha)}$
				$\mu > \mu_0$	$t_{obs} \geq t_{(V;\alpha)}$
				$\mu \neq \mu_0$	$\lvert t_{obs} \rvert \geq t_{\left(V;\frac{\alpha}{2}\right)}$
$P = P_0$ $(n \geq 30)$	———	BERNOULLI	$Z = \dfrac{\hat{P} - P_0}{\sqrt{\dfrac{P_0(1-P_0)}{n}}} \cap N(0,1)$	$P < P_0$	$Z_{obs} \leq -Z_{(\alpha)}$
				$P > P_0$	$Z_{obs} \geq Z_{(\alpha)}$
				$P \neq P_0$	$\lvert Z_{obs} \rvert \geq Z_{\left(\frac{\alpha}{2}\right)}$
$\sigma^2 = \sigma_0^2$	———	NORMAL	$\chi^2_{TESTE} = \dfrac{(n-1)S^2}{\sigma_0^2} \cap \chi^2_{(v=n-1)}$	$\sigma^2 < \sigma_0^2$	$\chi^2_{obs} \leq \chi^2_{(V;1-\alpha)}$
				$\sigma^2 > \sigma_0^2$	$\chi^2_{obs} \geq \chi^2_{(V;\alpha;)}$
				$\sigma^2 \neq \sigma_0^2$	$\chi^2_{obs} \leq \chi^2_{\left(V;1-\frac{\alpha}{2}\right)}$ *e* $\chi^2 \geq \chi^2_{\left(V;\frac{\alpha}{2}\right)}$

H_o	σ_1^2 *e* σ_2^2 CONHECIDOS?	TIPO DE POPULAÇÃO	ESTATÍSTICA TESTE	H_1	REJEITAR H_o SE
$\mu_1 = \mu_2$	SIM	NORMAIS	$Z = \dfrac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \cap N(0,1)$	$\mu_1 < \mu_2$	$Z_{obs} \leq -Z_{(\alpha)}$
				$\mu_1 > \mu_2$	$Z_{obs} \geq Z_{(\alpha)}$
				$\mu_1 \neq \mu_2$	$\left\lvert Z_{obs} \right\rvert \geq Z_{\left(\frac{\alpha}{2}\right)}$
	NÃO $\left(\sigma_1^2 = \sigma_2^2\right)$	NORMAIS	$t = \dfrac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}} \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \cap t_{(v = n_1 + n_2 - 2)}$	$\mu_1 < \mu_2$	$t_{obs} \leq -t_{(V = n_1 + n_2 - 2; \alpha)}$
				$\mu_1 > \mu_2$	$t_{obs} \geq t_{(V = n_1 + n_2 - 2; \alpha)}$
				$\mu_1 \neq \mu_2$	$\left\lvert t_{obs} \right\rvert \geq t_{\left(V = n_1 + n_2 - 2; \frac{\alpha}{2}\right)}$
	NÃO $\left(\sigma_1^2 \neq \sigma_2^2\right)$	NORMAIS	$t = \dfrac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \cap t_{\left[\left[v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}\right]\right]}$	$\mu_1 < \mu_2$	$t_{obs} \leq -t_{(V; \alpha)}$
				$\mu_1 > \mu_2$	$t_{obs} \geq t_{(V; \alpha)}$
				$\mu_1 \neq \mu_2$	$\left\lvert t_{obs} \right\rvert \geq t_{\left(V; \frac{\alpha}{2}\right)}$
	________	BERNOULLI	$Z = \dfrac{\left(\hat{P}_1 - \hat{P}_2\right) - \left(P_1 - P_2\right)}{\sqrt{\dfrac{P_1(1 - P_1)}{n_1} + \dfrac{P_2(1 - P_2)}{n_2}}} \cap N(0,1)$	$P_1 < P_2$	$Z_{obs} \leq -Z_{(\alpha)}$
				$P_1 > P_2$	$Z_{obs} \geq Z_{(\alpha)}$
				$P_1 \neq P_2$	$\left\lvert Z_{obs} \right\rvert \geq Z_{\left(\frac{\alpha}{2}\right)}$

H_o	σ^2 CONHECIDO?	TIPO DE POPULAÇÃO	ESTATÍSTICA TESTE	H_1	REJEITAR H_o SE
$\frac{\sigma_1^2}{\sigma_2^2} = \sigma_0^2$	———	NORMAIS	$F = \frac{S_1^2}{S_2^2}\frac{1}{\sigma_0^2} \cap F_{(v_1=n_1-1;v_2=n_2-1)}$	$\frac{\sigma_1^2}{\sigma_2^2} < \sigma_0^2$	$F_{OBS} \le F_{(v_1=n_1-1,v_2=n_2-1;1-\alpha)}$
				$\frac{\sigma_1^2}{\sigma_2^2} > \sigma_0^2$	$F_{OBS} \ge F_{(v_1=n_1-1,v_2=n_2-1;\alpha)}$
				$\frac{\sigma_1^2}{\sigma_2^2} \ne \sigma_0^2$	$F_{OBS} \le F_{\left(v_1=n_1-1,v_2=n_2-1;1-\frac{\alpha}{2}\right)}$ e $F_{OBS} \ge F_{\left(v_1=n_1-1,v_2=n_2-1;\frac{\alpha}{2}\right)}$

APÊNDICE D

TERMOS PRÓPRIOS OU DE USO CORRENTE NA APLICAÇÃO DA TEORIA DOS TESTES OU PROVAS DE HIPÓTESES OU DE SIGNIFICÂNCIA E EXERCÍCIOS RESOLVIDOS

Na estatística experimental, alguns termos de uso próprio ou corrente devem ser conhecidos pelo pesquisador a fim de que possa aplicar de forma mais eficiente a ferramenta estatística dos testes ou provas de hipóteses ou de significância, tais como:

Delineamento experimental: É o plano, desenho ou formato utilizado na experimentação de como se distribuem ou se aplicam os tratamentos ou fatores experimentais, nas unidades experimentais ou parcelas, em um amplo entendimento das análises a serem feitas quando todos os dados estiverem disponíveis.

Tratamento ou Fator experimental: É o método ou processo em que se comparam seus efeitos sobre as unidades experimentais.

Unidade experimental ou Parcela experimental: É a menor porção física do material experimental sobre o qual se aplicam os tratamentos.

Repetição: É o número de vezes em que se repetem os tratamentos nas unidades experimentais.

Casualização ou Aleatorização ou Sorteio dos tratamentos: É o processo de distribuição casual dos tratamentos nas unidades experimentais, mediante um dispositivo de sorteio, como pedaços de papeis numerados, ou dígitos em uma tabela de números aleatórios ou por meio de uma calculadora ou computador usando a função RANDOM, de tal forma que cada tratamento tenha a mesma probabilidade de ser aplicado nas unidades experimentais.

Controle local ou Bloqueamento: É o processo de formação de estratos homogêneos de unidades experimentais, restringindo a casualização, com o objetivo de se conseguir minimizar o grau de variabilidade de unidades experimentais mais próximas. Por exemplo, a formação de blocos no delineamento experimental em blocos completos casualizados.

Variável resposta ou Dependente: É o valor da medida, mensuração ou contagem das respostas dos efeitos dos tratamentos ensaiados ou comparados nas unidades experimentais, como, por exemplo, temos a produção leiteira de uma vaca em quilogramas por dia (kg/dia) depois de receber tratamento com ração enriquecida com proteínas à base de soja.

Tratamento controle ou Testemunha ou Placebo: É o tratamento ou fator padrão ou de referência, antigo ou quando as parcelas não recebem tratamento novo.

Unidade de medida (contagem ou mensuração): É a unidade utilizada para contar, enumerar ou mensurar a variável resposta obtida após a aplicação da prova ou teste de hipótese, como, por exemplo, gramas, centímetros, graus Celsius, tonelada por hectare, etc. Deve ser escolhida durante o período do planejamento experimental.

Arranjo experimental: Termo que se refere à combinação dos fatores ou tratamentos experimentais em seus diversos níveis em que podem e devem ser combinados, como, por exemplo, nos arranjos fatoriais 3 por 4, onde se tem três tratamentos em quatro níveis quantitativos ou qualitativos.

Simbologia ou nomenclatura usada nos testes de hipóteses:

De modo geral, as hipóteses são testadas em três níveis de significância, quais sejam: 1%, 5% e 10% e raramente 0,1%. Escolhe-se para apresentar a conclusão o menor valor de alfa (α) capaz de rejeitar a hipótese H_0. Se nenhum dos níveis conseguir rejeitar a hipótese H_0, diz-se que o teste é não significativo. Se o menor valor de alfa que rejeita a hipótese

H_0 é, por exemplo, 5%, diz-se que o teste é significativo ao nível de probabilidade de 5%. As decisões associadas a esses níveis de significância costumam ser assim classificadas conforme a Tabela abaixo: Por exemplo, se rejeitarmos a hipótese H_0 a 5% de probabilidade (o valor da estatística teste cai dentro da região de rejeição da hipótese H_0), coloca-se 1 (um) asterisco(*) no canto superior direito do valor da estatística teste; por outro lado, se aceitar a hipótese H_0(o valor da estatística teste cai dentro da região de aceitação da hipótese H_0), coloca-se o símbolo n.s. (não significativo) no canto superior direito do valor da estatística teste. Por exemplo, $Z_{TESTE} = +\,3{,}58^{*}$, ou ainda $t_{TESTE} = +\,1{,}32^{n.s.}$.

Tabela - Classificação dos níveis de significância alfa (α) ou p

Nível de significância alfa (α)	**Conclusão**	**Simbologia**
Valor de alfa menor ou igual a 0,1% $(\alpha \leq 0{,}1\%)$	Diferença muito altamente significante ou significativa	Três asteriscos (***)
Valor de alfa menor ou igual a 1% $(\alpha \leq 1\%)$	Diferença altamente significante ou significativa	Dois asteriscos (**)
Valor de alfa entre 1% e 5% $(1\% < \alpha < 5\%)$	Diferença significante ou significativa	Um asterisco (*)
Valor de alfa entre 5% e 10% $(5\% \leq \alpha \leq 10\%)$	Diferença provavelmente significante ou significativa	Um asterisco (*)
Valor de alfa maior que 10% $(\alpha > 10\%)$	Diferença não significante ou não significativa	n.s.

1. TIPOS DE TESTES PARAMÉTRICOS: TESTES DE HIPÓTESES ESTATÍSTICAS PARAMÉTRICOS: TESTES REFERENTES A MÉDIAS, VARIÂNCIAS E PROPORÇÃO

1.1 Teste de hipótese para a média ou de conformidade de uma média (μ) de uma população infinita, normal, com variância (σ^2) conhecida [amostra grande (n≥30) ou pequena (n<30)], ou variância (σ^2) desconhecida [amostra grande (n ≥ 30)]

1.1.1 O teste de hipótese

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$\mu = \mu_0$	$Z_{TESTE} = \dfrac{\overline{X} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}}$ ou $Z_{TESTE} = \dfrac{\overline{X} - \mu_0}{\dfrac{S}{\sqrt{n}}}$ com σ conhecido ou não	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$Z \leq -Z_{(\alpha)}$ $Z \geq Z_{(\alpha)}$ $Z \leq -Z_{\left(\frac{\alpha}{2}\right)}$ e $Z \geq Z_{\left(\frac{\alpha}{2}\right)}$

O teste de conformidade de uma média tem por objetivo verificar se a média μ de uma população é ou não igual a um valor dado μ_0. A hipótese nula é, portanto: $H_0 : \mu = \mu_0$, e rejeitamos, evidentemente, esta hipótese quando a média observada é muito diferente do valor teórico μ_0 .

1.1.2 Exercício de aplicação

Sabe-se que certa linhagem de camundongo da linhagem C57BL6J, alimentada com uma ração padrão, tem um aumento médio de peso igual a 64 gramas durante os três primeiros meses de vida. Um lote (amostra) de 81 camundongos dessa linhagem foi alimentado com uma nova ração, mantendo-se as condições ambientais padronizadas. O aumento médio de peso observado nos camundongos foi de $\overline{X} = 60,75$ g, com desvio padrão de S = 3,84 g. A nova ração tem a mesma eficiência alimentar da raça padrão? Use α = 0,05.

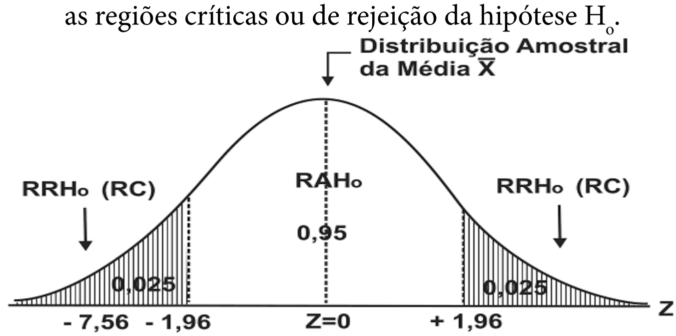
$$\mathbf{n = 81,\ \overline{X} = 60,75,\ S = 3,84,\ \mu = \mu_0 = 64}$$

a) $\begin{cases} H_0 : \mu = 64 \\ H_1 : \mu \neq 64 \end{cases}$

b) α = 0,05; $Z_{\left[\frac{0,05}{2}\right]} = \pm 1,96$

c)

Figura 8 - Curva do teste de hipótese Z Bilateral mostrando as regiões críticas ou de rejeição da hipótese H_o.



d) $Z_{Teste} = \dfrac{60,75 - 64,00}{\dfrac{3,84}{\sqrt{81}}} = \dfrac{-3,25}{0,43} = -7,56^{*}$

$$|Z_{Teste}| > \left| Z_{\left(\frac{0,05}{2}\right)} \right| \rightarrow, \ |-7,56| > |1,96| \rightarrow$$ Então, rejeita-se a hipótese H_0

Um Asterisco (*) significa que se forem conduzidos 100 experimentos desse tipo, utilizando a mesma metodologia, amostra, variável resposta, etc., assumindo-se que a hipótese de nulidade H_0 é verdadeira, assume-se um risco de 5% de probabilidade de rejeitá-la, em média.
v) Conclusões: De acordo com o teste Z, para a média "μ" ao nível de 5% de significância, pode-se concluir que a nova ração tem eficiência alimentar inferior à da ração padrão, de forma que a ração padrão é a melhor entre as duas ensaiadas.

1.2 Teste de hipótese para a média ou de conformidade de uma média (μ) de uma população infinita, aproximadamente normal com variância (σ^2) desconhecida e amostra pequena (n < 30)

1.2.1 O teste de hipótese

HIPÓTESE DE NULIDADE H_0	**ESTATÍSTICA TESTE**	**HIPÓTESE ALTERNATIVA H_1**	**REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_0]**
$\mu = \mu_0$	$t_{TESTE} = \dfrac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}}$, *com* $V = n - 1\ G.L.,$ **e com σ desconhecido**	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$t \leq -t_{(V;\alpha)}$ $t \geq t_{(V;\alpha)}$ $t \leq -t_{\left(V;\frac{\alpha}{2}\right)}$ *e* $t \geq t_{\left(V;\frac{\alpha}{2}\right)}$

O teste de conformidade de uma média tem por objetivo verificar se a média μ de uma população é ou não igual a um valor dado $\mu_{0.}$ A hipótese nula é, portanto, $H_0 : \mu = \mu_0$, e rejeitamos, evidentemente, esta hipótese quando a média observada é muito diferente do valor teórico μ_0.

1.2.2 Exercício de aplicação

Em um seringal, localizado no estado do Acre, no qual se utiliza o processo convencional de sangria em 2012, a produção média de borracha seca é de 26 g/árvore/corte. Tomou-se uma amostra ao acaso composta de 25 seringueiras, as quais foram sangradas usando-se um novo processo, tendo produzido 28 g em média, com desvio padrão de 4g. O novo processo é mais eficiente do que o convencional? Use α = 5%.

Sendo que $\mu = \mu_0 = 26$.

n=25; $\overline{X} = 28;\ S = 4$

V= n -1= 25 -1=24

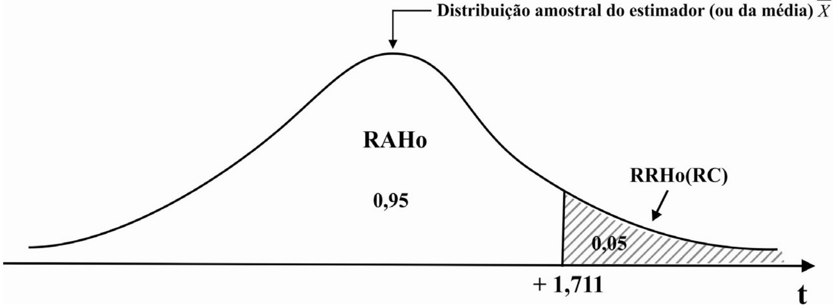
$t_{[24;0,05]} = +1,711$

a) $H_o: \mu = \mu_0 = 26$

$H_1: \mu > \mu_0 > 26$

b) $\alpha = 0,05 = 5\%;\ t_{[24;0,05]} = +1,711$

c)



4) $t_{TESTE} = t_{OBSERVADO} = t_{CALCULADO} = \dfrac{28 - 26}{\dfrac{4}{\sqrt{25}}} = \dfrac{2}{0,8} = 2,5^{*}$

Como o valor do t teste é maior do que o valor do t tabelado ou crítico, isto é: $2,5 > t_{[24;0,05]} = +1,711$, rejeita-se a hipótese H_{o}.

5) Conclusões: Baseado no teste "t" de "Student" para a média, com 24 graus de liberdade e ao nível de 5% de significância, pode-se concluir que o novo processo de sangria das seringueiras é mais eficiente do que o processo convencional, no que se refere a induzir uma maior produção média de borracha seca das seringueiras em g/árvore/corte.

1.3 Teste de hipótese para a diferença entre as médias $(\mu_1 - \mu_2)$ de duas populações infinitas, com variâncias $(\sigma_1^2 \; e \; \sigma_2^2)$ conhecidas [amostras grandes e independentes ($n_1 \geq 30$ e $n_2 \geq 30$) ou pequenas ($n_1 < 30$ e $n_2 < 30$)] ou com variâncias desconhecidas [amostras grandes ($n_1 \geq 30$ e $n_2 \geq 30$)].

1.3.1 O teste de hipótese

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$\mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$	$Z_{TESTE} = \dfrac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ $Z_{TESTE} = \dfrac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$, σ_1 ***e*** σ_2 ***conhecidos.*** ***ou não conhecidos.***	$\mu_1 < \mu_2$ $\mu_1 > \mu_2$ $\mu_1 \neq \mu_2$	$Z \leq -Z_{(\alpha)}$ $Z \geq Z_{(\alpha)}$ $Z \leq -Z_{\left(\frac{\alpha}{2}\right)}$ *e* $Z \geq Z_{\left(\frac{\alpha}{2}\right)}$

1.3.2 Exercício de aplicação

Um experimento aleatório, conduzido no Estado de São Paulo em 2012, consistiu em comparar os pesos vivos ao nascer, em quilogramas, de animais bovinos (bezerros) machos das raças Gir e Guzerá da estação experimental de zootecnia de Sertãozinho, do instituto de Zootecnia da secretária da Agricultura de São Paulo, ao nível de 5% de Probabilidade.

Pesos vivos ao nascer (kg) de bezerros machos zebu, onde:

$$n_1 = 40;\ \overline{X}_1 = 28,3\ Kg;\ S_1^2 = 108,37\ Kg^2;\ S_1 = 10,41\ Kg$$

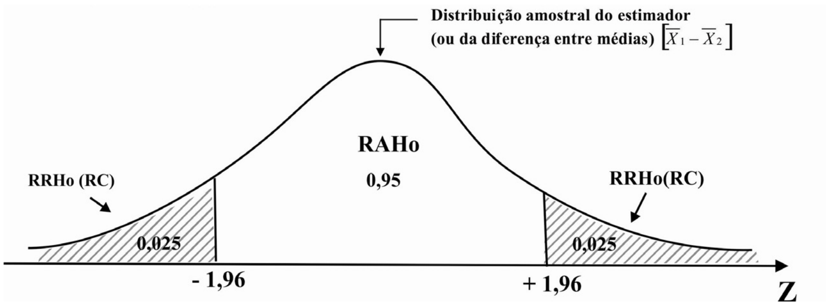
$$n_2 = 50;\ \overline{X}_2 = 23,5\ Kg;\ S_2^2 = 49,70\ Kg^2;\ S_2 = 7,05\ Kg$$

a) Ho: $\mu_1 = \mu_2 \therefore \mu_1 - \mu_2 = 0$

H_1: $\mu_1 \neq \mu_2$

b) α = 0,05 = 5%;

c) $Z_{\left[\frac{0,05}{2}\right]} = \pm 1{,}96$



d)

$$\boldsymbol{Z}_{TESTE} = \boldsymbol{Z}_{CALCULADO} = \frac{(28{,}3 - 23{,}5) - 0}{\sqrt{\frac{(10{,}41)^2}{40} + \frac{(7{,}05)^2}{50}}} = \frac{4{,}8 - 0}{\sqrt{2{,}7092 + 0{,}99405}} = \frac{4{,}8}{\sqrt{3{,}70325}} = \frac{4{,}8}{1{,}9244} = 2{,}5^{*}$$

Como o valor do Z teste é maior do que o valor do Z tabelado ou crítico, isto, é, 2,5 > $\boldsymbol{Z}_{[0,05]} = \pm\ 1{,}96$, rejeita-se a hipótese $H_{o.}$

e) Conclusões: De acordo com o teste Z para a diferença entre médias, ao nível de 5% de significância, conclui-se que os pesos vivos médios ao nascer, em quilogramas, dos bezerros machos Zebu foram estatisticamente ou significativamente diferentes. Os animais da raça Gir foram os de maior peso, portanto são animais com maior aptidão para formação de Plantel ou rebanho para abate, corte ou engorda.

1.4 Teste de hipótese para a diferença entre as médias $(\mu_1 - \mu_2)$ de duas populações infinitas, aproximadamente normais, com variâncias $(\sigma_1^2 \; e \; \sigma_2^2)$ desconhecidas e estatisticamente iguais $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$, ou seja, populações homocedásticas e amostras pequenas e independentes ($N_1 < 30$ e $n_2 < 30$).

1.4.1 O teste de hipótese

HIPÓTESE DE NULIDADE H_o	ESTATÍSTICA TESTE	HIPÓTESE ALTERNATIVA H_1	REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]
$\mu_1 = \mu_2$ $\mu_1 - \mu_2 = 0$	$t_{TESTE} = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_P \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$, $com\ V = n_1 + n_2 - 2, e$ $\sigma_1 = \sigma_2$ *mas desconhecidos, sendo* $S_P^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}, e\ S_P = \sqrt{S_P^2}$ $\dfrac{S^2_{MAIOR}}{S^2_{MENOR}} < 4$, e $V = v_1 + v_2 = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$	$\mu_1 < \mu_2$ $\mu_1 > \mu_2$ $\mu_1 \neq \mu_2$	$t \leq -t_{(V;\alpha)}$ $t \geq t_{(V;\alpha)}$ $t \leq -t_{\left(V;\frac{\alpha}{2}\right)}$ $e \quad t \geq t_{\left(V;\frac{\alpha}{2}\right)}$

1.4.2 Exercício de aplicação

Realizou-se um experimento aleatório em uma granja localizada em Mossoró, RN, em 2012, com 340 aves de pintos machos de um dia de idade, da linhagem Ross, com o objetivo de estudar o efeito da fonte de fósforo no ganho de peso das aves em gramas. Usou-se 10 repetições por tratamento e 17 pintos por parcela. Os resultados para ganho de peso médio por parcela em gramas estão mostrados na tabela abaixo.

Tabela 4: Dados de ganhos de peso de pintos machos de um dia de idade submetidos a tratamentos diferentes

Termofosfato Magnesiano (X_1)		**Fosfato Bicálcico (X_2)**	
63,0	60,9	90,0	91,8
62,5	62,5	90,6	92,0
61,9	63,8	91,7	92,3
60,7	62,0	90,9	90,6
63,2	61,8	90,3	91,2

Sendo que, $\overline{X}_{11} = \dfrac{x_1 + x_2 + x_3 + x_4 + \cdots + x_{17}}{17} = 63{,}0\ g$

Onde temos os seguintes dados:

$$n_1 = 10;\ \overline{X}_1 = 62{,}23\ g;\ S_1^2 = 0{,}96\ g^2;\ S_1 = 0{,}98\ g$$
$$n_2 = 10;\ \overline{X}_2 = 91{,}14\ g;\ S_2^2 = 0{,}61\ g^2;\ S_2 = 0{,}78\ g$$

a) Verificar se existe efeito significativo da fonte de fosfato, isto é, se os ganhos médios de pesos em gramas são estatisticamente iguais, ou ainda se existe igualdade entre os tratamentos ensaiados, considerando $\alpha = 0{,}05$.

$\dfrac{S^2_{MAIOR}}{S^2_{MENOR}} = \dfrac{0{,}96}{0{,}61} = 1{,}6 < 4.$ Portanto, as variâncias são iguais, $\sigma_1^2 = \sigma_2^2 = \sigma^2$

o que nos mostra populações homoscedásticas.

V = 10+10-2=18; $t_{\left[18;\frac{0{,}05}{2}\right]} = \pm\ 2{,}101$

$$S_C^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{(10-1)\,0{,}96 + (10-1)\,0{,}61}{10+10-2} = \frac{(9)\,0{,}96 + (9)\,0{,}61}{10+10-2} = \frac{14{,}13}{18} = 0{,}785,$$

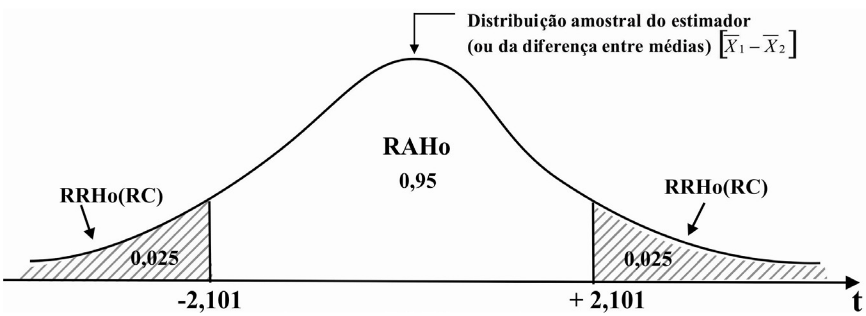
$$S_C = \sqrt{0{,}785}; \quad S_C = 0{,}886$$

a) **Ho:** $\mu_1 = \mu_2 \therefore \mu_1 - \mu_2 = 0$

H_1: $\mu_1 \neq \mu_2$

b) $\boldsymbol{\alpha = 0{,}05 = 5\%}$;

c) $t_{\left[18;\frac{0{,}05}{2}\right]} = \pm\, 2{,}101$



d)

$$t_{TESTE} = t_{OBSERVADO} = t_{CALCULADO} = \frac{(62{,}23 - 91{,}14) - 0}{0{,}886.\sqrt{\frac{1}{10} + \frac{1}{10}}} = \frac{-28{,}91 - 0}{0{,}886(0{,}447)} = \frac{-28{,}91}{0{,}396042} = -72{,}99^{*}$$

Como o valor do "t" teste ou t observado é maior do que o valor do "t" tabelado ou crítico, isto é, -72,99 < -2,101, ou $|\mathbf{-\,72{,}99}| > |\mathbf{-\,2{,}101}|$, quer dizer 72,99 > 2,101, então rejeita-se a hipótese $H_{o.}$ Em outras palavras, o "t" observado é significativo a 5% de probabilidade, razão pela qual se rejeita a hipótese H_o.

e) Conclusões: Baseado no teste "t" de "Student" para a diferença entre médias, com 18 graus de liberdade e ao nível de 5% de probabilidade, pode-se concluir que houve efeito significativo da fonte de fosfato, isto é, houve diferença estatística (significativa) entre os ganhos de peso vivo médio dos pintos machos de um dia de idade da Linhagem Ross, ou seja, a melhor fonte de fosfato é o fosfato bicálcico.

1.5 Teste de hipótese para a proporção (p) de uma população infinita, normal, e com amostra grande $(n \geq 30)$

1.5.1 O teste de hipótese

HIPÓTESE DE NULIDADE H_o	**ESTATÍSTICA TESTE**	**HIPÓTESE ALTERNATIVA H_1**	**REGIÃO CRÍTICA OU REGIÃO DE REJEIÇÃO DA HIPÓTESE H_0 [RRH_O]**
$P = P_0$	$Z_{TESTE} = \dfrac{\hat{P} - P_0}{\sqrt{\dfrac{P_0 q_0}{n}}}$ **com população infinita, e** $n\hat{P} > 5$, $n\hat{q} > 5$, $n \geq 30$, $\hat{P} = \dfrac{X}{n}$ **e** $q_0 = 1 - P_0$	$P < P_0$ $P > P_0$ $P \neq P_0$	$Z \leq -Z_{(\alpha)}$ $Z \geq Z_{(\alpha)}$ $Z \leq -Z_{\left(\frac{\alpha}{2}\right)}$ *e* $Z \geq Z_{\left(\frac{\alpha}{2}\right)}$

1.5.2 Exercício de aplicação

Em uma granja localizada no município de Mossoró, RN, selecionou-se em 2012 uma amostra aleatória de 800 coelhos da raça Nolfox, a qual apresentou 480 machos. Ao nível de 1% de significância, pode-se concluir que há prevalência (predominância) de coelhos machos nessa granja?

P(sucesso) = número de coelhos machos

$$\hat{P} = f = \frac{X}{n} = \frac{480}{800} = 0{,}60; \quad \hat{q} = 1 - 0{,}60 = 0{,}40$$

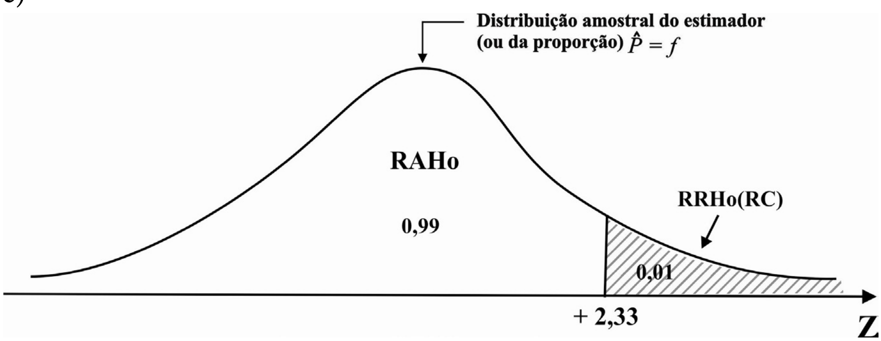
$P_0 = 0{,}50 = 50\%$ e $q_0 = 1 - 0{,}50 = 0{,}50 = 50\%$

a) **Ho:** $P = P_0,\ P = 50\% = 0{,}50$

H_1: $P > 50\%$

b) $\alpha = 0{,}01 = 1\%;\ Z_{[0{,}01]} = +2{,}33$

c)



d)

$$Z_{TESTE} = Z_{OBSERVADO} = Z_{CALCULADO} = \frac{0{,}60 - 0{,}50}{\sqrt{\frac{0{,}50(1-0{,}50)}{800}}} = \frac{0{,}1000000}{0{,}0176776} \cong 5{,}7^{**}$$

Como o valor do "Z" teste ou "Z" observado é maior do que o valor do "Z" tabelado ou "Z" crítico, isto é, 5,7 > + 2,33, rejeita-se a hipótese H_0.

e) Conclusões: De acordo com o teste "Z" para a proporção, ao nível de 1% de probabilidade ou significância, pode-se concluir que há prevalência de coelhos machos nessa granja, ou melhor, nessa população de coelhos da Raça Nolfox.

EdUFERSA
Editora Universitária da UFERSA
Av. Francisco Mota, 572
Compl.: Centro de Convivência (Campus Leste)
Costa e Silva - Mossoró/RN
CEP: 59.625-900

Tel: (84) 3317-8267
livraria.ufersa.edu.br
edufersa.ufersa.edu.br
edufersa@ufersa.edu.br