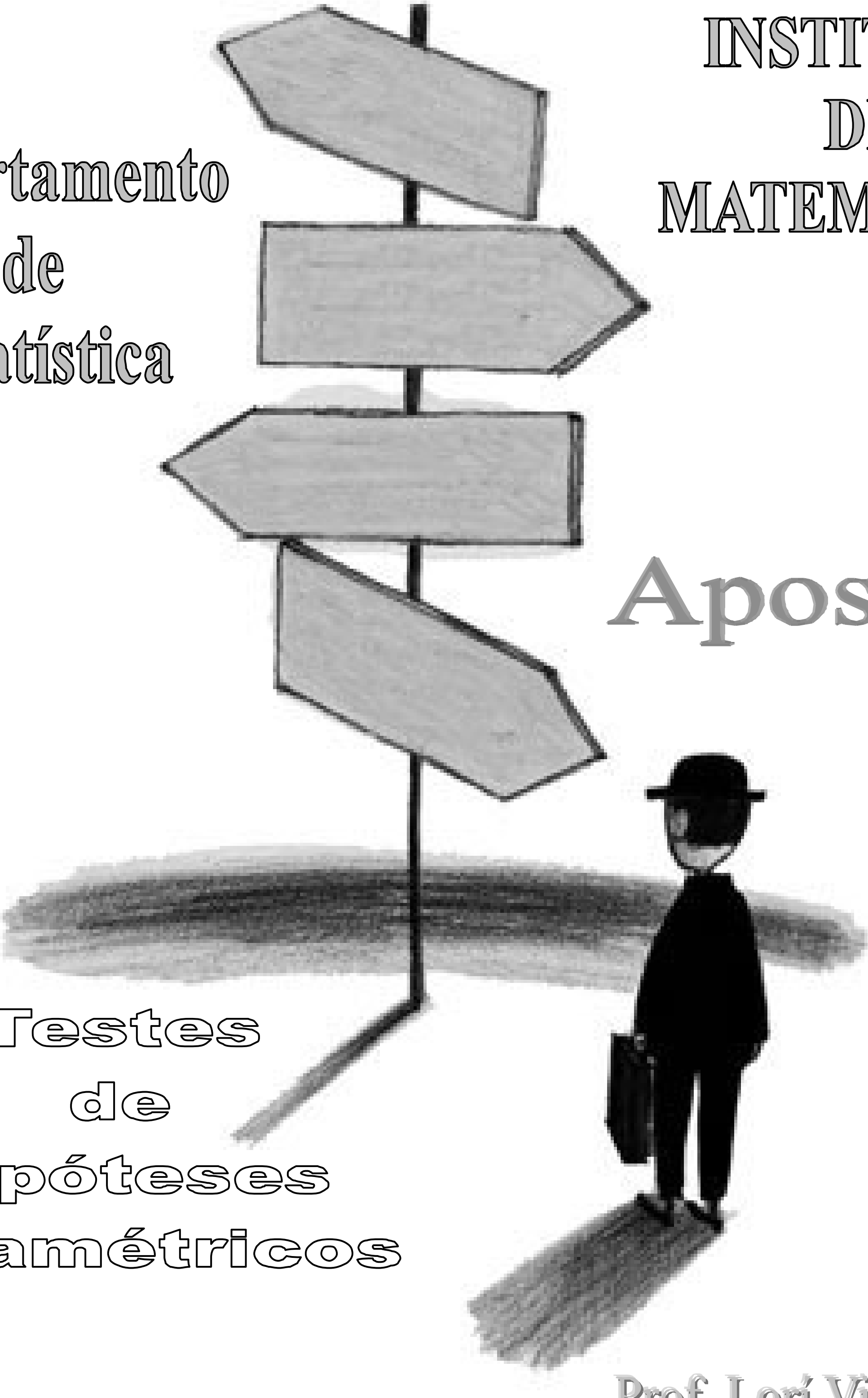
UFRGS

INSTITUTO
DE
MATEMÁTICA

Departamento
de
Estatística



Apostila

Testes de Hipóteses Paramétricos

Prof. Lorí Viali, Dr.

Porto Alegre, 2008

SUMÁRIO

1. INTRODUÇÃO

1.1. GENERALIDADES

Um dos principais assuntos da Estatística moderna é a *inferência estatística*. A inferência estatística é dividida em dois grandes tópicos: a estimação de parâmetros de uma população e os testes de hipóteses.

No desenvolvimento dos métodos da estatística moderna, as primeiras técnicas de inferência que apareceram foram as que faziam diversas hipóteses sobre a natureza da população da qual se extraíram os dados. Como os valores relacionados com a população são denominados "parâmetros", tais técnicas estatísticas foram denominadas de paramétricas.

1.2. METODOLOGIA DO TESTE DE HIPÓTESES

Nas ciências do comportamento, efetua-se levantamentos a fim de determinar o grau de aceitação de hipóteses baseadas em teorias do comportamento. Formulada uma determinada hipótese particular é necessário coletar dados empíricos e com base nestes dados decide-se então sobre a validade ou não da hipótese. A decisão sobre a hipótese pode levar a rejeição, revisão ou aceitação da teoria que a originou.

Para se chegar a conclusão que uma determinada hipótese deverá ser aceita ou rejeitada, baseado em um particular conjunto de dados, é necessário dispor de um processo objetivo que permita decidir sobre a veracidade ou falsidade de tal hipótese.

A objetividade deste processo deve ser baseada na informação proporcionada pelos dados, e como estes dados, em geral, envolvem apenas parte da população que se pretende atingir, no risco que se está disposto a correr de que a decisão tomada não esteja correta.

A metodologia para a decisão sobre a veracidade ou falsidade de uma determinada hipótese envolve algumas etapas.

1. Definir a hipótese de igualdade (H_0).
2. Escolher a prova estatística (com o modelo estatístico associado) para tentar rejeitar H_0.
3. Definir o nível de significância (α) e um tamanho de amostra (n).

4. Determinar (ou supor determinada) a distribuição amostral da prova estatística sob a hipótese de nulidade.
5. Definir a região de rejeição.
6. Calcular o valor da prova estatística, utilizando os valores obtidos na(s) amostra(s). Se tal valor estiver na região de rejeição, rejeitar, então a hipótese nula, senão a decisão será que a hipótese nula não poderá ser rejeitada ao nível de significância determinado.

1.3. AS HIPÓTESES

Uma hipótese estatística é uma suposição ou afirmação que pode ou não ser verdadeira, relativa a uma ou mais populações. A veracidade ou falsidade de uma hipótese estatística *nunca* é conhecida com certeza, a menos que, se examine toda a população, o que é impraticável na maior parte das situações.

Desta forma, toma-se uma amostra aleatória da população de interesse e com base nesta amostra é estabelecido se a hipótese é provavelmente verdadeira ou provavelmente falsa. A decisão de que a hipótese é provavelmente verdadeira ou falsa é tomada com base em distribuições de probabilidade denominadas de “distribuições amostrais”. Em estatística trabalha-se com dois tipos de hipótese.

A **hipótese nula** é a hipótese de igualdade. Esta hipótese é denominada de hipótese de nulidade e é representada por H_0 (lê-se h zero). A hipótese nula é normalmente formulada com o objetivo de ser rejeitada. A rejeição da hipótese nula envolve a aceitação de outra hipótese denominada de **alternativa**. Esta hipótese é a definição operacional da hipótese de pesquisa que se deseja comprovar. A natureza do estudo vai definir como deve ser formulada a hipótese alternativa. Por exemplo, se o teste é do tipo paramétrico, onde o parâmetro a ser testado é representado por θ, então a hipótese nula seria: $H_0 : \theta = \theta_0$ e as hipóteses alternativas seriam:

$H_1 : \theta = \theta_1$ (Hipótese alternativa simples) ou

$H_1 : \theta \neq \theta_0$; $\theta > \theta_0$ ou $\theta < \theta_0$. (Hipóteses alternativas compostas)

No primeiro caso, $H_1 : \theta \neq \theta_0$, diz-se que o teste é bilateral (ou bicaudal), se $H_1 : \theta > \theta_0$, diz-se que o teste é unilateral (ou unicaudal) à direita e se $H_1 : \theta < \theta_0$, então, diz-se que o teste é unilateral (ou unicaudal) à esquerda.

1.4. A ESCOLHA DO TESTE ESTATÍSTICO

Existem inúmeros testes estatísticos tanto paramétricos quanto não paramétricos. Alguns itens devem ser levados em conta na escolha da prova estatística para determinada situação. A maneira como a amostra foi obtida, a natureza da população da qual se extraiu a amostra e o tipo de mensuração ou escala empregado nas definições operacionais das variáveis envolvidas, isto é, o conjunto de valores numéricos e ainda o tamanho da amostra disponível.

Uma vez determinados à natureza da população e o método de amostragem ficará estabelecido o modelo estatístico. Associado a cada teste estatístico tem-se um modelo estatístico e condições de mensuração, o teste é válido sob as condições especificadas no modelo e pelo nível da escala de mensuração. Nem sempre é possível verificar se todas as condições do modelo foram satisfeitas e neste caso tem-se que admitir que estas condições foram satisfeitas. Estas condições do modelo estatístico são denominadas *suposições* ou *hipóteses* do teste. Qualquer decisão tomada através de um teste estatístico somente terá validade se as condições do modelo forem válidas.

É óbvio que quanto mais fracas forem as suposições do modelo mais gerais serão as conclusões. No entanto, as provas mais poderosas, isto é, as que apresentam maior probabilidade de rejeitar H_0 quando for falsa, são as que exigem as suposições mais fortes ou mais amplas.

1.5. CONCEITOS ADICIONAIS DO TESTE DE HIPÓTESES

Além dos conceitos já vistos para o teste de hipóteses é necessário ainda definir os erros envolvidos e as regiões de rejeição e de aceitação.

Para ilustrar estes conceitos será suposto o seguinte teste a ser feito: Dispõem-se de duas moedas com aparência idêntica, só que uma (M_1) é equilibrada, isto é, P(Cara) = P(Coroa) = 50%, enquanto que a outra (M_2) é viciada de tal forma que favorece cara na proporção de 80%, ou seja, P(Cara) = 80% enquanto que P(Coroa) = 20%. Supõem-se que uma das moedas é lançada e que com base na variável X = número de caras, deve-se decidir qual delas foi lançada. Neste caso o teste a ser feito envolve as seguintes hipóteses:

H_0: A moeda lançada é a equilibrada (M_1), ou seja, p = 50%

H_1: A moeda lançada é a viciada (M_2), ou seja p = 80%, onde "p" é a proporção de caras.

Tem-se que tomar a decisão de apontar qual foi a moeda lançada, baseado apenas em uma amostra, por exemplo 5 lançamentos, de uma população infinita de lançamentos possíveis. A decisão, é claro, estará sujeita a erros, pois se está tomando a decisão em condições de incerteza.

A decisão será baseada nas distribuições amostrais das duas moedas. A tabela 01 mostra as probabilidades de se obter os valores: 0, 1, 2, 3, 4 e 5, da variável X = número de caras, em 5 lançamentos de cada uma das moedas.

Tabela 01 - Probabilidades de se obter cara em 5 lançamentos de uma moeda

x	$P(X = x)$ sob H_0	$P(X = x)$ sob H_1
0	1/32 → 3,125%	1/3125 → 0,032%
1	5/32 → 15,625%	20/3125 → 0,640%
2	10/32 → 31,250%	160/3125 → 5,120%
3	10/32 → 31,250%	640/3125 → 20,480%
4	5/32 → 15,625%	1280/3125 → 40,960%
5	1/32 → 3,125%	1024/3125 → 32,768%
Total	**1 → 100%**	**1 → 100%**

Para poder aceitar ou rejeitar H_0 e como conseqüência, rejeitar ou aceitar H_1, é necessário estabelecer uma regra de decisão, isto é, é necessário estabelecer para que valores da variável X vai-se rejeitar $H_{0,}$ ou seja, afirmar H_1, e para que valores da variável X, vai-se aceitar H_0, ou seja, nesta situação particular, afirmar H_0.

Desta forma, estabelecendo-se que se vai rejeitar H_0, se a moeda lançada der um número de caras igual a 3, 4 ou 5, pode-se então determinar as probabilidades de tomar as decisões corretas ou as probabilidades dos erros envolvidos. Assim o conjunto de valores que levará a rejeição da hipótese nula será denominado de **região crítica (RC)** e, neste caso, este conjunto é igual a: RC = { 3, 4, 5 }

A faixa restante de valores da variável é denominada de **região de aceitação (RA)** e, neste caso, este conjunto vale: RA = { 0, 1, 2 }

Evidentemente esta regra como qualquer outra permitirá decidir sob a H_0, mas estará sujeita a erro. Está se tomando a decisão de aceitar ou rejeitar H_0 com base no número X de caras obtidas em 5 lançamentos, que é apenas uma amostra, muito pequena, do número infinito de lançamentos possíveis.

Com base em resultados amostrais, não é possível tomar decisões definitivamente corretas. Entretanto, pode-se calcular a probabilidade da decisão estar errada. Neste caso foi decidido rejeitar H_0

se X = "número de caras" assumir um dos valores do conjunto RC. No entanto, tais valores podem ocorrer sob H_0, isto é, tais valores podem ocorrer quando se lança a moeda M_1, conforme tabela. Então se H_0 for rejeitada porque X assumiu o valor 3, 4 ou 5, pode-se estar cometendo um erro. A probabilidade deste erro é igual a probabilidade de ocorrência destes valores sob H_0, isto é, quando a moeda M_1 é lançada, que é conforme tabela igual a:

$$10/32 + 5/32 + 1/32 = 16/32 = 50\%$$

Lembrando que rejeitar H_0 é apenas uma das duas situações possíveis num teste de hipóteses, tem-se que se X assumir um valor do conjunto RA se aceitará H_o. Mas tais valores podem ocorrer sob H_1, isto é, quando a moeda M_2 é lançada. Então se H_o for aceita porque X assumiu um dos valores: 1, 2 ou 3, pode-se estar cometendo um outro tipo de erro, cuja probabilidade é igual a da ocorrência destes valores sob H_1 que é de: 1/3125 + 20/3125 + 160/3125 = 181/3125 = 5,79%

A probabilidade de que a variável (número de caras) assuma um valor do conjunto RC é denominada de **nível de significância do teste**. O nível de significância do teste é, na realidade, a probabilidade de se rejeitar a hipótese nula, quando ela é verdadeira, sendo então a probabilidade de se cometer um erro. Como este é apenas um dos dois tipos de erro possível de ser cometido num teste de hipóteses, ele é denominado de **erro do tipo I.** O outro tipo de erro possível de ser cometido é aceitar H_0 quando ela é falsa e é denominado de **erro do tipo II**. Em resumo pode-se ter as seguintes situações em um teste de hipóteses:

Tabela 02 - Possibilidades envolvidas em um teste de hipóteses

Realidade \ Decisão	Aceitar H_0	Rejeitar H_0
H_0 é verdadeira	**Decisão correta** $1 - \alpha$ = P(Aceitar H_0 / H_0 é V) = P(H_0 / H_0)	**Erro do Tipo I** α = P(Erro do tipo I) = P(Rejeitar H_0 / H_0 é V) = Nível de significância do teste = P(H_1 / H_0)
H_0 é falsa	**Erro do Tipo II** β = P(Erro do tipo II) = = P(Aceitar H_0 / H_0 é falsa) = P(Aceitar H_0 /H_1 é V) = P(H_0 /H_1)	**Decisão correta** **$1 - \beta$** = P(Rejeitar H_0 / H_0 é falsa) = P(H_1 / H_1) = Poder do teste.

Pode-se, agora, determinar as probabilidades de se cometer os erros dos tipos I e II e como conseqüência as probabilidades de se tomar as decisões corretas. A probabilidade de se cometer erro do tipo II, pode ser determinada aqui, porque o teste é do tipo simples, isto é, a hipótese alternativa envolve um único valor (neste caso p = 80%). Geralmente, a hipótese alternativa é do tipo composto (p < 80% ou p > 80% ou ainda p ≠ 80%), e então a determinação do erro do tipo II só poderá ser feita mediante suposições à respeito dos valores que ela pode assumir. Existirão, na realidade, infinitas opções para o erro do tipo II. Para este caso, tem-se:

α = nível de significância do teste = P(Erro do tipo I) = P(rejeitar H_0 / H_0 é verdadeira) = P($x \in$ RC / p = 50%) = P($x \in \{ 3, 4, 5 \}$/ p = 50%) = 10/32 + 5/32 + 1/32 = 16/32 = 50%

1 - α = P(Decisão correta) = P(Aceitar H_0 / H_0 é verdadeira) = P($x \in$ RA / p = 50%) = P($x \in \{ 0, 1, 2 \}$/ p = 50%) = 1/32 + 5/32 + 10/32 = 16/32 = 50%

β = P(Erro do tipo II) = P(Aceitar H_0 / H_0 é falsa) = P($x \in$ RA / p = 80%) = P($x \in \{ 0, 1, 2 \}$/ p = 80%) = 1/3125 + 20/3125 + 160/3125 = 181/3125 = 5,69%

1 - β = Poder do teste = P(Decisão correta) = P(Rejeitar H_0 / H_0 é falsa) = P($x \in$ RC / p = 80%) = P($x \in \{ 3, 4, 5 \}$/ p = 80%) = 640/3125 + 1280/3125 + 1024/3125 = 2944/3125 = 94,31%

Por estes resultados pode-se verificar, que o erro do tipo II poderia ser aceitável, mas o erro do tipo I não, pois é um valor igual a probabilidade de se decidir corretamente. Neste caso, uma opção para diminuir o erro do tipo I seria mudar a região de rejeição. Se a região crítica escolhida tivesse sido RC = { 5 }, isto é, rejeitar a hipótese nula somente se em 5 lançamentos da moeda fosse obtida 5 caras as probabilidades acima ficariam:

α = nível de significância do teste = P(Erro do tipo I) = P(Rejeitar H_0 / H_0 é verdadeira) = P($x \in$ RC / p = 50%) = P($x \in \{ 5 \}$/ p = 50%) = 1/32 = 3,12%.

1 - α = 1 - P(Erro do tipo I) = P(Aceitar H_0 / H_0 é verdadeira) = P($x \in$ RA / p = 50%) = P($x \in \{ 0, 1, 2, 3, 4 \}$ / p = 50%) = 1/32 + 5/32 + 10/32 + 10/32 + 5/32 = 31/32 = 96, 88%.

β = P(Erro do tipo II) = P(Aceitar H_0 / H_0 é falsa) = P($x \in$ RA / p = 80%) = P($x \in \{ 0, 1, 2, 3, 4\}$/ p = 80%) = 1/3125 + 20/3125 + 160/3125 + 640/3125 + 1280/3125 = 2101/3125 = 67,33%.

1 - β = 1 - P(Erro do tipo II) = P(Rejeitar H_0 / H_0 é falsa) = P($x \in$ RC / p = 80%) = P($x \in \{ 5 \}$/ p = 80%) = 1024/3125 = 32,77% = Poder do teste.

Pode-se ver então que o erro do tipo I diminui sensivelmente, mas em compensação tivemos um aumento substancial do erro do tipo II. Isto sempre vai ocorrer. A única forma de reduzir os dois tipos de erro simultaneamente é pelo aumento do tamanho da amostra. Neste caso, está se considerando uma amostra de apenas 5 lançamentos dos infinitos possíveis. É natural que os erros associados sejam grandes, pois a amostra é muito pequena. Aumentado-se o tamanho da amostra é possível com a mesma região crítica diminuir sensivelmente os dois tipos de erro.

1.6. A DISTRIBUIÇÃO AMOSTRAL

A distribuição amostral é uma distribuição de probabilidade, isto é, é uma distribuição teórica que descreve o comportamento de uma determinada estatística ou estimador. As principais estatísticas utilizadas nos testes de hipóteses possuem modelos conhecidos. Têm-se a distribuição normal, a distribuição t (de Student) a distribuição χ^2 (qui-quadrado), a distribuição F (de Snedkor) como as principais.

1.7. TESTES ESTATÍSTICOS PARAMÉTRICOS

Em termos gerais, uma hipótese é uma conjectura sobre algum fenômeno ou conjunto de fatos. Em estatística inferencial o termo *hipótese* tem um significado bastante especifico. É uma conjectura sobre uma ou mais parâmetros populacionais. O teste de hipóteses paramétrico envolve fazer inferências sobre a natureza da população com base nas observações de uma amostra extraída desta população.

Em outras palavras, testar hipóteses envolve determinar a magnitude da diferença entre um valor observado de uma estatística, por exemplo, a proporção p, e o suposto valor do parâmetro (π) e então decidir se a magnitude da diferença justifica a rejeição da hipótese. O processo segue o esquema da figura 01.

1.8. ETAPAS DO TESTE DE HIPÓTESES

Qualquer teste de hipóteses paramétrico segue os seguintes passos:

1. Formular as hipóteses.

Estabelecer as hipóteses nula e alternativa. A construção de um teste de hipóteses pode ser colocado de forma geral do seguinte modo. Toma-se uma amostra da variável (ou das variáveis) X (no caso) de uma dada população, de onde se tem uma hipótese sobre um determinado parâmetro, por exemplo: θ. Esta hipótese é a hipótese nula ou hipótese de igualdade: H_0: $\theta = \theta_0$

Tendo formulado a hipótese nula é conveniente determinar qual será a hipótese aceita caso a hipótese nula seja rejeitada, isto é, convém explicitar a hipótese alternativa. A hipótese alternativa vai depender de cada situação mas de forma geral tem-se:

H_1: $\theta = \theta_2$ (hipótese simples), ou então o que é mais comum, hipóteses compostas:

H_1: $\theta > \theta_0$ (teste unilateral ou unicaudal à direita)

$\theta < \theta_0$ (teste unilateral ou unicaudal à esquerda)

$\theta \neq \theta_0$ (teste bilateral ou bicaudal) as hipóteses são do tipo composto.

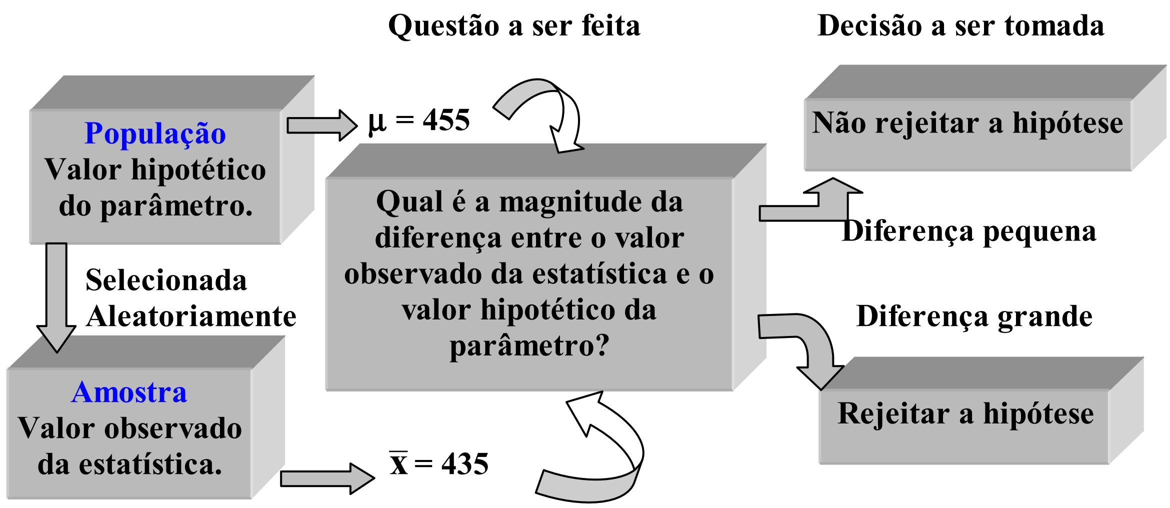


Figura 01 - A lógica do teste de hipóteses

2. Estabelecer a estatística (estimador) a ser utilizado.

Após fixar as hipóteses é necessário determinar se a diferença entre a estatística amostral e o suposto valor do parâmetro da população é suficiente para rejeitar a hipótese. A estatística utilizada deve ser definida e sua distribuição teórica determinada.

3. Fixar o nível de significância do teste.

Fixar a probabilidade de ser cometer erro do tipo I, isto é, estabelecer o nível de significância do teste. Fixado o erro do tipo I, é possível determinar o valor crítico, que é um valor lido na distribuição amostral da estatística considerada (tabela). Este valor vai separar a região de crítica (de rejeição) da região de aceitação.

4. Calcular a estatística teste (a estimativa).

Através da amostra obtida calcular a estimativa que servirá para aceitar ou rejeitar a hipótese nula. Dependendo do tipo de hipótese alternativa este valor servirá para aceitar ou rejeitar H_0. O procedimento é:

Teste estatístico = (Estatística - Parâmetro) / Erro padrão da Estatística

5. Tomar a decisão.

Se o valor da estatística estiver na região crítica rejeitar H_o, caso contrário, aceitar H_0.

6. Formular a conclusão.

Com base na aceitação ou rejeição da hipótese nula, enunciar qual a decisão a ser tomada na situação do problema.

2. TIPOS DE TESTES PARAMÉTRICOS

Os testes paramétricos podem ser divididos em testes para:

- Uma amostra
- Duas amostras independentes
- Duas amostras emparelhadas (dependentes)
- Várias amostras (Análise de Variância)

2.1. TESTES PARA UMA AMOSTRA

2.1.1. TESTE PARA A MÉDIA DE UMA POPULAÇÃO

(a) σ conhecido

O teste para a média de uma população pode ser executado com qualquer tamanho de amostra se soubermos que a população de onde for extraída a amostra segue uma distribuição normal. Se a distribuição da população não for conhecida então é necessário trabalhar com amostras grandes (pelo menos 30 elementos) para poder garantir a normalidade da média da amostra através do teorema central do limite.

As hipóteses são:

H_0: $\mu = \mu_0$ contra

H_1: $\mu = \mu_1$ ou então, o que é mais comum:

H_1: $\mu > \mu_0$

$\mu < \mu_0$

$\mu \neq \mu_0$

A estatística teste utilizada aqui é a média da amostra: $\overline{X}$. Esta média para ser comparada com o valor tabelado, determinado em função da probabilidade do erro do tipo I, (isto é, o nível de significância do teste), precisa ser primeiramente padronizada. Isto é feito, baseado no seguinte resultado:

Se X é uma variável aleatória normal com média μ e desvio padrão σ, então a variável:

$Z = (X - \mu) / \sigma$

Tem uma distribuição normal com média “0” e desvio padrão “1”. A variável resultante Z se encontra tabelada. Qualquer livro de Estatística traz esta tabela que fornece os valores desta variável, para z variando de -3,9 até 3,9 em intervalos de 0,1 (aproximação decimal), entre -3,9 e -3,0 e entre 3,0 e 3,9, e em intervalos de 0,01 (aproximação centesimal) para os valores entre -3,0 e 3,0.

Para $\overline{X}$ sabe-se que $\mu_{\overline{X}} = \mu$ (média das médias) que $\sigma_{\overline{X}} = \sigma/\sqrt{n}$ (erro padrão da média), então o valor padronizado de $\overline{X}$ será:

$$Z = (\overline{X} - \mu_{\overline{X}}) / \sigma_{\overline{X}} = (\overline{X} - \mu) / \sigma/\sqrt{n}$$

Supondo-se fixado um nível de significância de α = P(Erro do Tipo I), verifica-se na tabela qual o valor de z_α (no teste unilateral) ou $z_{\alpha/2}$ (teste bilateral). Rejeita-se H_0 (hipótese nula) se o valor de z calculado na expressão acima for:

(i) Maior do que z_α (no teste unilateral à direita);

(ii) Menor do $-z_\alpha$ (no teste unilateral à esquerda) e

(iii) Maior que $z_{\alpha/2}$ ou menor que $-z_{\alpha/2}$ (no teste bilateral).

Tabela 03 - Valores de z para alguns níveis de significância

	α = Nível de significância = P(Erro do Tipo I)		
	10%	**5%**	**1%**
Teste bilateral	1,64	1,96	2,57
Teste unilateral	1,28	1,64	2,33

Exemplo

A associação dos proprietários de indústrias metalúrgicas está preocupada com o tempo perdido em acidentes de trabalho, cuja média, nos últimos tempos, tem sido da ordem de 60 hora /homens por ano com desvio padrão de 20 horas/homem. Tentou-se um programa de prevenção de acidentes e, após o mesmo, tomou-se uma amostra de 9 indústrias e mediu-se o número de horas/homem perdidas por acidente, que foi de 50 horas. Você diria, ao nível de 5%, que há evidência de melhoria?

Solução

As hipóteses a serem testadas são:

H_0: μ = 60 hora/homens

H_1: $\mu < 60$ hora/homens

A evidência amostral para sugerir que a média baixou é dada através da amostra de n = 9 (elementos) que forneceu $\bar{x}$ = 50 horas/homens. Vamos testar se esta diferença de 10 horas/homens é ou não significativa ao nível de 5%. Para isto é necessário padronizar o resultado amostral.

$$Z = (\bar{X} - \mu_{\bar{X}}) / \sigma_{\bar{X}} = (\bar{X} - \mu) / \sigma/\sqrt{n} = (50 - 60) / 20/\sqrt{9} = -1{,}50$$

Para saber se este valor (-1,50) é pouco provável é necessário compará-lo com o valor crítico $-z_\alpha$ (pois se trata de um teste unilateral à esquerda), que neste caso vale -1,64, já que o nível de significância foi fixado em 5%. Vê-se portanto que o valor amostral não é inferior ao valor crítico, não estando portanto na região de rejeição. Isto quer dizer que a diferença apresentada na amostra não é suficientemente grande para provar que a campanha de prevenção deu resultado. Então a conclusão é:

"Não é possível ao nível de 5% de significância afirmar que a campanha deu resultado, isto é, rejeitar H_0. "

Convém lembrar que o fato de não rejeitar a hipótese nula, não autoriza a fazer afirmações a respeito da veracidade dela. Ou seja, não se provou H_0, pois no momento que se aceita a hipótese nula, o risco envolvido é o do Tipo II, e este neste caso não está fixado (controlado). O teste de hipóteses é feito para rejeitar a hipótese nula e sua força está na rejeição. Assim quando se rejeita se prova algo, mas quando se aceita, nada se pode afirmar.

(b) σ desconhecido

A distribuição t de Student

Quando o desvio padrão populacional (σ) é desconhecido é necessário estimá-lo através do desvio padrão da amostra (s). Mas ao substituir o desvio padrão da população na expressão:

$$Z = (\bar{X} - \mu_{\bar{X}}) / \sigma_{\bar{X}} = (\bar{X} - \mu) / \sigma/\sqrt{n}$$

não teremos mais uma distribuição normal.

De fato, conforme demonstrado por W. S. Gosset (Student) a distribuição da variável:

$$(\bar{X} - \mu_{\bar{X}}) / \hat{\sigma}_{\bar{X}} = (\bar{X} - \mu) / s/\sqrt{n}$$

Não é mais normal padrão. Ao substituir σ por s na expressão teremos uma distribuição parecida com a normal, isto é, simétrica em torno de zero, porém com uma variabilidade maior. Desta forma a distribuição "t" é mais baixa no centro do que a normal padrão, mas mais alta nas caudas.

Assim:

$(\overline{X} - \mu_{\overline{X}}) / \hat{\sigma}_{\overline{X}} = (\overline{X} - \mu) / s/\sqrt{n} = t_{n-1}$, onde “n - 1” indica a distribuição “t” considerada, pois cada tamanho de amostra produz uma distribuição de Student diferente.

A distribuição t de Student encontra-se tabelada em função de n = tamanho da amostra ou então em função de n - 1 denominado de graus de liberdade da distribuição. Neste caso cada linha de uma tabela se refere a uma distribuição particular e cada coluna da tabela a um determinado nível de significância. Conforme a tabela o nível de significância poderá ser unilateral ou bilateral. Em todo caso é necessário sempre ler no cabeçalho ou no rodapé da tabela as explicações sobre como ela está estruturada.

Desta forma a diferença entre o teste para a média de uma população com σ conhecido e um com σ desconhecido é que é necessário trocar a distribuição normal padrão pela distribuição “t “ de Student.

Exemplo

O tempo médio, por operário, para executar uma tarefa, tem sido 100 minutos. Introduziu-se uma modificação para diminuir este tempo, e, após certo período, sorteou-se uma amostra de 16 operários, medindo-se o tempo de execução gasto por cada um. O tempo médio da amostra foi 85 minutos com desvio padrão de 12 minutos. Este resultado evidencia uma melhora no tempo gasto para realizar a tarefa? Apresente as conclusões aos níveis de 5% e 1% de significância e diga quais as suposições teóricas necessárias que devem ser feitas para resolver o problema.

Solução

A suposição teórica necessária é admitir que a distribuição da população de onde foi extraída a amostra segue uma normal pois n < 30.

H_0: $\mu = 100$

H_1: $\mu < 100$

Considerando, então, um teste unilateral à esquerda e tendo $\alpha = 5\%$ ($\alpha = 1\%$) tem-se que a região de rejeição é constituída por RC = [-∞, -1,753].(RC = [-∞, -2,602])

O valor de teste é:

$$t_{15} = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{85 - 100}{12/4} = -5$$

Como este valor pertence as duas regiões críticas, pode-se rejeitar a hipótese nula, aos níveis de 5% e 1% de significância, isto é, neste caso, pode-se afirmar que a modificação diminuiu o tempo de execução da tarefa.

2.1.2. TESTE PARA A PROPORÇÃO

O teste para a proporção populacional é normalmente baseado na seguinte suposição: tem-se uma população e tem-se uma hipótese sobre a proporção π de elementos da população que possuem uma determinada característica. Esta proporção é supostamente igual a um determinado valor π_0. Assim a hipótese nula é:

$H_0 : \pi = \pi_0$

O problema fornece informações sobre a alternativa, que pode ser uma das seguintes:

$H_1 : \pi \neq \pi_0$

$H_1 : \pi > \pi_0$

$H_1 : \pi < \pi_0$

A estatística teste a ser utilizada é a proporção amostral "P", que para amostras grandes (n > 50) tem uma distribuição aproximadamente normal com média:

$\mu_P = \pi$, e desvio padrão

$$\sigma_P = \sqrt{\frac{\pi(1-\pi)}{n}}$$

<u>Exemplo</u>

As condições de mortalidade de uma região são tais que a proporção de nascidos que sobrevivem até 60 anos é de 0,60. Testar esta hipótese ao nível de 5% de significância se em 1000 nascimentos amostrados aleatoriamente, verificou-se 530 sobreviventes até os 60 anos.

<u>Solução</u>

H_1: $\pi = 0{,}60$

H_0: $\pi \neq 0{,}60$

Considerando, então, um teste bilateral e tendo α = 5% tem-se que a região de aceitação é constituída pelo intervalo RA = [-1,96, 196].

O valor de teste é:

$$z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{0{,}53 - 0{,}60}{\sqrt{\frac{060(1-0{,}60)}{1000}}} = -4{,}52.$$

Como este valor não pertence a região de aceitação, pode-se rejeitar a hipótese nula, ao nível de 5% de significância, isto é, neste caso, pode-se afirmar que a taxa dos que sobrevivem até os 60 anos é menor do que 60%. Neste caso, também poderia ser realizado um teste unilateral à esquerda. Este teste também rejeitaria a hipótese nula, pois para ele o valor crítico z_α = -1645.

2.1.3. TESTE PARA A VARIÂNCIA

Para aplicar o teste para a variância é necessário supor a normalidade da população de onde será extraída a amostra.

As hipóteses são:

H_0: $\sigma^2 = \sigma_0^2$ contra

H_1: $\sigma^2 \neq \sigma_0^2$

$\sigma^2 > \sigma_0^2$

$\sigma^2 < \sigma_0^2$

A estatística teste é $\frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{n-1}$

Quer dizer o quociente acima tem uma distribuição qui-quadrado com "n-1" graus de liberdade. A qui-quadrado é uma distribuição assimétrica positiva que varia de zero a mais infinito. Esta distribuição é tabelada também em função do número de graus de liberdade, isto é, cada grau de liberdade (n -1) representa uma distribuição diferente. As colunas das tabelas representam diferentes níveis de significância, isto é, área sob a curva acima do valor tabelado.

Em função do tipo de hipótese alternativa define-se a região de rejeição. No primeiro caso tem-se uma região de rejeição do tipo bilateral. Logo, fixado um nível de significância "α", a região crítica será:

RC = $[0, \chi_1^2] \cup [\chi_2^2, \infty)$

Desta forma, aceita-se a hipótese nula se a estatística teste, acima, pertencer ao intervalo $[\chi_1^2, \chi_2^2]$.

Exemplo

Uma das maneiras de controlar a qualidade de um produto é controlar a sua variabilidade. Uma máquina de empacotar café está regulada para encher os pacotes com desvio padrão de 10 g e média de 500g e onde o peso de cada pacote distribuiu-se normalmente. Colhida uma amostra de n = 16, observou-se uma variância de 169 g^2. É possível afirmar com este resultado que a máquina está desregulada quanto a variabilidade, supondo uma significância de 5%?

Solução

H_0: $\sigma^2 = 100$ contra

H_1: $\sigma^2 \neq 100$

$\chi^2_c = (15.169)/100 = 25{,}35$.

Como $\alpha = 5\%$ a região de aceitação é a região compreendida entre os valores: $[\chi^2_{97,5\%}, \chi^2_{2,5\%}] = [6{,}26, 27{,}49]$. Como o valor calculado pertence a esta região, aceita-se H_0, isto é, com esta amostra não é possível afirmar que a máquina está desregulada, ao nível de 5% de significância.

Supõem-se a existência de duas populações. Uma população X com média μ_X e desvio padrão σ_X e uma população Y com média μ_Y e desvio padrão σ_Y. Da população X é extraída uma amostra de tamanho "n" com média $\bar{X}$ e da população Y é extraída uma amostra de tamanho "m" com média $\bar{Y}$. Define-se a variável $\bar{D}$ como sendo a diferença entre as duas médias amostrais. Assim $\bar{D} = \bar{X} - \bar{Y}$ e tem-se:

$$\mu_{\bar{D}} = E(\bar{D}) = E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_X - \mu_Y$$

$$\sigma_{\bar{D}} = V(\bar{D}) = V(\bar{X} - \bar{Y}) = V(\bar{X}) + V(\bar{Y}) = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}.$$

2.2. TESTES PARA DUAS AMOSTRAS INDEPENDENTES

Neste tipo de teste são retiradas duas amostras de forma independente, isto é, as medidas são obtidas em unidades amostrais diferentes.

2.2.1. TESTE PARA A IGUALDADE ENTRE AS VARIÂNCIAS DE DUAS POPULAÇÕES

Supõem-se a existência de duas populações. Uma população X com média μ_X e desvio padrão σ_X e uma população Y com média μ_Y e desvio padrão σ_Y. Da população X é extraída uma amostra de tamanho "n" com média $\bar{X}$ e variância S_X^2 e da população Y é extraída uma amostra de tamanho "m" com média $\bar{Y}$ e variância S_Y^2.

As hipóteses são:

H_0: $\sigma_X^2 = \sigma_Y^2 = \sigma^2$

H_1: $\sigma_X^2 \neq \sigma_Y^2$

Nestas condições sabe-se que: $\dfrac{(n-1)S_X^2}{\sigma_X^2} : \chi^2_{n-1}$ e $\dfrac{(m-1)S_Y^2}{\sigma_Y^2} : \chi^2_{m-1}$

Sob a hipótese de H_0 ser verdadeira (isto é, $\sigma_X^2 = \sigma_Y^2$) tem-se:

$Q = \dfrac{S_X^2}{S_Y^2} = \dfrac{\dfrac{\sigma_X^2\chi^2_{n-1}}{n-1}}{\dfrac{\sigma_Y^2\chi^2_{m-1}}{m-1}} = F(n-1, m-1)$, isto é, o quociente entre as variâncias amostrais possui uma distribuição F (de Snedekor) com "n-1" graus de liberdade no numerador e "m - 1" graus de liberdade no denominador.

Como a distribuição F depende de dois parâmetros ν_1 e ν_2, uma tabela tridimensional será necessária para computar os valores de F correspondentes a diferentes probabilidades e valores de ν_1 e ν_2. Como conseqüência, somente os pontos da cauda à direita de 5% e 1% de área são tabelados, correspondendo a vários valores de ν_1 e ν_2, isto é, encontram-se tabelados os valores P(F > f) = 0,01 e P(F > f) = 0,05. Para poder se obter valores bilaterais da distribuição F é necessário usar a propriedade que se F é tal que tem uma distribuição com ν_1 e ν_2 graus de liberdade, então F' = 1 / F tem distribuição F' com ν_2 e ν_1 graus de liberdade. Assim a probabilidade de que F < f pode ser calculada por:

P(F < f) = P(1 / F > 1 / f) = P(F' > 1 / f)

Lembrando que só são fornecidos valores com as significâncias de 1% e 5%. Outro valor entre estes dois poderá ser obtido aproximadamente por interpolação.

Assim por exemplo dados $\nu_1 = 5$ (graus de liberdade do numerador) e $\nu_2 = 8$ (graus de liberdade do denominador), o valor de f de F(5, 8) tal que P(F > f) = 5% é f = 3,69. Então o valor f' de F(5, 8) tal que P(F < f') = 5% é dado por: 1 / F(8, 5) = 1 / 4,82 = 0,21.

Fixado um nível de significância α a região crítica RC é encontrada através de dois valores F_1 e F_2 da distribuição F tais que:

$P(F \in RC) = P(F < F_1 \text{ ou } F > F_2) = \alpha$, onde F_1 e F_2 são encontrados na tabela de modo a satisfazer a igualdade: $P(F < F_1) = P(F > F_2) = \alpha/2$.

Exemplo: (BUS81 - pg. 275)

Quer se verificar se duas máquinas produzem peças com a mesma homogeneidade quanto à resistência à tensão. Para tal, sorteiam-se duas amostras de 6 peças de cada uma das máquinas e observa-se as resistências. Os resultados estão na tabela.

Máquina X	145	127	136	142	141	137
Máquina Y	143	128	132	138	142	132

Solução:

Como n = m = 6, tem-se que:

$$Q = \frac{S_X^2}{S_Y^2} = F(5, 5) = 5,05$$

A região crítica RC será: RC = (0; 1/5,05) U (5,05; ∞) = (0; 0,20) U (5,05; ∞)

As amostras fornecem:

$S_X^2 = 40$ e $S_Y^2 = 37$, portanto a distribuição do quociente Q calculado será:

$$Q_c = \frac{S_X^2}{S_Y^2} = 40 / 37 = 1,08.$$

Por estes resultados não é possível rejeitar a hipótese de igualdade entre as variâncias a um nível de significância de 10%. (Como o teste é bilateral, ele envolve uma área de 5% em cada cauda da distribuição, logo a significância total é de 10%).

2.2.2. TESTE PARA A DIFERENÇA ENTRE DUAS MÉDIAS POPULACIONAIS

(a) Supondo as variâncias (σ_X^2 e σ_Y^2) conhecidas

As hipóteses são:

H_0: $\mu_X - \mu_Y = \Delta$ contra

H_1: $\mu_X - \mu_Y \neq \Delta$ ou

$\mu_X - \mu_Y > \Delta$ ou ainda

$\mu_X - \mu_Y < \Delta$

Se $\Delta = 0$, então $\mu_X - \mu_Y = 0$, isto é, $\mu_X = \mu_Y$.

Como as variâncias são conhecidas, tem-se então que, para n, m $\geq$ 30 ou para amostras extraídas de populações normais, que a variável $\bar{D} = \bar{X} - \bar{Y}$ terá uma distribuição aproximadamente normal com média $E(\bar{D}) = \mu_X - \mu_Y$ e variância $V(\bar{D}) = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}$.

A variável teste será, então:

$$z = \frac{\bar{X} - \bar{Y} - \Delta}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

Assim fixando o nível de significância "α", a hipótese nula será rejeitada se:

$|z| > z_{\alpha/2}$ no teste bilateral;

$z > z_\alpha$, no teste unilateral à direita e

$z < z_\alpha$ no teste unilateral à esquerda.

Exemplo:

Um fabricante produz dois tipos de pneus. Para o pneu do tipo A o desvio padrão é de 2500 km e para o pneu do tipo B é de 3000 km. Uma cia de táxis testou 50 pneus do tipo A e 40 do tipo B, obtendo 24000 km de média para o "A" e 26000 para o tipo "B". Adotando α = 4% testar a hipótese de que a duração média dos dois tipos é a mesma.

Solução:

As hipóteses são:

H_0: $\mu_A - \mu_B = 0$ ($\mu_A = \mu_B$) contra

H_1: $\mu_A - \mu_B \neq 0$ ($\mu_A \neq \mu_B$)

Como α = 4%, então $z_{\alpha/2} = -2{,}05$.

O valor da variável teste será:

$$z = \frac{24000 - 26000}{\sqrt{\frac{2500^2}{50} + \frac{3000^2}{40}}} = -3,38$$

Portanto, rejeita-se a hipótese de igualdade entre as durações médias dos dois tipos de pneus. Com base nestas amostras, pode-se afirmar, ao nível de 4% de significância, que os dois tipos de pneus diferem quanto a durabilidade média.

(b) Variâncias σ_X^2 e σ_Y^2 desconhecidas, mas supostamente iguais

Vamos supor que as duas populações tenham a mesma variância $\sigma^2 = \sigma_X^2 = \sigma_Y^2$, porém desconhecidas.

As hipóteses são:

H_0: $\mu_X - \mu_Y = \Delta$ contra

H_1: $\mu_X - \mu_Y \neq \Delta$ ou

$\mu_X - \mu_Y > \Delta$ ou ainda

$\mu_X - \mu_Y < \Delta$

A variável teste anterior, para esta situação, será:

$Z = \dfrac{\bar{X} - \bar{Y} - \Delta}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$, mas neste caso $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ (por suposição), então:

$Z = \dfrac{\bar{X} - \bar{Y} - \Delta}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} = \dfrac{\bar{X} - \bar{Y} - \Delta}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}} = \dfrac{\bar{X} - \bar{Y} - \Delta}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}}$, como o valor σ^2 não é conhecido, deverá ser substituído por um estimador não-tendencioso. Como S_X^2 e S_Y^2 são estimadores não tendenciosos do mesmo parâmetro σ^2, então, a média ponderada:

$S^2 = \dfrac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$, também será um estimador não-tendencioso de σ^2.

Logo a expressão acima poderá ser escrita como:

$\dfrac{\bar{X} - \bar{Y} - \Delta}{S\sqrt{\frac{1}{n} + \frac{1}{m}}}$, que terá uma distribuição não mais normal mas sim "t" com "n + m – 2" graus de liberdade, desde que n, m sejam maiores ou iguais a 30, ou então que as amostras tenham sido extraídas de populações que tenham distribuições normais.

Desta forma, a expressão para testar a diferença entre duas médias populacionais, nesta situação será:

$$t_c = t_{n+m-2} = \frac{\overline{X} - \overline{Y} - \Delta}{S\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

Assim fixando o nível de significância "α", a hipótese nula será rejeitada se:

$|t_c| > t_{\alpha/2}$ no teste bilateral;

$t_c > t_\alpha$, no teste unilateral à direita e

$t_c < t_\alpha$ no teste unilateral à esquerda.

Exemplo:

As resistências de dois tipos de concreto foram medidas, mostrando os resultados da tabela. Fixado um nível de significância de 5%, existe evidência de que o concreto do tipo A seja mais resistente do que o concreto do tipo B.

Tipo A	54	55	58	51	57
Tipo B	50	54	56	52	53

Solução:

Antes de mais nada vamos testar se as duas populações possuem a mesma variância. Para tanto aplica-se o teste de igualdade de variâncias, utilizando as amostras acima e uma significância de 5%.

Tem-se: Graus de liberdade: 4 (numerador), 4 (denominador)

$F = 7{,}5/5{,}0 = 1{,}50.$

$F_{2,5\%} = 0{,}10$

$F_{97,5\%} = 9{,}60$

Significância do resultado obtido: 35,20%.

Neste caso, não é possível afirmar que as variâncias populacionais são diferentes.

As hipóteses são:

H_0: $\mu_A - \mu_B = 0$ ($\mu_A = \mu_B$) contra

H_1: $\mu_A - \mu_B > 0$ ($\mu_A > \mu_B$)

Os dados obtidos da tabela são:

$\overline{X} = 55{,}0$ e $\overline{Y} = 53{,}0$

$S_X^2 = 7{,}50$ e $S_Y^2 = 5{,}0$, então $S^2 = \dfrac{(n-1)S_X^2+(m-1)S_Y^2}{n+m-2} = \dfrac{(5-1).7{,}5+(5-1).5{,}0}{5+5-2} = 6{,}25$.

O valor da variável teste será:

$$t_c = \frac{55-53}{2{,}50.\sqrt{\frac{1}{5}+\frac{1}{5}}} = 1{,}265$$

Como $\alpha = 5\%$, e o grau de liberdade n - m - 2 = 10 - 2 = 8, então o valor de "t" tabelado será: 1,86.

Neste caso, com estas amostras não é possível afirmar que o concreto do tipo A seja mais resistente do que o concreto do tipo B.

(c) Variâncias σ_X^2 e σ_Y^2 desconhecidas e supostamente desiguais

As hipóteses são:

H_0: $\mu_X - \mu_Y = \Delta$ contra

H_1: $\mu_X - \mu_Y \neq \Delta$ ou

$\mu_X - \mu_Y > \Delta$ ou ainda

$\mu_X - \mu_Y < \Delta$

Como as variâncias são desconhecidas é necessária estimá-las através das variâncias amostrais S_X^2 e S_Y^2. Neste caso, ao se substituir as variâncias populacionais pelas amostrais na expressão:

$\dfrac{\overline{X}-\overline{Y}-\Delta}{\sqrt{\dfrac{\sigma_X^2}{n}+\dfrac{\sigma_Y^2}{m}}}$ não se terá mais uma distribuição normal, mas sim uma distribuição "t" com o grau de liberdade fornecido pela seguinte expressão:

$$\nu = \frac{\left(\dfrac{S_X^2}{n}+\dfrac{S_Y^2}{m}\right)^2}{\dfrac{\left(\dfrac{S_X^2}{n}\right)^2}{n-1}+\dfrac{\left(\dfrac{S_Y^2}{m}\right)^2}{m-1}}$$

desde que n, m sejam maiores ou iguais a 30, ou então que as amostras tenham sido extraídas de populações que tenham distribuições normais.

Assim fixando o nível de significância “α“, a hipótese nula será rejeitada se:

$|t_c| > t\alpha_{/2}$ no teste bilateral;

$t_c > t\alpha$, no teste unilateral à direita e

$t < t\alpha$ no teste unilateral à esquerda, onde $t = \dfrac{\overline{X} - \overline{Y} - \Delta}{\sqrt{\dfrac{S_X^2}{n} + \dfrac{S_Y^2}{m}}}$

Exemplo:

As resistências de dois tipos de concreto foram medidas, mostrando os resultados da tabela. Fixado um nível de significância de 5%, existe evidências de que o concreto do tipo A seja mais resistente do que o concreto do tipo B.

Tipo A	54	55	58	50	61
Tipo B	51	54	55	52	53

Solução:

Antes de mais nada vamos testar se as duas populações possuem a mesma variância. Para tanto aplica-se o teste de igualdade de variâncias, utilizando as amostras acima e uma significância de 10%.

Tem-se: Graus de liberdade: 4 (numerador), 4 (denominador).

F = 17,3/2,5 = 6,92.

Significância do resultado obtido: 4,38%.

F crítico: 6,39.

Neste caso, é possível afirmar que as variâncias populacionais **são** diferentes.

As hipóteses são:

H_0: $\mu_A - \mu_B = 0$ ($\mu_A = \mu_B$) contra

H_1: $\mu_A - \mu_B > 0$ ($\mu_A > \mu_B$)

Os dados obtidos da tabela são:

$\overline{X} = 55,6$ e $\overline{Y} = 53,0$

$S_X^2 = 17{,}3$ e $S_Y^2 = 2{,}5$

O valor da variável teste será:

$$t = \frac{55{,}6 - 53{,}0}{\sqrt{\frac{17{,}3}{5} + \frac{2{,}5}{5}}} = 1{,}31$$

Com α = 5%, e o grau de liberdade $\nu = \dfrac{\left(\frac{S_X^2}{n} + \frac{S_Y^2}{m}\right)^2}{\frac{\left(\frac{S_X^2}{n}\right)^2}{n-1} + \frac{\left(\frac{S_Y^2}{m}\right)^2}{m-1}} = \dfrac{\left(\frac{17{,}3}{5} + \frac{2{,}5}{5}\right)^2}{\frac{\left(\frac{17{,}3}{5}\right)^2}{4} + \frac{\left(\frac{2{,}5}{5}\right)^2}{4}} = \dfrac{6{,}25}{0{,}8125} = 5{,}48 \cong$ 5,

então o valor de "t" tabelado será: 2,57.

Neste caso, com estas amostras não é possível afirmar que o concreto do tipo A seja mais resistente do que o concreto do tipo B.

2.3. DUAS AMOSTRAS RELACIONADAS (DEPENDENTES)

Quando se compara as médias de duas populações, pode ocorrer uma diferença significativa por causa de fatores externos não-controláveis. Um modo de contornar este problema é coletar observações aos pares, de modo que os dois elementos de cada par sejam homogêneos em todos os sentidos, exceto naquele que se quer comparar.

Por exemplo, para testar dois métodos de ensino A e B, pode-se usar pares de gêmeos, sendo que um recebe o método de ensino A e o outro o método de ensino B. Este procedimento controla a maioria dos fatores externos que afetam a aprendizagem e se houver diferença deve-se realmente ao método.

Outra forma é fazer as observações das duas amostras no mesmo indivíduo. Por exemplo, medindo uma característica do indivíduo antes e depois dele ser submetido a um tratamento.

A exemplo da comparação de duas médias com amostras independentes, neste caso, tem-se duas amostras: $X_1, X_2, ..., X_n$ e $Y_1, Y_2, ..., Y_n$, só que agora as observações estão emparelhadas, isto é, a amostra é formada pelos pares:

$(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$

Define-se a variável D = X - Y.

Como resultado tem-se a amostra: $D_1, D_2, ..., D_n$

Supõem-se que D segue uma $N(\mu_D, \sigma_D)$. Então: $S_D^2 = \frac{1}{n}\sum_{i=1}^{n} D_i = \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) = \overline{X} - \overline{Y}$

Terá uma distribuição: $N(\mu_D, \frac{\sigma_D}{\sqrt{n}})$. Definindo:

$$S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(D_i - \overline{D}\right)^2 = \frac{\sum_{i=1}^{n} D_i - n\overline{D}^2}{n-1}$$, tem-se que a estatística:

$$t = \frac{\overline{D} - \mu_D}{S_D/\sqrt{n}}$$, tem uma distribuição “t” com “n - 1” graus de liberdade.

Exemplo:

Cinco operadores de máquinas são treinados em duas máquinas de diferentes fabricantes, para verificar qual delas apresentava maior facilidade de aprendizagem. Mediu-se o tempo que cada um dos operadores gastou na realização de uma mesma tarefa com cada um dos dois tipos de máquinas. Os resultados estão na tabela ao lado. Ao nível de 10% é possível afirmar que a tarefa realizada na máquina X demora mais do que na máquina Y?

Solução:

Operador	Fabricante 1	Fabricante 2
1	80	75
2	72	70
3	65	60
4	78	72
5	85	78

As hipóteses são:

H_0: $\mu_X - \mu_Y = 0$ ($\mu_X = \mu_Y$) contra

H_1: $\mu_X - \mu_Y > 0$ ($\mu_X > \mu_Y$)

Pela tabela vê-se que:

d_i: 5, 2, 5, 6 e 7

Logo: $\overline{d} = 5$ e $S_D = 1{,}8708$, logo t = 5,98.

Como $\alpha = 10\%$, então $t\alpha = 1{,}54$, pois o número de graus de liberdade é n - 1 = 4.

Portanto, rejeita-se a hipótese nula, isto é, a 10% de significância pode-se afirmar que com a máquina X se demora mais do que com a máquina Y.

2.3.1. Teste para a diferença entre duas proporções

As hipóteses são:

H_0: $\pi_1 - \pi_2 = \pi$ contra

H_1: $\pi_1 - \pi_2 \neq \pi$ ou

$\pi_1 - \pi_2 > \pi$ ou ainda

$\pi_1 - \pi_2 < \pi$

Se $\pi = 0$, então $\pi_1 - \pi_2 = 0$, isto é, $\pi_1 = \pi_2$.

Extraídas uma amostra de cada uma das duas populações a variável $P_1 - P_2$ terá uma distribuição aproximadamente normal com média $E(P_1 - P_2) = \pi_1 - \pi_2$ e variância $\sigma^2_{P_1-P_2} = \frac{\pi_1(1-\pi_1)}{n} + \frac{\pi_2(1-\pi_2)}{m}$, desde que $nP_1 > 5$ e $mP_2 > 5$.

A variável teste será, então: $z = \dfrac{P_1 - P_2 - \pi}{\sqrt{\dfrac{\pi_1(1-\pi_1)}{n} + \dfrac{\pi_2(1-\pi_2)}{m}}}$

Como os valores de π_1 e π_2 não são conhecidos, deve-se utilizar suas estimativas P_1 e P_2. Desta forma, o valor de z será:

$$z = \frac{P_1 - P_2 - \pi}{\sqrt{\dfrac{P_1(1-P_1)}{n} + \dfrac{P_2(1-P_2)}{m}}}$$

Assim fixando o nível de significância "α", a hipótese nula será rejeitada se:

$|z| > z_{\alpha/2}$ no teste bilateral;

$z > z_\alpha$, no teste unilateral à direita e

$z < z_\alpha$ no teste unilateral à esquerda.

Exemplo:

Em uma pesquisa de opinião, 32 dentre 80 homens declararam apreciar certa revista, acontecendo o mesmo com 26 dentre 50 mulheres. Ao nível de 5% de significância os homens e as mulheres apreciam igualmente a revista?

Solução:

As hipóteses são:

H_0: $\pi_1 - \pi_2 = 0$ ($\pi_1 = \pi_2$) contra

H_1: $\pi_1 - \pi_2 \neq 0$ ($\pi_1 \neq \pi_2$)

Tem-se que $P_1 = 32 / 80 = 0{,}40$ e $P_2 = 26 / 50 = 52\%$

O valor da variável teste será:

$$z = \frac{0{,}40 - 0{,}52}{\sqrt{\frac{0{,}40.0{,}60}{80} + \frac{0{,}52.0{,}48}{50}}} = -1{,}34$$

Como α = 5%, então $z\alpha_{/2}$ = -1,96.

Portanto, aceita-se a hipótese de igualdade entre as preferências de homens e mulheres, isto é, a este nível de significância não é possível afirmar que exista diferença entre as preferências de homens e mulheres quanto à revista.

3. REFERÊNCIAS

BUSSAB, Wilton O, MORETTIN, Pedro A. *Estatística Básica*. São Paulo: Atual, 1986. 3. ed.

DOWNING, Douglas, CLARK, Jeff. *Statistics the Easy Way*. Hauppauge (NY): Barron's Educational Series, 1989.

HINKLE, Dennis E., WILLIAM, Wiersma, JURS, Stephen G. *Applied Statistics for the Behavioral Sciences*. Boston(MA): Houghton Mifflin, 1988.

HOFFMAN, Rodolfo. *Estatística para Economistas*. São Paulo: Pioneira, 1980.

NETO, Pedro Luiz de Oliveira Costa. *Estatística*. São Paulo: Edgard Blücher, 1977.

MASON, Robert D., DOUGLAS, Lind A. *Statistical Techniques in Business And Economics*. Boston: Irwin, 1990.

MEYER, Paul L. *Probabilidade: aplicações à Estatística.* Rio de Janeiro: Livros Técnicos e Científicos, 1978

WLKOWITZ, Joan, EWEN, Robert B., COHEN, Jacob. *Introductory Statistics for the Behavioral Sciences.* Orlando(FL): Hartcourt Brace Javanovich, 1982.

WONNACOTT, Ronald J., WONNACOTT, Thomas. *Fundamentos de Estatística*. Rio de Janeiro: Livros Técnicos e Científicos Editora , 1985.