Material Didático

Série

Estatística Básica

Texto IV

Testes de Hipóteses

Enfoque: Exatas

Prof. Lori Viali, Dr.

SUMÁRIO

TESTES DE HIPÓTESES PARAMÉTRICOS

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 a 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 $\rightarrow$ 3,125%	1/3125 $\rightarrow$ 0,032%
1	5/32 $\rightarrow$ 15,625%	20/3125 $\rightarrow$ 0,640%
2	10/32 $\rightarrow$ 31,250%	160/3125 $\rightarrow$ 5,120%
3	10/32 $\rightarrow$ 31,250%	640/3125 $\rightarrow$ 20,480%
4	5/32 $\rightarrow$ 15,625%	1280/3125 $\rightarrow$ 40,960%
5	1/32 $\rightarrow$ 3,125%	1024/3125 $\rightarrow$ 32,768%
Total	**1 $\rightarrow$ 100%**	**1 $\rightarrow$ 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 $p \neq 80\%$), 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.

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

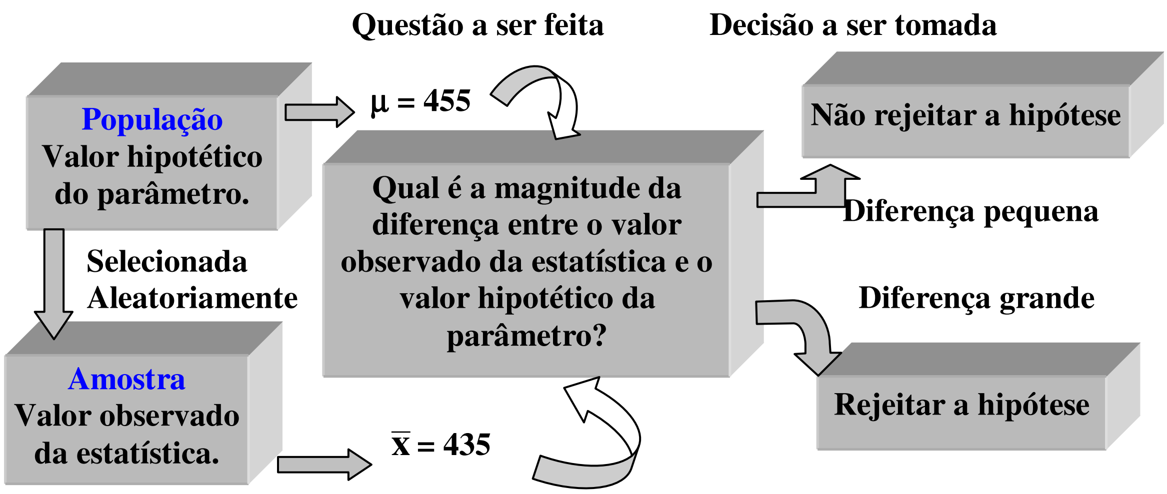

Figura 01 - A lógica do teste de hipóteses

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.

5. 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: $\mu > \mu_0$ ou $\mu < \mu_0$ ou $\mu \neq \mu_0$

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

<u>Exemplo</u>

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?

<u>Solução</u>

As hipóteses a serem testadas são:

H_0: μ = 60 hora/homens

H_1: μ < 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 = $[-\infty, -1,753]$.(RC = $[-\infty, -2,602]$)

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:

$H_1 : \pi \neq \pi0$ ou $H1 : \pi > \pi_0$ ou $H_1 : \pi < \pi_0$

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: $\sigma^2 \neq \sigma_0^2$ ou $\sigma^2 > \sigma_0^2$ ou $\sigma^2 < \sigma_0^2$

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 dos 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, χ_1^2] U [χ_2^2, ∞). Desta forma, aceita-se a hipótese nula se a estatística teste, acima, pertencer ao intervalo [χ_1^2, χ_2^2].

<u>Exemplo</u>

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 distribuí-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%?

<u>Solução</u>

H_0: $\sigma^2 = 100$ contra

H_1: $\sigma^2 \neq 100$

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

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 $\overline{X}$ e da população Y é extraída uma amostra de tamanho "m" com média $\overline{Y}$. Define-se a variável $\overline{D}$ como sendo a diferença entre as duas médias amostrais. Assim $\overline{D} = \overline{X} - \overline{Y}$ e tem-se:

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

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

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 $\overline{X}$ e variância S_X^2 e da população Y é extraída uma amostra de tamanho "m" com média $\overline{Y}$ 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: $\frac{(n-1)S_X^2}{\sigma_X^2} : \chi^2_{n-1}$ e $\frac{(m-1)S_Y^2}{\sigma_Y^2} : \chi^2_{m-1}$

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

$Q = \frac{S_X^2}{S_Y^2} = \frac{\frac{\sigma_X^2 \chi^2_{n-1}}{n-1}}{\frac{\sigma_Y^2 \chi^2_{m-1}}{m-1}} = F(n-1, m-1)$, 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:

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

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:

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

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 ≥ 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 $V(\bar{D}) = \dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}$.

A variável teste será, então:

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

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 $\alpha = 4\%$ 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 $\alpha = 4\%$, 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:

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

$Z = \dfrac{\overline{X} - \overline{Y} - \Delta}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}} = \dfrac{\overline{X} - \overline{Y} - \Delta}{\sqrt{\dfrac{\sigma^2}{n} + \dfrac{\sigma^2}{m}}} = \dfrac{\overline{X} - \overline{Y} - \Delta}{\sigma\sqrt{\dfrac{1}{n} + \dfrac{1}{m}}}$, 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:

$\dfrac{\overline{X} - \overline{Y} - \Delta}{S\sqrt{\dfrac{1}{n} + \dfrac{1}{m}}}$, 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:

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

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

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.

<u>Exemplo:</u>

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?

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

Solução:

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: $\bar{d} = 5$ 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:

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

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 $\alpha = 5\%$, 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. EXERCÍCIOS

(01) Pretende-se lançar uma moeda 5 vezes e rejeitar a hipótese de que a moeda é não-tendenciosa, isto é, pretende-se rejeitar H_o: $\pi = 0,50$, se em 5 (cinco) jogadas ocorrerem 5 coroas ou 5 caras. Qual é a probabilidade de se cometer erro do tipo I?

(02) (Bussab, pg. 249) Se, ao lançarmos 3 vezes uma moeda, supostamente equilibrada, aparecerem 3 caras decide-se rejeitar a hipótese de que a moeda é "honesta", qual a probabilidade de se cometer erro do tipo I? Se a moeda favorece cara em 80% das vezes, qual a probabilidade de se cometer erro do tipo II?

(03) Você suspeita que um dado é viciado, isto é, você suspeita que a probabilidade de obter face 6 é maior do que 1/6. Você decide testar a hipótese de que o dado é não-viciado, jogando-o cinco vezes e rejeitando essa hipótese se ocorrer a face 6 (seis), 4 ou 5 vezes. Qual o nível de significância do teste?

(04) Nas faces de dois tetraedros regulares, aparentemente idênticos, estão marcados os valores: 0, 1, 2 e 3. Ao lançar um destes tetraedros o resultado observado é o valor da face que fica em contato com a superfície. Os dois tetraedros são "chumbados", de tal maneira que, ao jogá-los, as probabilidades de cada uma das faces ficar em contato com a superfície são as da tabela. Tomando ao acaso um dos tetraedros tem-se duas hipóteses: H_0 : Trata-se do tetraedro A; H_1 : Trata-se do tetraedro B.

Face	Tetraedro A	Tetraedro B
0	0,40	0,20
1	0,20	0,20
2	0,20	0,20
3	0,20	0,40
Total	**1**	**1**

(04.1) Para testar H_0 contra H_1, o tetraedro escolhido é lançado duas vezes. Adota-se a seguinte regra de decisão: rejeitar H_0 se a soma dos resultados dos dois lançamentos for maior ou igual a 5. Determinar o nível de significância e o poder do teste.

(04.2) Determinar o nível de significância e o poder do teste se a regra de decisão for: rejeitar H_0 se sair o valor 3 (três) em ao menos um dos lançamentos e o outro resultado não for o valor 0 (zero).

(05) Em cada uma das quatro faces de dois tetraedros regulares, aparentemente idênticos, estão marcados os valores: 1, 2, 3 e 4. Entretanto, um dos tetraedros é feito de material homogêneo (tetraedro A) , de maneira que, ao lançá-lo a probabilidade de qualquer uma das 4 faces fique em contato com a superfície é 0,25. O outro tetraedro (tetraedro B) é "chumbado", de tal maneira que, ao jogá-lo, a face com o valor 4 (quatro) tem probabilidade de 0,50 de ficar em contato com a superfície, enquanto que qualquer uma das outras três tem probabilidade igual a 1/6. Suponha que um dos tetraedros é lançado 48 vezes, para testar a hipótese H_0 de que foi lançado o tetraedro A, contra a hipótese H_1 de que foi lançado o tetraedro B. Supõem-se ainda a seguinte regra de decisão: "se nos 48 lançamentos, a face com o valor 4 (quatro), for obtida 20 ou mais vezes, rejeita-se H_0 em favor de H_1. Determine o nível de significância e o poder do teste.

(06) Uma urna contém 6 fichas, das quais θ são brancas e $6 - \theta$ são pretas. Para testar a hipótese de nulidade de que $\theta = 3$, contra a alternativa de que $\theta \neq 3$, são retiradas 2 (duas) fichas da urna ao acaso e sem reposição. Rejeita-se a hipótese nula se as duas fichas forem da mesma cor.

(06.1) Determine P(Erro do Tipo I).

(06.2) Determine o poder do teste para os diferentes valores de θ.

(06.3) Considere, agora, que a segunda ficha é retirada após a reposição da primeira. Calcule, novamente, o nível de significância e os valores do poder do teste.

(06.4). Compare os dois procedimentos (com e sem reposição da segunda ficha retirada). Qual a conclusão?

(07) Para decidirmos se os habitantes de uma ilha são descendentes da civilização A ou B, iremos proceder da seguinte forma:

(i) Selecionamos uma amostra aleatória de 100 moradores adultos da ilha e determinamos a altura média;

(ii) Se a altura média for superior a 176 cm, diremos que os habitantes são descendentes de B, caso contrário, admitiremos que são descendentes de A.

Os parâmetros das duas civilizações são: A: μ_A = 175 cm e σ_A = 10 cm e B: μ_B = 177 cm e σ_B = 10 cm. Define-se ainda: erro do tipo I como sendo "dizer que os habitantes são descendentes de B quando, na realidade, são de A" e erro do tipo II "dizer que os habitantes são de A quando, na realidade, são descendentes de B".

(07.1) Qual a probabilidade de erro do tipo I e do tipo II?

(07.2) Se $\sigma_A = \sigma_B = 5$, como ficariam os valores dos erros do tipo I e II?

(07.3) Qual deve ser a regra de decisão se quisermos fixar a a probabilidade de Erro I em 5%. Qual a probabilidade de erro II neste caso?

(07.4) Quais as probabilidades de Erro II, se as médias forem: μ_A = 178 e se μ_B = 180?

(08) Fazendo o teste H_0: $\mu = 1150$ ($\sigma = 150$) contra H_1: $\mu = 1200$ ($\sigma = 200$) e com $n = 100$, estabeleceu-se a seguinte região crítica: RC = $[1170, +\infty)$.

(08.1) Qual a probabilidade α de rejeitar H_0 quando verdadeira?

(08.2) Qual a probabilidade β de Aceitar H_0 quando H_1 é verdadeira?

(09) Dados os valores: 4, 6, 3, 6 e 6, de uma amostra aleatória de 5 (cinco) observações de uma variável X, estime a média e a variância de X e admitindo que X tenha uma distribuição normal, teste, a 5%, a hipótese de que a média da população é 1 (um), contra a hipótese alternativa de que é maior do que 1 (um).

(10) Sabe-se que o consumo mensal per capita de determinado produto tem distribuição normal, com desvio padrão de 2 kg. A diretoria da empresa que fabrica esse produto resolveu que retiraria o produto da linha de produção se a média de consumo per capita fosse menor do que 8 kg, caso contrário, continuaria a fabricá-lo. Foi realizado uma pesquisa de mercado, tomando-se uma amostra aleatória de 25 pessoas e verificou-se um consumo total de 180 kg do produto.

(10.1) Construa um teste de hipótese adequado para verificar a hipótese acima a um nível de significância de 5% e diga qual deve ser a decisão a ser adotada pela empresa?

(10.2) Qual a probabilidade β de a empresa tomar a decisão errada se, na realidade, o consumo médio mensal populacional é de 7,80 kg?

(10.3) Se a diretoria tivesse fixado uma significância de 1%, a decisão seria a mesma?

(10.4) Se o desvio padrão populacional fosse de 4 kg, qual seria a decisão a ser tomada com base na amostra mencionada acima?

(11) A associação dos proprietários de indústrias metalúrgicas está preocupada com o tempo perdido com acidentes de trabalho, cuja média, nos últimos tempos, tem sido da ordem de 60 homens/hora por ano, com desvio padrão de 20 homens/hora. Tentou-se um programa de prevenção de acidentes e, após o mesmo, tomou-se uma amostra aleatória de 16 indústrias e verificou-se que o tempo perdido baixou para 50 homens /hora ano. Você diria que, ao nível de 5% de significância, o programa surtiu efeito?

(12) Está-se desconfiado de que a média das receitas municipais, per capita, das cidades pequenas (menos de 20 mil habitantes) é maior do que a média da receita estadual que é de 1229 unidades monetárias. Para testar a hipótese é realizada uma amostragem com 10 pequenas cidades que forneceram os seguintes resultados (em termos de receitas médias):

1230, 582, 576, 2093, 2621, 1045, 1439, 717, 1838, 1359

Verifique que não é possível rejeitar a hipótese de que as receitas municipais são iguais as do estado, aos níveis usuais de significância. Como isto se justifica, já que a média da amostra obtida é bem maior do que a média do estado!

(13) Medidos os diâmetros de 31 eixos de um lote aleatório, produzido pela empresa "Sofazredondo S.A." obteve-se a distribuição abaixo:

Diâmetros (em mm)	56,5	56,6	56,7	56,8	56,9	57,0	57,1	57,2	57,3
Número de eixos	1	2	2	4	10	5	4	2	1

Ao nível de significância de 5%, há evidência de que o diâmetro médio dos eixos esteja fora da especificação de uma média de 57 mm?

(14) Um fabricante garante que 90% das peças que fornece a um cliente estão de acordo com as especificações exigidas. O exame de uma amostra aleatória de 200 destas peças revelou 25 fora das especificações. Verifique se as níveis de 5% e 1% de significância há exagero na afirmativa do fabricante.

(15) Suponha que a experiência tenha mostrado que dos alunos submetidos a determinado tipo de prova, 20% são reprovados. Se de uma determinada turma de 100 alunos, são reprovados apenas 13, pode-se concluir, ao nível de significância de 5%, que estes alunos, são melhores?

(16) Um exame é composto de 100 testes do tipo certo-errado. (a) Determine o número mínimo de testes que um aluno deve acertar para que se possa, ao nível de significância de 5%, rejeitar a hipótese de que o aluno nada sabe sobre a matéria e respondeu ao acaso, em favor da hipótese de que o alunos sabia alguma coisa sobre a matéria do teste? (b) Qual seria este mínimo, se fosse adotado o nível de significância de 1%?

(17) O rótulo de uma caixa de sementes informa que a taxa de germinação é de 90%. Entretanto, como a data de validade está vencida, acredita-se que a taxa de germinação seja inferior a este número. Faz-se um experimento e de 400 sementes, tomadas ao acaso, 350 germinam. Qual a conclusão ao nível de 5% de significância?

(18) Observou-se a produção mensal de uma indústria durante alguns anos e verificou-se que ela obedecia a uma distribuição normal com variância igual a 300 u^2. Foi adotada então uma nova técnica de produção e durante um período de 24 meses observou-se a produção mensal. Após este período constatou-se que a variância foi de 400 u^2. Há motivos para se acreditar que houve alteração na variância ao nível de 10%?

(19) Numa linha de produção é importante que o tempo gasto numa determinada operação não varie muito de empregado para empregado. Em operários bem treinados a variabilidade fica em 100 u^2. A empresa colocou 11 novos funcionários para trabalhar na linha de produção, supostamente bem treinados, e observou os seguintes valores, em segundos:

125 135 115 120 150 130 125 145 125 140 130

Testar se a tempo despendido por estes funcionários pode ser considerado mais variável do que os demais funcionários. Utilize 5% de significância.

(20) O departamento de psicologia fez um estudo comparativo do tempo médio de adaptação de uma amostra de 50 homens e outra de 50 mulheres, tomados ao acaso, de um grande complexo industrial que mostrou os seguintes resultados da tabela. É possível afirmar, ao nível de 5% de significância que as mulheres desta empresa levam mais tempo para se adaptarem?

Estatísticas	Homens	Mulheres
Média	3,2 meses	3,7 meses
Desvio padrão	0,8 meses	0,9 meses

(21) Diversas políticas, em relação às filiais de uma rede de supermercados, estão associadas ao gasto médio dos clientes em cada compra. Deseja-se comparar estes parâmetros de duas novas filiais, através de duas amostras de 50 clientes, selecionados ao acaso, de cada uma das novas filiais. As médias obtidas foram 62 e 71 unidades monetárias. Supondo que os desvios padrões sejam idênticos e iguais a 20 um, teste a hipótese de que o gasto médio dos clientes não é o mesmo nas duas filiais. Utilize uma significância de 2,5%?

(22) Uma fábrica de embalagens para produtos químicos está estudando dois processos diferentes de combate a corrosão nas latas usadas para embalagem. Para verificar o efeito dos dois processos foram utilizadas duas amostras aleatórias que apresentaram os valores da tabela, quanto a variável "duração da embalagem (em meses) antes da primeira mancha de corrosão aparecer". Ao nível de significância de 5% é possível afirmar que um tratamento é melhor do que o outro?

Processo	Tamanho da amostra	Média	Desvio padrão
A	**15**	**48**	**10**
B	**12**	**52**	**15**

(23) Você recebe a informação de que a diferença entre duas médias amostrais é "estatisticamente significativa ao nível de 1%". Dizer se as afirmações abaixo estão certas ou erradas e justificar.

(23.1) Há pelo menos 99% de probabilidade de existir uma diferença real entre as médias das duas populações.

(23.2) Se não houvesse diferença entre as médias das duas populações, a probabilidade de detectar uma tal diferença (ou diferença maior) entre as médias amostrais seria de 1% ou menos.

(23.3) A informação constituí uma evidência sólida de que realmente exista diferença entre as médias populacionais. Todavia, por si só, não constituí evidência suficiente de que tal diferença seja suficientemente grande para ter importância prática. Isto ilustra a diferença entre os conceitos "significância estatística" e "significância prática".

(23.4) O valor da estatística teste (valor calculado) é exatamente 1%.

(23.5) A probabilidade de que as médias das duas amostras sejam diferentes é de 1%.

(24) Foram levantadas quatro hipóteses sobre a média salarial anual de engenheiros mecânicos e civis:

(i) Engenheiros mecânicos e civis ganham em média o mesmo salário.

(ii) Os engenheiros mecânicos ganham, em média R$ 500 a mais do que os civis.

(iii) Os engenheiros mecânicos ganham, em média R$ 1000 a mais do que os civis.

(iv) Os engenheiros mecânicos ganham, em média R$ 2000 a mais do que os civis.

Para testar a hipótese foram extraídas duas amostras aleatórias dos salários dos dois tipos de profissionais que apresentaram os valores da tabela. Com base, nos valores, responda, justificando, as seguintes questões:

Engenheiros	Tamanho da amostra	Média	Desvio padrão
Mecânicos	250	38000	8000
Civis	200	36000	10000

(24.1) Sem quaisquer, cálculos detalhados, podemos verificar imediatamente, qualquer uma das hipóteses.

(24.2) Se aplicarmos um teste bilateral a cada uma das hipóteses, quais seriam rejeitadas ao nível de 5%?

(24.3) Se aplicarmos um teste unilateral a cada uma das hipóteses, quais seriam rejeitadas ao nível de 5%?

(24.4) Várias hipóteses foram consideradas aceitáveis, ao nível de 5% de significância. Se você tivesse que escolher apenas uma delas para publicar como conclusão do estudo, por qual optaria? Por quê?

(25) Calculadoras eletrônicas utilizam dois métodos diferentes de entrada e processamento numérico. Vamos denominar um dos métodos de "método algébrico" (MA) e o outro de "método polonês" (MP). Para comparar qual deles é mais eficaz é feito um teste com 20 usuários sem experiência prévia com calculadoras, onde 10 vão utilizar calculadoras de um tipo e o outros 10 as de outro tipo. A tabela mostra o tempo em segundos que cada operador gastou para realizar um conjunto padrão de cálculos. Testar a hipótese de que não existe diferença entre os dois métodos no que se refere ao tempo de operação, utilizando uma significância de 5%.

Operador	1	2	3	4	5	6	7	8	9	10
MA	12	16	15	13	16	10	15	17	14	12
MP	10	17	18	16	19	12	17	15	17	14

(26) Num ensaio para testar a proteção de dois tipos de tinta em superfícies metálicas, 55 painéis foram pintados com a tinta PK12 e 75 com a tinta PK15. Decorridos dois anos de exposição dos painéis ao ar livre, verificou-se que, dos painéis pintados com PK12, 6 apresentaram problemas enquanto que dos 75 painéis pintados com PK15, 19 apresentaram problemas. Pode-se concluir, destes valores, com 5% de significância, que as duas marcas de tintas diferem quanto a capacidade de proteção?

(27) Um psicólogo defende a idéia de que a autorização para dirigir só deve ser dada a maiores do que 21 anos de idade. Para tanto argumentou que os jovens entre 18 e 21 causam no mínimo 15% a mais acidentes dos que os de mais de 21 anos. Suas conclusões são baseadas em uma amostra de 150 pessoas entre os 18 e 21 anos, dos quais 60 já haviam se envolvido em algum tipo de acidente. Já entre os motoristas maiores de 21 anos de 200 observados, 30 já haviam se envolvido em algum tipo de acidente. (a) Teste a argumentação do psicólogo a um nível de 5% de significância. (b) Qual o problema que as amostras coletadas pelo psicólogo apresentam?

(28) Em dois anos consecutivos foi feito um levantamento de mercado sobre a preferência dos consumidores pelo por um determinado produto. No primeiro ano o produto era anunciado com freqüência semanal nos veículos de comunicação e no segundo ano com freqüência mensal. No levantamento foram utilizados duas amostras independentes de 400 consumidores cada. No primeiro ano o percentual de compradores ficou em 33% e no segundo ano em 29%. Considerando o nível de significância de 5%, teste a hipótese de que a freqüência do anúncio tem influência na manutenção da fatia de mercado.

(29) Uma das maneiras de medir o grau de satisfação dos empregados de uma mesma categoria quanto a política salarial é através do desvio padrão de seus salários. A fábrica A diz ser mais coerente na política salarial do que a fábrica B. Para verificar essa afirmação, sorteou-se uma amostra de 10 funcionários não especializados de A e 15 de B, obtendo-se os desvios padrões: $s_A = 1,0$ s.m. e $s_B = 1,6$ s.m. Qual a sua conclusão a um nível de 5% de significância?

(30) (BUSSAB - pg. 277) Deseja-se comparar a qualidade de um produto produzido por dois fabricantes. Esta qualidade está sendo medida pela uniformidade com que é produzido o produto por cada fábrica. Tomaram-se duas amostras, uma de cada fábrica, medindo-se o comprimento dos produtos. A qualidade da produção das duas fábricas é a mesma a um nível de 5%?

Estatísticas	Fábrica A	Fábrica B
Amostra	21	17
Média	21,15	21,12
Variância	0,0412	0,1734

4. RESPOSTAS

(01) RC = { 0, 5} α = P(RC) = P{ X = 0 ou X = 5 / π = 0,50} = $(1/2)^5 + (1/2)^5$ = 1/16 = 6,25%

(02) RC = { 3 } α = P({ 3 }) = $(1/2)^3$ = 1/8 = 12,50%

β = P(Ac. H_0 / H_0 é Falsa} = P(X = 0, 1, 2 / π = 0,8} = $(1/5)^3 + 3(4/5)^1 (1/5)^2 + 3(4/5)^2 (1/5)^1$ = 48,80%

(03) RC { 4, 5} α = P(RC) = P({ X = 4 ou X = 5 / π = 1/6}) = 13/3888 = 0, 33%

(04) **(04.1)** RC = { (2, 3), (3, 2), (3, 3) } α = P(RC) = 0,20.0,20 + 0,20.0,20 + 0,20.0,20 = 12%

Poder do Teste = 1 - β = P(Rej. H_0 / H_0 é Falsa} = 0,20.0,40 + 0,20.0,40 + 0,40.0,40 = 32%

(04.2) RC = { (3, 1), (3, 2), (3, 3), (1, 3), (2, 3) } α = P(RC) = 5.0,04 = 0,20 = 20%

Poder do Teste = 1 - β = P(Rej. H_0 / H_0 é Falsa} = 4.0,08 + 0,16 = 0,48 = 48%

(05) RC = { X $\geq$ 20 / Tetraedro A) α = P(RC) = P({ X $\geq$ 20 / Tet. A}) $\cong$ P(Z $\geq$ (19,5 - 12) / 3) = 0,62%

β = P(Ac. H_0 / H_0 é Falsa} = P(X < 20 / Tet. B} = 9,68% Poder = 1 - β = 100% - 9,68% = 90,32%

(06) **(06.1)** n = 2 S/R RC = {BB, PP} α = P(RC) =(3/6).(2/5) + (3/6).(2/5) = 1/5 + 1/5 = 0,40 = 40%

(06.2) n = 2 S/R 1 - β = P(Rejeitar H_0 / H_0 é falsa)

θ = 0 ou θ = 6 1 - β = P(RC / θ = 0) = 1 = 100% = P(R / θ = 6)

θ = 1 ou θ = 5 1 - β = P(RC / θ = 1) = (1/6).(0/5) + (5/6).(4/5) = 2/3 = 66,67% = P(RC / θ = 5)

θ = 2 ou θ = 4 1 - β = P(RC / θ = 2) = (2/6).(1/5) + (4/6).(3/5) = 7/15 = 46,67% = P(RC / θ = 4)

(06.3) n = 2 C/R RC = {BB, PP} α = P(RC) =(3/6).(3/6) + (3/6).(3/6) = 1/4 + 1/4 = 0,50 = 50%

θ = 0 ou θ = 6 1 - β = P(RC / θ = 0) = 0 + 1 = 100% = P(RC / θ = 6)

θ = 1 ou θ = 5 1 - β = P(RC / θ = 1) = (1/6).(1/6) + (5/6).(5/6) = 13/18 = 72,22% = P(RC / θ = 5)

θ = 2 ou θ = 4 1 - β = P(RC / θ = 2) = (2/6).(2/6) + (4/6).(4/6) = 5/9 = 55,56% = P(RC / θ = 4)

(06.4) Com reposição o NS (α) é maior do que SR. Por outro lado, repondo o poder do teste é maior ou igual a quando não se faz reposição.

(07) **(07.1)** P(Erro I) = P($\overline{X}_A$ > 176) = P(Z > 176 - 175) = P(Z > 1) = 15,87%

P(Erro II) = P($\overline{X}_B$ < 176) = P(Z < 176 - 177) = P(Z < -1) = 15,87%

(07.2) P(Erro I) = P($\overline{X}_A$ > 176) = P[Z > (176 - 175)/0,5] = P(Z > 2) = 2,28%

P(Erro II) = P($\overline{X}_B$ < 176) = P[Z < (176 - 177)/2] = P(Z < -2) = 2,28%

(07.3) 5% = P(Erro I) = P($\overline{X}_A$ > 176) = P(Z > 176 - 175) ⇒ P(Z > x - 175) = 5% ⇒ x = 176,645. Neste caso, deve-se rejeitar H_0 somente se a média for superior a 176,645.

P(Erro II) = P($\overline{X}_B$ < 176,645 - 177) = P(Z < -0,36) = 35,94%

(07.4) μ_B = 178 P(Erro II) = P($\overline{X}_B$ < 176 - 178) = P(Z < -2) = 2,28%

μ_B = 180 P(Erro II) = P($\overline{X}_B$ < 176 - 180) = P(Z < -4) = 0,00%

(08) **(08.1)** α = P(Rej. H_0 / H_0 é V) = P($\overline{X}$ > 1170 / μ = 1150) = P[Z > (1170 - 1150) / 15)] = P(Z > 1,33) = 9,18%

(08.2) β = P(Ac H_0 / H_1 é V) = P($\overline{X}$ < 1170 / μ = 1200) = P[Z < (1170 - 1200) / 20)] = P(Z < -1,50) = 6,68%

(08.3) P[Z > (x - 1150) / 15)] = P[Z < (x - 1200) / 20)] ⇒ (x - 1150) / 15 = -(x - 1200) / 20 ⇒ x = 1171,43

(09) $\overline{x}$ = 5, s^2 = 2 t = 6,32 > $t_{5\%}$ = 2,132, portanto rejeita H_0

(10) **(10.1)** H_0: μ = 8 kg contra H_1: μ < 8 kg. Como α = 5%, z_α = -1,645 e z_c = -2. Logo rejeitar H_0

(10.2) β = P(Ac. H_0 / H_1 é V) = P($\overline{X}$ > 7,34 / μ = 7,80) = P(Z < 1,14) = 87,29%

(10.3) H_0: μ = 8 kg contra H_1: μ < 8 kg. Como α = 1%, z_α = -2,33 e z_c = -2. Não rejeita H_0

(10.4) Aceitar H_0 tanto ao nível de 5% quanto ao de 1% de significância.

(11) Como α = 5%, z_α = -1,645 e z_c = -2. Rejeita-se H_0, isto é, pode-se dizer que o programa surtiu efeito.

(12) Como t_c = -0,566, não é possível rejeitar a hipótese aos níveis de 1%, 5% e mesmo 10%. Isto se justifica devido a grande variabilidade da amostra que apresenta um desvio padrão igual a 675,82.

(13) H_0: μ = 57mm contra H_1: μ ≠ 57 mm Como t_c = -2,557 e t_t = -2,042, rejeita-se H_0.

(14) H_0: π = 10% contra H_1: π > 10%. Como z_c = 1,18. Logo não se pode rejeitar H_0.

(15) H_0: π = 20% contra H_1: π < 20%. Como z_c = -1,75 e $z_{5\%}$ = -1,645 . Logo pode-se rejeitar H_0.

(16) H_0: π = 50% contra H_1: π > 50%. Como $z_{5\%}$ = -1,645 o número mínimo de acertos é: 50% + 1,645.σ_P ≅ 59

Como $z_{1\%}$ = -2,33 o número mínimo de acertos é: 50% + 2,33.σ_P ≅ 62

(17) H_0: π = 90% contra H_1: π < 90%. Como z_c = -1,667 e $z_{5\%}$ = -1,645 . Logo pode-se rejeitar H_0.

(18) Não, pois χ^2 = 30,67 está na região de aceitação que é: RA = [13,09; 35,17]

(19) Não, pois χ^2 = 11,41 está na região de aceitação que é: RA = [0; 18,3]

(20) H_0: $\mu_H = \mu_M$ contra H_1: $\mu_H < \mu_M$. Como α = 5%, t_α = -1,645 e t_c = -2,936. Rejeitar H_0.

(21) H_0: $\mu_1 = \mu_2$ contra H_1: $\mu_1 \neq \mu_2$. Como α = 2,5%, t_α = -2,24 e t_c = -2,25. Rejeitar H_0.

(22) H_0: $\mu_1 = \mu_2$ contra H_1: $\mu_1 \neq \mu_2$. Como $\alpha = 5\%$, $t_{25} = -2{,}06$ e $t_c = -0{,}79$. Não rejeitar H_0

(23) **(23.1).** Errada. **(23.2)** Correta. **(23.3)** Errada. **(23.4)** Errada. **(23.5)** Errada.

(24) **(24.1)** Sim a quarta. **(24.2)** Somente a (i) **(24.3)** Somente a (i) e a (ii).

(24.4) A (i) que pode ser confirmada tanto no teste unilateral quanto no bilateral (mais rigoroso)

(25) H_0: $\mu_1 = \mu_2$ contra H_1: $\mu_1 \neq \mu_2$. Como $\alpha = 5\%$, $t_\alpha = 2{,}26$ e $t_c = -2{,}42$. Não rejeitar H_0, supondo amostras emparelhadas.

(26) H_0:$\pi_1 = \pi_2$ contra H_1: :$\pi_1 \neq \pi_2$. Como $z_c = 2{,}20$ e $z_{5\%} = 1{,}96$. Pode-se afirmar que as duas tintas diferem.

(27) **(a)** H_0:$\pi_1 - \pi_2 = 15\%$ contra H_1: :$\pi_1 - \pi_2 < 15\%$ Como $z_c = -2{,}11$ e $z_{5\%} = -1{,}645$. Logo pode-se afirmar que os jovens causam pelo menos 15% a mais de acidentes.

(b) O problema é que as amostras tem um vício de origem, pois fica difícil de saber se esta diferença é devida a imprudência ou ao fato de que os motoristas são menos experientes.

(28) H_0:$\pi_1 = \pi_2$ contra H_1: :$\pi_1 > \pi_2$ Como $z_c = 1{,}22$ e $z_{5\%} = 1{,}645$. Logo não se pode rejeitar H_0

(29) Não se pode afirmar que não são iguais, pois $F_C = 2{,}56$ e a RA = [0,38; 2,65]

(30) Pode-se afirmar que a qualidade difere, pois $F_c = 4{,}21$ e RA = [0,37; 2,54]

5. REFERÊNCIAS

[BUS86] BUSSAB, Wilton O, MORETTIN, Pedro A. *Estatística Básica*. 3. ed. São Paulo, Atual, 1986.

[DOW89] DOWNING, Douglas, CLARK, Jeff. *Statistics The Easy Way*. Hauppauge (New York): Barron's Educational Series, Inc, 1989.

[HIN88] HINKLE, Dennis E., WILLIAM, Wiersma, JURS, Stephen G. *Applied Statistics for the Behavioral Sciences*. Boston: Houghton Mifflin Co., 1988.

[HOF80] HOFFMAN, Rodolfo. *Estatística para Economistas*. São Paulo. Livraria Pioneira Editora, 1980.

[NET77] NETO, Pedro Luiz de Oliveira Costa. *Estatística*. São Paulo, Edgard Blücher, 1977.

[MAS90] MASON, Robert D., DOUGLAS, Lind A. *Statistical Techniques in Business And Economics*. IRWIN, Boston, 1990.

[MEY78] MEYER, Paul L. *Probabilidade: aplicações à Estatística.* Tradução do Prof. Ruy C. B. Lourenço Filho. Rio de Janeiro, Livros Técnicos e Científicos Editora S.A., 1978

[RES93] Research & Education Association. *The Statistics Problem Solver*. Piscataway (New Jersey): 1993.

[STE81] STEVENSON, William J. *Estatística Aplicada à Administração*. São Paulo. Editora Harbra, 1981.

[WEL82] WLKOWITZ, Joan, EWEN, Robert B., COHEN, Jacob. *Introductory Statistics for the Behavioral Sciences.* Orlando(FL): Hartcourt Brace Javanovich, 1982.

[WON85] WONNACOTT, Ronald J., WONNACOTT, Thomas. *Fundamentos de Estatística*. Rio de Janeiro. Livros Técnicos e Científicos Editora S. A., 1985.