Universidade Federal
de Mato Grosso

UNIDADE II – TESTE DE HIPÓTESE

PROF.: RÔMULO MÔRA
romulomora.webnode.com

CUIABÁ, MT
2015/1

1. INTRODUÇÃO

A inferência estatística preocupa-se em determinar se existe alguma significância estatística acoplada aos resultados de uma pesquisa.

Inferência Estatística
- **A identificação da população para a qual a inferência se aplica**
- **A confiabilidade da inferência**

Têm como objetivo principal obter informações da população a partir de amostras retiradas destas, num processo de amostragem adequado.

μ - Média
σ^2 - Variância
σ - Desvio padrão

→ Parâmetros da População

$\bar{x}$ - Média
s^2 - Variância
s - Desvio padrão

→ Estimativas da amostra

Estatísticas

2. TESTE DE HIPÓTESE

Tomada de decisão acerca da **população**, ou seja, feita uma afirmação sobre uma população, usualmente sobre um **parâmetro** desta, desejamos saber se os resultados de uma amostra contrariam ou não tal afirmação.

É uma regra decisória que nos permite aceitar ou rejeitar uma **hipótese estatística** previamente formulada, com base em informações obtidas de uma amostra representativa da população.

Exemplo

O gerente de produção de eucalipto de uma dada empresa designa um técnico para vistoriar a população de plantas com relação ao ataque de lagartas.

O objetivo principal é saber se a infestação deste inseto ultrapassa um nível de controle que acima do qual ocorre prejuízo econômico.

Hipótese: “O ataque de lagartas está abaixo do nível de controle.

Decisão: controlar ou não a praga.

Teste paramétrico

- ✓Variável tenha distribuição normal;
- ✓ A variâncias são homogêneas;
- ✓ Os efeitos dos fatores de variação são aditivos e independentes.

TESTE DE HIPÓTESES

Testes não paramétricos

Variável não apresentam distribuição normal.

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

2.1. CURVAS DE DISTRIBUIÇÃO NORMAL

Características

✓Curva Padrão

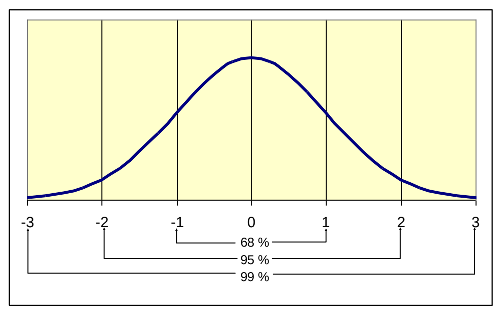
- •Simétrica
- •Uni-modal
- •Possuí forma de Sino
- •Média, mediana e moda são iguais

✓Curva Normal Padrão

$$z_i = \frac{(X_i - \mu)}{\sigma}$$

- •Curva padrão possuí média igual a zero e variância igual a um
- •Valor Padrão

• Áreas sob a curva com $\pm 1\sigma$, $\pm 2\sigma$ e $\pm 3\sigma$



2.2. HIPÓTESE ESTATÍSTICA

H_0 = Hipótese da Nulidade – É a afirmativa que se faz sobre um parâmetro a ser testado. É a hipótese a ser testada

H_0: μ = 1,66 H_0: μ = 50 H_0: σ = 50

H_1 ou H_a = Hipótese Alternativa - É aquela que contraria a hipótese H_0

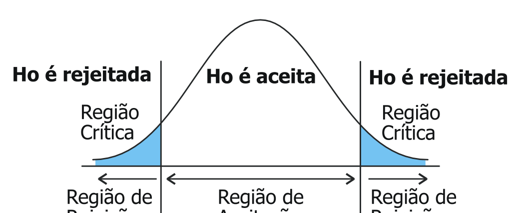
H_1: $\mu_1 > \mu_2$
H_1: $\mu_1 < \mu_2$ } Teste unilateral

H_1: $\mu_1 \neq \mu_2$ ➡ Teste bilateral

O Método Científico pressupõe que a Hipótese da Nulidade é sempre VERDADEIRA!

Região crítica

É a faixa de valores da variável teste correspondente à rejeição de H_0.



Tipos de erros

a) Erro tipo I ou α: É a probabilidade de rejeitar H_0, sendo H_0 verdadeira. É denominado nível de significância.

b) Erro tipo II ou β: É a probabilidade de aceitar H_0, sendo H_0 falsa.

OBS:

- Se aumentarmos o número de amostras diminuímos o erro.

Realidade	Decisão	
	Aceitar H_0	Rejeitar H_0
H_0 verdadeira	**Correta** **P= 1-α**	Erro tipo I ou α
H_0 falsa	Erro tipo II ou β	**Correta** **P= 1-β**

Tipos de testes:

5.1. Testes unilaterais:

5.1.1. Teste unilateral à direita:

A partir de um valor c, rejeita-se H_0 se x for $\geq$ c.

5.1.2. Teste unilateral à esquerda:

A partir de um valor c, rejeita-se H_0 se x for $\leq$ c.

5.1.3. Teste bilateral:

A partir de um valor c, rejeita-se H_0 se x for $\leq$ c ou x $\geq$ c.

- Unilateral à direita *:*

Ho: $\mu = 50$

H1:: $\mu > 50$

- Unilateral à esquerda :

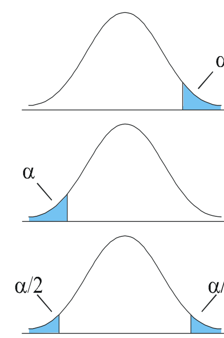
Ho: : $\mu = 50$

H1: : $\mu < 50$

- Bilateral:

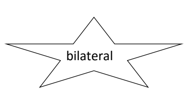
Ho: : $\mu = 50$

H1:: $\mu \neq 50$

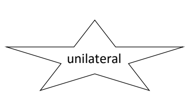


Por exemplo, podemos formular a hipótese que a produtividade é diferente de 2,5 peças/hora. Formalmente isso é escrito como:

$$H_0: \ \mu = 2{,}5\, peças/hora$$
$$H_a: \ \mu \neq 2{,}5\, peças/hora$$



$$H_0: \ \mu = 2{,}5\, peças/hora$$
$$H_a: \ \mu < 2{,}5\, peças/hora$$



Procedimentos para realização de um teste de hipótese

a) Enunciar as hipóteses H_0 e H_a;

b) Fixar o nível de significância α e identificar a estatística do teste;

c) Determinar a região crítica e a região de aceitação em função de α;

d) Calcular o valor da estatística do teste, por meio de dados amostrais;

e) Concluir o teste;

3. TESTES DE SIGNIFICÂNCIA

Teste para duas amostras

σ² conhecida | **σ² desconhecida**

Teste para uma amostra

Teste para duas amostras independentes

Teste para duas variâncias

Teste para dados pareados ou emparelhados

4. TESTE PARA UMA AMOSTRA

1º passo – Definir as hipóteses

H_0: $\mu_1 = \mu$

H_a: $\mu_1 \neq \mu$ ou H_a: $\mu_1 > \mu$ ou H_a: $\mu_1 < \mu$

2º passo – Definir o nível de significância (α)

3º passo – Calcular a estatística do teste

$$t_{calc} = \frac{\overline{X} - \mu_o}{S/\sqrt{n}} \sim t\,(n-1)$$ **σ² desconhecida**

$$Z_{calc} = \frac{\overline{X} - \mu_o}{\sigma/\sqrt{n}} \sim N(0,1)$$ **σ² conhecida**

4º passo – Definir o valor de Z ou t tabelado de acordo com o nível de significância e a região crítica

✓ Teste bilateral

$$Z_{tab} = Z_{\alpha/2}$$

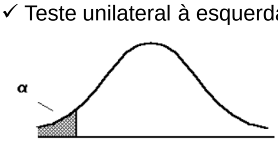
$$t_{tab} = t_{\alpha}(n-1)$$

✓ Teste unilateral à direita

$$Z_{tab} = Z_{\alpha}$$

$$t_{tab} = t_{\alpha}(n-1)$$

✓ Teste unilateral à esquerda



$$Z_{tab} = Z_{\alpha}$$

$$t_{tab} = t_{\alpha}(n-1)$$

Importante !!!! – Cálculo do valor tabelado para o teste t

Tabela bilateral
- Teste bilateral: entrar com α
- Teste unilateral: entrar com 2α

Tabela unilateral
- Teste bilateral: entrar com α/2
- Teste unilateral: entrar com α

<u>5º passo</u> – Regra de decisão

✓ Teste Z

$|Z_{calc}| \geq Z_{tab}$ - rejeita-se a hipótese H_0

$|Z_{calc}| < Z_{tab}$ - não rejeita-se a hipótese H_0

✓ Teste t

$|t_{calc}| \geq t_{tab}$ - rejeita-se a hipótese H_0

$|t_{calc}| < t_{tab}$ - não rejeita-se a hipótese H_0

<u>6º passo</u> – Conclusão

Exemplo 1:

Sabe-se que resistência à tração da madeira produzida pela máquina de tração é em média de 72 kg/mm^2 e com desvio padrão de 2,0 kg/mm^2. Recentemente, a máquina foi ajustada. A fim de verificar a média de resistência a tração, 10 amostras foram testadas apresentando os respectivos valores de resistência a tração:

74,2 75,3 72,4 73,7 77,6 72,4 73,7 72,2 73,3 74,2

Podemos concluir que teve diferença entre os valores da resistência à tração da madeira? (Adote um nível de significância de 5%)

Distribuição Normal Padrão - P(0 < Z < z tabela)

z (Unidade e primeira casa decimal)	Segunda casa decimal de z: 0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998

5. TESTE PARA DUAS AMOSTRAS

5.1 TESTES PARA DUAS AMOSTRAS INDEPENDENTES

✓ **TESTE F PARA DUAS VARIÂNCIAS**

<u>1º passo</u> - Definir as hipóteses

H_0: $\sigma^2_1 = \sigma^2_2$

H_a: $\sigma^2_1 > \sigma^2_2$ ou H_a: $\sigma^2_1 < \sigma^2_2$

<u>2º passo</u> – Definir o nível de significância (α)

<u>3º passo</u> – Estatística do teste

$$F_{calc} = \frac{s_1^2}{s_2^2} = \frac{s_{maior}^2}{s_{menor}^2}$$

<u>4º passo</u> – Definir o valor de F tabelado

$$F_{tab} = F_\alpha\left[(n_1 - 1), (n_2 - 1)\right]$$

Tem distribuição de Fisher com (n_1-1) e (n_2-1) graus de liberdade

<u>5º passo</u> – Regra de decisão e Conclusão

$F_{calc} \geq F_{tab}$ rejeita-se a hipótese H_0, logo as variâncias são diferentes e portanto heterogêneas. Desse modo calcula-se o teste para duas médias considerando variâncias heterogêneas.

$F_{calc} < F_{tab}$ não rejeita-se a hipótese H_0, logo as variâncias são iguas e portanto homogêneas. Desse modo calcula-se o teste para duas médias considerando variâncias homogêneas

6.1.1 TESTE PARA DUAS AMOSTRAS INDEPENDENTES CONSIDERANDO VARIÂNCIAS HOMOGÊNEAS

1º passo – Definir as hipóteses

H_0: $\mu_1 = \mu_2$

H_a: $\mu_1 \neq \mu_2$ ou H_a: $\mu_1 > \mu_2$ ou H_a: $\mu_1 < \mu_2$

2º passo – Definir o nível de significância (α)

3º passo – Estatística do Teste

$$t_{cal} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_c^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t\,(n_1 + n_2 - 2)$$

✓ Calculando a variância comum

$$s_c^2 = \frac{(n_1 - 1)\,s_1^2 + (n_2 - 1)\,s_2^2}{n_1 + n_2 - 2}$$

4º passo – Definir o t tabelado de acordo com o nível de significância

✓ Teste bilateral

$$t_{tab} = t_{\alpha}(n_1 + n_2 - 2)$$

✓ Teste unilateral

$$t_{tab} = t_{\alpha}(n_1 + n_2 - 2)$$

5º passo – Regra de decisão

$|t_{calc}| \geq t_{tab}$ - rejeita-se a hipótese H_0.

$|t_{calc}| < t_{tab}$ - não rejeita-se a hipótese H_0

6º passo – Conclusão

Exemplo 2:

Um grupo de árvores teve seus DAP's medidos com auxílio de uma suta (A) e uma fita diamétrica (B). Prevendo que a suta produz superestimativas dos diâmetros, verifique ao nível de 5% de probabilidade se em média os diâmetros medidos pela suta são superiores aos medidos com fita.

Suta	26,4	23,4	20,1	22,0	23,7	28,0
Fita	26,0	23,3	19,8	22,0	23,1	26,0

6.1.2 TESTE PARA DUAS AMOSTRAS INDEPENDENTES CONSIDERANDO VARIÂNCIAS HETEROGÊNEAS

1º passo – Definir as hipóteses

H_0: $\mu_1 = \mu_2$

H_a: $\mu_1 \neq \mu_2$ ou H_a: $\mu_1 > \mu_2$ ou H_a: $\mu_1 < \mu_2$

2º passo – Definir o nível de significância (α)

3º passo – Estatística do Teste

$$t_{cal} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}} \sim t\,(n^*)$$

4º passo – Definir o t tabelado de acordo com o nível de significância

✓ Teste bilateral

$$t_{tab} = t_{\alpha}(n^*)$$

✓ Teste unilateral

$$t_{tab} = t_{\alpha}(n^*)$$

Cálculo do valor do número de graus de liberdade (n*) conforme proposto por SATTERTWAITE

$$n^* = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

5º passo – Regra de decisão

$|t_{calc}| \geq t_{tab}$ - rejeita-se a hipótese H_0.

$|t_{calc}| < t_{tab}$ - não rejeita-se a hipótese H_0

6º passo – Conclusão

Exemplo 3:

Altura de árvores de dois grupos de uma mesma floresta tropical foram mensuradas com auxilio de um hipsômetro VERTEX. Sabendo que o grupo A e B apresentaram as seguintes médias e variâncias mostradas no quadro a seguir:

	$\overline{X}$	S^2	n
Grupo A	32,8	18,3	500
Grupo B	31	30,3	500

Verificar ao nível de 5% de probabilidade se existe diferença entre as alturas dos dois grupos analisados.

5.2 TESTE PARA DUAS AMOSTRAS DEPENDENTES, PAREADAS OU EMPARELHADAS

1º passo – Definir as hipóteses

H_0: $\mu_d = 0$

H_a: $\mu_d \neq 0$ ou H_a: $\mu_d > 0$ ou H_a: $\mu_d < 0$

2º passo – Definir o nível de significância (α)

3º passo – Estatística do Teste

$$t_{cal} = \frac{\overline{X}_d}{S_d / \sqrt{n}} \sim t\,(n-1)$$

4º passo – Definir o t tabelado de acordo com o nível de significância

✓ Teste bilateral

$t_{tab} = t_{\alpha}(n-1)$

✓ Teste unilateral

$t_{tab} = t_{\alpha}(n-1)$

5º passo – Regra de decisão

$|t_{calc}| \geq t_{tab}$ - rejeita-se a hipótese H_0.

$|t_{calc}| < t_{tab}$ - não rejeita-se a hipótese H_0

6º passo – Conclusão

Exemplo 4:

Em processo de dinâmica natural, foram medidos os diâmetros de seis fustes em um ano (n) e no ano subsequente (n+1) a fim de verificar se houve aumento do DAP no período de 1 ano. Utilize alfa 1%.

Ano 1	19,10	17,55	15,20	24,75	17,53	17,15
Ano 2	20,75	19,05	17,40	26,68	18,48	18,00
Dif= (2)-(1)	1,65	1,50	2,20	1,93	0,95	0,85

FIM