

XMAC02

Métodos Matemáticos para Análise de Dados

Aula 16 – Testes Z e t de duas amostras

Teste Z de duas amostras

2

□ Condições

- ▣ Amostra aleatória
- ▣ Observações independente
- ▣ Distribuição amostral deve se aproximar de uma Distribuição Normal
 - População possui uma distribuição normal e o desvio padrão da população é conhecido
 - OU
 - Tamanho da amostra ≥ 30

O desvio padrão amostral pode substituir o desvio padrão populacional.

Teste Z

3

Uma amostra

$$H_0: \mu = 150\text{cc}$$

$$H_a: \mu \neq 150\text{cc}$$

$$Z_{cal} = \frac{(\bar{x} - \mu)}{\sigma / \sqrt{n}}$$

Duas amostras

$$\text{Hipótese nula: } H_0: \mu_1 = \mu_2$$

$$\text{ou } H_0: \mu_1 - \mu_2 = 0$$

$$\text{Hipótese alternativa: } H_a: \mu_1 \neq \mu_2$$

Média de cada amostra no X1 E X2

$$Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Teste Z de duas amostras

4

- Exemplo: Duas máquinas fabricam perfumes. Foram coletadas 100 amostras de cada uma.
 - ▣ Máquina 1: Média = 151,2 ml / dp = 2,1 ml
 - ▣ Máquina 2: Média = 151,9 ml / dp = 2,2 ml
- Há diferença entre essas duas máquinas? (nível de confiança = 95%)

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

$$Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Teste Z de duas amostras

5

- Exemplo: Duas máquinas fabricam perfumes. Foram coletadas 100 amostras de cada uma.
 - ▣ Máquina 1: Média = 151,2 ml / dp = 2,1 ml
 - ▣ Máquina 2: Média = 151,9 ml / dp = 2,2 ml
- Há diferença entre essas duas máquinas? (nível de confiança = 95%)

$$Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(151,2 - 151,9)}{\sqrt{\frac{2,1^2}{100} + \frac{2,2^2}{100}}} = -0,7/0,304 = -2,30$$

Teste Z de duas amostras

6

- Exemplo: Duas máquinas fabricam perfumes. Foram coletadas 100 amostras de cada uma.
 - ▣ Máquina 1: Média = 151,2 ml / dp = 2,1 ml
 - ▣ Máquina 2: Média = 151,9 ml / dp = 2,2 ml
- Há diferença entre essas duas máquinas? (nível de confiança = 95%)

$$Z_{\text{cal}} = -2,30$$

$$Z_{\text{crítico}} = ?$$

7

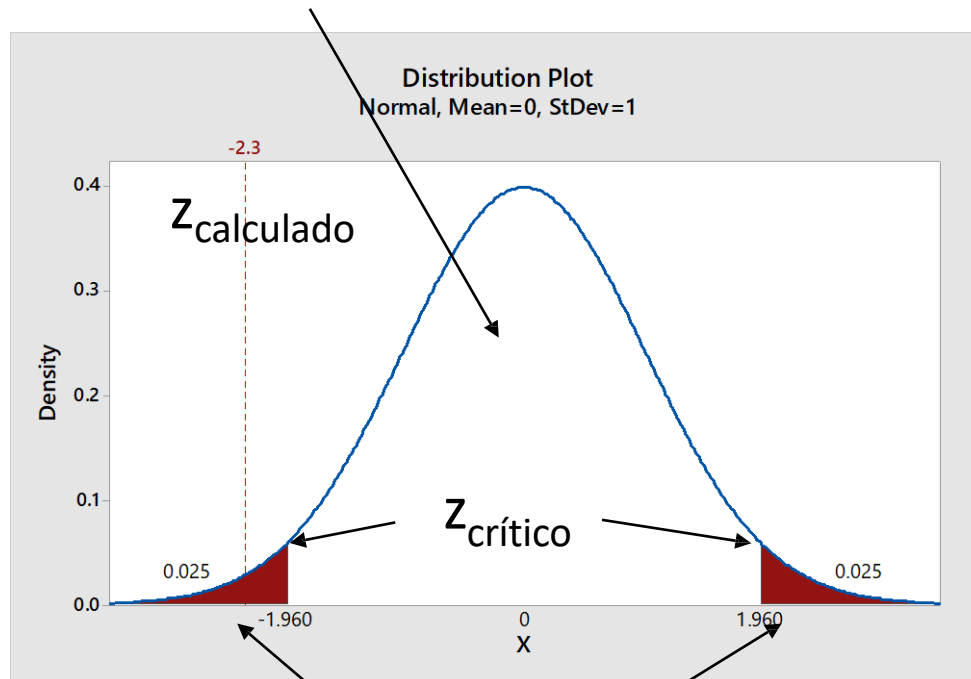
- ▣ $\alpha = 0,05$, duas caudas
 - ▣ 0,025 em cada cauda
 - ▣ Z crítico = 1,96

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Teste Z de duas amostras

8

Falha em rejeitar H_0



$$Z_{\text{cal}} = -2,30$$

$$Z_{\text{crítico}} = 1,96$$

Conclusão:
Rejeitar hipótese nula
 $H_0: \mu_1 = \mu_2$

Teste t de duas amostras

9

□ Condições

- ▣ Amostra aleatória
- ▣ Observações independente
- ▣ Distribuição amostral deve se aproximar de uma Distribuição Normal

Fazer quando você não consegue fazer o teste z, ou seja, não possui o desvio padrão ou não tem uma amostra maior que 30

- População possui uma distribuição normal e o desvio padrão da população **não** é conhecido

E

- Tamanho da amostra < 30

Tipos de Teste t

10

- ❑ Teste t de duas amostras
 - ▣ Conjuntos de dados são independentes
 - Valores de uma amostra não revelam informação a respeito da outra amostra
 - Exemplo: Duas máquinas de envase de perfume

- ❑ Teste t pareado
 - ▣ Conjuntos de dados são dependentes
 - Valores de uma amostra afetam valores da outra amostra
 - Exemplo: Pressão arterial antes e depois de tomar um remédio

Teste t de duas amostras

11

Variância igual para as duas amostras?

Calcula o valor de pulled variance para as duas amostras

Dps calcula o tcal com o pulled variance(sp^2)

Sim

Não

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

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

$$df = n_1 + n_2 - 2$$

Degree of freedom tem que calcular

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

Teste t de duas amostras

Mesma variância

12

- ❑ Exemplo: Amostras envasadas por duas máquinas têm apresentam os seguintes volumes:
 - ❑ Máquina A: 150, 152, 154, 152, 151
 - ❑ Máquina B: 156, 155, 158, 155, 154
 - ❑ A média é diferente? Calcular com 95% de confiança
-
- | | | |
|-----------------------|-----------------------|---|
| ❑ $n_1 = 5$ | ❑ $n_2 = 5$ | $H_0: \mu_A = \mu_B$
$H_a: \mu_A \neq \mu_B$ |
| ❑ $s_1 = 1,48$ | ❑ $s_2 = 1,52$ | |
| ❑ $\bar{x}_1 = 151,8$ | ❑ $\bar{x}_2 = 155,6$ | |

S é o desvio padrão

Teste t de duas amostras

Mesma variância

$$\square n_1 = 5$$

$$\square s_1 = 1,48$$

$$\square \bar{x}_1 = 151,8$$

$$\square n_2 = 5$$

$$\square s_2 = 1,52$$

$$\square \bar{x}_2 = 155,6$$

$$H_0: \mu_A = \mu_B$$

$$H_a: \mu_A \neq \mu_B$$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

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

Teste t de duas amostras

Mesma variância

$$\square n_1 = 5$$

$$\square n_2 = 5$$

$$\square s_1 = 1,48$$

$$\square s_2 = 1,52$$

$$\square \bar{x}_1 = 151,8$$

$$\square \bar{x}_2 = 155,6$$

$$H_0: \mu_A = \mu_B$$

$$H_a: \mu_A \neq \mu_B$$

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

$$s_p^2 = \frac{(5-1)*1,48^2 + (5-1)*1,52^2}{5 + 5 - 2}$$

$$s_p = 1,50$$

Teste t de duas amostras

Mesma variância

$$\square n_1 = 5$$

$$\square n_2 = 5$$

$$\square s_1 = 1,48$$

$$\square s_2 = 1,52$$

$$\square \bar{x}_1 = 151,8$$

$$\square \bar{x}_2 = 155,6$$

$$H_0: \mu_A = \mu_B$$

$$H_a: \mu_A \neq \mu_B$$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{151,8 - 155,6}{1,5 \sqrt{\left(\frac{1}{5} + \frac{1}{5}\right)}} = -4,01$$

df	TAIL PROBABILITY P											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646

❖ $\alpha = 0.05$ Two Tails

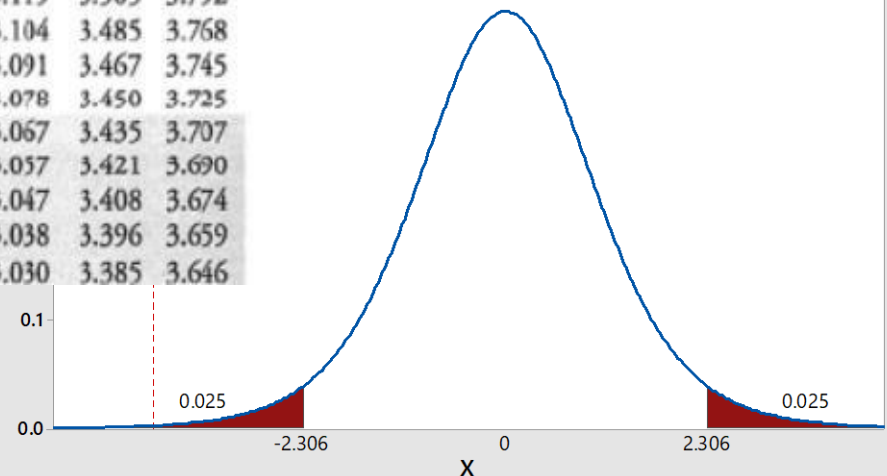
❖ $df = 8$

❖ $t_{\text{crítico}} = 2,306$

Ele pegou o errado

df = degree of freedom

Distribution Plot
T, df=8



Teste t de duas amostras

Mesma variância

17

❑ Exemplo: Amostras envasadas por duas máquinas têm apresentam os seguintes volumes:

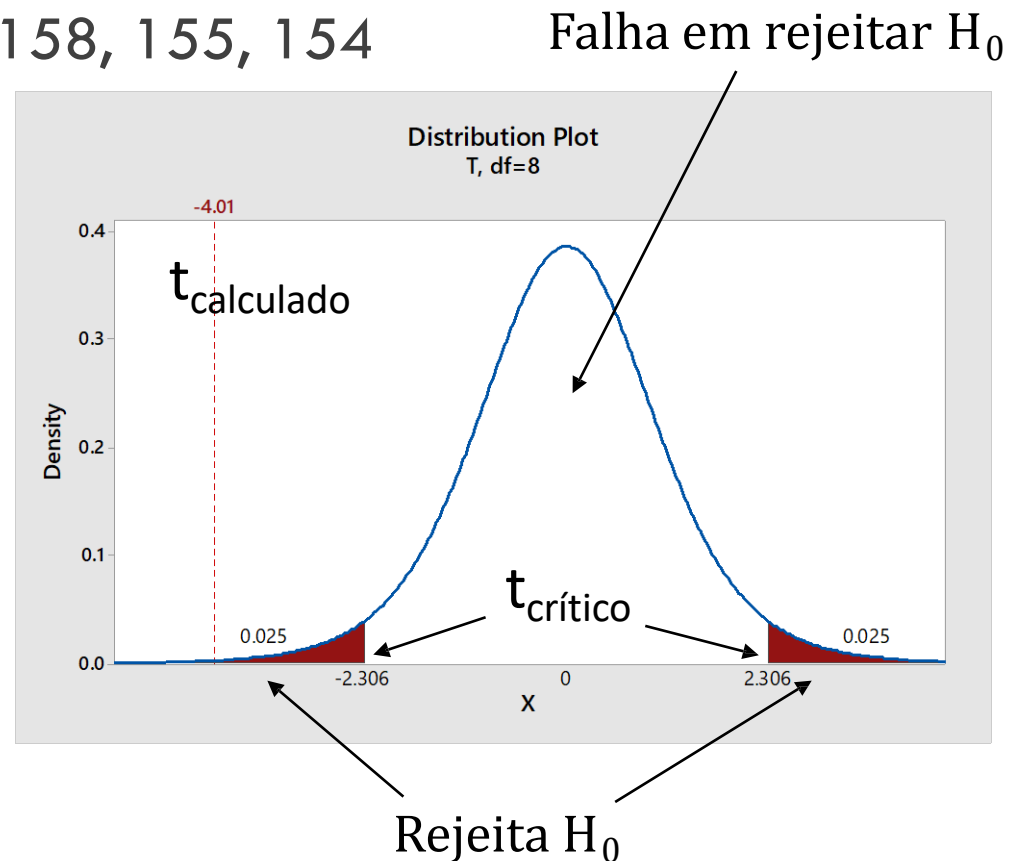
❑ Máquina A: 150, 152, 154, 152, 151

❑ Máquina B: 156, 155, 158, 155, 154

❑ A média é diferente?

❑ $t_{\text{cal}} = -4,01$

❑ $t_{\text{crítico}} = 2,306$



Teste t de duas amostras

Variâncias diferentes

18

- ❑ Exemplo: Amostras envasadas por duas máquinas têm apresentam os seguintes volumes:
 - ❑ Máquina A: 150, 152, 154, 152, 151
 - ❑ Máquina C: 144, 162, 177, 150, 140
 - ❑ A média é diferente? Calcular com 95% de confiança
-
- | | | |
|-----------------------|-----------------------|---|
| ❑ $n_1 = 5$ | ❑ $n_2 = 5$ | $H_0: \mu_A = \mu_B$
$H_a: \mu_A \neq \mu_B$ |
| ❑ $s_1 = 1,48$ | ❑ $s_2 = 15,0$ | |
| ❑ $\bar{x}_1 = 151,8$ | ❑ $\bar{x}_2 = 154,6$ | |

Teste t de duas amostras

Variâncias diferentes

19

$$\square n_1 = 5$$

$$\square s_1 = 1,48$$

$$\square \bar{x}_1 = 151,8$$

$$\square n_2 = 5$$

$$\square s_2 = 15,0$$

$$\square \bar{x}_2 = 154,6$$

$$H_0: \mu_A = \mu_B$$

$$H_a: \mu_A \neq \mu_B$$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}} = \frac{151,8 - 154,6}{\sqrt{\left(\frac{1,48^2}{5} + \frac{15,0^2}{5}\right)}} = -0,41$$

Teste t de duas amostras

Variâncias diferentes

20

$$\square n_1 = 5$$

$$\square n_2 = 5$$

$$\square s_1 = 1,48$$

$$\square s_2 = 15,0$$

$$\square \bar{x}_1 = 151,8$$

$$\square \bar{x}_2 = 154,6$$

$$H_0: \mu_A = \mu_B$$

$$H_a: \mu_A \neq \mu_B$$

$$df = \frac{\left[\frac{1}{n_1} + \frac{2}{n_2} \right]}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}} = 4,078 = 4$$

df	TAIL PROBABILITY P											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646

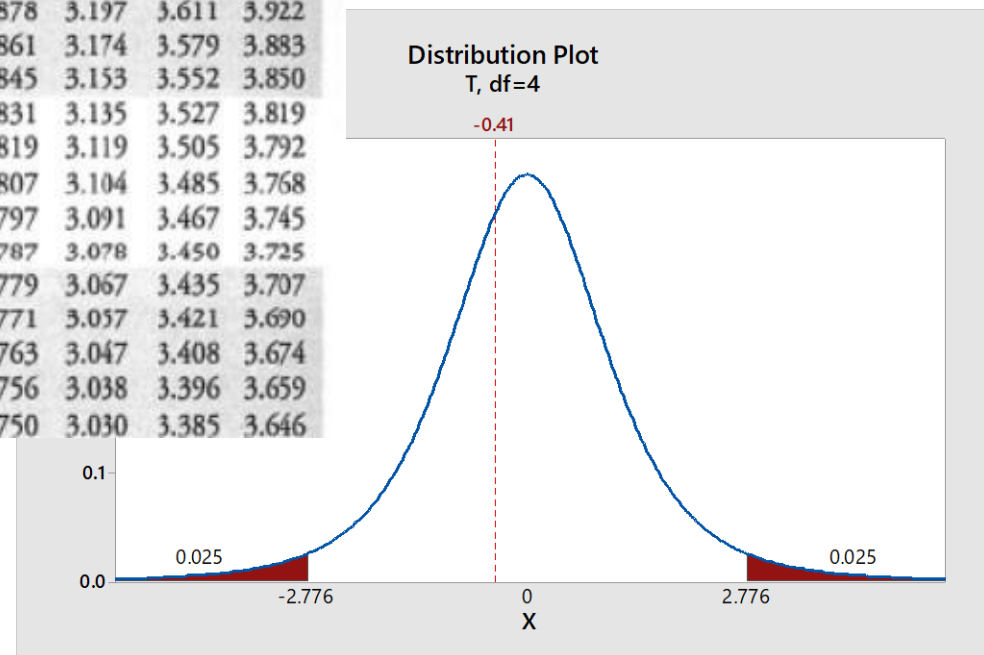
❖ $\alpha = 0.05$ Two Tails

❖ $df = 4$

❖ $t_{\text{crítico}} = 2,776$

Pegou o errado dnv

df = degree of freedom



Teste t de duas amostras

Variâncias diferentes

22

- Exemplo: Amostras envasadas por duas máquinas têm apresentam os seguintes volumes:

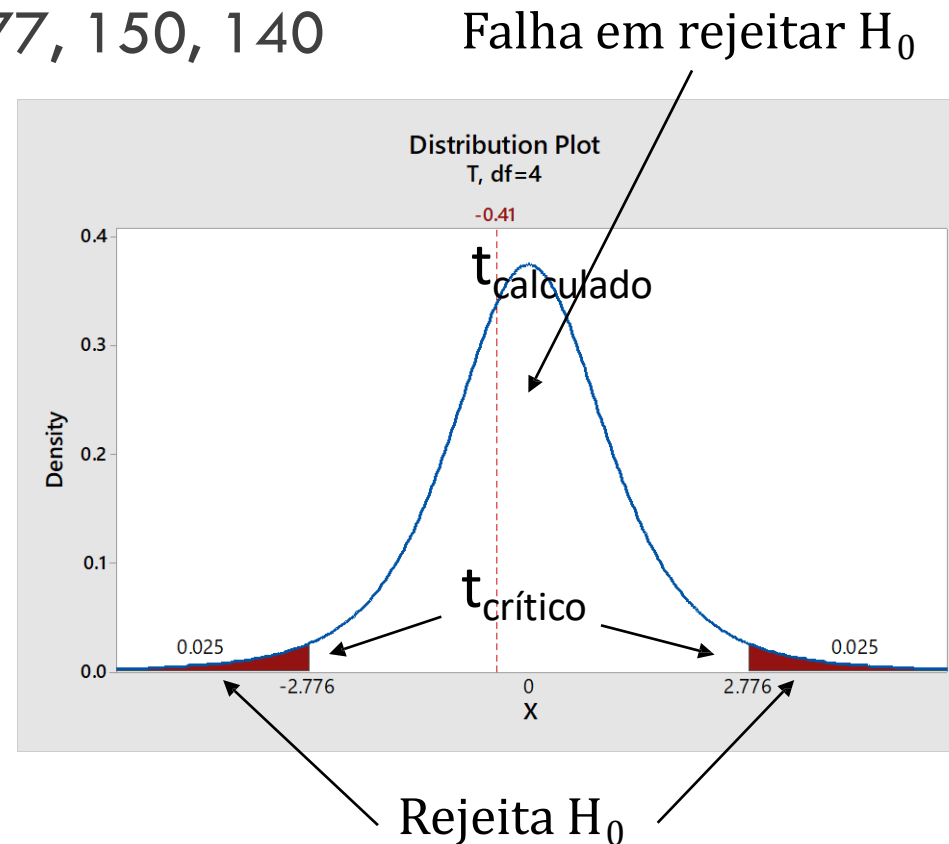
- Máquina A: 150, 152, 154, 152, 151

- Máquina C: 144, 162, 177, 150, 140

- A média é diferente?

- $t_{\text{cal}} = -0,41$

- $t_{\text{crítico}} = 2,776$



Teste t de duas amostras

Interpretação dos resultados

23

Machine A	Machine B	Machine C
150	156	144
152	155	162
154	158	177
152	155	150
151	154	140
$\bar{x}_A = 151.8$	$\bar{x}_B = 155.6$	$\bar{x}_C = 154.6$

