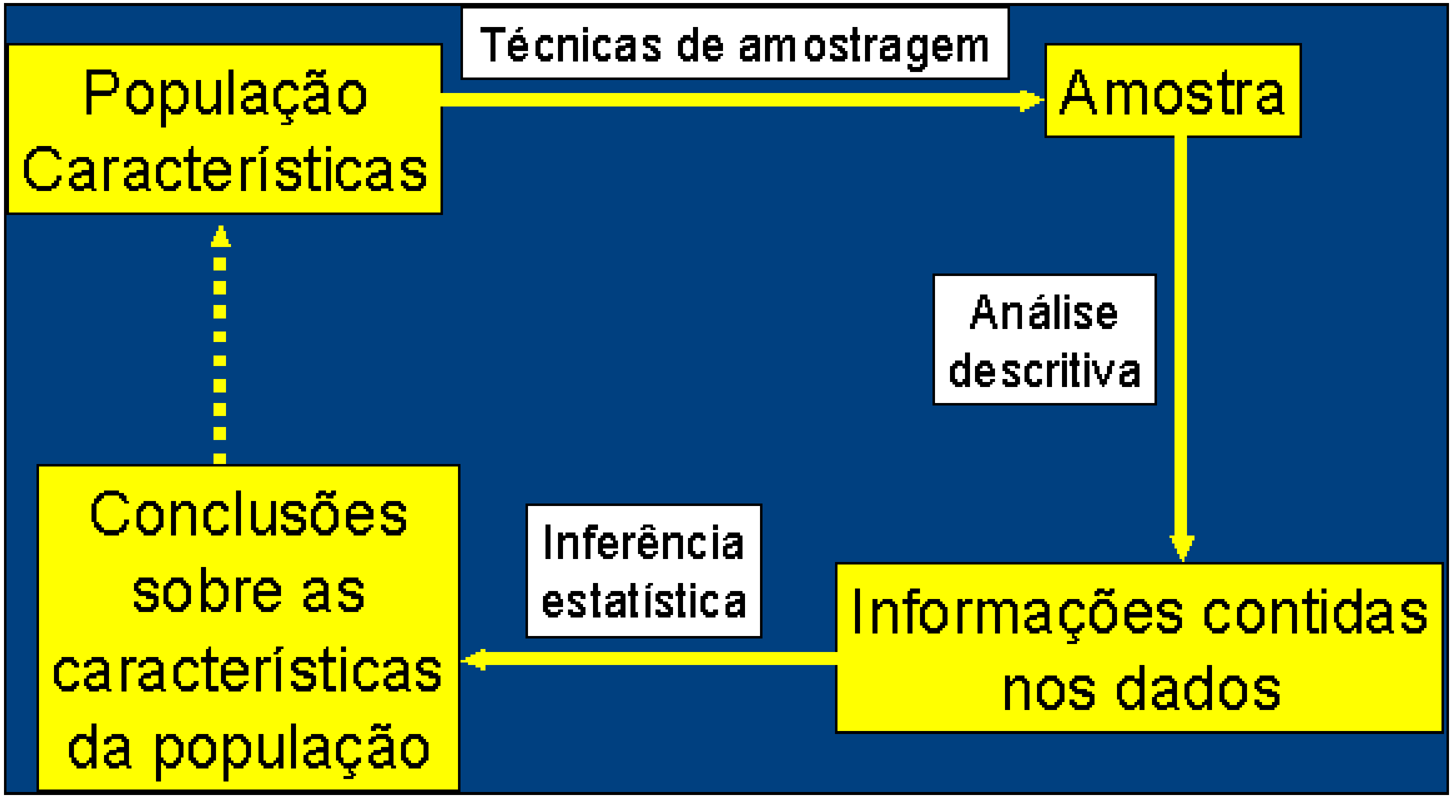
INFERÊNCIA ESTATÍSTICA

Teste de Hipóteses

Ana Amélia Benedito Silva

Etapas da Analise Estatística



ANÁLISE DESCRITIVA

- conjunto de técnicas que tem como objetivo descrever uma amostra extraída de uma população.
  - Tabelas
  - Gráficos
  - Medidas-resumo
    - medidas de tendência central
      - média, mediana, moda
    - medidas de dispersão
      - amplitude, desvio-padrão, erro-padrão
    - medidas separatrizes
      - percentis, quartis, decis

INFERÊNCIA ESTATÍSTICA

- Conjunto de técnicas que tem como objetivo estudar uma população através de evidências fornecidas por uma amostra.
  - Teste de hipóteses
  - Estimação por parâmetros ou intervalo de confiança

- Permite ao pesquisador ir além da descrição dos dados

Inferência estatística

<u>Estimação</u>

- Qual é a probabilidade de "cara" no lançamento de uma moeda?
- Qual é a media da altura dos brasileiros?
- Qual é a porcentagem de votos que o candidato A vai receber nas eleições?
- Qual é a porcentagem de adultos que já tomaram as 4 doses de vacina pra COVID-19 no Brasil?

<u>Teste de hipóteses</u>

- A moeda é honesta?
- Será que a média da altura dos brasileiros é maior que 1,65m?
- O candidato A vencerá as eleições?
- Será que pelo menos 50% dos adultos já tomou as 4 doses de vacina para COVID-19?

TESTE DE HIPÓTESES

Será que a média da altura dos brasileiros é maior que 1,65m?

- Para responder a esta questão escolhe-se estrategicamente uma amostra $(x_1,x_2,...,x_n)$ que seja representativa da população de adultos brasileiros e verifica-se se $\mu>1{,}65$m, com alta probabilidade.

TESTE DE HIPÓTESES

PERGUNTA

↓

HIPÓTESES ESTATÍSTICAS

⇒

média da altura dos brasileiros > 1,65m?

↓

$H_0 : \mu = 1{,}65m$

$H_1 : \mu > 1{,}65m$

HIPÓTESES ESTATÍSTICAS

H_0 : Hipótese de igualdade ou nulidade
H_1 : Hipótese alternativa

- Aplicar um teste de hipóteses significa calcular as probabilidades de errar ao se aceitar ou rejeitar a hipótese de nulidade H_0
- A decisão é sempre tomada em relação à H_0:

Aceita-se ou rejeita-se H_0

Exemplo 1 – pacotes de café

Exemplo 1 – pacotes de café

- **Situação**

Uma máquina automática enche pacotes de café.

Sabe-se que a distribuição de probabilidade do peso destes pacotes segue uma **normal** com média de 500g e desvio-padrão de 20g.

Deseja-se verificar se a máquina está calibrada sem interromper a produção.

- **Evidência amostral**

Para testá-la um técnico colhe uma amostra com 16 pacotes a cada 30 minutos.

Suponha que as médias das amostras de café sejam iguais à 490g.

A máquina está descalibrada ou a diferença encontrada foi devida ao acaso?

Região crítica

- Conjunto de valores assumidos pela variável dependente ou estatística do teste para os quais a hipótese H_0 é rejeitada

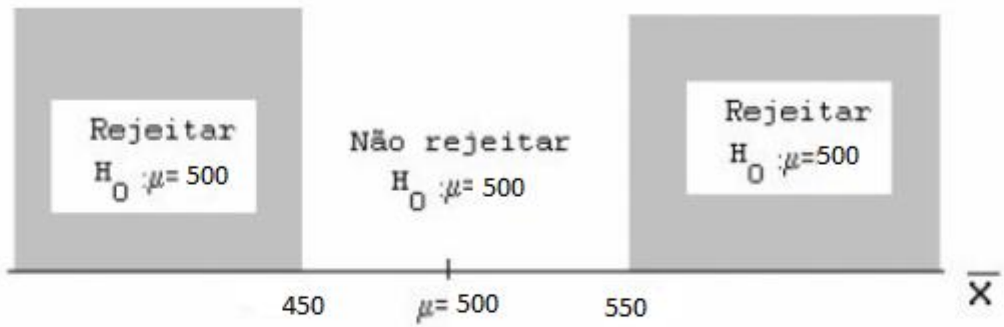
- Se a máquina estiver descalibrada, isto é, se a média for diferente de 500g, espera-se que a média amostral $\bar{x}$ seja inferior ou superior a 500g

- Suponha que a equipe técnica tenha decidido adotar a seguinte regra:
  rejeitar H_0 se a média amostral $\bar{x}$ for maior que 550g e/ou menor que 450g.

- $R_c = \{\bar{x} > 550$ <u>ou</u> $\bar{x} < 450\}$
  → Região de rejeição de Ho

- $R_a = \{450 <= \bar{x} <= 550\}$
  → Região de aceitação de Ho

Região crítica



Procedimento (teste)

$$Se\ \ \bar{x} \in R_c \Rightarrow \text{Rejeita - se } \mathrm{H}_0$$

$$Se\ \ \bar{x} \notin R_c \Rightarrow \textit{Ace}\text{ita - se } \mathrm{H}_0$$

Tipos de erro num teste estatístico

	Realidade no lote	
DECISÃO DO TÉCNICO	**H_0 é verdadeira:** Máquina está calibrada	**H_0 é falsa:** Máquina não está calibrada
H_0 é verdadeira: Máquina está calibrada	**Decisão correta** Probabilidade= 1- α	**Decisão errada** **Erro β**
H_0 é falsa: Máquina não está calibrada	**Decisão errada** **Erro α**	**Decisão correta** Probabilidade= 1- β

α = P (Erro tipo I) = chamado de nível de significância (em geral 5%)
risco máximo aceitável de errar ao dizer que H_0 é falsa
β = P (Erro tipo II)
risco máximo aceitável de errar ao dizer que H_0 é verdadeira

Tipos de erro num teste estatístico

P(Erro tipo I)=α (**nível de significância**)

$$\alpha = P(\text{Rejeitar } \mathrm{H}_0 \mid H_0 \text{ verdadeira})$$

$$P(\mathit{Erro\ II}) = \beta = P(\text{Não rejeitar} \mathrm{H}_0 \mid H_0 \text{ falso}).$$

$$1-\beta = P(\text{Rejeitar} \mid H_0 \text{ é falso}).$$ $\rightarrow$**Poder do teste**

Tipos de erro num teste estatístico

- No exemplo dos pacotes de café, selecionamos uma amostra de 16 pacotes e obtivemos uma média de 490g.
- Essa média da amostra é compatível com a média suposta de 500g?
- E se selecionarmos uma amostra com média 450g? Ou uma outra com média 550g?

- **Pergunta:**
  - Quão distante da média populacional = 500g precisa a média amostral se localizar antes que possamos concluir que esta amostra refere-se à outra população de pacotes de café?

Tipos de erro num teste estatístico

- Supondo H_0 verdadeira, se a probabilidade for “pequena” da média da amostra ser igual à 490g ou 450g ou 550g, rejeitamos H_0, ou seja estas amostras não são compatíveis com uma população com média = 500g.

- Em consequência, poder-se-ia concluir que a média da população de pacotes não pode ser 500g.

Tipos de erro num teste estatístico

O que é uma probabilidade “pequena”?

- Na maioria das aplicações, utiliza-se $\alpha = 0{,}05$.
- Mais conservativos, escolhem $\alpha = 0{,}01$.
- Menos conservativos, escolhem $\alpha = 0{,}10$.

A probabilidade α que escolhemos (0,05; 0,01; 0,10...) é conhecida como *nível de significância do teste* de hipótese.

Tipos de erro num teste estatístico

O que é o p-value?

- É chamado de nível descritivo (p-value).
- O p-value é comparado ao α pré-determinado, para decidir se a H_0 deve ser rejeitada ou não.
- É a probabilidade de se obter uma média igual ou mais extrema do que a média da amostra observada, supondo H_0 verdadeira.

Tipos de erro num teste estatístico

- Como decidir?
  - Se p-value ≤ α → rejeitamos H_0
  - Se p-value > α → aceitamos H_0

- Para conduzir um teste de hipótese usamos a distribuição amostral da média.
- Quando a população é normal com desvio-padrão conhecido ou *n* suficientemente grande, utilizamos *a estatística z* (*segue uma distribuição z)*.
- Quando o desvio-padrão da população não é conhecido, substituímos pelo valor da amostra *s*; e se a população original seguir uma distribuição normal, utilizamos *a estatística t*. (*segue uma distribuição t*)

Passos para realizar teste de hipóteses

- Passo 1 – Determinar as hipóteses
- Passo 2 - Escolha da estatística do teste
- Passo 3 - Determinação da **Região Crítica** para α=5%
- Passo 4 – Calcular a estatística do teste para os dados amostrais
- Passo 5 – Concluir pela aceitação ou rejeição de H0, comparando o valor obtido no Passo 4 com a Região de Aceitação ou Região Crítica.

Voltando aos pacotes de café

Abordagem pela região de aceitação

Passo 1 - Determinação das hipóteses

$H_0 : \mu = 500g$

$H_1 : \mu \neq 500g$

μ representa a média do peso da população de pacotes

O técnico deve determinar a probabilidade de se observar uma diferença tão grande quanto 10g ao acaso se a média populacional da máquina for de fato igual à 500g.

Abordagem pela região de aceitação

Passo 2 - Escolha da estatística do teste é:

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

Passo 3 - Determinação da Região crítica para α=5%

$$R_c = \{ z \in Z \mid |z| \geq 1{,}96 \}$$

$$\mathbf{z_{\alpha=0,025} = -1{,}96}$$

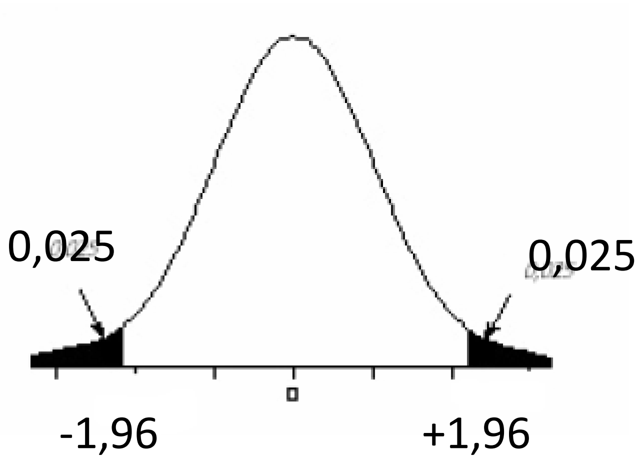


Tabela distribuição normal

Cada casa na tabela dá a proporção sob a curva inteira entre z – 0 e um valor positivo de z. As áreas para os valores de z negativos são obtidas por simetria.

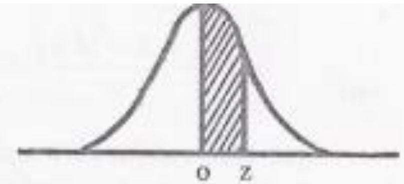
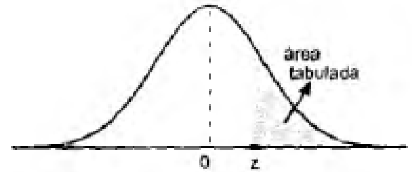


Tabela A-2 – Áreas de uma distribuição normal padrão

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3261	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

TABELA IV Distribuição normal padrão.



	segunda decimal de z									
z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2842	0,2810	0,2776
0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0722	0,0708	0,0694	0,0681
1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
1,8	0,0359	0,0352	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
2,9	0,0019	0,0018	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
3,0	0,00135									
3,5	0,000 233									
4,0	0,000 031 7									
4,5	0,000 003 40									
5,0	0,000 000 287									

Abordagem pela região de aceitação

Passo 4 – Calcular a estatística do teste para os dados amostrais

$$Z = \frac{(\overline{X} - \mu)}{\sigma / \sqrt{n}} = \frac{(490 - 500)}{20/\sqrt{16}} = -2 < z_{\alpha=(0,05/2)} = -1,96$$

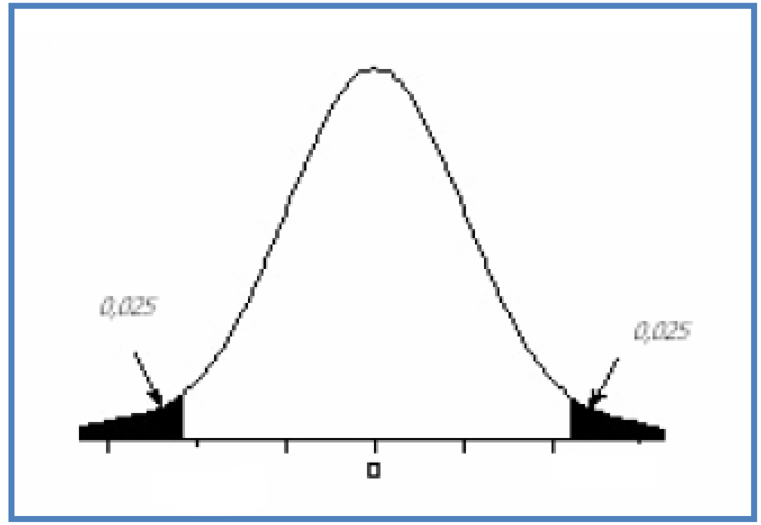
Passo 5 – Conclusão

z_{obs} caiu fora da região de aceitação de H_0

A máquina está descalibrada, a um nível de significância de 5%.

Abordagem pelo nível descritivo (p_value)

Como foi dito a média amostral tem distribuição normal
média populacional = μ
desvio-padrão amostral= $\frac{\sigma}{\sqrt{n}}$



Neste exemplo da máquina de café:
media amostral $\bar{x}$= 490g
desvio-padrão amostral = σ= 20/ $\sqrt{16}$ = 5

p-value = P{$\bar{x}$ ≠490} = P {z ≠ ((490-500)/5)} = P {z ≠ -2} = 0,023

Fixa-se α = 0,05
P < α , ou seja, a Probabilidade P de se aceitar H_o é inferior a 0,05

Logo a máquina está descalibrada!!!

Tabela distribuição normal

Cada casa na tabela dá a proporção sob a curva inteira entre z – 0 e um valor positivo de z. As áreas para os valores de z negativos são obtidas por simetria.

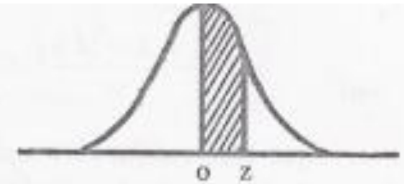
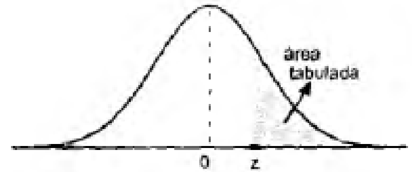


Tabela A-2 – Áreas de uma distribuição normal padrão

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3261	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

TABELA IV Distribuição normal padrão.



z	segunda 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	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2842	0,2810	0,2776
0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0722	0,0708	0,0694	0,0681
1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
1,8	0,0359	0,0352	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
2,9	0,0019	0,0018	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
3,0	0,00135									
3,5	0,000 233									
4,0	0,000 031 7									
4,5	0,000 003 40									
5,0	0,000 000 287									

Exemplo 2 – teste vocacional
teste t de Student

Exemplo 2

- Os registros dos últimos anos de um colégio atestam para os calouros admitidos uma nota média num teste de QI = 115.
- Para testar a hipótese de que a média de uma nova turma é a mesma das turmas anteriores, retirou-se, ao acaso, uma amostra de 20 notas, obtendo-se média 118 e desvio-padrão 20.

- Dados populacionais:

  μ = 115; σ = desconhecido

- Dados amostrais:

  $\bar{x}$ = 118; s = 20; n = 20

Passos para realizar teste de hipóteses

- Passo 1 – Determinar as hipóteses
- Passo 2 - Escolha da estatística do teste
- Passo 3 - Determinação da Região Crítica para α=5%
- Passo 4 – Calcular a estatística do teste para os dados amostrais
- Passo 5 – Concluir pela aceitação ou rejeição de H_0, comparando o valor obtido no Passo 4 com a RA ou RC.

Passos para realizar teste de hipóteses

- Passo 1 – Determinar as hipóteses

$$H_0 : \mu = 115$$
$$H_1 : \mu \neq 115$$

μ representa a média da nota da população dos últimos anos

Passos para realizar teste de hipóteses

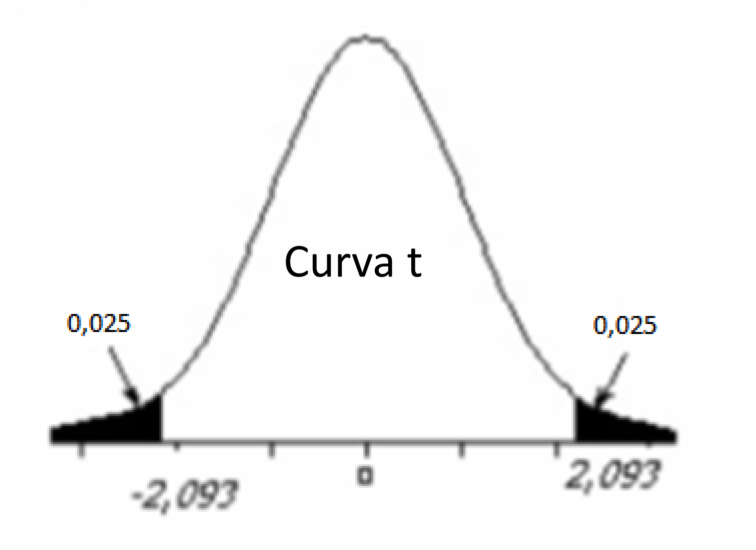
- Passo 2 - Escolha da estatística do teste

Como não conhecemos o desvio padrão populacional, utilizamos uma estatística t ao invés de uma estatística z.

$$T = \frac{\overline{X} - 115}{S/\sqrt{n}} \underset{sob\ H_0}{\sim} t(n-1)$$

Passos para realizar teste de hipóteses

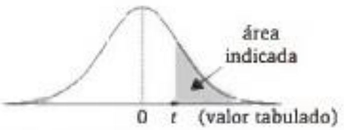
- Passo 3 - Determinação da Região crítica para α=5%



$$R_c = \{ t \in T \mid |T| \geq 2{,}093 \}$$

graus de liberdade = n - 1

Tabela 4 *Distribuição t de Student.*



gl	Área na cauda superior								
	0,25	**0,10**	**0,05**	**0,025**	**0,01**	**0,005**	**0,0025**	**0,001**	**0,0005**
1	1,000	3,078	6,314	12,71	31,82	63,66	127,3	318,3	636,6
2	0,816	1,886	2,920	4,303	6,965	9,925	14,09	22,33	31,60
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,21	12,92
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,894	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819

Passos para realizar teste de hipóteses

- Passo 4 – Calcular a estatística do teste para os dados amostrais

$$T_{obs} = \frac{118 - 115}{20/\sqrt{20}} = 0,67$$

- Passo 5 – Concluir pela aceitação ou rejeição de H_0, comparando o valor obtido no Passo 4 com a RA ou RC.

  T_{obs} = 0,67 valor que pertence à Região de Aceitação de H_o

  Logo concluímos que a média da nova turma é a mesma das turmas anteriores

Exercício para fazer na aula

A média da concentração de colesterol no sangue para a população de homens de 20 a 74 anos é 211 mg/100ml.

Suponhamos que a distribuição da concentração de colesterol no sangue para a população de homens fumantes hipertensos é aproximadamente normal (média desconhecida e desvio-padrão = 46mg/100ml)

Selecionamos uma amostra de 12 homens desse grupo de fumantes hipertensos e o colesterol foi de 217 mg/100ml.

Essa média da amostra é compatível com a média populacional de 211 mg/100ml?

Solução

Abordagem pela região de aceitação

Passo 1 - Determinação das hipóteses

$H_0 : \mu = 211$ mg/100ml

$H_1 : \mu \neq 211$ mg/100ml

μ representa a média do colesterol na população de homens de 20 a 74 anos é 211 mg/100ml.

Abordagem pela região de aceitação

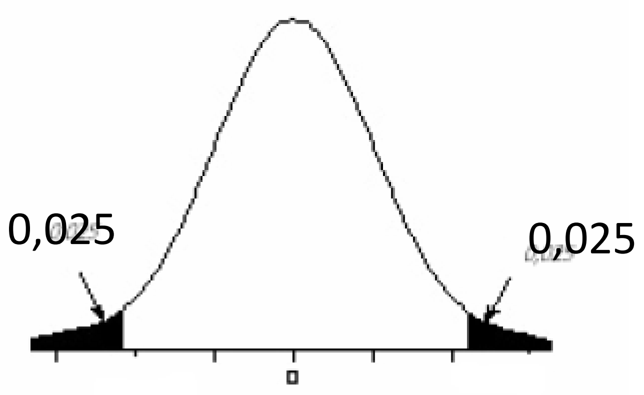
Passo 2 - Escolha da estatística do teste é:

Como conhecemos o desvio padrão populacional, utilizamos a estatística z.

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

Passo 3 - Determinação da Região crítica para α=5%

$z_{\alpha=0,025} = \pm\ 1,96$
(da tabela da curva normal)

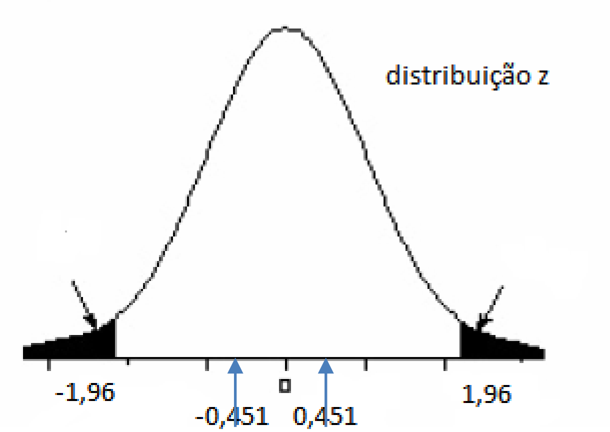


Abordagem pela região de aceitação

Passo 4 – Calcular a estatística do teste para os dados amostrais

$$Z = \frac{(\bar{X} - \mu)}{\sigma / \sqrt{n}} = \frac{(217 - 211)}{46/\sqrt{12})}$$

$$z_{observado} = 0{,}451$$



Passo 5 – Conclusão

$z_{observado}$ caiu dentro da região de aceitação de H_0

Logo → aceitamos H_0

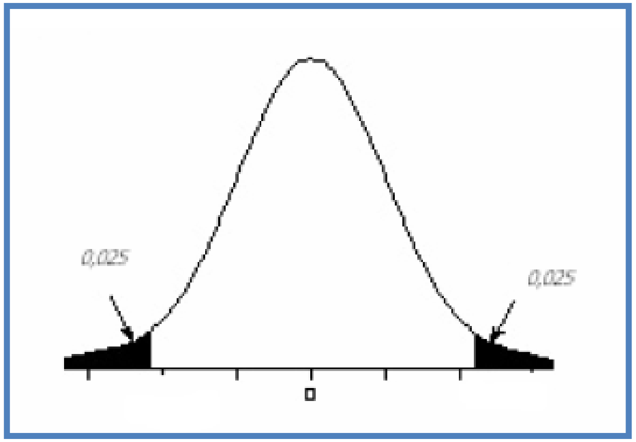
ou seja, a evidência observada na amostra é insuficiente para concluir que o nível médio de colesterol da população de fumantes hipertensos é diferente de 211 mg/100ml.

Abordagem pelo nível descritivo (p_value)

Como foi dito a média amostral tem distribuição normal

média populacional= μ

desvio-padrão amostral= $\dfrac{\sigma}{\sqrt{n}}$



media amostral = 217 mg/100ml

desvio-padrão amostral = $46/\sqrt{12}$ = 13,279

$Z_{observado}$= (217 – 211)/(13,279) = 0,451

Na tabela da normal: p-value = P(z>0,451 ou z<-0,451) = 2x0,3264 = 0,6528

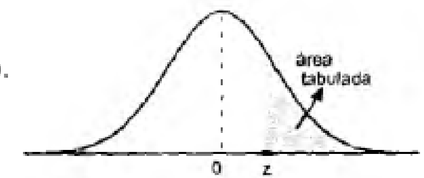
Se α = 0,05

p-value > α, ou seja, a Probabilidade P (p-value) de se aceitar H_o é maior que 0,05

Logo → aceitamos H_0

Tabela distribuição normal

TABELA IV Distribuição normal padrão.



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
	segunda decimal de z									
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2842	0,2810	0,2776
0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0722	0,0708	0,0694	0,0681
1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
1,8	0,0359	0,0352	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
2,9	0,0019	0,0018	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
3,0	0,00135									
3,5	0,000 233									
4,0	0,000 031 7									
4,5	0,000 003 40									
5,0	0,000 000 287									

Exemplo 3 – insônia
teste do quiquadrado

Exemplo 3 - insônia

- Numa pesquisa sobre queixas de insônia, observou-se dentre 50 entrevistados, que 31 eram mulheres e 19 eram homens. Pode-se afirmar que a proporção entre homens e mulheres é 1:1 nesta população?

Passos para realizar teste de hipóteses

- Passo 1 – Determinar as hipóteses
- Passo 2 - Escolha da estatística do teste
- Passo 3 - Determinação da Região crítica para α=5%
- Passo 4 – Calcular a estatística do teste para os dados amostrais
- Passo 5 – Concluir pela aceitação ou rejeição de H0, comparando o valor obtido no Passo 4 com a Região de Aceitação ou a Região Crítica.

Passos para realizar teste de hipóteses

- Passo 1 – Determinar as hipóteses

H_o : proporção de homens = proporção de mulheres
(segue uma distribuição 1:1)
H_1 : proporção de homens ≠ proporção de mulheres
(não segue uma distribuição 1:1)

$\alpha = 5\%$

Passos para realizar teste de hipóteses

- Passo 2 - Escolha da estatística do teste

Determinação da variável dependente
variável dependente: sexo com 2 categorias

Tipo da variável dependente
sexo é uma variável qualitativa nominal

N° de Amostras
1 amostra

Relacionamento entre as amostras
não se aplica (é uma amostra apenas)

TABELA DE ORIENTAÇÃO NA ESCOLHA DE TESTES ESTATÍSTICOS

	Uma variável					Duas variáveis
Tipo da variável dependente	Uma amostra	Duas amostras		Mais de duas amostras		Medidas de correlação
		relacionadas	independentes	relacionadas	independentes	
Qualitativa nominal ou ordinal	binomial ou X^2	McNemar	X^2 ou Fischer	Prova Q de Cochran	X^2 para várias amostras	coeficiente de contigência C
Quantitativa discreta ou contínua (dados não seguem curva de Gauss)	Kolmogorov Smirnov	Wilcoxon ou Prova dos sinais	Mann-Whitney Ou Prova da Mediana	Prova de Friedman	Kruskal-Wallis ou Prova da mediana	correlação de Spearman
Quantitativa discreta ou contínua (dados seguem curva de Gauss)	teste de proporções	teste t de Student pareado	teste t de Student para amostras independentes	ANOVA para medidas repetidas	ANOVA para grupos independentes	correlação de Pearson

Passos para realizar teste de hipóteses

- Passo 2 - Escolha da estatística do teste
  - Para comparar 2 categorias (F e M) de uma variável qualitativa com apenas uma amostra utiliza-se o teste do qui-quadrado
  - O teste do qui-quadrado segue uma distribuição chamada distribuição qui-quadrado

Passos para realizar teste de hipóteses

- Passo 3 - Determinação da Região crítica para α=5%

Para obtenção da região crítica precisamos calcular os graus de liberdade

gl = 2-1 = 1
Na Tabela: $x^2_{crítico}$ = 3,84

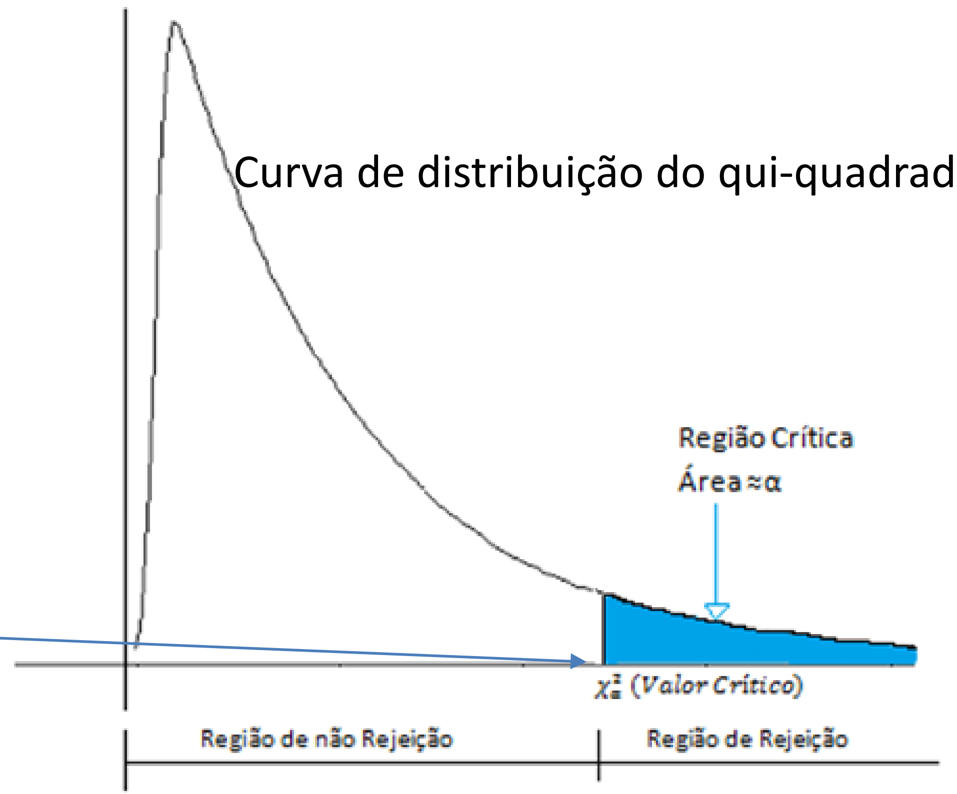
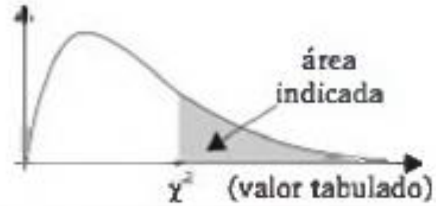


Tabela do qui-quadrado

Tabela 5 (*Continuação*).



gl	Área na cauda superior							
	0,25	**0,10**	**0,05**	**0,025**	**0,01**	**0,005**	**0,0025**	**0,001**
1	1,32	2,71	3,84	5,02	6,63	7,88	9,14	10,83
2	2,77	4,61	5,99	7,38	9,21	10,60	11,98	13,82
3	4,11	6,25	7,81	9,35	11,34	12,84	14,32	16,27
4	5,39	7,78	9,49	11,14	13,28	14,86	16,42	18,47
5	6,63	9,24	11,07	12,83	15,09	16,75	18,39	20,51
6	7,84	10,64	12,59	14,45	16,81	18,55	20,25	22,46
7	9,04	12,02	14,07	16,01	18,48	20,28	22,04	24,32
8	10,22	13,36	15,51	17,53	20,09	21,95	23,77	26,12
9	11,39	14,68	16,92	19,02	21,67	23,59	25,46	27,88
10	12,55	15,99	18,31	20,48	23,21	25,19	27,11	29,59
11	13,70	17,28	19,68	21,92	24,73	26,76	28,73	31,26
12	14,85	18,55	21,03	23,34	26,22	28,30	30,32	32,91
13	15,98	19,81	22,36	24,74	27,69	29,82	31,88	34,53

Passos para realizar teste de hipóteses

- Passo 4 – Calcular a estatística do teste para os dados amostrais

Etapas para o cálculo do qui-quadrado:

1. Determinar valores esperados

	Sexo F	Sexo M
Observado	31	19
Esperado	25	25

Passos para realizar teste de hipóteses

- Passo 4 – Calcular a estatística do teste para os dados amostrais

Etapas para o cálculo do qui-quadrado:

2. Calcular o χ^2

– $\chi 2_{\text{observado}} = \sum_{i=1}^{k}(\text{Observado}_i - \text{Esperadoi})2/\text{Esperadoi})$

– $\chi 2_{\text{observado}} = (31\text{-}25)^2/25 + (19\text{-}25)^2/25 = 2{,}88$

	Sexo F	Sexo M
Observado	31	19
Esperado	25	25

Passos para realizar teste de hipóteses

- Passo 5 – Concluir pela aceitação ou rejeição de H_0

Como $\chi^2_{\,observado} = 2{,}88 < \chi^2_{\,tabelado} = 3{,}84$,

$\chi^2_{\,observado}$ caiu na região de aceitação de H_0

Logo, devemos, a um nível de significância de 5%, aceitar H_0, ou seja, não há evidências que haja diferença na proporção de homens e mulheres com insônia nesta população.

Exercício - fazer durante a aula

Situação

Segundo estatísticas do Ministério da Saúde, a distribuição de doenças mentais em hospitais psiquiátricos brasileiros é a seguinte:

30% de esquizofrenia

30% de dependência de drogas

10% de depressão

30% de outros diagnósticos

Evidencia amostral

Em um hospital psiquiátrico recém-inaugurado, há 200 pacientes psiquiátricos internados com os seguintes diagnósticos:

64 com esquizofrenia

78 com dependência de drogas

16 com depressão

42 com outros diagnósticos

Pergunta: Pode-se considerar que neste hospital a distribuição dos diagnósticos é a preconizada pelo Ministério da Saúde?

Passos para realizar teste de hipóteses

- Passo 1 – Determinar as hipóteses
- Passo 2 - Escolha da estatística do teste
- Passo 3 - Determinação da Região crítica para α=5%
- Passo 4 – Calcular a estatística do teste para os dados amostrais
- Passo 5 – Concluir pela aceitação ou rejeição de H0, comparando o valor obtido no Passo 4 com a RA ou RC.

Passos para realizar teste de hipóteses

- Passo 1 - Determinação das hipóteses:

H_o : proporção de diagnósticos segue o Ministério da Saúde

H_1 : proporção de diagnósticos não segue o Ministério da Saúde

$\alpha = 5\%$

Passos para realizar teste de hipóteses

- Passo 2 – Escolha da estatística do teste

1) Determinação da variável dependente

tipos de doenças psiquiátricas com 4 categorias

2) Tipo da variável dependente

tipo de doença é uma variável qualitativa nominal

3) N° de Amostras

1 amostra

4) Relacionamento entre as amostras

não se aplica

Passos para realizar teste de hipóteses

- Passo 2 – Escolha da estatística do teste

TABELA DE ORIENTAÇÃO NA ESCOLHA DE TESTES ESTATÍSTICOS

	Uma variável					**Duas variáveis**
Tipo da variável dependente	**Uma amostra**	**Duas amostras**		***Mais de duas amostras***		**Medidas de correlação**
		relacionadas	***independentes***	***relacionadas***	***independentes***	
Qualitativa nominal ou ordinal	binomial ou X^2	McNemar	X^2 ou Fischer	Prova Q de Cochran	X^2 para várias amostras	coeficiente de contigência C
Quantitativa discreta ou contínua (dados não seguem curva de Gauss)	Kolmogorov Smirnov	Wilcoxon ou Prova dos sinais	Mann-Whitney Ou Prova da Mediana	Prova de Friedman	Kruskal-Wallis ou Prova da mediana	correlação de Spearman
Quantitativa discreta ou contínua (dados seguem curva de Gauss)	teste de proporções	teste t de Student pareado	teste t de Student para amostras independentes	ANOVA para medidas repetidas	ANOVA para grupos independentes	correlação de Pearson

Passos para realizar teste de hipóteses

- Passo 2 – Escolha da estatística do teste

- Para comparar 4 categorias de uma variável qualitativa nominal, com uma amostra. utiliza-se o *teste do qui-quadrado*
- O teste do qui-quadrado segue uma distribuição chamada distribuição qui-quadrado

Passos para realizar teste de hipóteses

- Passo 3 - Determinação da Região crítica para α=5%

Para obtenção da região crítica precisamos calcular os graus de liberdade

gl = 4-1 = 3
Na Tabela: $x^2_{\text{crítico=tabelado}}$ = 7,81

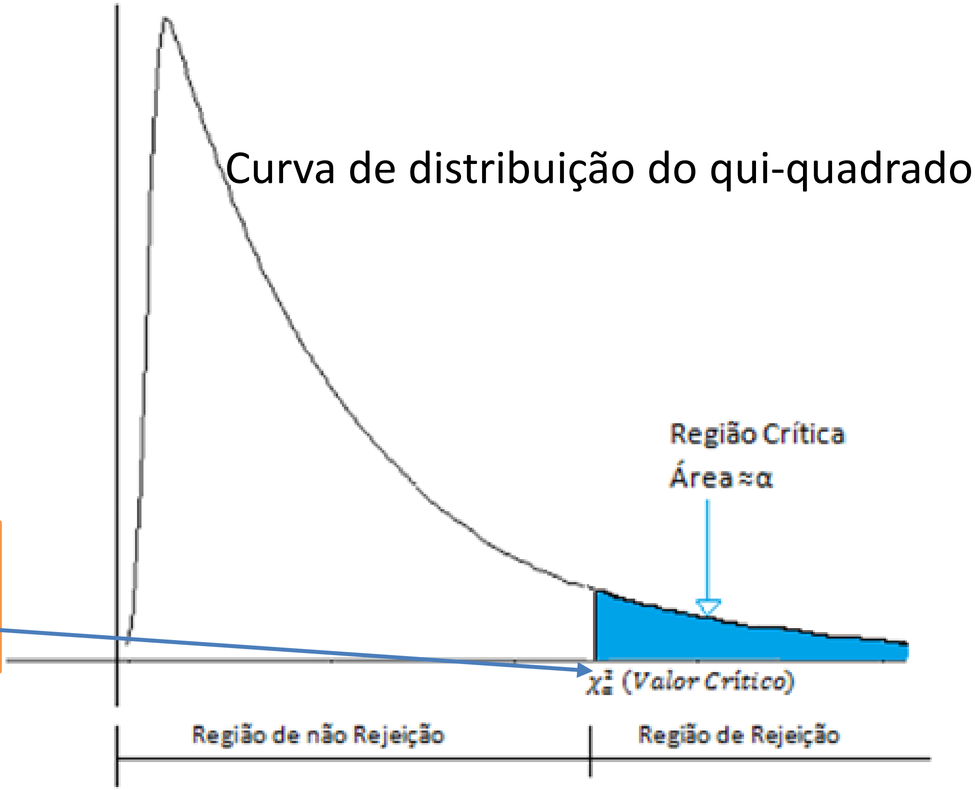
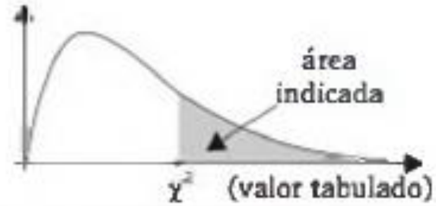


Tabela do qui-quadrado

Tabela 5 (*Continuação*).



gl	Área na cauda superior							
	0,25	**0,10**	**0,05**	**0,025**	**0,01**	**0,005**	**0,0025**	**0,001**
1	1,32	2,71	3,84	5,02	6,63	7,88	9,14	10,83
2	2,77	4,61	5,99	7,38	9,21	10,60	11,98	13,82
3	4,11	6,25	7,81	9,35	11,34	12,84	14,32	16,27
4	5,39	7,78	9,49	11,14	13,28	14,86	16,42	18,47
5	6,63	9,24	11,07	12,83	15,09	16,75	18,39	20,51
6	7,84	10,64	12,59	14,45	16,81	18,55	20,25	22,46
7	9,04	12,02	14,07	16,01	18,48	20,28	22,04	24,32
8	10,22	13,36	15,51	17,53	20,09	21,95	23,77	26,12
9	11,39	14,68	16,92	19,02	21,67	23,59	25,46	27,88
10	12,55	15,99	18,31	20,48	23,21	25,19	27,11	29,59
11	13,70	17,28	19,68	21,92	24,73	26,76	28,73	31,26
12	14,85	18,55	21,03	23,34	26,22	28,30	30,32	32,91
13	15,98	19,81	22,36	24,74	27,69	29,82	31,88	34,53

Passos para realizar teste de hipóteses

- Passo 4 – Escolha da estatística do teste

Etapas para o cálculo do qui-quadrado:

1. Determinar valores esperados

	esquizofrenia	dependência	depressão	outros	TOTAL
Observado	**64**	**78**	**16**	**42**	**200**
Esperado	**60**	**60**	**20**	**60**	**200**

30% de esquizofrenia
30% de dependência de drogas
10% de depressão
30% de outros diagnósticos

Passos para realizar teste de hipóteses

- Passo 4 – Escolha da estatística do teste

Etapas para o cálculo do qui-quadrado:

2. Calcular o χ^2

$\chi 2_{\text{observado}} = \sum_{i=1}^{k}(\text{Observado}_i - \text{Esperadoi})2/\text{Esperadoi})$

$\chi 2_{\text{observado}} = (64\text{-}60)^2/60 + (78\ \text{-}60)^2/60 + (16\text{-}20)^2/20 + (42\text{-}60)^2/60$

$\chi 2_{\text{observado}} = 11{,}87$

	esquizofrenia	dependência	depressão	outros	TOTAL
Observado	64	78	16	42	200
Esperado	60	60	20	60	200

Passos para realizar teste de hipóteses

- Passo 5 – Conclusão

Como $\chi^2_{observado} = 11{,}87 > \chi^2_{tabelado} = 7{,}82$,

há evidências, a um nível de significância (alfa) de 5%, que haja diferença na distribuição dos tipos de doenças no hospital em relação ao preconizado pelo Ministério da Saúde, ou seja, rejeito H_0.

Exemplo 4 – dieta
teste t para amostras pareadas

Exemplo 4 – dieta

- **Situação**

  Um médico acredita que uma dieta de emagrecimento consegue produzir bons resultados em 2 meses.

- **Evidência amostral**

  Para verificar se a dieta é eficiente, foram selecionados 9 pacientes aleatoriamente e pediu a elas que seguissem a dieta por 2 meses.

  Antes de começar a dieta o médico pesou cada paciente e **depois** de 2 meses, pesou-as novamente.

Exemplo 4 – dados amostrais

paciente	Peso antes	Peso depois
1	77	80
2	62	58
3	61	61
4	80	76
5	90	79
6	72	69
7	86	90
8	59	51
9	88	81

Passos para realizar teste de hipóteses

- Passo 1 – Determinar as hipóteses
- Passo 2 - Escolha da estatística do teste
- Passo 3 - Determinação da Região crítica
- Passo 4 – Calcular a estatística do teste para os dados amostrais
- Passo 5 – Concluir pela aceitação ou rejeição de H_0, comparando o valor obtido no Passo 4 com a Região de Aceitação ou com a Região Crítica.

Passos para realizar teste de hipóteses

- Passo 1 – Determinar as hipóteses

$$\begin{cases} H_0 : \mu_{\text{antes}} = \mu_{\text{depois}} \\ H_1 : \mu_{\text{antes}} \neq \mu_{\text{depois}} \end{cases}$$

Fixa-se $\alpha = 0{,}01$

μ_{antes} = média do peso da população de mulheres antes da dieta

μ_{depois} = média do peso da população de mulheres depois da dieta

Passos para realizar teste de hipóteses

- Passo 2 - Escolha da estatística do teste

1)Variável dependente

peso (Kg)

2)Tipo da variável dependente

quantitativa contínua

3)Relacionamento entre as amostras

Dependentes ou relacionadas

4)N° de Amostras

Duas

TABELA DE ORIENTAÇÃO NA ESCOLHA DE TESTES ESTATÍSTICOS

Tipo da variável dependente	Uma variável					Duas variáveis
	Uma amostra	Duas amostras		Mais de duas amostras		Medidas de correlação
		relacionadas	*independentes*	*relacionadas*	*independentes*	
Qualitativa nominal ou ordinal	binomial ou X^2	McNemar	X^2 ou Fischer	Prova Q de Cochran	X^2 para várias amostras	coeficiente de contigência C
Quantitativa discreta ou contínua (dados não seguem curva de Gauss)	Kolmogorov Smirnov	Wilcoxon ou Prova dos sinais	Mann-Whitney Ou Prova da Mediana	Prova de Friedman	Kruskal-Wallis ou Prova da mediana	correlação de Spearman
Quantitativa discreta ou contínua (dados seguem curva de Gauss)	teste de proporções	teste t de Student pareado	teste t de Student para amostras independentes	ANOVA para medidas repetidas	ANOVA para grupos independentes	correlação de Pearson

Passos para realizar teste de hipóteses

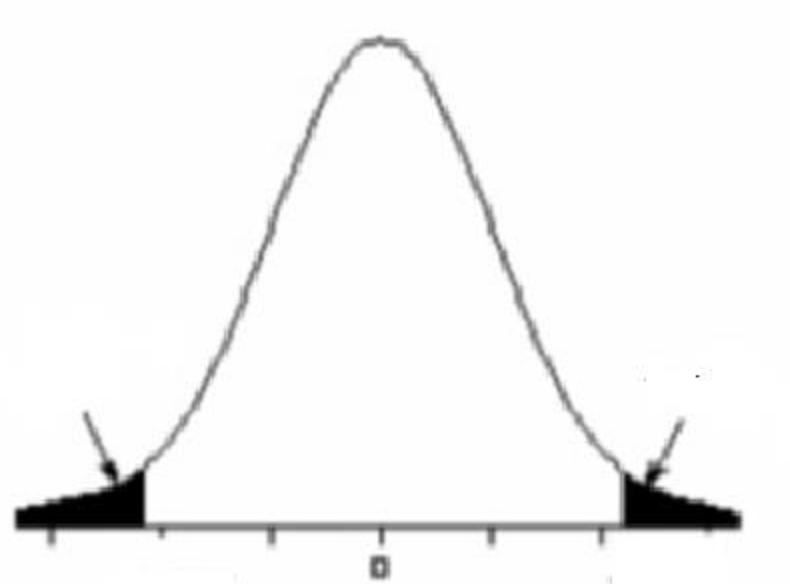
- Passo 2 - Escolha da estatística do teste

- Para comparar as médias de 2 amostras pareadas utiliza-se o teste t pareado
- O teste t pareado segue uma distribuição chamada distribuição t de Student

Passos para realizar teste de hipóteses

- Passo 3 - Determinação da Região crítica para α=1%

Para obtenção da região crítica precisamos calcular os graus de liberdade



Graus de liberdade = 9 – 1 = 8

$t_{crítico}$ = 3,36

Table A.2 The *t*-distribution

Value of t for a confidence interval of Critical value of $\|t\|$ for P values of number of degrees of freedom	90% 0.10	95% 0.05	98% 0.02	99% 0.01
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.89	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
12	1.78	2.18	2.68	3.05
14	1.76	2.14	2.62	2.98
16	1.75	2.12	2.58	2.92
18	1.73	2.10	2.55	2.88
20	1.72	2.09	2.53	2.85
30	1.70	2.04	2.46	2.75
50	1.68	2.01	2.40	2.68
∞	1.64	1.96	2.33	2.58

Atenção

The critical values of $|t|$ are appropriate for a *two*-tailed test. For a *one*-tailed test the value is taken from the column for *twice* the desired P-value, e.g. for a one-tailed test, $P = 0.05$, 5 degrees of freedom, the critical value is read from the $P = 0.10$ column and is equal to 2.02.

Passos para realizar teste de hipóteses

- Passo 4 – Escolha da estatística do teste

  Etapas para o cálculo do t pareado

1) **calcular a diferença entre os valores de cada um dos n pares:**

$$d = x_2 - x_1$$

2) **calcular a média das diferenças**

$$\bar{d} = \frac{\sum d}{n}$$

3) **Calcular a variância**

$$s^2 = \frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1}$$

4) **calcular o valor de t**

$$t = \frac{\bar{d}}{\sqrt{\frac{s^2}{n}}}$$

Passos para realizar teste de hipóteses

- Passo 4 – Escolha da estatística do teste

d = depois - antes
3
-4
0
-4
-11
-3
4
-8
-7

$$\sum d = -30 \text{ kg}$$

$$\bar{d} = -3{,}3 \text{ kg}$$

$$s^2 = 25 \text{kg}^2$$

$$\mathbf{t = -2{,}0}$$

Passos para realizar teste de hipóteses

- Passo 5 – Conclusão

Após obter-se o valor de t, compara-se este valor ao t da tabela

$$t_{\alpha=1\%;\ g.l.=\ 9-1=8} = 3{,}36$$

Como $t_{observado} < t_{tabelado}$ não rejeito H_0

Logo. há evidências, a um nível de significância de 1%, que a dieta não foi eficiente.

Exemplo 5 – técnico de voleibol

Exemplo 5 – técnico de voleibol

- **Situação**

  Um técnico de voleibol acredita que o consumo máximo de oxigênio de jogadores de voleibol e nadadores seja diferente.

- **Evidência amostral**

  Para testar a hipótese, o técnico de voleibol mediu o consumo máximo de oxigênio de 10 atletas da sua equipe e de 10 atletas de uma equipe de natação.

Exemplo 5 – dados amostrais

	Consumo Máximo de Oxigênio (mL/(Kg.min))	
	Voleibol	**Natação**
	52,0	40,0
	49,0	38,0
	47,0	42,0
	48,0	40,0
	52,0	37,0
	53,0	36,0
	46,0	40,0
	54,0	36,0
	60,0	39,0
	59,0	37,0
média	**52,0**	**38,5**
desvio-padrão	**4,8**	**2,0**

Passos para realizar teste de hipóteses

- Passo 1 - Determinação das Hipóteses estatísticas

$$H_0 : \mu_{\text{voleibol}} = \mu_{\text{natação}} \rightarrow \text{consumos iguais}$$
$$H_1 : \mu_{\text{voleibol}} \neq \mu_{\text{natação}} \rightarrow \text{consumos diferentes}$$

Passos para realizar teste de hipóteses

- Passo 2 – Escolha da estatística do teste

1) Variável dependente

consumo máximo de oxigênio

2) Tipo da variável dependente

quantitativa contínua

3) Relacionamento entre as amostras

Independentes – amostra de dados do voleibol e amostra de dados da natação

4) N° de Amostras

2 independentes

Passos para realizar teste de hipóteses

- Passo 2 – Escolha da estatística do teste

TABELA DE ORIENTAÇÃO NA ESCOLHA DE TESTES ESTATÍSTICOS

	Uma variável					Duas variáveis
Tipo da variável dependente	Uma amostra	Duas amostras		Mais de duas amostras		Medidas de correlação
		relacionadas	independentes	relacionadas	independentes	
Qualitativa nominal ou ordinal	binomial ou X^2	McNemar	X^2 ou Fischer	Prova Q de Cochran	X^2 para várias amostras	coeficiente de contigência C
Quantitativa discreta ou contínua (dados não seguem curva de Gauss)	Kolmogorov Smirnov	Wilcoxon ou Prova dos sinais	Mann-Whitney Ou Prova da Mediana	Prova de Friedman	Kruskal-Wallis ou Prova da mediana	correlação de Spearman
Quantitativa discreta ou contínua (dados seguem curva de Gauss)	teste de proporções	teste t de Student pareado	teste t de Student para amostras independentes	ANOVA para medidas repetidas	ANOVA para grupos independentes	correlação de Pearson

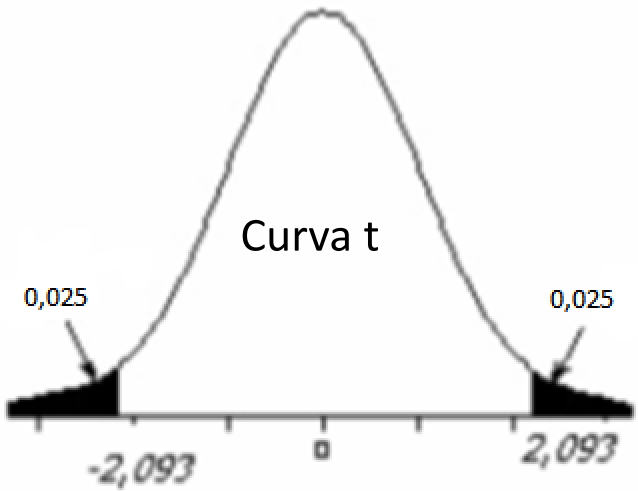
Passos para realizar teste de hipóteses

- Passo 2 – Escolha da estatística do teste

- Para comparar as médias de 2 amostras independentes utiliza-se o **teste t de Student**
- O teste t segue uma distribuição chamada distribuição t

Passos para realizar teste de hipóteses

- Passo 3 - Determinação da Região crítica para α=5%



$$R_c = \{\, t \in T \;|T|\geq 2{,}093 \,\}$$

graus de liberdade = $n_1 - 1 + n_2 - 1$
graus de liberdade = 18

Testes de Significância - Teste *t* de Student

Table A.2 The *t*-distribution

Value of *t* for a confidence interval of Critical value of $\|t\|$ for *P* values of number of degrees of freedom	90% 0.10	95% 0.05	98% 0.02	99% 0.01
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.89	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
12	1.78	2.18	2.68	3.05
14	1.76	2.14	2.62	2.98
16	1.75	2.12	2.58	2.92
18	1.73	2.10	2.55	2.88
20	1.72	2.09	2.53	2.85
30	1.70	2.04	2.46	2.75
50	1.68	2.01	2.40	2.68
∞	1.64	1.96	2.33	2.58

The critical values of $|t|$ are appropriate for a *two*-tailed test. For a *one*-tailed test the value is taken from the column for *twice* the desired *P*-value, e.g. for a one-tailed test, $P = 0.05$, 5 degrees of freedom, the critical value is read from the $P = 0.10$ column and is equal to 2.02.

Passos para realizar teste de hipóteses

- Passo 4 – Calcular a estatística do teste para os dados amostrais

Etapas para o cálculo da estatística t

1)Calcular a média para as 2 amostras $\bar{x}_1$ e $\bar{x}_2$

2)Calcular a diferença entre as médias

3)Calcular a variância conjunta

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

s_1 : desvio-padrão amostra1; n_1 : tamanho amostra1;
s_2 : desvio-padrão amostra2; n_2 : tamanho amostra2

$$T = \frac{\overline{x1} - \overline{x2}}{\sqrt{S_p^2\left(\frac{1}{n} + \frac{1}{m}\right)}} \underset{sob\ H_0}{\sim} t(n+m-2)$$

Passos para realizar teste de hipóteses

- Passo 4 – Calcular a estatística do teste para os dados amostrais

4)Calcular o valor da estatística T

$$T = \frac{\overline{x1} - \overline{x2}}{\sqrt{S_p^2\left(\frac{1}{n} + \frac{1}{m}\right)}} \underset{sob\ H_0}{\sim} t(n+m-2)$$

Numerador = (52,0 - 38,5) = 13,5

S_p^2= [(10−1)∗23,04+(10−1)∗*4,0*)]/(20−2) = 13,52

Denominador = $\sqrt{13{,}52x(\frac{2}{10}}$)= 1,64

$t_{observado}$ = 13,5/1,64= **8,22**

	Consumo de Oxigênio (mL/(Kg.min))	
	Voleibol	Natação
	52,0	40,0
	49,0	38,0
	47,0	42,0
	48,0	40,0
	52,0	37,0
	53,0	36,0
	46,0	40,0
	54,0	36,0
	60,0	39,0
	59,0	37,0
média	**52,0**	**38,5**
desvio-padrão	**4,8**	**2,0**
variância	**23,04**	**4,0**

Numerador = (52,0 - 38,5) = 13,5

$S_p^2 = [(10-1)*4,8^2+(10-1)*2,0^2)]/(20-2) = 13,52$

Denominador = $\sqrt{13,52x(\frac{2}{10})} = 1,64$

Logo: $t_{observado}$ = 13,5/1,64 = 8,22

Testes de Significância - Teste *t* de Student

Table A.2 The *t*-distribution

Value of *t* for a confidence interval of Critical value of $|t|$ for *P* values of number of degrees of freedom	90% 0.10	95% 0.05	98% 0.02	99% 0.01
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.89	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
12	1.78	2.18	2.68	3.05
14	1.76	2.14	2.62	2.98
16	1.75	2.12	2.58	2.92
18	1.73	2.10	2.55	2.88
20	1.72	2.09	2.53	2.85
30	1.70	2.04	2.46	2.75
50	1.68	2.01	2.40	2.68
∞	1.64	1.96	2.33	2.58

The critical values of $|t|$ are appropriate for a *two*-tailed test. For a *one*-tailed test the value is taken from the column for *twice* the desired *P*-value, e.g. for a one-tailed test, $P = 0.05$, 5 degrees of freedom, the critical value is read from the $P = 0.10$ column and is equal to 2.02.

Passos para realizar teste de hipóteses

- Passo 5 – Conclusão

Após obter-se o valor de t, compara-se este valor ao t da tabela

$$t_{\alpha=6\%;\ g.l.=18} = 2,10$$

Como $t_{observado} = 8,22 > t_{tabelado} = 2,10$ devo rejeitar H_0

Logo, há evidências, a um nível de significância de 5%, que os consumos de oxigênio não são iguais.

obrigada