Testes de Hipóteses

VPS126

Teste de hipóteses

- Ferramenta estatística para auxiliar no acúmulo de evidências sobre uma questão

- "Média de glicemia de um grupo de animais é diferente do esperado?"
- "Qual o melhor tipo de dieta para cães diabéticos?"
- "Proporção de crianças daltônicas em uma cidade é a esperada?"
- "Qual o melhor método para inseminação artificial?"
- "Qual a relação entre peso de ovelhas e sua circunferência abdominal?"

Lógica dos testes de hipótese

- Elaborar uma Hipótese Nula (também chamada ***H0***), com a qual é possível prever a probabilidade de amostras aleatórias apresentarem uma certa característica.

- Calcular a probabilidade da amostra analisada ser obtida, dado que a hipótese nula é verdadeira *(valor de p, p-valor)* .

- Comparar essa probabilidade (***valor de p***) com um valor pré-definido (*alfa, nível crítico, Erro Tipo I*).

- Caso a probabilidade seja baixa, **menor** que o valor pré-definido (***alfa***), temos evidências de que a Hipótese Nula seja falsa.

- Caso a probabilidade seja alta, **maior** que o valor pré-definido (***alfa***), não temos evidências de que a hipótese nula seja falsa
  - Atenção: Isso não quer dizer, necessariamente, que H0 é verdadeira.

Valor de *p* ou *P*-valor

- Dos dados pode-se calcular o valor da **estatística do teste** (expressão algébrica para a hipótese que está sendo testada).

  Há uma probabilidade relacionada a este valor da estatística do teste que se chama **valor de *p***.

- O valor de *p* (nível descritivo) descreve a chance de obter o resultado observado (ou um mais extremo) **se a hipótese nula for verdadeira**.

Valor de *p*

- Se o valor de *p* for *muito pequeno*, então é **pouco provável** que tenhamos obtido os resultados observados sendo H_0 verdadeira, então rejeitamos H_0.
- Se o valor de *p* for *muito grande*, então há uma **grande chance** de termos obtido os dados observados sendo H_0 verdadeira, então não rejeitamos H_0.

Significância em alguns programas estatísticos

- muito altamente significante (*** representa $p<0{,}001$)
- altamente significante (** representa $0{,}001<p<0{,}01$)
- significante (* representa $0{,}01<p<0{,}05$)
- não-significante (NS representa $p>0{,}05$)

Cuidado! Esse critério é arbitrário e deve considerado com precaução. A decisão com base no valor de *p* deve ser tomada em função do problema analisado.

Erros Tipo I e Tipo II

A decisão de **rejeitar** ou **não rejeitar** H_0 pode estar incorreta

Erro tipo I: rejeitar H_0 quando H_0 é verdadeira.
[probabilidade de rejeitar H_0 de forma incorreta $\Rightarrow \alpha$]

Erro tipo II: não rejeitar H_0 quando H_0 é falsa.
[probabilidade de cometer um erro tipo II $\Rightarrow \beta$]

Erros Tipo I e Tipo II

Conclusão do teste (baseada na amostra)	"Realidade"	
	H_0 verdadeira	H_0 falsa
Rejeitar H_0	erro tipo I (α)	decisão correta
Não rejeitar H_0	decisão correta	erro tipo II (β)

α e β

Probabilidade de cometer um erro tipo I : **nível de significância do teste (α)**

Ex. Se α =0,05 , há uma chance de 1 em 20 de rejeitar H_0 quando H_0 é verdadeira

Escolha de α:

H_0 **será rejeitada** se $p \leq \alpha$

H_0 **não será rejeitada** se $p > \alpha$

p: Valor de *p* ou P-valor é o nível descritivo (veja adiante)

α e β

Probabilidade de cometer um erro tipo II (β)
(probabilidade de não rejeitar H_0 quando H_0 é falsa)

- Normalmente se pensa em 1- β (poder do teste): probabilidade de rejeitar H_0 quando H_0 é falsa

Escolha do teste

- Como escolher o teste de hipótese mais adequado para a pergunta a ser respondida?
- Diferentes testes de hipóteses foram desenvolvidos para lidar com diferentes situações.
- É necessário checar em quais situações cada teste é aplicável, e verificar se os dados atendem às premissas do teste escolhido.

Lembre-se

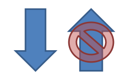
- Sempre dê preferências aos testes Paramétricos: Quando as premissas desses são satisfeitas, eles possuem maior poder (menor erro Tipo II) do que os Não-paramétricos.

Lembre-se

- Variável nominal pode ser utilizada para dividir o conjunto de dados em grupos

Transformação de variáveis

Quantitativa
Qualitativa Ordinal
Qualitativa Nominal



Atenção!

- O que é *estatisticamente significante* pode não ser *biológica* ou *clinicamente significante* e vice-versa.
- *Ex.* Métodos de inseminação artificial (uma economia de 1 ou 2% pode ser uma diferença econômica grande, mas estatisticamente difícil de se obter)
- *Ex.* Dois diferentes anestésicos (pequenas variações na pressão sangüínea; a diferença pode ser estatisticamente significante, mas de pequena importância biológica)

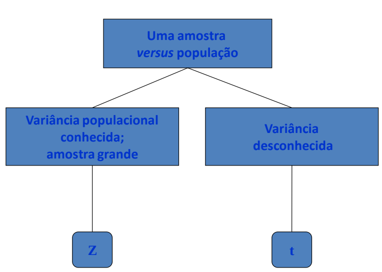
Dois modos de se testar H_0

- Calcula-se a estatística (fórmula) do teste e o valor de *p*

  $\Rightarrow$ rejeita-se H_0 se *p* for pequeno
- Calcula-se IC 95%

  $\Rightarrow$ rejeita-se H_0 se o valor do parâmetro ficar fora dos limites de confiança (para um nível de 5%)

Inferência sobre média de uma amostra de dados com distribuição Normal



Adaptado de Fisher LD, Van Belle G. "Biostatistics: a Methodology for the Health Sciences", Wiley, 1993.

A distribuição dos dados é Normal?

- Se a distribuição dos dados não for Normal, há dois modos de se prosseguir na análise dos dados:
  - Transformar os dados para se aproximar da Normalidade (ex. transformação logarítmica)
  - Teste não-paramétrico (que não faz nenhuma hipótese sobre a distribuição)

Implicações do tamanho da amostra

- amostras pequenas (< 6 observações): é difícil dizer qual a distribuição da variável; podem ser pouco representativas da população
- amostras pequenas (< 30 observações): distribuição de *t* de Student para dados que se distribuem de modo Normal
- amostras grandes: distribuição do teste é Normal (Teorema do Limite Central)

Observação: teste Z e teste t

- No curso, nos casos em que o teste *Z* seria adequado, utilizaremos o teste *t*, que fornece resultados equivalentes.

Teste *t* para uma amostra

- Investigar se a média de um grupo de observações assume um certo valor.
- Exemplo (Petrie e Watson, 1999):
- Questão: Deseja-se saber se suínos em crescimento de um certo lote de uma granja apresentam uma conversão alimentar **média** diária consistente com o ganho **médio** esperado para aquela granja (607 g/dia).

Procedimento do teste

1) Especificar a hipótese nula e a hipótese alternativa:

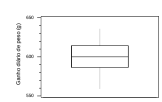
$$H_0 : \mu = 607\ g/dia$$
$$H_1 : \mu \neq 607\ g/dia$$

2) Estatística descritiva e gráfico para verificar a distribuição dos dados (diagrama de pontos, *boxplot*, histograma)

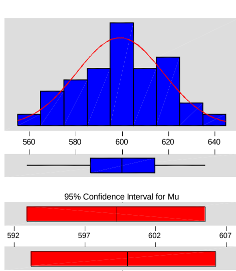
Descriptive Statistics: suinos

Variable	N	Mean	Median	TrMean	StDev	SE Mean
suinos	36	599,19	600,00	599,38	18,66	3,11

Variable	Minimum	Maximum	Q1	Q3
suinos	559,00	636,00	586,50	614,25



Descriptive Statistics



Variable: suinos

Anderson-Darling Normality Test

A-Squared:	0,094
P-Value:	0,997
Mean	599,194
StDev	18,656
Variance	348,047
Skewness	-1,3E-01
Kurtosis	-4,1E-01
N	36
Minimum	559,000
1st Quartile	586,500
Median	600,000
3rd Quartile	614,250
Maximum	636,000

95% Confidence Interval for Mu

592,882 605,507

95% Confidence Interval for Sigma

15,132 24,336

95% Confidence Interval for Median

593,206 606,265

Testa a Normalidade da distribuição de dados

Teste de Anderson-Darling:

H_0: Distribuição é Normal

H_1: Distribuição não é Normal

Para α = 0,05 = 5%:

Como p=0,997, p > $\alpha \Rightarrow$ Não se rejeita H_0, ou seja, assumimos que a distribuição seja Normal

3) Calcular a estatística (fórmula) do teste: $t = \dfrac{\bar{x} - \mu}{s_m}$, onde $s_m = \dfrac{s}{\sqrt{n}}$

t=-2,51

4) Obter o valor de *p* :

p=0,017 . Como p<2%, há uma chance de menos de 2% de se obter um ganho médio diário de 599,2 g/dia se H_0 for verdadeira.

5) Decidir se rejeita ou não a hipótese nula H_0 :

É pouco provável que H_0 seja verdadeira. Ou seja, os dados são inconsistentes com um ganho médio diário de 607 g. Para α=0,05: como p<α, rejeitamos H_0 para um nível de significância de 5%.

6) Determinar, se quiser, o intervalo de confiança de 95%

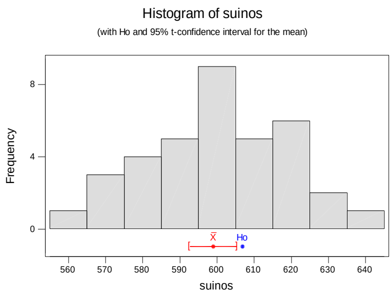
IC 95% : (592,88 ; 605,51). O IC95% não contém o valor testado (607g/dia), confirmando a rejeição de H_0.

One-Sample T: suinos

Test of mu = 607 vs mu not = 607

Variable	N	Mean	StDev	SE Mean
suinos	36	599,19	18,66	3,11

Variable	95,0% CI	T	P
suinos	(592,88; 605,51)	-2,51	0,017



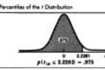


TABLE E Percentiles of the t Distribution

d.f.	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	3.078	6.3138	12.706	31.821	63.657
2	1.886	2.9200	4.3027	6.965	9.9248
3	1.638	2.3534	3.1825	4.541	5.8409
4	1.533	2.1318	2.7764	3.747	4.6041
5	1.476	2.0150	2.5706	3.365	4.0321
6	1.440	1.9432	2.4469	3.143	3.7074
7	1.415	1.8946	2.3646	2.998	3.4995
8	1.397	1.8595	2.3060	2.896	3.3554
9	1.383	1.8331	2.2622	2.821	3.2498
10	1.372	1.8125	2.2281	2.764	3.1693
11	1.363	1.7959	2.2010	2.718	3.1058
12	1.356	1.7823	2.1788	2.681	3.0545
13	1.350	1.7709	2.1604	2.650	3.0123
14	1.345	1.7613	2.1448	2.624	2.9768
15	1.341	1.7530	2.1315	2.602	2.9467
16	1.337	1.7459	2.1199	2.583	2.9208
17	1.333	1.7396	2.1098	2.567	2.8982
18	1.330	1.7341	2.1009	2.552	2.8784
19	1.328	1.7291	2.0930	2.539	2.8609
20	1.325	1.7247	2.0860	2.528	2.8453
21	1.323	1.7207	2.0796	2.518	2.8314
22	1.321	1.7171	2.0739	2.508	2.8188
23	1.319	1.7139	2.0687	2.500	2.8073
24	1.318	1.7109	2.0639	2.492	2.7969
25	1.316	1.7081	2.0595	2.485	2.7874
26	1.315	1.7056	2.0555	2.479	2.7787
27	1.314	1.7033	2.0518	2.473	2.7707
28	1.313	1.7011	2.0484	2.467	2.7633
29	1.311	1.6991	2.0452	2.462	2.7564
30	1.310	1.6973	2.0423	2.457	2.7500
35	1.3062	1.6896	2.0301	2.438	2.7239
40	1.3031	1.6839	2.0211	2.423	2.7045
45	1.3007	1.6794	2.0141	2.412	2.6896
50	1.2987	1.6759	2.0086	2.403	2.6778
60	1.2959	1.6707	2.0003	2.390	2.6603
70	1.2938	1.6669	1.9945	2.381	2.6480
80	1.2922	1.6641	1.9901	2.374	2.6388
90	1.2910	1.6620	1.9867	2.368	2.6316
100	1.2901	1.6602	1.9840	2.364	2.6260
120	1.2887	1.6577	1.9799	2.358	2.6175
140	1.2876	1.6558	1.9771	2.353	2.6114
160	1.2869	1.6545	1.9749	2.350	2.6070
180	1.2863	1.6534	1.9733	2.347	2.6035
200	1.2858	1.6525	1.9719	2.345	2.6006
∞	1.282	1.645	1.96	2.326	2.576

Distribuição *t*

Mas, e se o teste fosse monocaudal?

- A hipótese alternativa deve ser especificada antes da coleta dos dados e deve ser independente deles. Quando houver conhecimento prévio para dizer que a diferença ocorre em uma dada direção (maior ou menor), aplicamos o teste monocaudal.

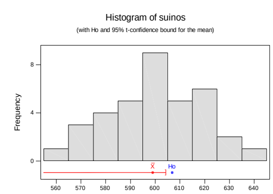
$$H_0 : \mu = 607\ g / dia$$
$$H_1 : \mu < 607\ g / dia$$

Teste *t* monocaudal

One-Sample T: suinos
Test of mu = 607 vs mu < 607

Variable	N	Mean	StDev	SE Mean
suinos	36	599,19	18,66	3,11

Variable	95,0% Upper Bound	T	P
suinos	604,45	-2,51	0,008



Conclusão: H_0 é rejeitada, porque *p*=0,8% é menor que um nível de significância de 5%.

Observe que este valor de *p* é a metade do valor obtido no teste bicaudal.

Cuidado: É mais fácil rejeitar H_0 quando o teste é monocaudal. No entanto, lembre-se que a hipótese nula, neste caso, deve ser feita *a priori* com base em conhecimentos prévios.