TESTES DE HIPÓTESES - Generalidades

Uma hipótese estatística é uma conjectura sobre uma característica da população.

Um teste de hipóteses é um procedimento estatístico que averigua se os dados sustentam uma hipótese.

As hipóteses: Num teste de hipóteses há sempre duas hipóteses:

Hipótese Nula — H_0 $\quad vs \quad$ Hipótese alternativa — H_1

Exemplo:

$H_0 : \mu = 3 \quad vs \quad H_1 : \mu < 3$ (População Normal)

Tipos de erros:

		Hipótese verdadeira	
		H_0	H_1
Decisão do teste	Rejeito H_0	Erro de tipo I	$\surd$
	Não rejeito H_0	$\surd$	Erro de tipo II

$$P(\textbf{Erro de tipo I}) = \alpha \qquad P(\textbf{Erro de tipo II}) = \beta.$$

Tamanho do teste ou Nível de significância:

$$\alpha = P(\text{Erro de tipo I}) = P(\text{rejeitar } H_0 | H_0 \text{ verdadeiro}).$$

Potência do teste:

$$1 - \beta = 1 - P(\text{Erro de tipo II}) = P(\text{rejeitar } H_0 | H_1 \text{ verdadeiro}).$$

Exemplos de hipóteses

1. $H_0 : \mu = 0 \quad vs \quad H_1 : \mu \neq 0$ (População Normal)
2. $H_0 : \mu = 3 \quad vs \quad H_1 : \mu < 3$ (População Normal)
3. $H_0 : \sigma^2 = 1 \quad vs \quad H_1 : \sigma^2 > 1$ (População Normal)
4. $H_0 : \mu = 4 \quad vs \quad H_1 : \mu = 7$ (População Normal)
5. $H_0 : \mu > 1 \quad vs \quad H_1 : \mu \leq 1$ (População Normal)
6. $H_0 : X \frown Normal \quad vs \quad H_1 : X \frown$ outra distribuição.
7. $H_0 : X \frown Normal(10, 0.4) \quad vs \quad H_1 : X \frown$ outra distribuição.

Os dois últimos exemplos designam-se habitualmente por testes de ajustamento.

Tipos de hipóteses: As várias hipóteses podem ser **simples** ou **compostas**. Uma hipótese simples apenas contempla uma possibilidade (sinal de $=$).

Iremos apenas considerar testes em que H_0 é simples.

Tipos de testes: Os testes podem ser **unilaterais** ou **bilaterais**.

Nos testes unilaterais a hipótese alternativa apenas contempla possibilidades à direita ou à esquerda da hipótese nula.

Exemplos de testes unilaterais à direita:

$$H_0 : \mu = 1 \quad vs \quad H_1 : \mu > 1$$

$$H_0 : \mu = 4 \quad vs \quad H_1 : \mu = 7$$

Exemplos de um testes unilaterais à esquerda:

$$H_0 : \mu = 1 \quad vs \quad H_1 : \mu < 1$$

$$H_0 : \mu = 4 \quad vs \quad H_1 : \mu = 2$$

Nos testes bilaterais a hipótese alternativa contempla possibilidades à direita e à esquerda da hipótese nula.

Exemplo de um teste bilateral:

$$H_0 : \mu = 1 \quad vs \quad H_1 : \mu \neq 1$$

Ingredientes:

Estatística de teste: é uma estatística calculada a partir da amostra e que é usada para tomar a decisão acerca de rejeitar ou não a hipótese nula. Costuma-se representar por T.

Região de rejeição (ou **região crítica, RC**):é o conjunto de valores da estatística de teste que nos levam a rejeitar a hipótese nula.

Região de não rejeição: é o complementar da região de rejeição.

Se T pertencer à região crítica rejeita-se H_0 a favor de H_1. Caso contrário não se rejeita H_0.

(Dada a incerteza associada a um teste de hipóteses não se costuma dizer que se aceita H_0, mas sim que não se rejeita H_0.)

Procedimentos para a realização de um teste de hipóteses de tamanho α:

1-Procedimento com base na região de rejeição

1. Identificar o parâmetro de interesse e especificar as hipóteses H_0, H_1.

2. Escolher uma estatística de teste, T, com distribuição conhecida (admitindo que H_0 é verdadeira).

3. Identificar a região de rejeição.

4. Calcular t_{obs} que é o valor que T assume para os dados observados.

5. Tomar uma decisão.

6. Concluir.

Concluir sempre referindo-se à hipótese alternativa. Desta forma as hipóteses (e tipo de teste) ficam sempre bem identificadas.

Exemplo 1: Teste bilateral

Os tubarões são por natureza peixes de água salgada. No entanto, em alguns rios já foram vistos tubarões e várias pessoas foram violentamente agredidas por esses animais. O rio Ganges é um desses rios onde vários indianos foram já violentamente agredidos. Dada a raridade destas ocorrências não se sabe ao certo qual a espécie que se aventura a subir as águas do rio. Os biólogos crêem que se trata de uma só espécie mas não sabem ao certo qual. As suspeitas populares incidem sobre o tubarão branco, um animal que atinge os 5 m de comprimento, mas os biólogos têm dúvidas. Sabe-se que quando um tubarão ataca uma presa por vezes deixa cair um ou mais dos seus dentes. Assim, os biólogos procuraram o leito do rio e encontraram 4 dentes de tubarão. O comprimento médio destes dentes foi de 3.31cm e o respectivo desvio padrão $s = 0.20$. Sabendo que os dentes do tubarão branco apresentam um comprimento médio (populacional) de 3.6cm, será de considerar que estamos perante uma amostra de dentes de tubarão branco? Considere $\alpha = 0.05$.

$$H_0 : \mu = 3.6 \quad vs \quad H_1 : \mu \neq 3.6$$

onde μ representa a média da distribuição dos comprimentos dos dentes.

Exemplo 2: Teste unilateral à esquerda

Os biólogos desconfiam que a espécie de tubarões que invade o Ganges não é o tubarão branco mas sim o tubarão touro. Este tubarões têm menor porte mas são igualmente agressivos. Devido ao seu menor porte os seus dentes têm comprimento médio inferior aos do tubarão branco. Com base nesta amostra será de considerar que estamos perante uma amostra de dentes de tubarão branco ou de outra espécie com dentes mais curtos, como é o caso do tubarão touro? Considere $\alpha = 0.05$.

$$H_0 : \mu = 3.6 \quad vs \quad H_1 : \mu < 3.6.$$

Exemplo 3: Teste unilateral à direita Um viveiro produz trutas salmonadas. Os gestores do viveiro estão a introduzir alterações na dieta das trutas com vista a aumentar o peso dos animais em estado de serem capturados para consumo humano. De acordo com a dieta regular dos viveiros as trutas pesam em média 760g e apresentam um desvio padrão de 40 (valores populacionais). Um grupo de 25 trutas recém-nascidas foi sujeito à nova dieta. Quando atingiram a idade necessária para poderem ser abatidas pesaram-se e obteve-se $\bar{x} = 784$. Supondo que a dieta não altera o desvio padrão e que os dados são bem modelados por um distribuição Normal, será de acreditar que a nova dieta faz aumentar o peso médio das trutas salmonadas ($\alpha = 0.05$)?

$$H_0 : \mu = 760 \quad vs \quad H_1 : \mu > 760$$

onde μ representa o peso médio das trutas sujeitas à nova dieta (populacional)

2-Procedimento alternativo com base nos intervalos de confiança (válido apenas para testes bilaterais)

1. Identificar o parâmetro de interesse e especificar as hipóteses H_0, H_1.

2. Construir um intervalo de confiança para o parâmetro.

3. Rejeitar H_0 se o valor do parâmetro especificado em H_0 não pertencer ao intervalo de confiança. (O intervalo de confiança fornece uma região de não rejeição do teste.)

3-Procedimento com base no p-value (SPSS)

***p*-value do teste** é a **probabilidade de observar um valor da estatística de teste tanto ou mais afastado que o valor observado na amostra, assumindo que H_0 é verdadeira**.

1. Identificar o parâmetro de interesse e especificar as hipóteses H_0, H_1.

2. Determinar o p-*value* do teste. (O software estatístico fornece o resultado.)

3. Rejeitar H_0 se p-*value* $\leq \alpha$. Não rejeitar H_0 se p-*value* $> \alpha$.

4. Concluir.

Exemplo 4.11: Pretende-se averiguar se o diâmetro médio dos discos de travão é significativamente diferente de 322 mm. Isto corresponde a testar as hipóteses seguintes

$$H_0 : \mu = 322 \quad vs \quad H_1 : \mu \neq 322.$$

No SPSS: `Analyze / Compare Means / One-Sample T Test`, `Test value`$=$ 322

One-Sample Statistics

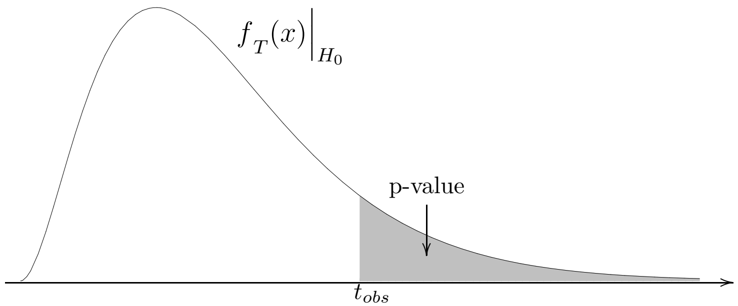
	N	Mean	Std. Deviation	Std. Error Mean
Diametro do disco de travao	128	322,00201	,0108224	,0009566

One-Sample Test

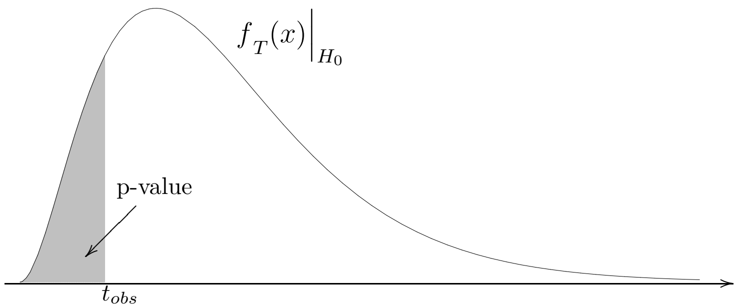
	Test Value = 322					
					99% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Diametro do disco de travao	2,101	127	,038	,002009	-,000492	,004511

Como determinar o p-value

- Quando a região de rejeição é da forma $T > c$ (rejeitar para valores elevados da estatística de teste), o p-*value* é igual a $P(T > t_{obs}|H_0)$.



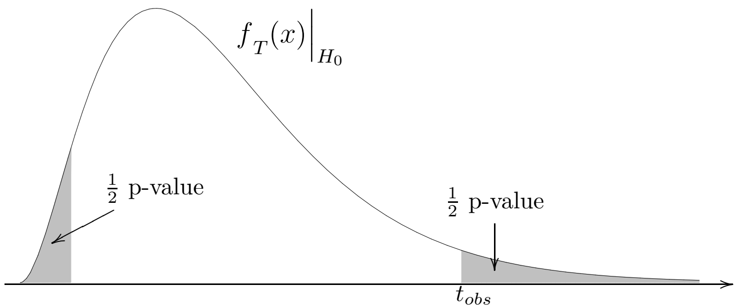
- Quando a região de rejeição é da forma $T < c$ (rejeitar para valores reduzidos da estatística de teste), o p-*value* é igual a $P(T < t_{obs}|H_0)$.

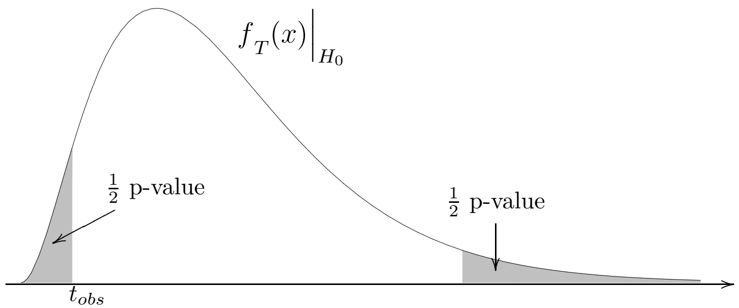


- Quando a região de rejeição é da forma $T < c_1$ ou $T > c_2$ (com igual probabilidade para os dois casos), o p-*value* é igual a

$$\begin{cases} 2P(T < t_{obs}|H_0) & \text{se } t_{obs} \text{ for reduzido} \\ 2P(T > t_{obs}|H_0) & \text{se } t_{obs} \text{ for elevado} \end{cases}.$$

Dizer que t_{obs} é reduzido (elevado) significa dizer que a estimativa que se obtém para o parâmetro a testar é inferior (superior) ao valor especificado em H_0.





Exemplo 4.7: Recordemos os Exemplos 4.1 e 4.2 onde o Rocha durante uns quantos dias registou a quantidade de álcool no sangue dos clientes à saída do bar. A sua amostra tinha 125 valores com $\bar{x} = 0.6$ e $s_c^2 = 0.041$ e podia ser considerada como proveniente de uma população Normal. A GNR tem interesse em saber se os clientes do bar conduzem ultrapassando o valor estabelecido pela lei, ou seja $0.5mg/l$. Estabeleçamos para tamanho do teste o valor típico $\alpha = 0.05$.

1. $$H_0 : \mu = 0.5 \quad vs \quad H_1 : \mu > 0.5$$
   onde μ representa a média da distribuição que modela a quantidade de álcool no sangue dos clientes do bar onde o Rocha trabalha.

2. Estatística de teste:
$$T = \frac{\bar{X} - 0.5}{S_c/\sqrt{125}} \underset{sob\ H_0}{\frown} t_{124}.$$

3. Uma vez que se trata de um teste unilateral à direita, rejeita-se H_0 para valores elevados de T, ou seja, se $T > c$.

4. $t_{obs} = \frac{0.6-0.5}{\sqrt{0.041/125}} = 5.52$

5.
$$p\text{-}value = P(T > t_{obs}|H_0) = P(T > 5.52|H_0) = 9.4 \times 10^{-8} \simeq 0.$$

6. Como p-*value*$<$ 0.05 rejeita-se H_0. Aliás, para este valor do p-*value*, rejeita-se H_0 para qualquer nível de significância usual.

   Conclui-se, ao nível de significância 0.05, que os clientes conduzem, em média, ultrapassando os limites impostos pela lei.

Exemplo 4.6 (continuação): Suponhamos que uma certa amostra forneceu $\bar{x} = 20.04$ e $s_c = 0.2$. Com base nesta amostra pretende-se realizar um teste de hipóteses. O tamanho de teste é pre-estabelecido, $\alpha = 0.001$ (neste tipo de aplicações não se podem permitir erros com probabilidade elevada).

1. $$H_0 : \mu = 20 \quad vs \quad H_1 : \mu \neq 20,$$
   onde μ representa a verdadeira média da quantidade de Fluoxetina contida nas cápsulas.

2. Estatística de teste:
   $$T = \frac{\bar{X} - 20}{S_c/\sqrt{100}} \underset{sob\ H_0}{\frown} t_{99}.$$

3. Rejeita-se H_0 se T for muito reduzido ou muito elevado, ou seja se $T < c_1$ ou $T > c_2$.

4. Para este exemplo $\bar{x} = 20.04$ que é superior ao valor especificado em

$H_0 : \mu = 20$. Assim, vamos ter um valor eventualmente "elevado" para a estatística de teste T. O p-*value* para este teste é dado por

$$2\,P(T > t_{obs}|H_0) = 2\,P(T > 2|H_0) = 0.048$$

5. Como este valor é superior a $\alpha = 0.001$ não se rejeita H_0. Logo, ao nível de significância de 0.001, não há razões para acreditar que as cápsulas não contenham 20mg de Fluoxetina, em média.

Procedimento para transformação de p-*values* bilaterais em unilaterais.

Por vezes o software estatístico apenas fornece o valor do p-value bilateral e nós estamos interessados em realizar um teste unilateral. Nesses casos há que transformar o p-value bilateral em unilateral.

- se a(s) amostra(s) aponta(m) no sentido da hipótese alternativa deve-se dividir o p-*value* por 2 e tomar esse valor como o p-*value* do teste unilateral, $p\text{-}value_{uni} = p\text{-}value_{bil}/2$;
- se a(s) amostra(s) não aponta(m) no sentido da hipótese alternativa, então o p-*value* do teste unilateral é igual a $p\text{-}value_{uni} = 1 - p\text{-}value_{bil}/2$.