
Bioestatística
Mestrado Farmácia
T4

Carina Silva
carina.silva@estesl.ipl.pt

Sumário

Testes de Hipóteses.

Testes de hipóteses para μ .

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 técnicas de inferência estatística tem sempre por base, uma **amostra recolhida aleatoriamente**. Os testes de hipóteses, muito utilizados na investigação, tem por objetivo testar uma conjectura acerca de uma característica de uma população. O investigador que recorre a estas técnicas não deve de abdicar da sua experiência, pois a decisão não deve ser tomada unicamente em função de métodos estatísticos, até porque poderão fazer sentido estatisticamente mas biologicamente não fazer qualquer sentido.

Como realizar um Teste de Hipóteses

Uma correcta formulação das hipóteses estatísticas (⚠ não confundir com hipóteses de investigação), que se obtêm a partir das hipóteses de investigação, conduz a uma boa realização de um teste de hipóteses. As hipóteses devem ser formuladas antes da recolha da amostra. Deve-se ter em conta que o que se pretende é tomar uma boa decisão ao mesmo tempo que se avaliam os riscos envolvidos.

A escolha de um determinado teste estatístico depende de um conjunto de fatores: objetivos, características dos dados, pressupostos, etc. Apresentam-se os principais passos para a realização de um teste de hipóteses.

1º Passo – Tipo de teste de hipóteses

As características dos dados: natureza dos dados/escala de medição, distribuição da população de onde foi retirada a amostra, dimensão da amostra, independência das amostras (quando temos mais do que uma amostra), etc.

Os Testes de Hipóteses dividem-se em dois grandes grupos:

TESTES DE HIPÓTESES PARAMÉTRICOS: O MODELO ESPECIFICA CONDIÇÕES SOBRE OS PARÂMETROS DAS POPULAÇÕES. A escala dos dados é métrica.

TESTES DE HIPÓTESE NÃO-PARAMÉTRICOS: O MODELO NÃO ESPECIFICA CONDIÇÕES SOBRE PARÂMETROS DE POPULAÇÕES. A escala de medição pode ser qualquer.

2º Passo – Formular as hipóteses

Um teste envolve duas hipóteses: Hipótese nula (H_0), que é a hipótese que vai ser testada e a Hipótese alternativa (H_1), que exprime as convicções do investigador e que vai definir a região de rejeição. Vai ser apenas abordado o caso de hipóteses simples, i.e., a hipótese nula é definida por uma igualdade.

3º Passo – Escolha da estatística de teste

A estatística de teste (é uma v.a.) é utilizada para tomar uma decisão relativamente à hipótese nula (rejeitar ou não rejeitar eis a questão!).



4º Passo – Definição da regra de teste

Para tomar uma decisão, para além do valor da estatística de teste, tem de se fixar o nível de significância (α) e ter em consideração a hipótese alternativa, H_1 , pois é esta que define a zona de rejeição.

α , é o valor máximo para a probabilidade de cometer erro de tipo I:

$$P(\text{REJEITAR } H_0 | H_0 \text{ É VERDADEIRO}) = \alpha$$

Ao tomar uma decisão pode-se cometer dois tipos de erro: erro de tipo I, probabilidade de rejeitar H_0 quando é verdadeira e erro de tipo II, probabilidade de não rejeitar H_0 quando H_1 é verdadeiro.



		SITUAÇÃO VERDADEIRA	
		H_0 VERDADEIRA	H_1 VERDADEIRA
DECISÃO	NÃO REJEITAR H_0	DECISÃO CORRECTA	ERRO TIPO II β
	REJEITAR H_0	ERRO TIPO I α	DECISÃO CORRECTA

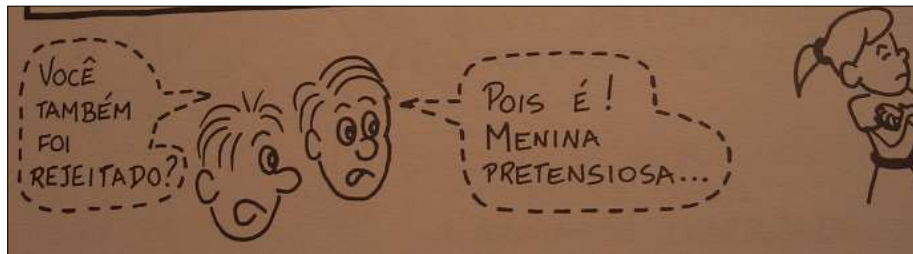
Na escolha do teste, o nosso objectivo é controlar o erro de tipo I. É comum adoptar-se os níveis usuais de significância: 0.01, 0.05, 0.1. Mas agora com o apoio informático é possível tomar qualquer um entre 0.01 e 0.1.

A partir do erro de tipo II, (β), calcula-se a **potência** de um teste estatístico:

$$\text{Potência } 1 - \beta$$

Ou seja é a probabilidade de rejeitar a hipótese nula quando ela é falsa.
A potência de um teste aumenta quando a dimensão da amostra aumenta.

Um processo alternativo ao anterior é calcular a probabilidade p (valor-P) que é o menor valor para o nível de significância que levaria à rejeição da hipótese nula.



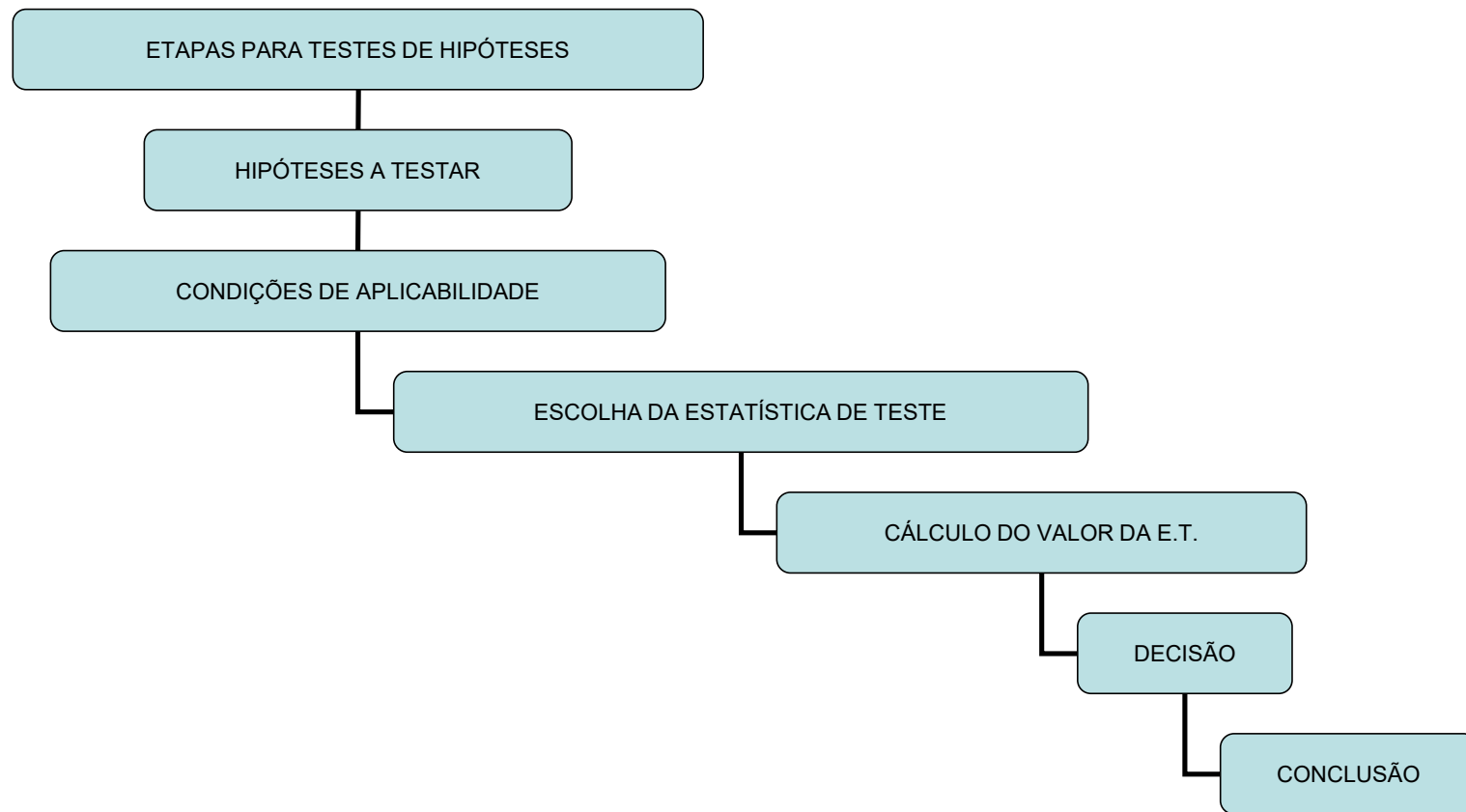
5º Passo – Decisão

Se se optou pela definição clássica toma-se a decisão em função da região de rejeição. Se a opção for o valor-P (mais vulgarmente utilizada na investigação) a regra de decisão é:

se $\text{valor-P} \leq \alpha$ então rejeita-se H_0 ao n.s. α

6º Passo – Conclusão

Se se rejeita a H_0 , conclui-se que a hipótese alternativa é provavelmente verdadeira, se não se rejeita H_0 conclui-se que os dados não fornecem evidência ou prova suficiente para apoiar H_1 . Não podemos concluir que H_0 é verdadeira.



Testes de Hipóteses para o valor médio - μ

$$\text{Hipóteses: } \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \begin{cases} \mu \neq \mu_0 & (\text{teste bilateral}) \\ \mu < \mu_0 & (\text{teste uni. esq.}) \\ \mu > \mu_0 & (\text{teste uni. dir.}) \end{cases} \end{cases}$$

Valor da Estatística de Teste sob as condições de H_0 :

- **$n > 30$ (grandes amostras) qualquer que seja a distribuição de X**
⇒ se **σ conhecido** e X é uma variável aleatória com **distribuição arbitrária**

$$z_0 = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

⇒ se σ desconhecido e X é uma variável aleatória com **distribuição Normal** ou **arbitrária**

$$z_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

• $n \leq 30$ (pequenas amostras), e X é uma variável aleatória. com **distribuição Normal**

⇒ se σ conhecido

$$z_0 = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

⇒ se σ desconhecido

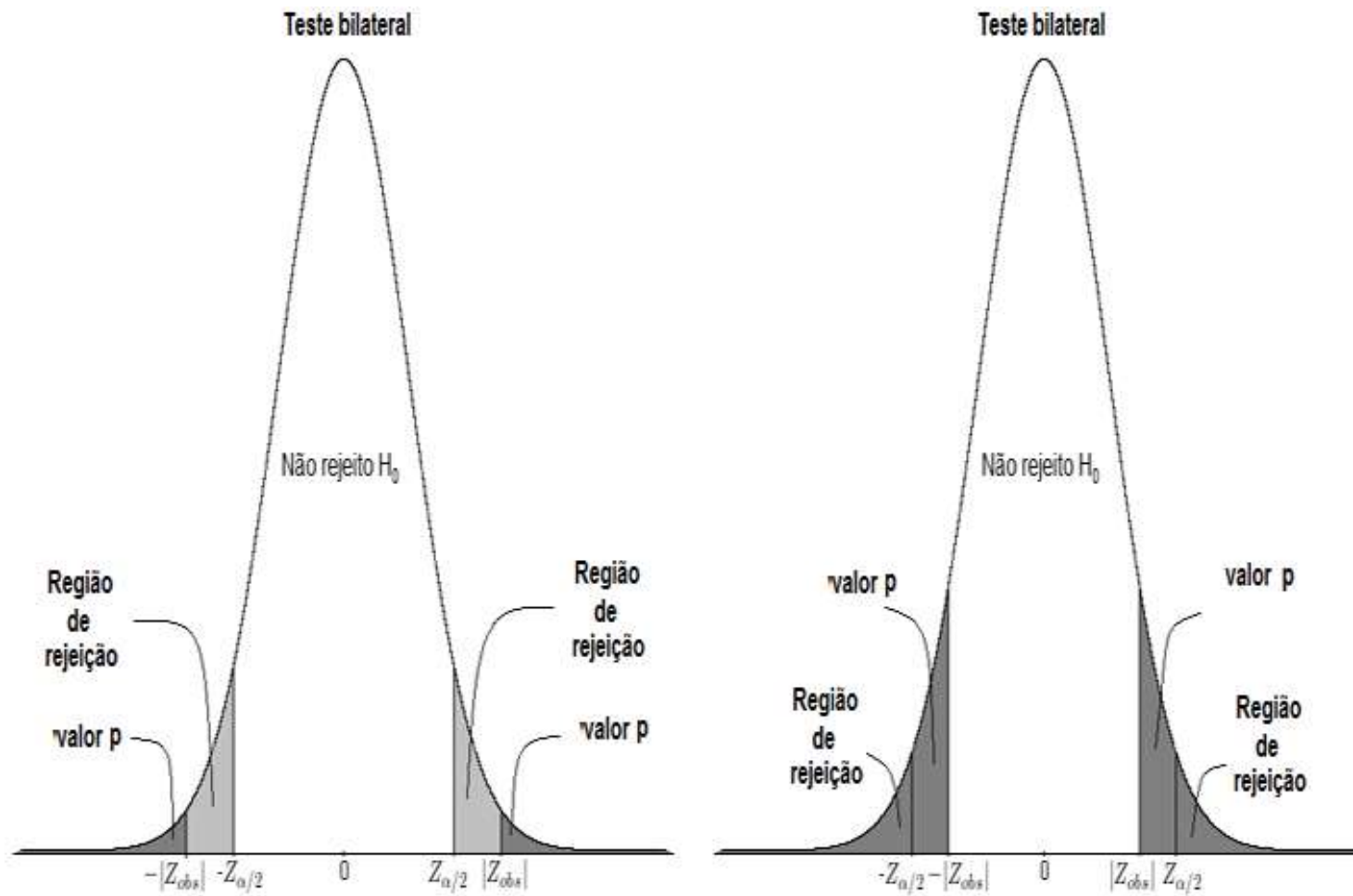
$$t_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Valor-p

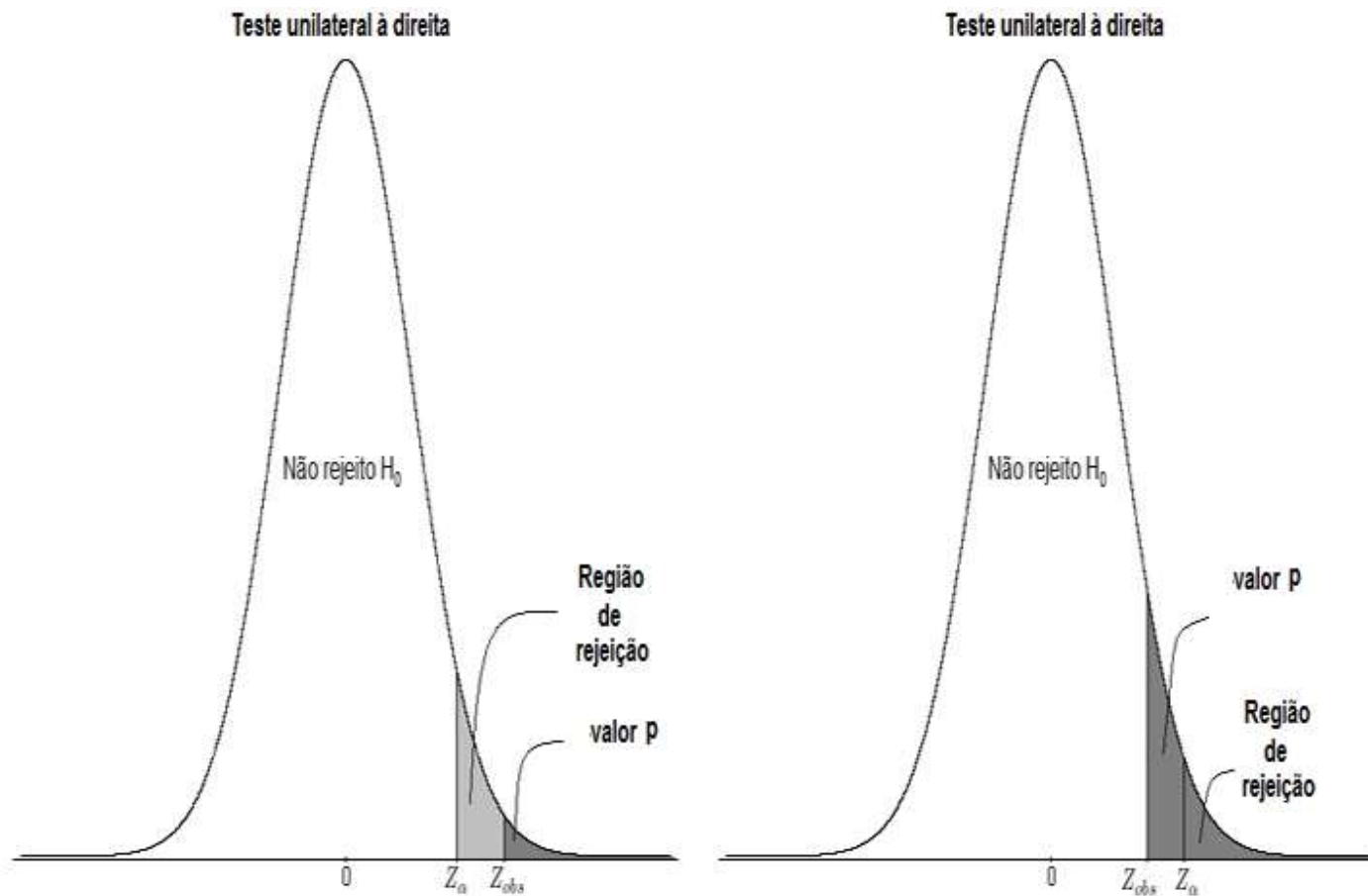
O valor-p, também denominado nível descritivo do teste, é a probabilidade de que a estatística do teste (como variável aleatória) tenha valor extremo em relação ao valor observado (estatística) quando a hipótese é verdadeira.

As figuras a seguir representam, respectivamente, o valor-p nos casos em que temos um teste de hipóteses bilateral com rejeição da hipótese nula e sem rejeição da hipótese nula.

Valor-p



Valor-p



Exercicio 1

20. Num estudo sobre o efeito do monóxido de carbono produzido pelos cigarros no crescimento dos fetos de mães fumadoras, mediu-se a percentagem de hemoglobina acoplada a monóxido de carbono como carboxi-hemoglobina (cohb) em 30 fumadoras grávidas, antes e depois de fumarem um cigarro. Após a recolha dos dados os resultados obtidos em SPSS foram

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
INCREMENTOS NA PERCENTAGEM DE COHb	30	4.060	1.4792	.2701

One-Sample Test

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
INCREMENTOS NA PERCENTAGEM DE COHb	3,925	29	,000	1.060	.508	1.612

Admitindo normalidade dos dados, poder-se-á dizer que o incremento médio não difere significativamente de 3. Obtenha conclusões analisando os “outputs” anteriores. (1.5 valor)

Exercício 2

Os resultados obtidos no SPSS, referem-se ao mercúrio metílico ingerido por uma amostra aleatória de 54 sujeitos que consumiram peixe contaminado com mercúrio metílico:

Descriptives			Statistic	Std. Error
VAR00001	Mean		329,02	19,44
	95% Confidence	Lower Bound	290,03	
	Interval for Mean	Upper Bound	368,00	
	5% Trimmed Mean		320,68	
	Median		287,50	
	Variance		20400,132	
	Std. Deviation		142,83	
	Minimum		123	
	Maximum		693	
	Range		570	
	Interquartile Range		115,25	
	Skewness		1,187	,325
	Kurtosis		1,130	,639

Um investigador suspeita que o nível médio do mercúrio ingerido é significativamente superior a 300. Para um nível de significância de 1% verifique se o investigador tem razão.