

Intervalos de confiança e teste t à média

Francisco Caramelo - fcaramelo@fmed.uc.pt

Neste exercício pretende-se determinar intervalo de confiança para 95% ($\alpha = 0.05$) e realizar testes t à média.

Para o exercício vamos usar dados de um dos conjuntos existentes na package "datasets". No caso concreto iremos usar os dados "iris".

```
library(datasets)
head(iris)
```

##	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
## 1	5.1	3.5	1.4	0.2	setosa
## 2	4.9	3.0	1.4	0.2	setosa
## 3	4.7	3.2	1.3	0.2	setosa
## 4	4.6	3.1	1.5	0.2	setosa
## 5	5.0	3.6	1.4	0.2	setosa
## 6	5.4	3.9	1.7	0.4	setosa

Os dados referem-se a medidas em centímetros de sépalas e pétalas de três espécies de flores.

1. Quais são os nomes dessas espécies?
2. Qual é o valor do desvio padrão do comprimento da sépala para a espécie "setosa"?

Para determinar o intervalo de confiança é necessário calcular erro padrão e também o valor do ponto de corte da distribuição Z ou t. Em R o valor correspondente a uma determinada área numa distribuição de probabilidade é dada pelos comandos "qnorm" e "qt", respectivamente para a distribuição normal e a distribuição t-Student.

O valor de Z_p correspondente à probabilidade $P(Z < Z_p) = 0.975$ é, assim, dado por:

```
qnorm(0.975)
```

```
## [1] 1.959964
```

Já para a distribuição t com 20 graus de liberdade o valor seria

```
qt(0.975, 20)
```

```
## [1] 2.085963
```

3. Qual o valor do erro padrão do comprimento da sépala para a espécie "setosa"?
4. Determine o intervalo de confiança para a média do comprimento da sépala para a espécie "setosa"?

Teste t à média

Conjectura-se que a média do comprimento das sépalas da espécie 'setosa' é de 5 cm. Por intermédio de um teste t à média deseja-se verificar a veracidade daquela assumção.

5. Qual é a hipótese nula neste caso?

A realização de um teste t à média pode ser feito de formas distintas, destacando-se duas possibilidades: i) determinar o valor de teste e comparar com os valores críticos; ii) calcular o valor de p;

No primeiro caso calcula-se o valor dado pela expressão $t = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$. Se este valor se encontrar dentro do intervalo $[-t_{n-1, 1-\alpha/2}, t_{n-1, 1-\alpha/2}]$ diz-se que se encontra dentro da zona de aceitação e, portanto, não se rejeita H_0 .

6. Qual é o valor do teste para a média do comprimento da sépala para a espécie "setosa"?
7. Qual a conclusão relativamente ao teste t?

No segundo caso, determina-se o valor de teste como anteriormente e, a partir deste, calcula-se o valor de p. Suponha-se que o valor de teste obtido é 1.35, então o valor de p, considerando uma amostra com 21 elementos será dado por:

```
2*(1-pt(1.35, 20))  
## [1] 0.1920977
```

8. Qual o valor de p para o teste à média realizado?