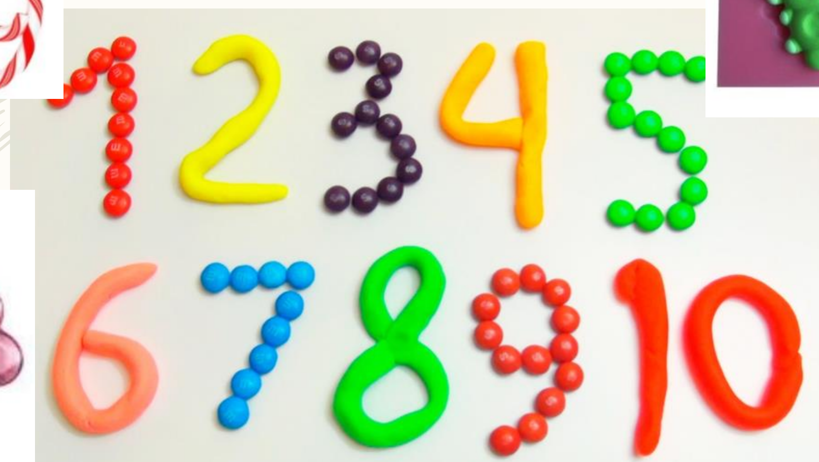
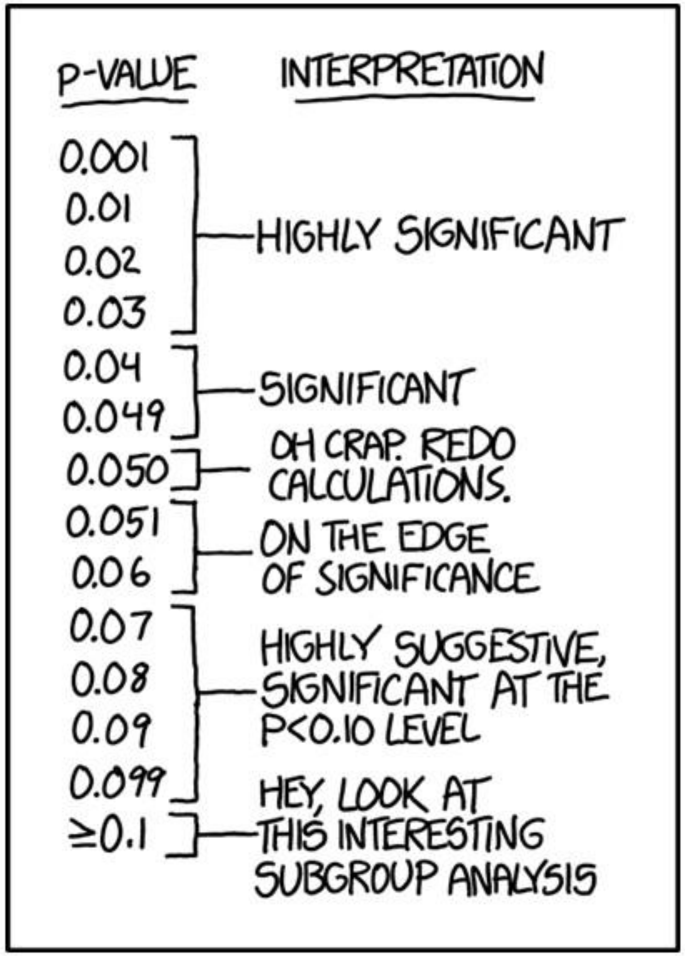
Goodies*



* Goodies related to animals, plants and numbers…

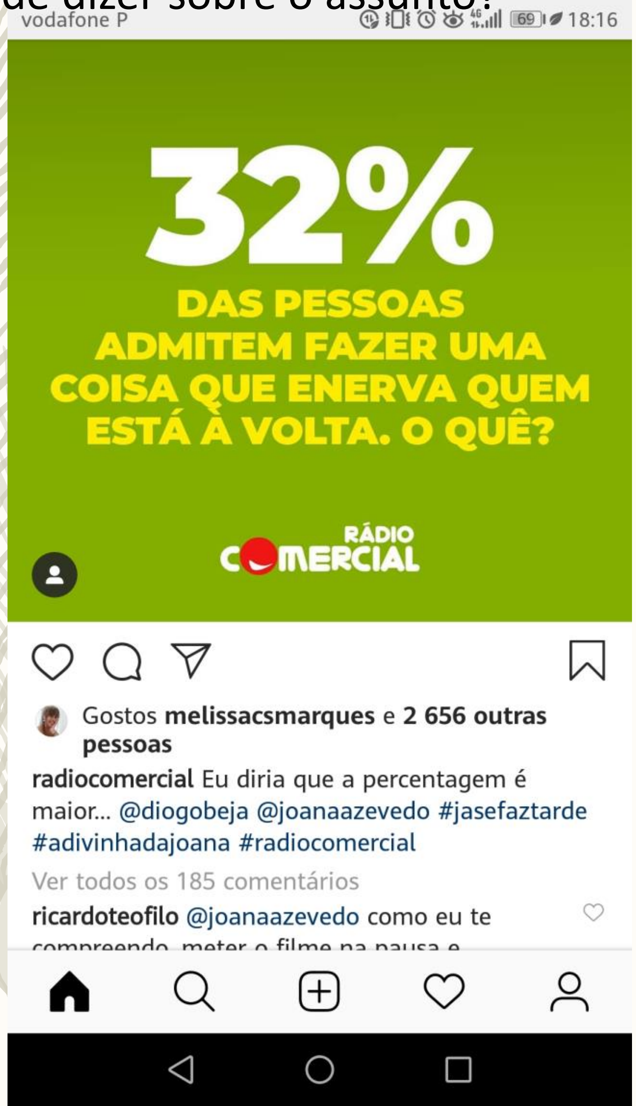


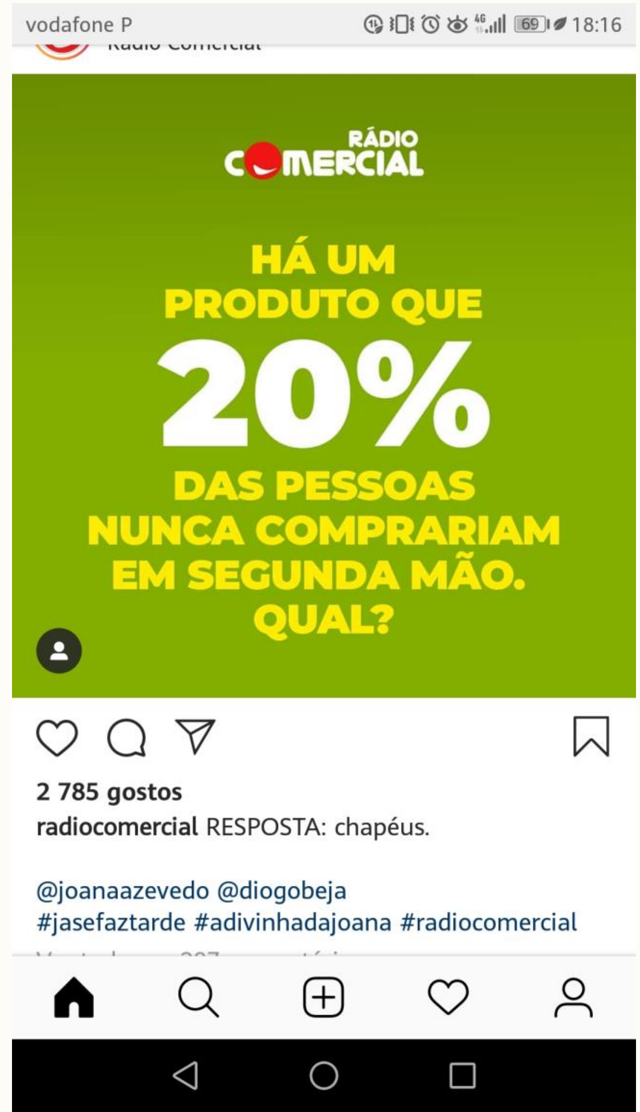
In [Turning Data into Stories](#) we pointed out that numbers have no inherent meaning…

https://verstaresearch.com/blog/a-statistically-significant-cartoon/

– JOE, COULD YOU GET A SIGNIFICANT P-VALUE OUT OF ALL THIS?

Boa tarde professor, ao visitar a página de Instagram da rádio comercial, deparei-me com alguns post sobre estatística, como poderá ver pelos anexos.
A minha questão é : onde é que as vão buscar, porque não consigo perceber que base de dados será esta, nem o propósito do seu estudo inicial.
Penso que seja apenas para ter comentários nos posts.
O que me pode dizer sobre o assunto?
Obrigada.





Devemos ter extremo cuidado com os números que vemos na Comunicação Social.

Pensar sempre em qual poderá ser o valor destes números.

Em geral, há muitos erros que são propagados, quer de forma não intencional, quer de forma intencional.

No caso d'"Adivinha da Joana" o valor é nulo, é recreativo e serve apenas para quem não tem nada para fazer se divertir. Mas concordo que o fundamento está errado, as pessoas deviam poder confiar num órgão de Comunicação Social de referência, mesmo quando estão a brincar. Radio Comercial should know better!

Better Data.
Better Lives.

Dia Europeu da Estatística

20 de outubro de 2019

Programa:

- Comunicações sobre aplicações da Estatística em diferentes áreas de investigação;
    - "O preço de um voto: deseconomia em eleições proporcionais" - Hygor Piaget, CFTC - Centro de Física Teórica e Computacional
    - "A Estatística ao serviço da prevenção e combate dos incêndios rurais em Portugal" - Carlos da Câmara, IDL - Instituto Dom Luís
    - "Statistics and evolutionary biology: inferring the history of species from genomes" - Vitor Sousa, cE3c - Centro de Ecologia, Evolução e Alterações Ambientais

- Apresentação de colaborações desenvolvidas entre o CEAUL e outras unidades de I&D da FCUL;

- Discussão aberta com café e bolinhos.

Data: 21 de outubro de 2019
Hora: 14h30 às 16h30
Local: Auditório da Fciências.ID

FCT Fundação para a Ciência e a Tecnologia
Projeto UID/MAT/00006/2019

CEAUL Centro de Estatística e Aplicações Universidade de Lisboa

Obrigado à Lara Martins pela sua contribuição para o material das aulas de Ecologia Numérica

Resolução do desafio da Aula Teórica 5

Um exemplo prático:

		4							
				11			5		
		1							
				8				4	
						6			
	3					9			
			2						
	4			1				3	
				3		5			
		0							
1									
						0			
		1							1

Estrato	Ni	ni	$\overline{xi}$	si
1	70	9	51/9	3.16
2	60	7	18/7	1.72
3	30	4	3/4	0.5
Total = N	160			

Estrato	fi	wi	w^2*s^2/n	w^2*s^2/n*1-f
1	0.128	0.438	0.2787	0.25
2	0.117	0.375	0.047	0.041
3	0.133	0.188	0.0017	0.0015
sX				0.54

Um exemplo prático:

		4							
				11			5		
		1							
				8				4	
						6			
	3					9			
			2						
	4			1				3	
				3		5			
		0							
1									
						0			
		1							1

Estrato	Ni	ni	$\overline{xi}$	si
1	70	9	51/9	3.16
2	60	7	18/7	1.72
3	30	4	3/4	0.5
Total = N	160			

Desvio padrão:

```
>sd(c(3,4,1,11,8,6,9,5,4))
[1] 3.162278
>sd(c(4,0,2,1,3,5,3))
[1] 1.718249
> sd(c(0,1,1,1))
[1] 0.5
```

Exercício:

		14							
				11			15		
		11							
				18				14	
						16			
	13					19			
			2						
	4			1				3	
				3		5			
		0							
1									
						0			
		1							1

Estrato	Ni	ni	$\overline{x}$	s
1	70	9	?/9	
2	60	7	18/7	
3	30	4	3/4	
Total = N	160			

Fazer este segundo exemplo, e comparar os resultados. Em particular, como se comparam os resultados, entre este e o primeiro exemplo de amostragem estratificada, correspondentes com os que seriam obtidos assumindo amostragem aleatória simples.

Exercício:

		14							
				11			15		
		11							
				18				14	
						16			
	13					19			
			2						
	4			1				3	
				3		5			
		0							
1									
						0			
		1							1

Tamanho do estrato

Un amostra

Estrato	Ni	ni	$\overline{x}$	s
1	70	9	131/9	
2	60	7	18/7	
3	30	4	3/4	
Total = N	160	20		

1-(ni/Ni)

ni/N

Estrato	fi	wi	w^2*s^2/n	w^2*s^2/n*1-f
1	9/70=0.128	70/160 =0.438		
2	7/60=0.117	60/160 =0.375		
3	4/30=0.133	30/160 =0.188		
sX				

Exercício:

		14							
				11			15		
		11							
				18				14	
						16			
	13					19			
			2						
	4			1				3	
				3		5			
		0							
1									
						0			
		1							1

Tamanho do estrato

Un amostra

Estrato	Ni	ni	$\bar{x}$	s
1	70	9	131/9	2.79
2	60	7	18/7	1.72
3	30	4	3/4	0.5
Total = N	160	20		

```
>sd(c(13,14,11,11,18,16,19,15,
    14))
[1] 2.788867
>sd(c(4,0,2,1,3,5,3))
[1] 1.718249
> sd(c(0,1,1,1))
[1] 0.5
```

Exercício:

		14							
				11			15		
		11							
				18				14	
						16			
	13					19			
			2						
	4			1				3	
				3		5			
		0							
1									
						0			
		1							1

Estrato	fi	wi	w^2*s^2/n	w^2*s^2/n*1-f
1	9/70=0.128	70/160 =0.438	0.0093	
2	7/60=0.117	60/160 =0.375	0.0026	
3	4/30=0.133	30/160 =0.188	5.5e-05	
sX				

```
> ((0.438^2)*(2.79)^2)/160
[1] 0.009333331
> ((0.375^2)*(1.72)^2)/160
[1] 0.002600156
> ((0.188^2)*(0.5)^2)/160
[1] 5.5225e-05
```

Exercício:

		14							
				11			15		
		11							
				18				14	
						16			
	13					19			
			2						
	4			1				3	
				3		5			
		0							
1									
						0			
		1							1

Estrato	fi	wi	w^2*s^2/n	w^2*s^2/n*1-f
1	9/70=0.128	70/160 =0.438	0.0093	0.0107
2	7/60=0.117	60/160 =0.375	0.0026	0.0029
3	4/30=0.133	30/160 =0.188	5.5e-05	6.4e-05
sX				

```
> ((0.438^2)*(2.79)^2)/(160*(1-0.128))
[1] 0.01070336
> ((0.375^2)*(1.72)^2)/(160*(1-0.117))
[1] 0.002944684
((0.188^2)*(0.5)^2)/(160*(1-0.133))
[1] 6.369666e-05
```

Exercício:

		14							
				11			15		
		11							
				18				14	
						16			
	13					19			
			2						
	4			1				3	
				3		5			
		0							
1									
						0			
		1							1

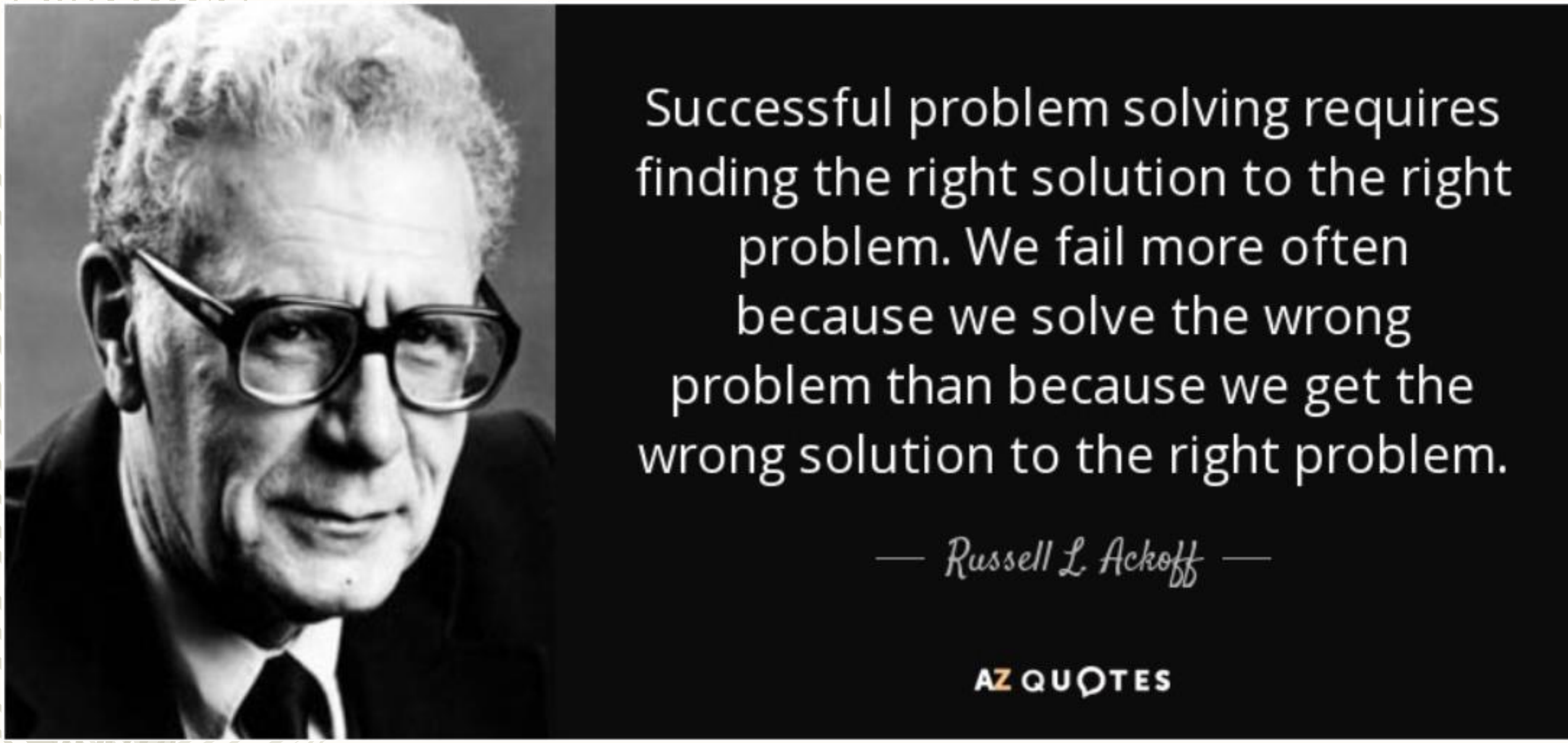
Estrato	fi	wi	w^2*s^2/n	w^2*s^2/n*1-f
1	9/70=0.128	70/160 =0.438	0.0093	0.0107
2	7/60=0.117	60/160 =0.375	0.0026	0.0029
3	4/30=0.133	30/160 =0.188	5.5e-05	6.4e-05
sX				0.117

sX:

```
>sqrt(sum(c(0.01070336,0.002944684,6.369666e-05)))
[1] 0.1170971
```

Solução da ficha TP3.pdf, realizada na aula TP 5

Ecologia Numérica - Aula Teórica 10 – 21-10-2018



Testes de hipóteses

		Hipótese nula	
		Aceitar	*Rejeitar*
Hipótese nula	*Verdadeira*	Não há erro	Erro tipo I
	Falsa	Erro tipo II	Não há erro

Testes de hipóteses

Hipótese nula	Hipótese nula: Aceitar	Hipótese nula: Rejeitar
Verdadeira	Não há erro	α
Falsa	β	1-β

significância

potência

Testes de hipóteses

O ideal seria minimizar α e β simultaneamente, mas...

α estão β negativamente correlacionados!

Uma analogia do dia-a-dia:

1. para não ter inocentes na prisão, temos de estar dispostos a ter alguns culpados em liberdade
2. para termos todos os culpados presos, temos de aceitar ter alguns inocentes na prisão

A única forma de minimizar simultaneamente α e β é aumentando a <u>dimensão da amostra</u>.

(que na analogia da prisão, seria ter tanta informação que no limite conhecíamos a verdade)

Testes de hipóteses

nível de significância do teste $\rightarrow$ α

Definido à priori

Representa o risco de cometer um erro de tipo I que estamos dispostos a corre

Como tal, deve ser decidido em consciência, e não cegamente adotando um critério generalista.

Testes de hipóteses

Potência do teste (*power*) → $1 - \beta$

é a probabilidade de rejeitar a hipótese nula quando esta é de facto falsa.

Com baixo tamanho amostral, um verdadeiro efeito baixo ou α baixo, a potência do teste é baixa.

Testes unilaterais vs bilaterais

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

vs

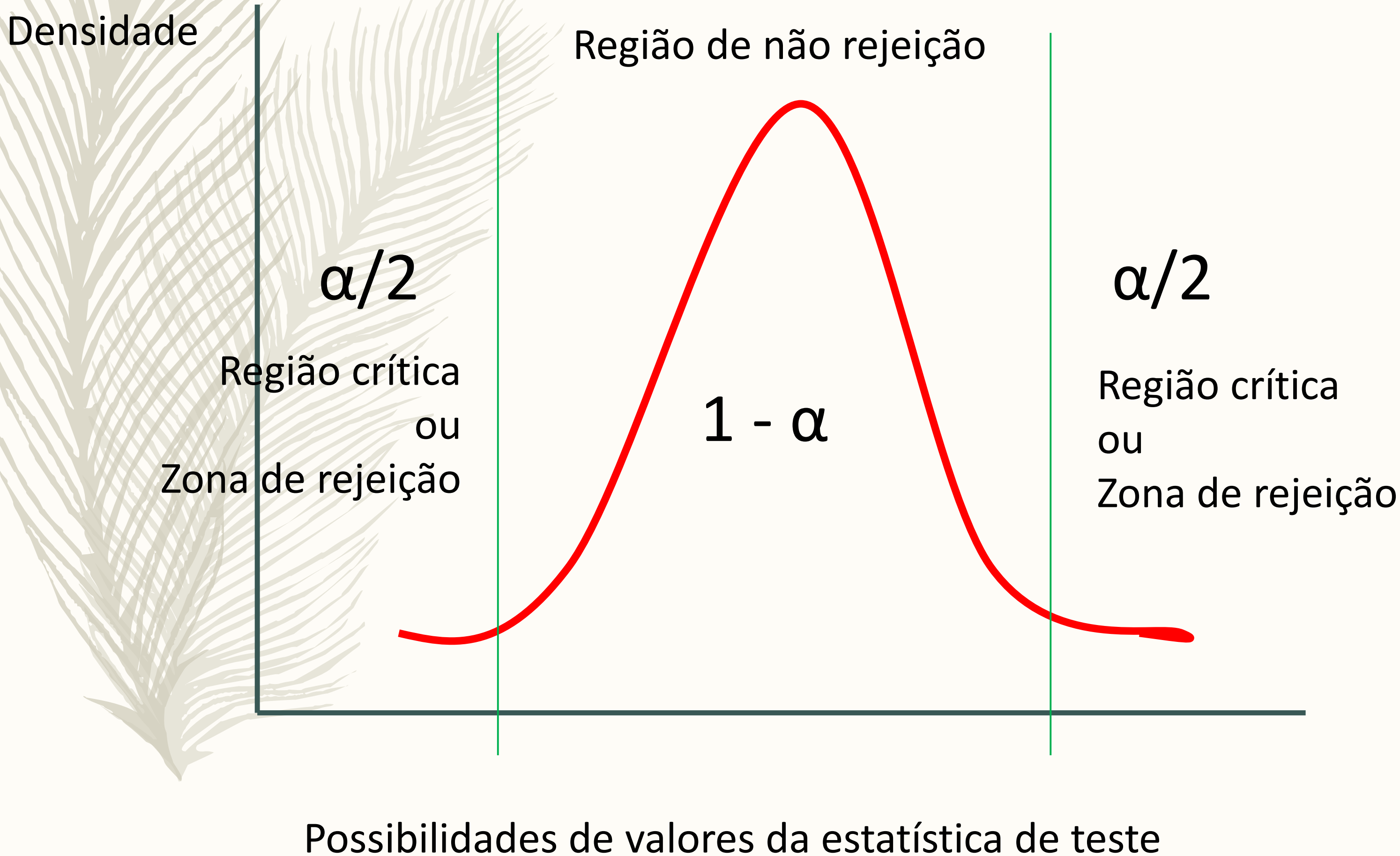
H_0: $\mu \geq X$
H_1: $\mu < X$

H_0: $\mu \leq X$
H_1: $\mu > X$

testes de hipóteses

$$H_0: \mu = X \quad \text{vs.} \quad H_1: \mu \neq X$$

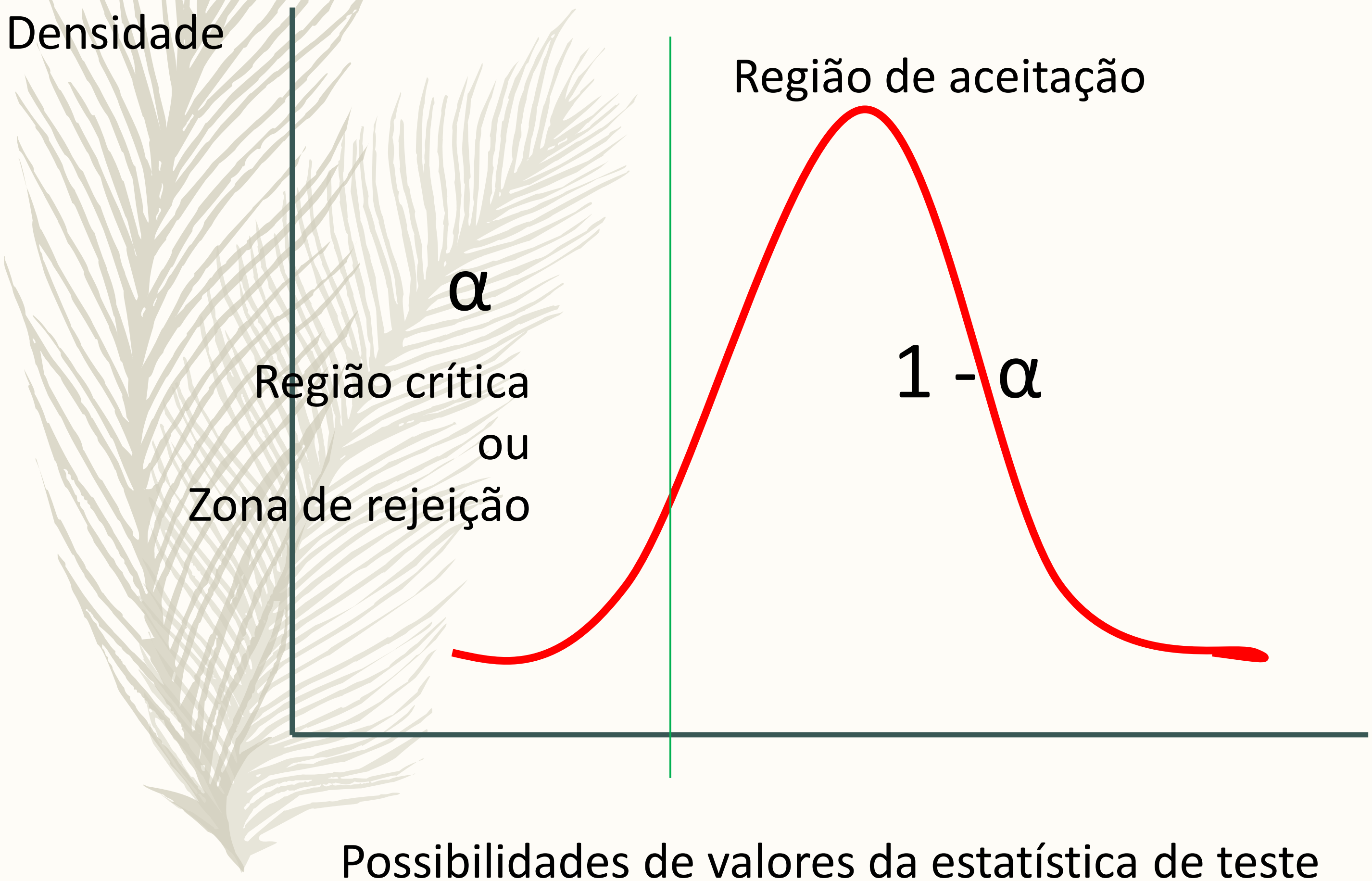
Teste bilateral



testes de hipóteses

$$H_0: \mu \geq X \text{ vs. } H_1: \mu < X$$

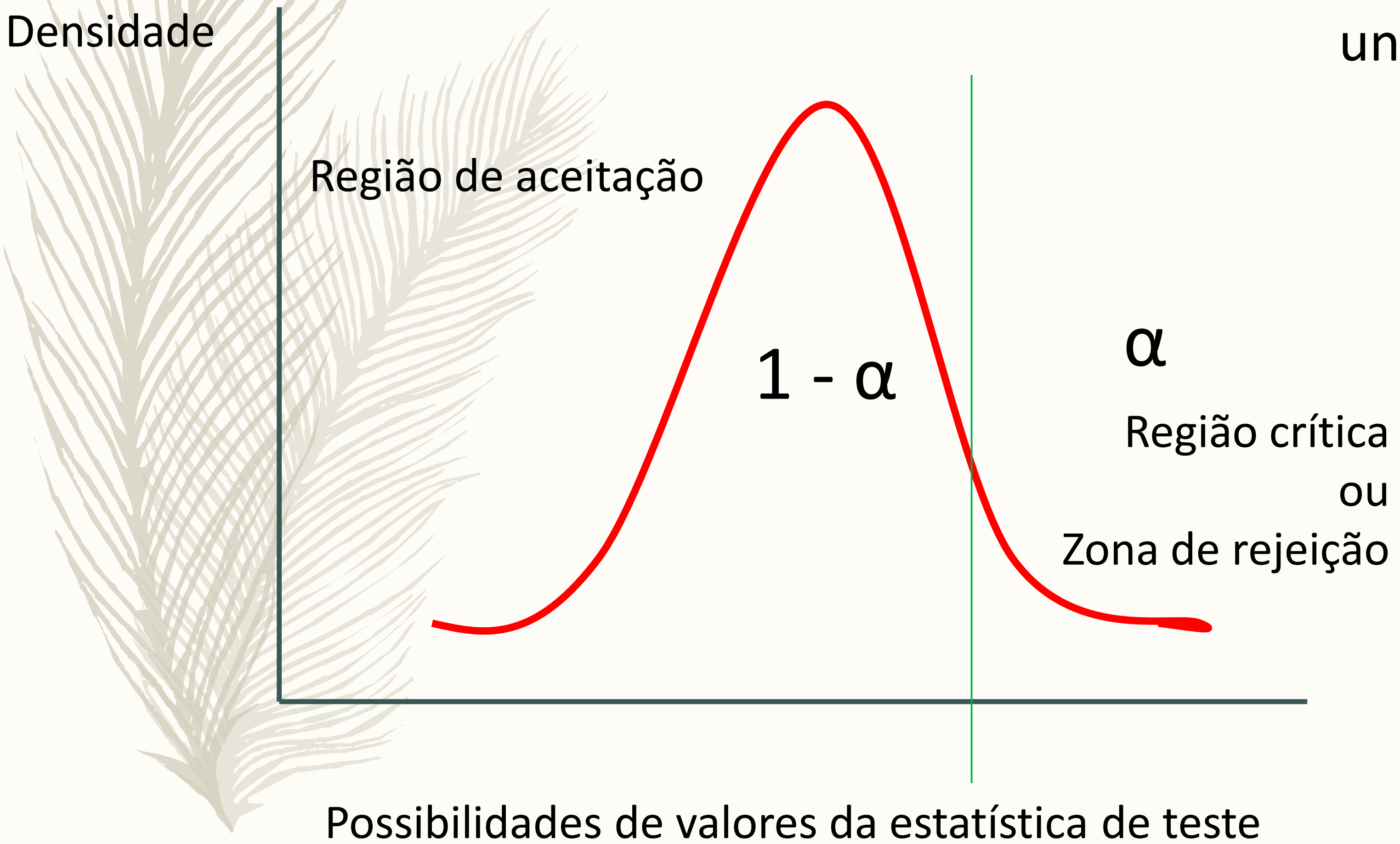
Teste unilateral



testes de hipóteses

$$H_0: \mu \leq X \text{ vs. } H_1: \mu > X$$

Teste unilateral



Valor-p ou *p-value*

O p-value é a probabilidade, sendo H_0 verdade, de obter um valor da estatística de teste igual ou mais extremo que o observado com base numa determinada amostra.

O p-value é dado por:

$P(X>x \mid H_0)$ — teste unilateral "à direita"

$P(X>x \mid H_0)$ ou $P(X<x \mid H_0)$ — teste bilateral

$P(X<x \mid H_0)$ — teste unilateral "à esquerda"

Murtaugh, P. A. 2014 In defense of P values *Ecology* **95**: 611-617

O que o p-value **não** é!

1. The P value is the probability that the test hypothesis is true; for example, if a test of the null hypothesis gave P = 0.01, the null hypothesis has only a 1 % chance of being true; if instead it gave P = 0.40, the null hypothesis has a 40 % chance of being true.
2. The P value for the null hypothesis is the probability that chance alone produced the observed association; for example, if the P value for the null hypothesis is 0.08, there is an 8 % probability that chance alone produced the association.
3. A significant test result (P < 0.05) means that the test hypothesis is false or should be rejected.
4. A nonsignificant test result (P > 0.05) means that the test hypothesis is true or should be accepted.
5. A null-hypothesis P value greater than 0.05 means that no effect was observed, or that absence of an effect was shown or demonstrated.
6. A large P value is evidence in favor of the test hypothesis.
7. The P value is the chance of our data occurring if the test hypothesis is true; for example, P = 0.05 means that the observed association would occur only 5 % of the time under the test hypothesis.
8. If you reject the test hypothesis because P < 0.05, the chance you are in error (the chance your ‘‘significant finding’’ is a false positive) is 5 %.

Greenland, S.; Senn, S. J.; Rothman, K. J.; Carlin, J. B.; Poole, C.; Goodman, S. N. & Altman, D. G. 2016 Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations *European Journal of Epidemiology* **31**: 337-350

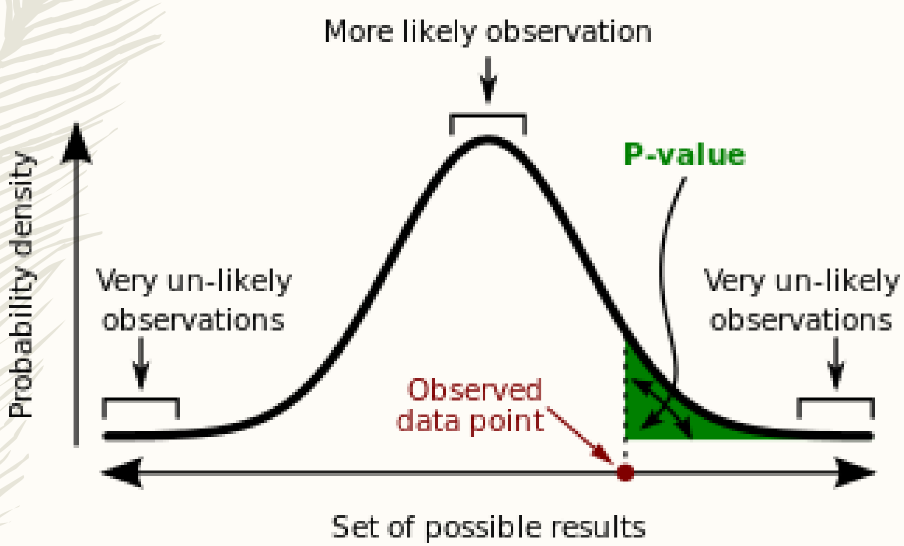
testes de hipóteses

Important:

Pr (observation | hypothesis) ≠ Pr (hypothesis | observation)

The probability of observing a result given that some hypothesis is true is *not equivalent* to the probability that a hypothesis is true given that some result has been observed.

Using the p-value as a "score" is committing an egregious logical error: **the transposed conditional fallacy.**



A **p-value** (shaded green area) is the probability of an observed (or more extreme) result assuming that the null hypothesis is true.

Decisão:

Se p-value < ou igual ao nível de significância desejado, então rejeitamos H0

Se p-value > que nível de significância desejado, então não Rejeitamos H0

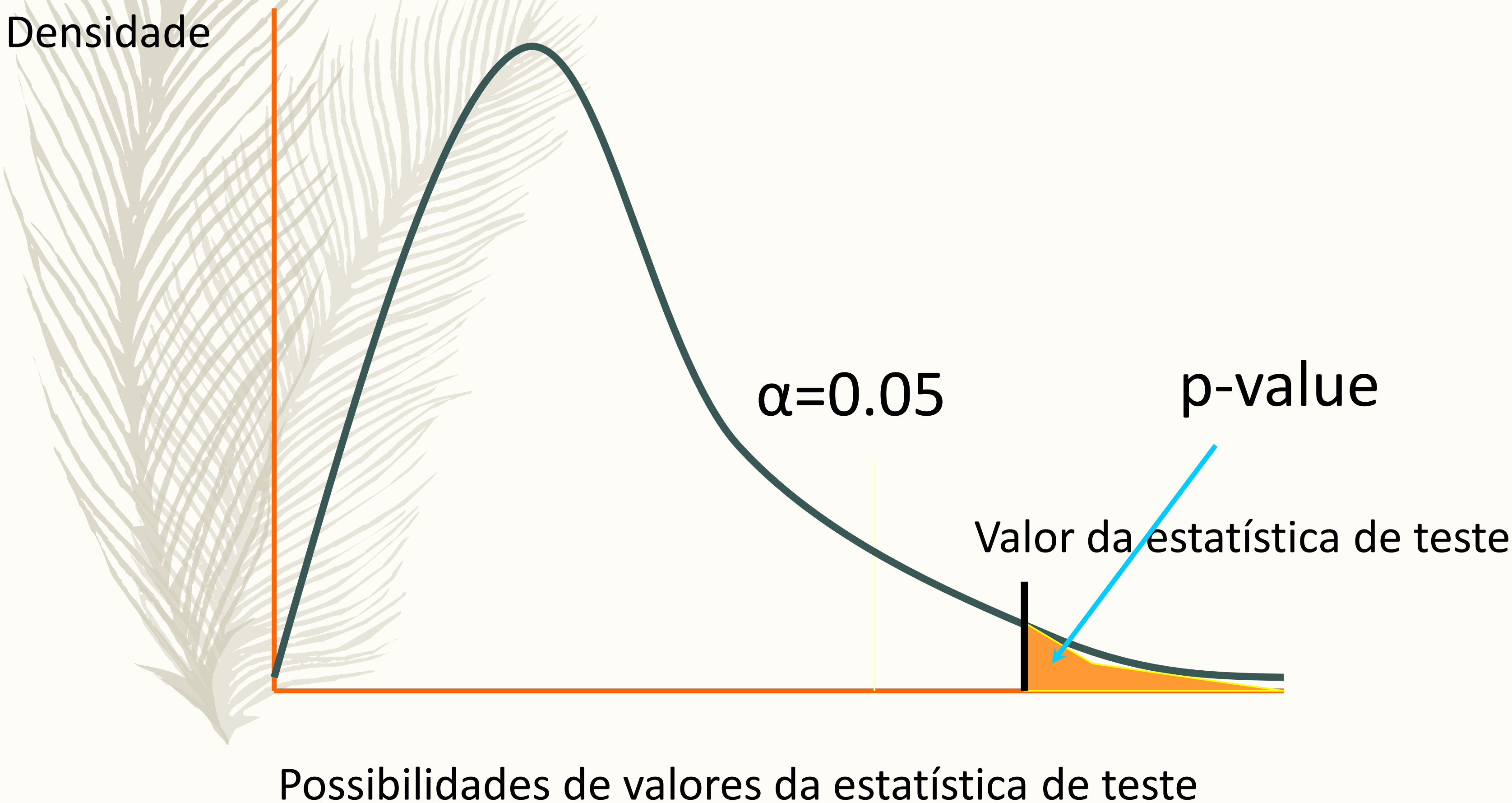
Não dizemos que “aceitamos H0”, apenas que não temos evidências para a rejeitar

It’s the same as when a person is found not-guilty, we should not say they are innocent.

Não dizemos que “aceitamos H1”, apenas que temos evidencias para rejeitar H0

Mas de facto em geral dizemos que uma pessoa que vai para a prisão é culpada, mas se calhar o que deveríamos dizer é que foi considerada culpada.
Era inocente até prova em contrário.

Acerca dos p-values



Acerca dos p-values

Alternativamente à determinação dum valor crítico, podemos estimar a probabilidade de obter um resultado tão extremo para a estatística de teste como o observado, condicional a que H0 seja verdadeira

Este valor é usualmente denominado p-value.

O p-value deve ser comparado com o valor de significância considerado *a priori* (α).

Acerca dos p-values

Alguns autores consideram o p-value em relação a um valor pré-definido de “corte”: a estatística de teste “cai” acima ou abaixo de α.

Outros consideram o p-value como uma medida de distanciamento em relação a H0. Assim, vários níveis de significância usuais podem ser utilizados para fazer referência ao p-value, e.g. 0.05, 0.01, 0.001 (*, **, ***).

O nível de significância 0.05 é quase um dogma, mas cada vez mais controverso (e na realidade, não é especial!).

Podemos efectuar testes de hipóteses sem ter conhecimento prévio da nossa variável?

NÃO!

Todos os testes exigem o cumprimento de certos pressupostos que têm que ser verificados.

Porquê?

Porque para que seja possível a inferência estatística temos que associar a nossa estatística de teste a distribuições de variáveis aleatórias de lei conhecida, para podermos utilizar modelos probabilísticos teóricos na avaliação de se o valor observado da estatística de teste é improvável sob H0, ou não.

Testes Paramétricos vs. Não Paramétricos

teste paramétrico – um teste que tem pressupostos distribucionais em relação à(s) população(ões) de onde provêm a(s) amostra(s) – e.g. t-test, ANOVA

teste não-paramétrico – um teste que não tem pressupostos distribucionais em relação à(s) população(ões) de onde provêm a(s) amostra(s) e.g. teste de Wilcoxon

Testes Paramétricos vs. Não Paramétricos

As estatísticas não-paramétricas podem ser utilizadas em casos de variáveis medidas em escalas ordinais e intervaladas ou de razão.

(os testes paramétricos precisam de escalas, no mínimo, intervaladas)

Testes Paramétricos vs. Não Paramétricos

Os testes não-paramétricos não têm pressupostos?

Têm!

No mínimo, embora se desconheça a distribuição das variáveis, assume-se que as observações são independentes.

Como decidir?

Avaliar se os pressupostos são cumpridos

Não

Transformação dos dados

Não

Sim

Testes não paramétricos

Para a maioria das situações a potência é elevada, ca. 90% ou mais, quando comparada com os testes equivalentes paramétricos

Testes paramétricos

Quando os pressupostos são satisfeitos, a potência dos testes paramétricos é quase sempre superior à dos equivalentes não paramétricos

avaliação de pressupostos

Avaliar se os pressupostos são cumpridos

Não

Sim

Transformação dos dados

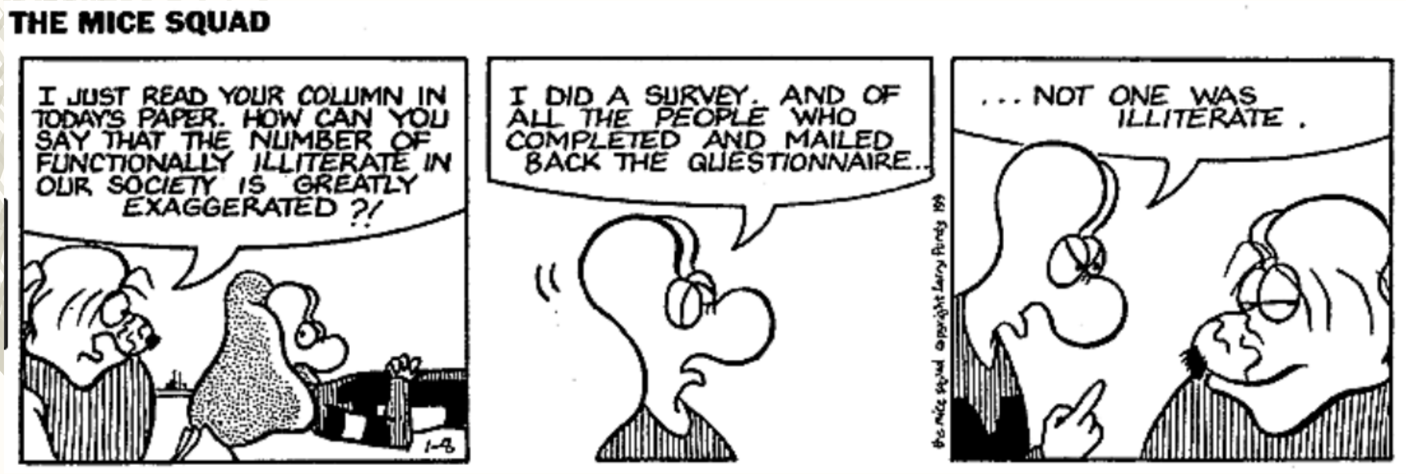
Não

Testes não paramétricos

Testes paramétricos

Testes que não tenham os mesmos pressupostos distribucionais

Mais uma vez, o exemplo para percebermos o que é um teste de hipóteses

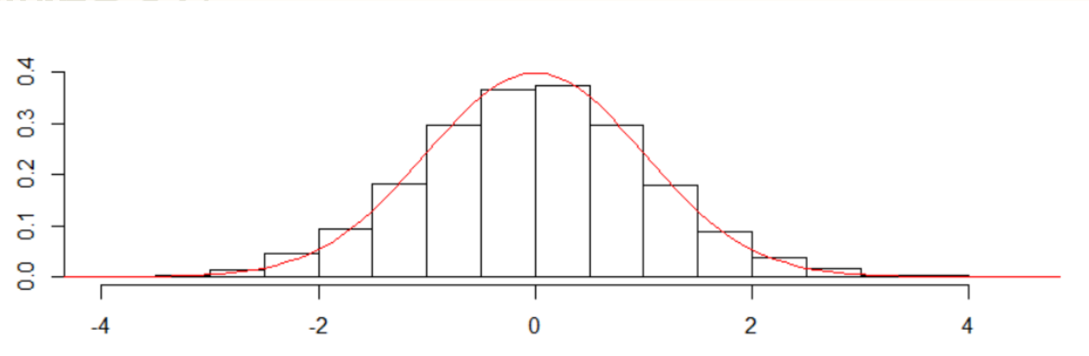


Como funciona um teste de hipóteses:

H0: hipótese nula

H1: hipótese alternativa

Estatística de teste:
uma função da amostra
que, sob H0, tem uma
distribuição conhecida

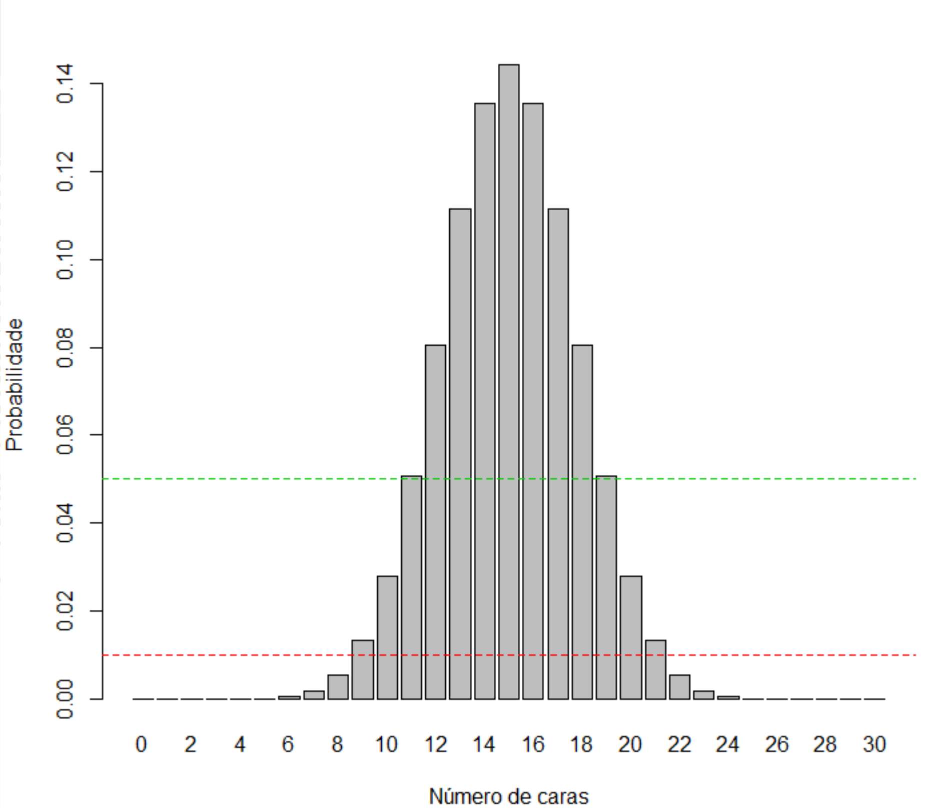


Se o valor da estatística de teste observado for extremo, a probabilidade de observar dados como os observados se H0 fosse verdadeira é baixa, logo tendemos a rejeitar H0

Caso contrário, não há evidências suficientes que nos permitam rejeitar H0

O exemplo das moedas (a.k.a teste dos sinais)

```
#distribution of the number of tails in n throws of a coin
n=30
barplot(dbinom(0:n,size=n,prob=0.5),names.arg=0:n,xlab="Número
de caras",ylab="Probabilidade")
abline(h=c(0.01,0.05),lty=2,col=2:3)
```



O exemplo das moedas (a.k.a teste dos sinais)

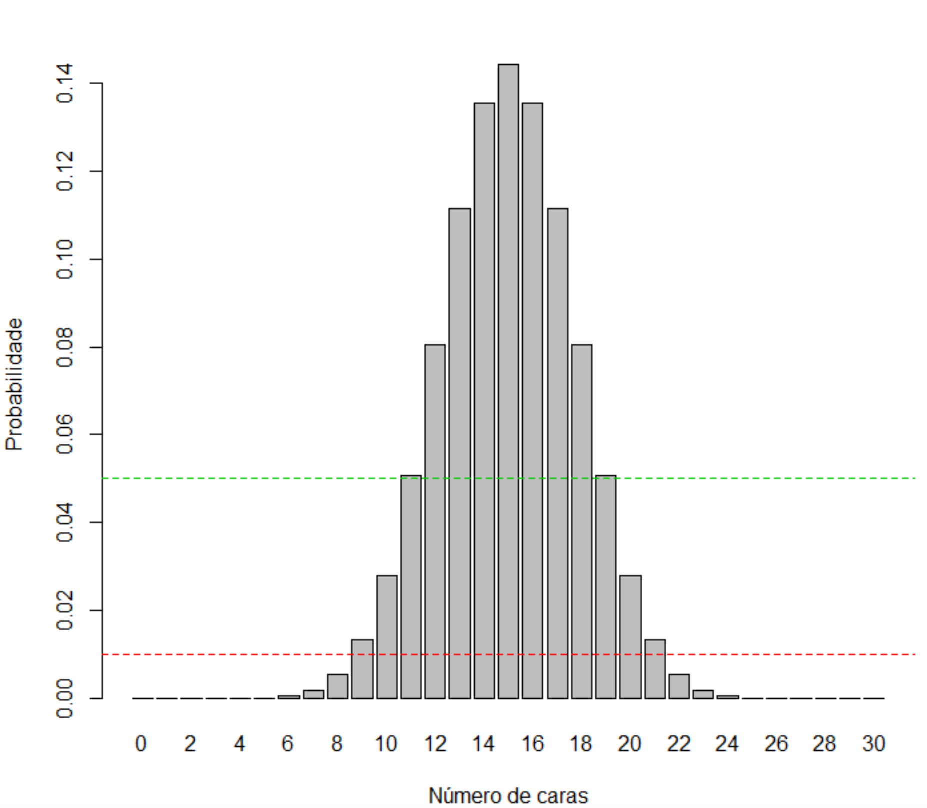
Exemplo:
testar se sex ratio é de 1:1 ou não em ninhadas de *Passer domesticus*

H0: Sex Ratio = 1:1
H1: Sex ratio ≠ 1:1

Amostra:
30 ninhos de *Passer domesticus*

Estatística de teste:
Z= nº de ninhos em que há mais machos que fêmeas

(nota: esta não seria a melhor forma de testar isto, mas o exemplo é pedagógico)



Observar menos de Z<9 ou mais de Z>21, rejeitar H0 (para um nivel de significância de 0.04)!

O exemplo das moedas (a.k.a teste dos sinais)

```
#the distribution of the number of tails in n throws of a coin
n=30
barplot(pbinom(0:n,size=n,prob=0.5),names.arg=0:n,xlab="Número de caras",ylab="Probabilidade comulativa")
abline(h=c(0.01,0.05),lty=2,col=2:3)
```

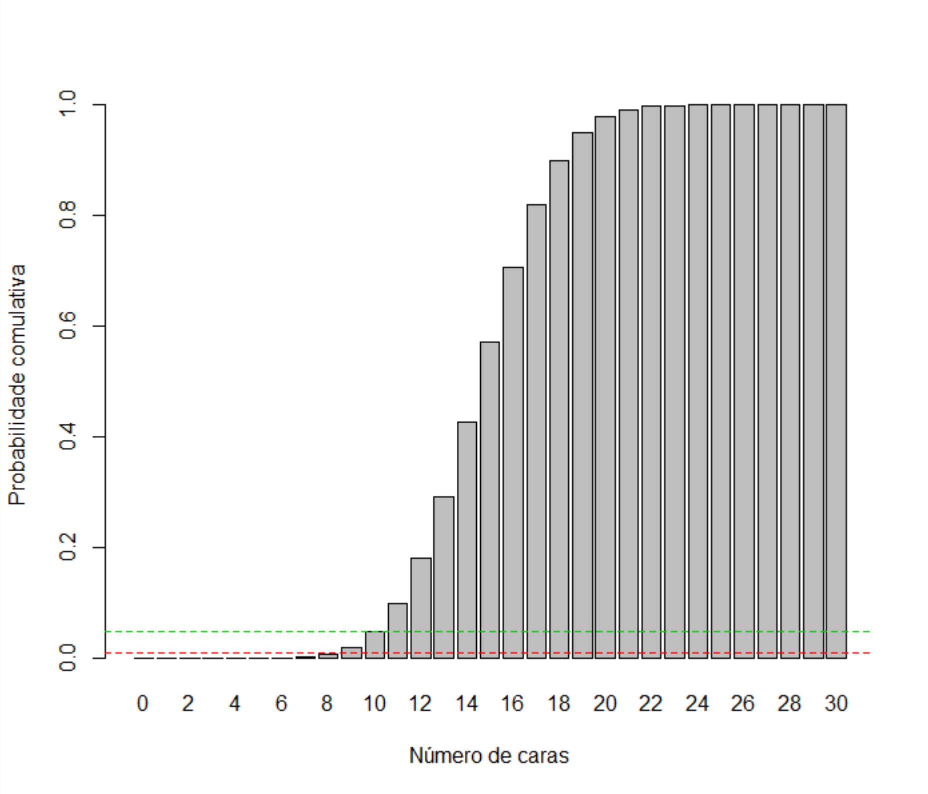
Exemplo: testar se sex ratio é de 1:1 ou não em ninhadas de *Passer domesticus*

H0: Sex Ratio = 1:1
H1: Sex ratio ≠ 1:1

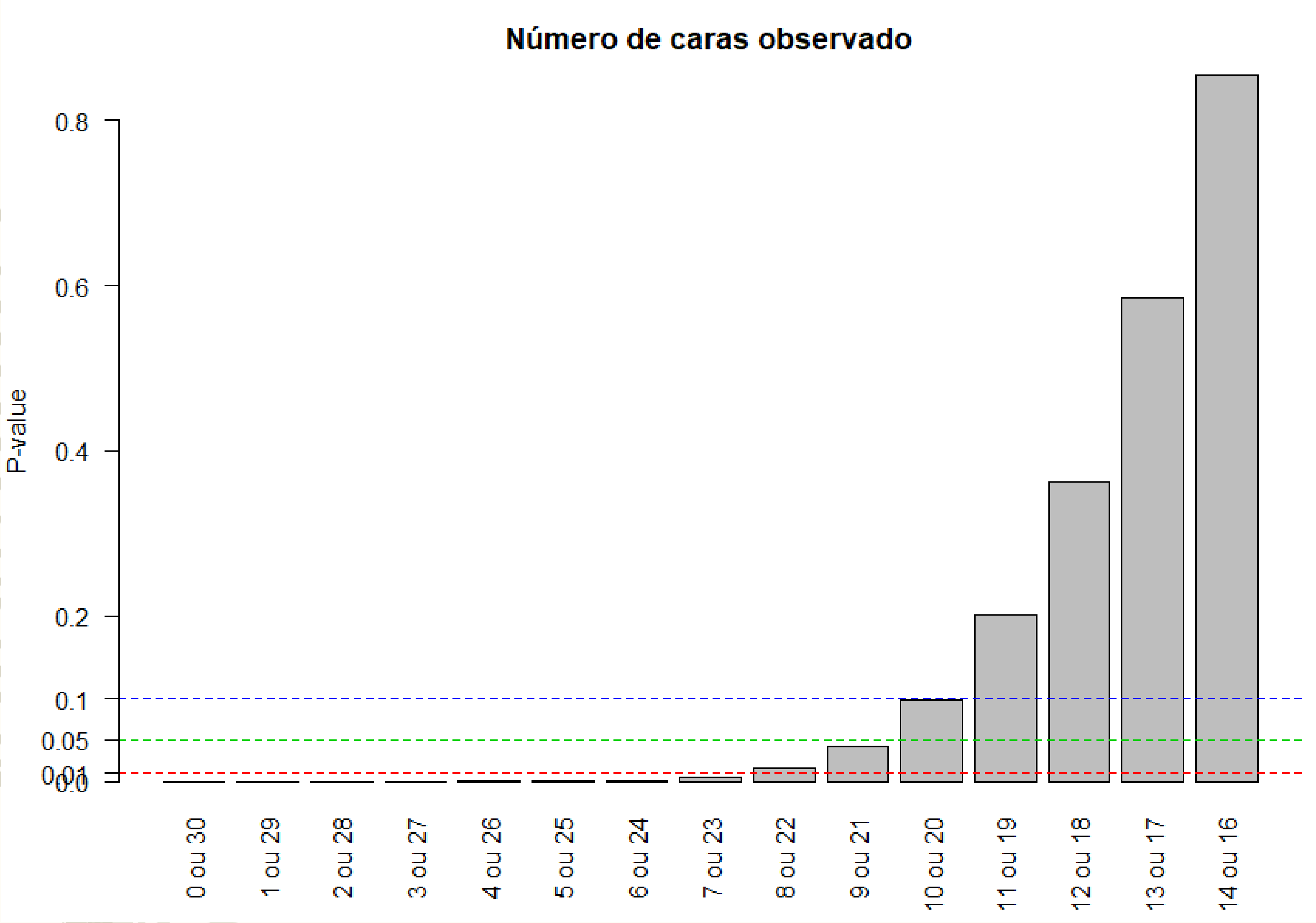
Amostra: 30 ninhos de *Passer domesticus*

Estatística de teste: Z= nº de ninhos em que há mais machos que fêmeas

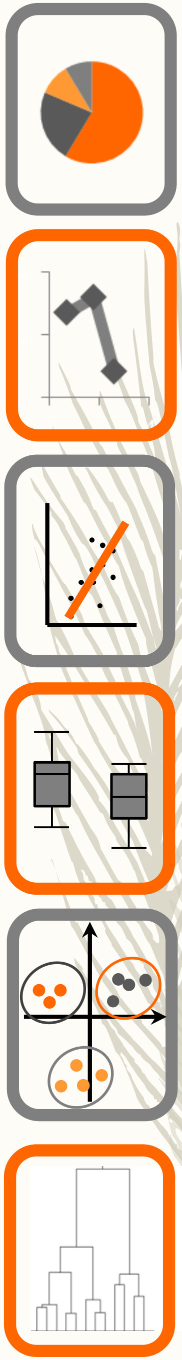
(nota: esta não seria a melhor forma de testar isto, mas o exemplo é pedagógico)



Observar menos de Z<9 ou mais de Z>21, rejeitar H0!



Número de caras observado
P-value
0.8
0.6
0.4
0.2
0.1
0.05
0.01
0.0
0 ou 30
1 ou 29
2 ou 28
3 ou 27
4 ou 26
5 ou 25
6 ou 24
7 ou 23
8 ou 22
9 ou 21
10 ou 20
11 ou 19
12 ou 18
13 ou 17
14 ou 16



Ecologia Numérica

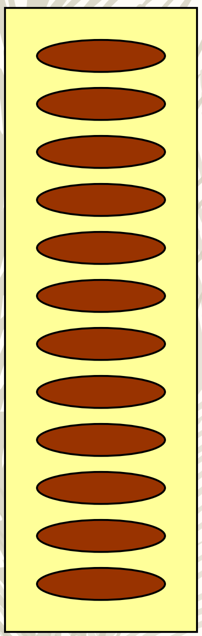
testes de hipóteses
a 1 ou 2 amostras

testes de hipóteses a 1 ou duas amostras

- Quais os testes mais correntes para situações de 1 ou duas amostras?
- Que testes utilizar?
- Quais as condições para a sua aplicação?
- Como interpretar os seus resultados?

testes de hipóteses a 1 amostra

A



Quando temos uma única amostra proveniente duma população e queremos fazer inferência sobre um qualquer parâmetro da população, por exemplo, que o valor médio é igual a um qualquer valor em particular, e.g.

$$\mu=24.3$$

Teste *t-student* para 1 amostra

Hipóteses (exº):

H_0: μ=24.3

H_1: μ≠24.3

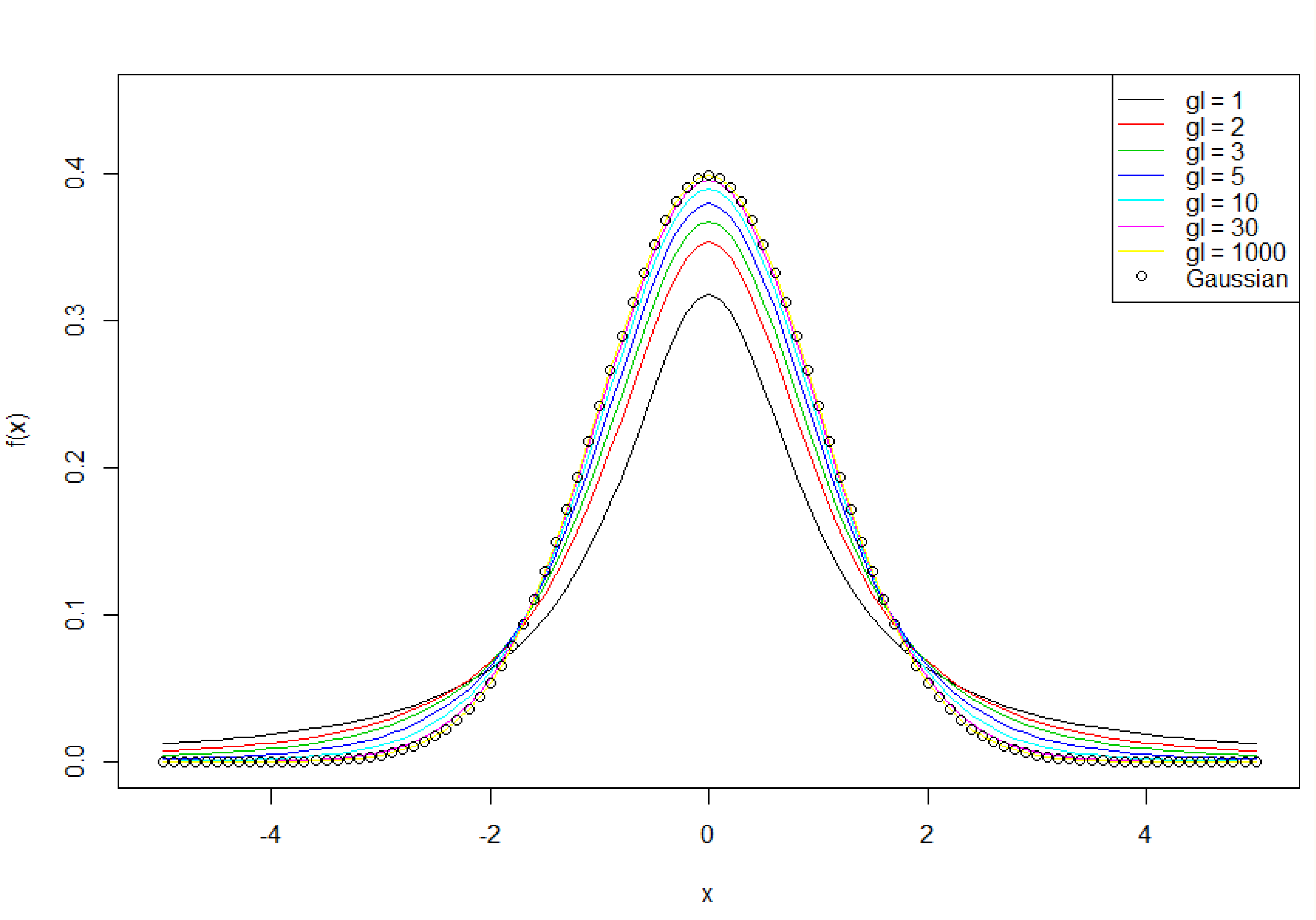
Estatística de teste:

$$t = \frac{\bar{X} - \mu}{s_{\bar{X}}}$$

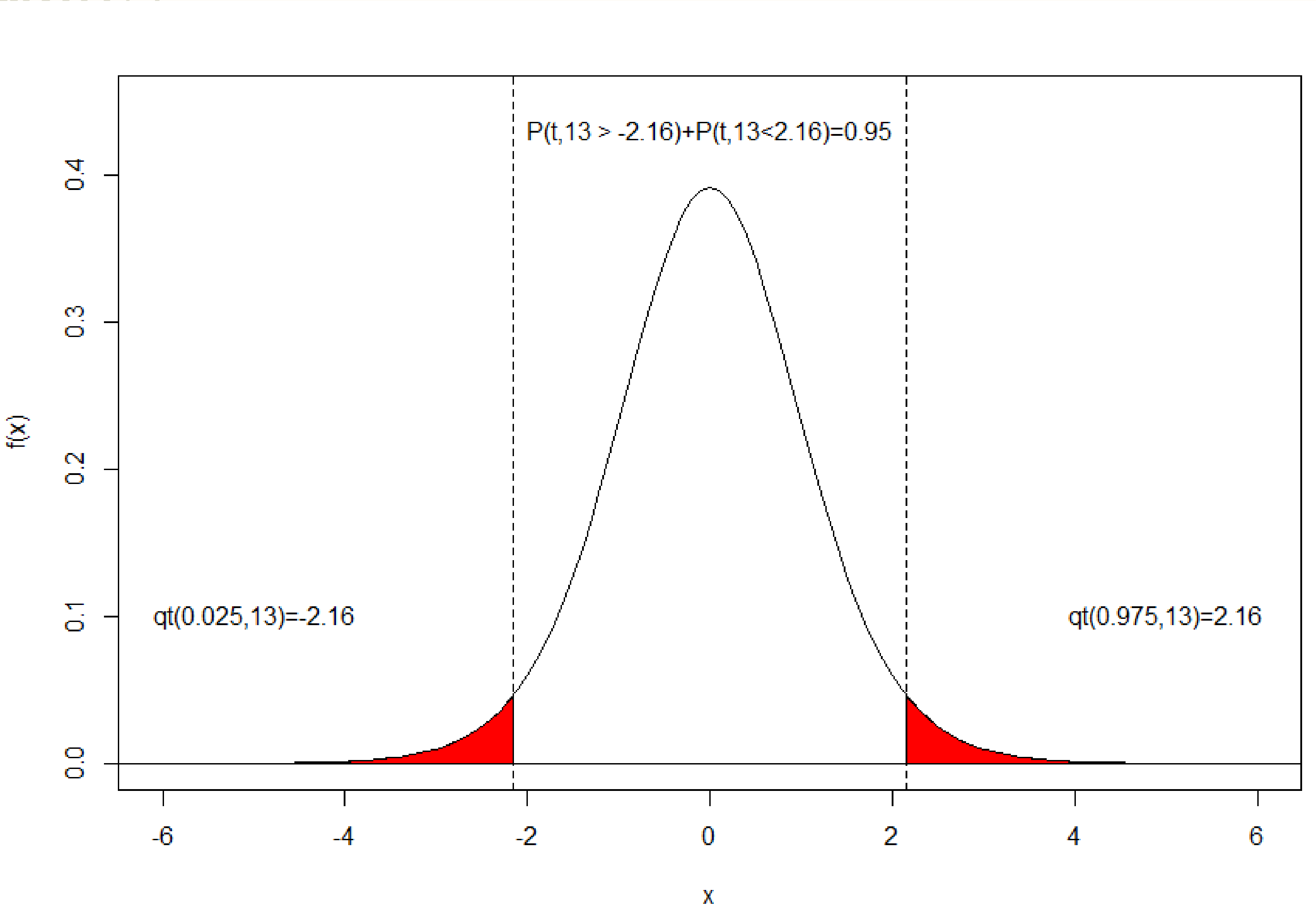
onde

$$s_{\bar{X}} = \sqrt{\frac{s^2}{n}}$$

Distribuição t



Distribuição t (exemplo de região de rejeição bilateral)



Teste t para 1 amostra

Hipóteses (exº):

H_0: μ=24.3
H_1: μ≠24.3

Estatística de teste:

$$t = \frac{\overline{X} - \mu}{s_{\overline{X}}}$$

Valor crítico:

$$t_{\alpha/2,\upsilon}$$

Critério de decisão:

Rejeitar H_0 se: $|t| > t_{\alpha/2,\upsilon}$

Não rejeitar H_0, caso contrário

Teste t para 1 amostra

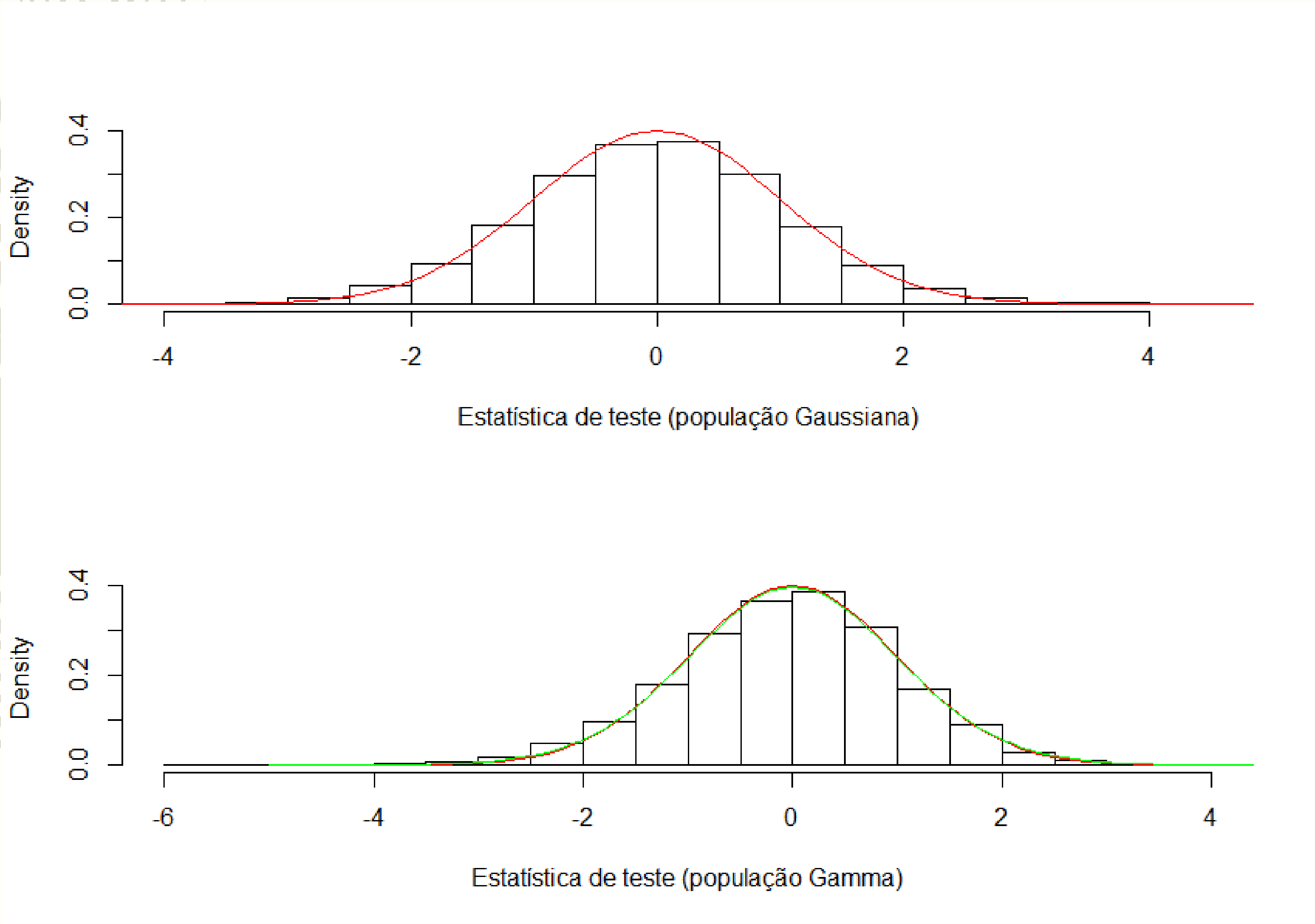
Pressupostos do teste *t*:

Independência das observações

Os dados são provenientes duma população com distribuição Gaussiana.

No entanto, o teste *t* é bastante robusto: a sua validade não é grandemente afectada por desvios moderados do pressuposto de Gaussianidade.

Um exemplo: população parente Gaussiana vs. população parente Gamma; distribuição da estatística de teste, com n=30



Teste t para 1 amostra

Hipóteses (exº):

H_0: $\mu \geq 0$

H_1: $\mu < 0$

Estatística de teste:

$$t = \frac{\bar{X} - \mu}{s_{\bar{X}}}$$

Valor crítico:

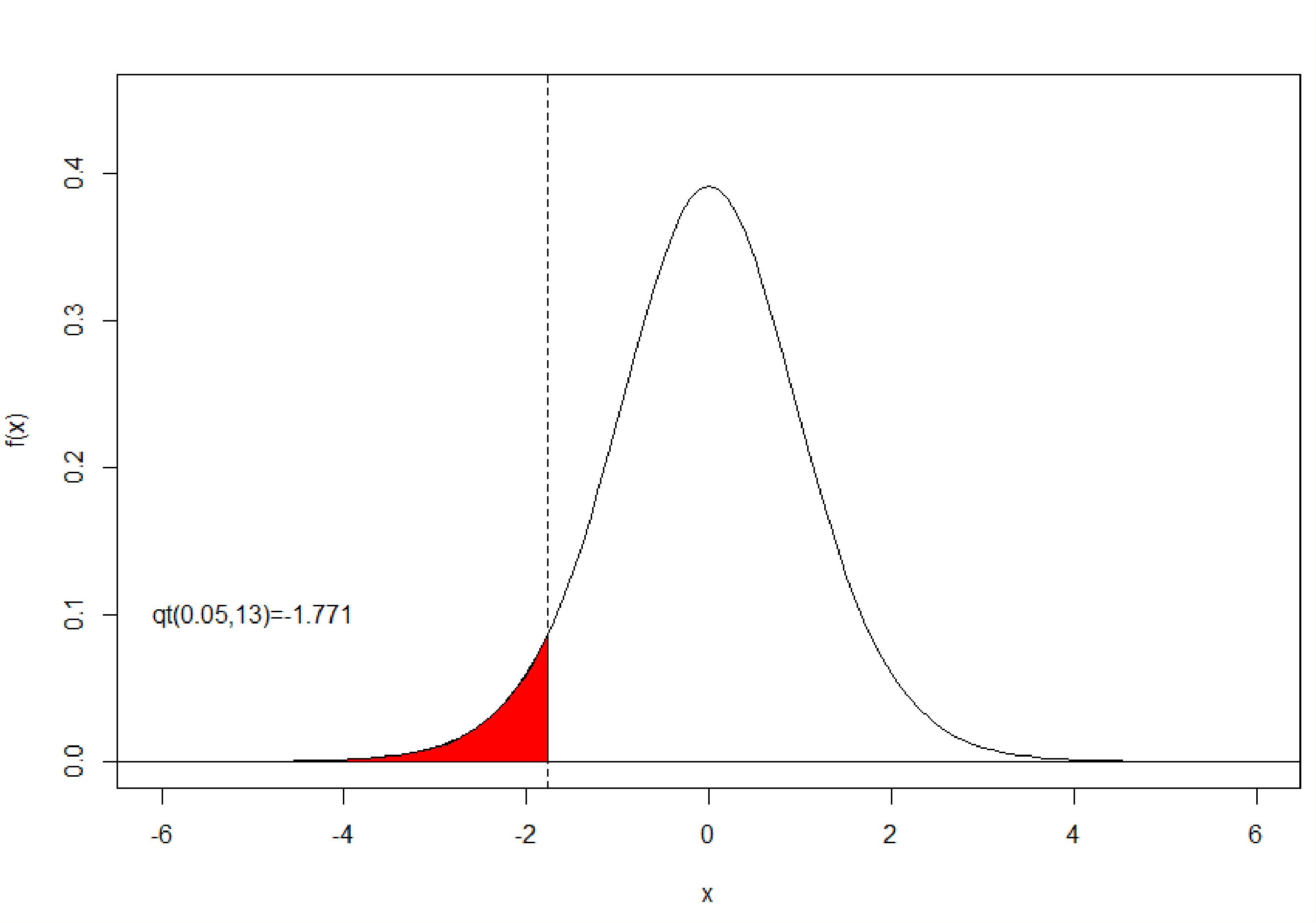
$$-t_{\alpha(1),\upsilon}$$

Critério de decisão:

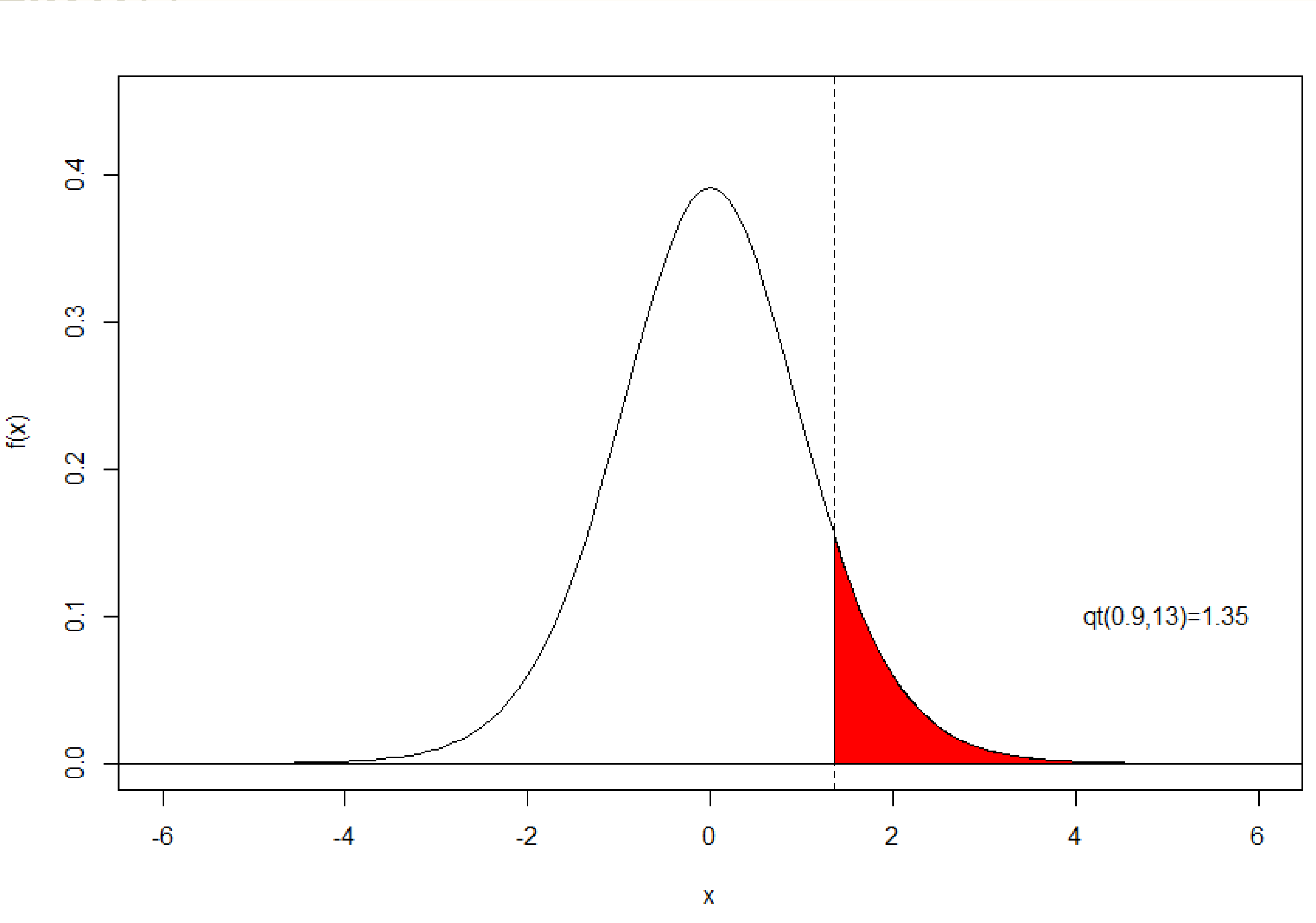
Rejeitar H_0 se: $t < -t_{\alpha(1),\upsilon}$

Não rejeitar H_0, caso contrário

Teste t para 1 amostra, unilateral à esquerda, α 0.05



Teste t para 1 amostra, unilateral à direita, α 0.1



Para garantir os mesmos resultados em qualquer computador

```
> set.seed(123)
> t.test(rnorm(50),mu=0)

        One Sample t-test

data:  rnorm(50)
t = 0.26275, df = 49, p-value = 0.7938
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.2287258  0.2975329
sample estimates:
 mean of x
0.03440355

> t.test(rnorm(50),mu=1)

        One Sample t-test

data:  rnorm(50)
t = -6.6661, df = 49, p-value = 2.184e-08
alternative hypothesis: true mean is not equal to 1
95 percent confidence interval:
 -0.1109170  0.4037335
sample estimates:
mean of x
0.1464083

> 2*pt(-6.6661,49)
[1] 2.183586e-08
```