
BIOESTATÍSTICA

TESTES DE HIPÓTESES

O QUE É HIPÓTESE?

- PALAVRA GREGA (*ὑπόθεσις*) QUE SIGNIFICA **BASE, FUNDAMENTO, PRINCÍPIO DE ALGO, PROPOSIÇÃO.**
- **DEFINIÇÃO:** “PROPOSIÇÃO QUE SE ADMITE, INDEPENDENTEMENTE DO FATO SE SER VERDADEIRA OU FALSA, MAS UNICAMENTE A TÍTULO DE UM PRINCÍPIO A PARTIR DO QUAL SE PODE DEDUZIR UM DETERMINADO CONJUNTO DE CONSEQÜÊNCIAS” (DICIONÁRIO HOUAISS)

EXEMPLOS DE HIPÓTESES

- $\mu = 100$, μ É A MÉDIA DO TEOR DE FLUOR NO AR
- $\pi > 0,1$, ONDE π É A PROPORÇÃO DE ANIMAIS DA ESPÉCIE A NA COMUNIDADE.

TESTE DE HIPÓTESES

- AO TESTAR HIPÓTESES SEMPRE EXISTIRÁ DUAS HIPÓTESES CONTRADITÓRIAS EM CONSIDERAÇÃO.
- $\mu = 100$ E OUTRA $\mu \neq 100$
- $\pi = 0,1$ E OUTRA $\pi > 0,1$
- O OBJETIVO É DECIDIR QUAL A HIPÓTESE CORRETA
- PORTANTO, TESTE DE HIPÓTESES É O MÉTODO USADO PARA DECIDIR QUAL DAS DUAS PROPOSIÇÕES CONTRADITÓRIAS É A MAIS CORRETA.

DUAS HIPÓTESES

- HIPÓTESE NULA, REPRESENTADA POR H_0 , É A PROPOSIÇÃO QUE É CONSIDERADA INICIALMENTE VERDADEIRA. A OUTRA HIPÓTESE É CHAMADA DE ALTERNATIVA E REPRESENTADA POR H_A .
- PARA TESTAR HIPÓTESES USA-SE TESTES ESTATÍSTICOS.

TESTES ESTATÍSTICOS

- SÃO PROCEDIMENTOS ATRAVÉS DOS QUAIS SE TESTAM HIPÓTESES DA PESQUISA CONTRA A(S) HIPÓTESE(S) ALTERNATIVA(S) BASEANDO-SE NOS DADOS COLETADOS.
- SE OS DADOS SÃO EXTREMOS EM RELAÇÃO A HIPÓTESE NULA ESTA É REJEITADA.
- UM TESTE ESTATÍSTICO CONCLUI SOBRE A VALIDADE DA HIPÓTESE NULA.
- H_0 PODE SER VERDADEIRA OU FALSA E PODE SER ACEITA OU REJEITADA.

ERROS DO TESTE ESTATÍSTICO

<i>TIPO DE ERRO</i>		
	NÃO REJEITA H_0	REJEITA H_0
H_0 É VERDADEIRA	CORRETA	ERRO TIPO I
H_0 É FALSA	ERRO TIPO II	CORRETA

<i>DECISÕES</i>		
	NÃO REJEITA H_0	REJEITA H_0
H_0 É VERDADEIRA	CORRETA	INCORRETA
H_0 É FALSA	INCORRETA	CORRETA

PROBABILIDADE DOS ERROS

<i>PROBABILIDADES</i>		
	NÃO REJEITA H_0	REJEITA H_0
H_0 É VERDADEIRA	$1 - \alpha$	α
H_0 É FALSA	β	$1 - \beta$

REGRAS DE DECISÃO (NÍVEL DE SIGNIFICÂNCIA)

- O PESQUISADOR ESTABELECE *A PRIORI* QUAL A PROBABILIDADE DE REJEIÇÃO DE H_0 .
- SE OS DADOS COLETADOS SÃO MUITO DIFERENTES DA HIPÓTESE NULA DAÍ REJEITA H_0 .
- A PROBABILIDADE DE REJEIÇÃO DE H_0 PODE SER 0,05 OU 0,01 OU SEJA 5% OU 1%.

VALOR DE p

- O TESTE ESTATÍSTICO MEDE O VALOR DE p .
- O VALOR p DE UM TESTE ESTATÍSTICO É A PROBABILIDADE CALCULADA DA HIPÓTESE NULA.
- SE O VALOR p FOR MENOR QUE A PROBABILIDADE PREVIAMENTE DEFINIDA PELO PESQUISADOR, REJEITA-SE A HIPÓTESE NULA.

INTERPRETAÇÃO DO VALOR p

<i>VALOR p</i>	<i>EVIDÊNCIA CONTRA HIPÓTESE NULA</i>
$p \leq 0,01$	MUITO FORTE
$0,01 < p \leq 0,05$	FORTE
$0,05 < p \leq 0,10$	MODERADA
$0,10 < p$	POUCA OU NENHUMA

TIPOS DE TESTES

- BILATERAL

$$H_0: \mu = 100$$

$$H_A: \mu \neq 100$$

- UNILATERAL

$$H_0: \mu = 100$$

$$H_A: \mu > 100$$

- O **SAS** USA A TABELA UNILATERAL

- $t = \text{tinv}(0.95, \text{g.l.})$; PARA 5% - UNILATERAL

- $t = \text{tinv}(0.975, \text{g.l.})$; PARA 5% - BILATERAL

O TESTE t (STUDENT)

- COMPARA DUAS POPULAÇÕES (DUAS MÉDIAS).
- TESTA SE AS MÉDIAS DE DOIS GRUPOS SÃO ESTATISTICAMENTE DIFERENTES.
- EX.: AMOSTRAS COLHIDAS EM DOIS LOCAIS DIFERENTES
- EXPERIMENTOS COM DOIS TRATAMENTOS
- TIPOS DE TESTE: PAREADO E NÃO PAREADO.

STUDENT



WILLIAM SEALEY GOSSET

NASCEU EM 1876 E FALECEU EM 1937 NA INGLATERRA.

ESTUDOU QUÍMICA E MATEMÁTICA EM OXFORD.

EM 1899 FOI CONTRATADO PELA CERVEJARIA GUINNESS, EM DUBLIN, IRLANDA . DURANTE O SEU TRABALHO NA GUINNESS DESENVOLVEU IMPORTANTE TRABALHO EM ESTATÍSTICA E É CONSIDERADO O PRECURSOR DA ESTATÍSTICA MODERNA.

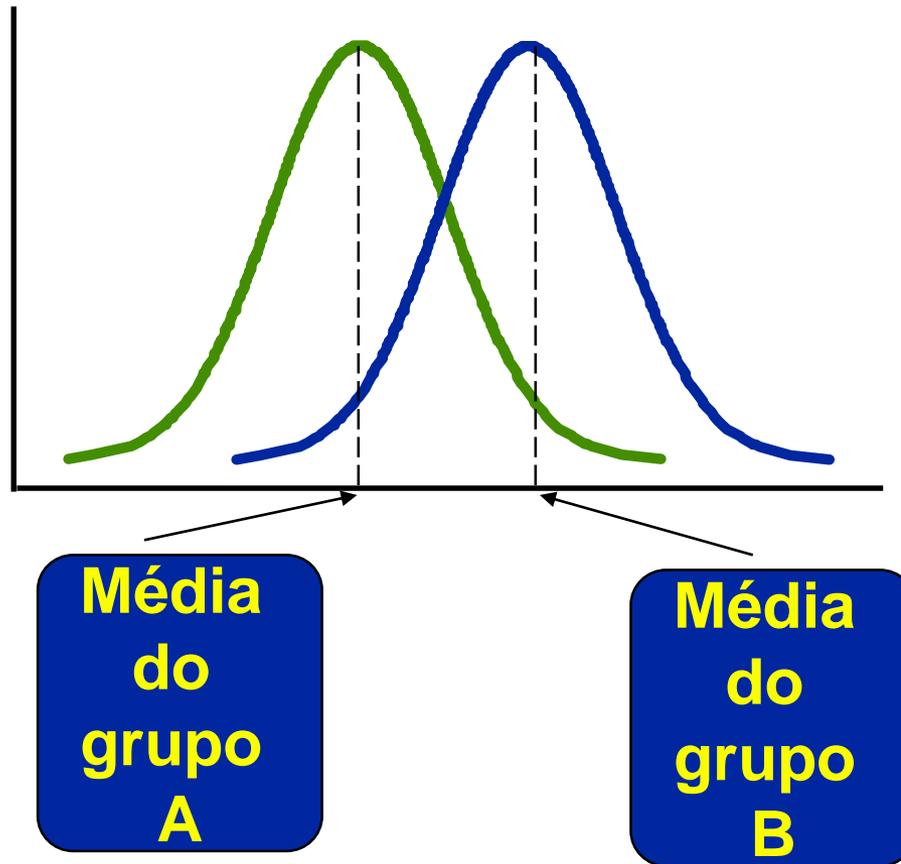
PROPÔS O TESTE t E ESCREVEU TRABALHOS COM O PSEUDÔNIMO DE STUDENT.

TROCOU CORRESPONDÊNCIAS COM FISHER, NEYMAN E PEARSON

Student (William Sealy Gosset)

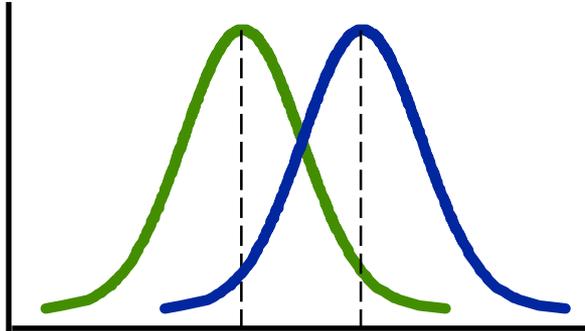


Há diferença estatística entre as médias dos dois grupos?

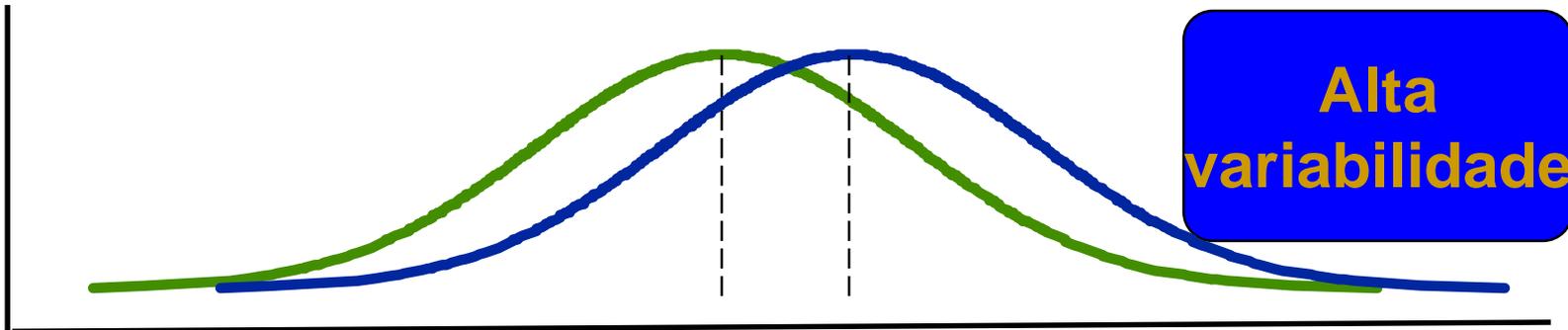


O que a diferença significa?

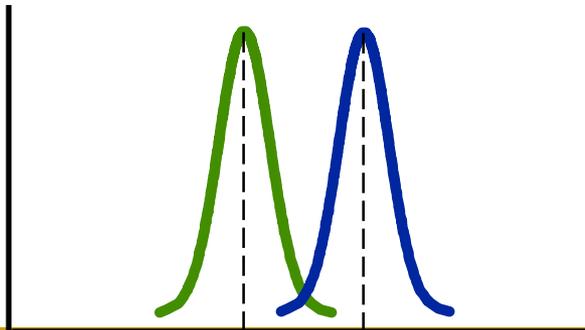
Variabilidade média



Alta variabilidade



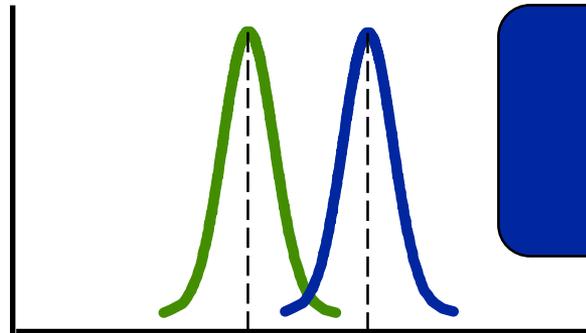
Baixa variabilidade



O que a diferença significa?

- A diferença estatística é uma função da diferença entre as médias em relação a variabilidade.
- Uma pequena diferença entre médias com alta variabilidade pode ser devido ao acaso.

**Baixa
variabilidade**



**Mostra a
maior diferença**

O que nós estimamos?

Sinal

Ruído

Diferença entre as médias dos dois grupos

Variabilidade dos grupos

=

=

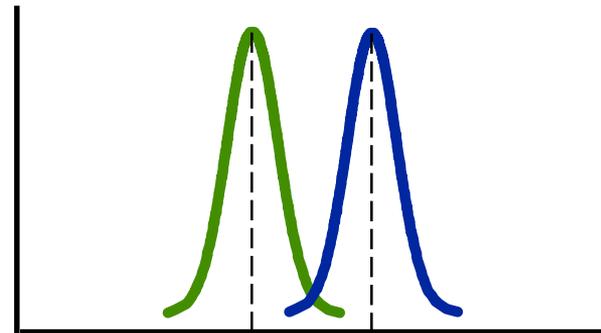
=

$$\bar{X}_T - \bar{X}_C$$

$$\text{DesvPad}(\bar{X}_T - \bar{X}_C)$$

Valor de t

**Baixa
variabilidade**



TESTE t NÃO PAREADO

■ AMOSTRAS DE TAMANHOS IGUAIS

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} \implies s_{\bar{x}_1 - \bar{x}_2} = \sqrt{2 \frac{s_p^2}{n}}$$

$$s_p^2 = \frac{SQ1 + SQ2}{2(n-1)} \implies g.l. = 2(n-1)$$

$$SQ1 = \sum x_{1i}^2 - \frac{(\sum x_{1i})^2}{n}$$

$$SQ2 = \sum x_{2i}^2 - \frac{(\sum x_{2i})^2}{n}$$

TESTE t NÃO PAREADO

- AMOSTRAS DE TAMANHOS DIFERENTES (MUDA O VALOR DA VARIÂNCIA)

$$s_p^2 = \frac{SQ1 + SQ2}{(n_1 - 1) + (n_2 - 1)}$$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

O PROC TTEST

- O **SAS** UTILIZA O PROCEDIMENTO **TTEST** PARA REALIZAR O TESTE t NÃO PAREADO.
- COMPARA DUAS MÉDIAS E APRESENTA CORREÇÃO PARA HETEROSCEDASTICIDADE (HETEROGENEIDADE DA VARIÂNCIA)

```
TITLE1'**** B I O E S T A T Í S T I C A ****';
TITLE3'**** Análise de dados pelo teste t-Student não pareado ****';
TITLE5'***** Setembro de 2019 ****';
TITLE7'***** Alimentação de teiú em cativeiro *****';
FOOTNOTE1'TRATAMENTO A=10 % DE PROTEÍNA';
FOOTNOTE2'          B=30 % DE PROTEÍNA';
DATA ALIMEN;
INPUT TRAT $ CRESC;
DATALINES;
A 23
A 19
A 21
A 9
A 25
A 33
A 18
A 22
B 34
B 32
B 19
B 17
B 35
B 26
B 16
B 29
;
ODS PDF FILE='C:\Arquivos2019\Bioestatística2019\RESULTADO_TEIU.PDF';
PROC TTEST DATA=ALIMEN;
CLASS TRAT;
VAR CRESC;
RUN;
ODS PDF CLOSE;
```

**Programa SAS para
análise de dados
não pareados –
PROC TTEST**

**** *BIOESTATÍSTICA* ****

**** *Análise de dados pelo teste t-Student não pareado* ****

***** *Setembro de 2019* ****

***** *Alimentação de teiú em cativeiro* ****

The TTEST Procedure

Variable: CRESC

TRAT	N	Mean	Std Dev	Std Err	Minimum	Maximum
A	8	21.2500	6.7771	2.3961	9.0000	33.0000
B	8	26.0000	7.7460	2.7386	16.0000	35.0000
Diff (1-2)		-4.7500	7.2777	3.6388		

TRAT	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
A		21.2500	15.5842	26.9158	6.7771
B		26.0000	19.5242	32.4758	7.7460
Diff (1-2)	Pooled	-4.7500	-12.5545	3.0545	7.2777
Diff (1-2)	Satterthwaite	-4.7500	-12.5675	3.0675	

**Resultado do
Programa
SAS**

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	14	-1.31	0.2128
Satterthwaite	Unequal	13.757	-1.31	0.2132

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	7	7	1.31	0.7333

TRATAMENTO A=10 % DE PROTEÍNA

B=30 % DE PROTEÍNA

Exercício

- Foi realizado um estudo para determinar se havia influência de um gene sobre a resistência a geadas de plantas de uma determinada espécie. Foram produzidas 10 plantas com o gene e 10 sem. Os resultados estão na tabela a seguir. Defina as hipóteses e faça um programa SAS para testá-las.

Dados: Biotecnologia Crescimento mm/semana

Com	Sem
12	9
11	5
9	7
8	12
14	7
12	10
12	9
11	12
10	6
9	4

TESTE t PAREADO

- EM UMA FLORESTA TOMAMOS 20 ÁRVORES AO ACASO E EM 10 DELAS MEDIMOS A ALTURA COM UM APARELHO E NAS OUTRAS 10 ESTIMAMOS VISUALMENTE A ALTURA DAS ÁRVORES.
- EM OUTRA FLORESTA TOMAMOS 10 ÁRVORES AO ACASO E EM CADA ÁRVORE MEDIMOS A ALTURA COM UM APARELHO E NAS MESMAS ÁRVORES ESTIMAMOS VISUALMENTE AS ALTURAS.
- QUAL DOS DOIS DELINEAMENTOS É MAIS ADEQUADO, PARA COMPARAR OS DOIS MÉTODOS DE MEDIÇÃO?

t PAREADO

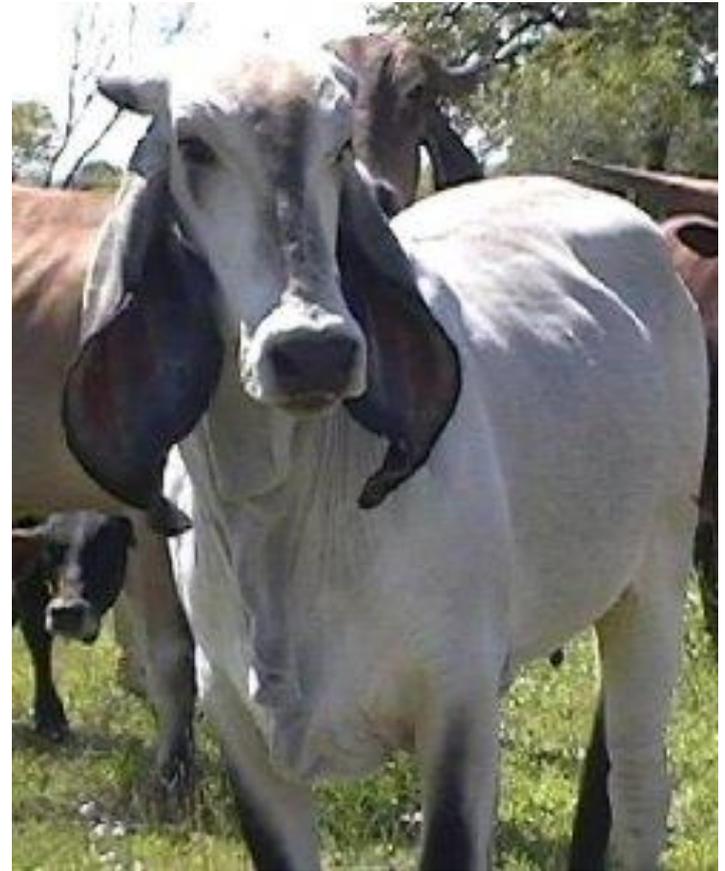
- CADA PAR DE OBSERVAÇÕES É INDEPENDENTE DOS DE MAIS.
- AS DIFERENÇAS PROVÉM DE UMA DISTRIBUIÇÃO NORMAL.
- EX. 1: INFLUÊNCIA DE FENÓIS NA ATIVIDADE EM RATOS.
- EX. 2: AVALIAÇÃO DO NÚMERO DE MOSCA DO CHIFRE EM BOVINOS.

MOSCA DO CHIFRE

Haematobia irritans



COMO AMOSTRAR MOSCA DO CHIFRE?



RATOS DE LABORATÓRIO



**ESTUDO DE FENÓIS NO AR E
ATIVIDADE: CONSUMO DE
ALIMENTO, ÁGUA E
CRESCIMENTO**

TESTE t PAREADO: FÓRMULAS

$$t = \frac{\bar{d}}{s_{\bar{d}}}, \implies d_i = \text{diferença entre os valores}$$

$$g.l. = n - 1$$

$$s_d^2 = \frac{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}{n - 1}$$

$$s_{\bar{d}} = \sqrt{\frac{s_d^2}{n}}$$

```
TITLE2'**** B I O E S T A T Í S T I C A ****';
TITLE4'**** Análise de dados pelo teste t pareado ****';
TITLE6'**** SETEMBRO DE 2019 ****';
TITLE8'*** Consumo de água por ratos ***';
FOOTNOTE1'TRATAMENTO A=RESPIRANDO FENÓIS';
FOOTNOTE2' B=SEM FENÓIS NO AR';
```

```
DATA A;
```

```
INPUT C_FENOL S_FENOL;
```

```
DIF=C_FENOL-S_FENOL;
```

```
DATALINES;
```

```
14 12
```

```
19 17
```

```
26 20
```

```
12 13
```

```
11 12
```

```
21 20
```

```
17 14
```

```
;
```

```
ODS PDF FILE='C:\BIOESTATISTICA2019\RESULTADO_RATOS.PDF';
```

```
PROC UNIVARIATE NORMAL DATA=A;
```

```
VAR DIF;
```

```
RUN;
```

```
ODS PDF CLOSE;
```

PROGRAMA SAS PARA O TESTE t PAREADO

The SAS System
**** *BIOESTATÍSTICA* ****

**** *Análise de dados pelo teste t pareado* ****

**** *SETEMBRO DE 2019* ****

*** *Consumo de água por ratos* ***

The UNIVARIATE Procedure
Variable: DIF

Moments			
N	7	Sum Weights	7
Mean	3	Sum Observations	21
Std Deviation	1.63299316	Variance	2.66666667
Skewness	0.96448659	Kurtosis	1.1625
Uncorrected SS	79	Corrected SS	16
Coeff Variation	54.4331054	Std Error Mean	0.6172134

Basic Statistical Measures			
Location		Variability	
Mean	3.000000	Std Deviation	1.63299
Median	3.000000	Variance	2.66667
Mode	2.000000	Range	5.00000
		Interquartile Range	2.00000

Note: The mode displayed is the smallest of 2 modes with a count of 2.

**Resultado
Programa
SAS**

Tests for Location: $\mu_0=0$				
Test	Statistic		p Value	
Student's t	t	4.860556	Pr > t	0.0028
Sign	M	3.5	Pr >= M	0.0156
Signed Rank	S	14	Pr >= S	0.0156

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.932528	Pr < W	0.5726
Kolmogorov-Smirnov	D	0.214286	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.049105	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.305288	Pr > A-Sq	>0.2500

TRATAMENTO A=RESPIRANDO FENÓIS
B=SEM FENÓIS NO AR

Resultado do Teste t pareado

Resultado do Teste de Normalidade. Não normalidade = teste não paramétrico.

Exercício

- Um estudo para avaliar a influência de um estímulo visual sobre a pressão sistólica em homens foi realizado com 12 indivíduos. Com os dados a seguir definir as hipóteses e testá-las através de um programa SAS.

Pressão sistólica em homens

Indivíduo	Antes	Depois
1	122	136
2	130	141
3	125	136
4	136	141
5	110	131
6	118	130
7	133	140
8	124	130
9	141	146
1	129	139
11	122	131
12	126	139

OBRIGADO !!!

ATÉ A
PRÓXIMA !!!