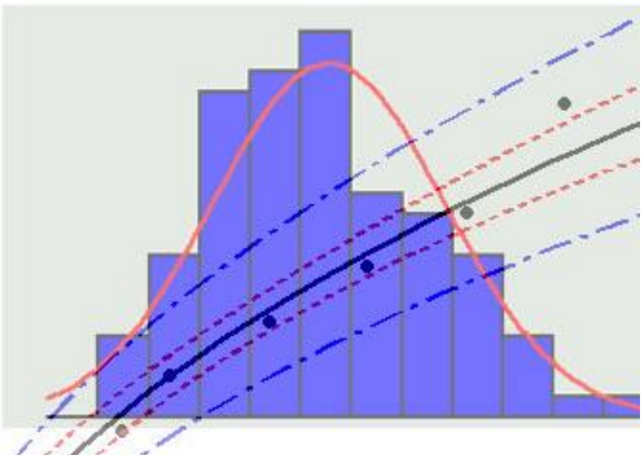
Teste de Hipóteses
(ou Teste de Significância)



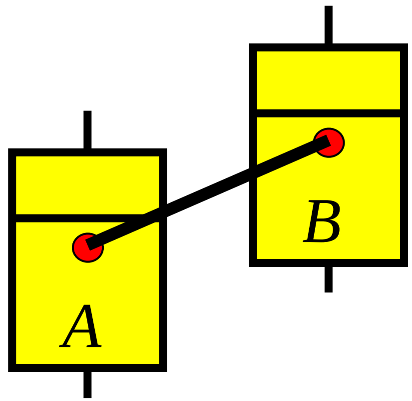
Pedro Paulo Balestrassi

www.pedro.unifei.edu.br

ppbalestrassi@gmail.com

35-36291161 / 999012304 (cel)

São diferenças estatisticamente significantes?



$$H_0 : \mu_A = \mu_B$$

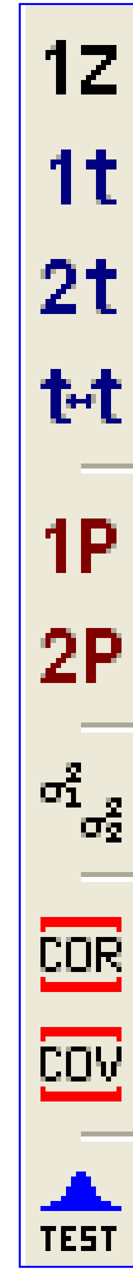
Rejeita-se H_0
P value<0.05

São diferenças Estatisticamente Significantes?

A perda em um processo caiu de uma **proporção** de 10% para 5%.

Dois operadores tem em **média** tempos de 34 e 40 minutos, respectivamente para desenvolver uma atividade.

Quanto maior o número de horas-extras maior a insatisfação dos trabalhadores (**correlação** de 0.40)



As hipóteses e seus erros

•Na afirmação: "Uma pessoa é considerada inocente até que se prove o contrário pois é um erro maior condenar um inocente do que libertar um culpado.", defina:

- •Erros Tipo I e Tipo II
- •Hipóteses Nula e Alternativa

RC		**Situação Real**	
		Ho	**H1**
Decisão	**Ho**	Correta	Erro II β
	H1	Erro I α	Correta

$$\alpha = P(\text{Erro Tipo I}) \quad \beta = P(\text{Erro Tipo II})$$

Procura-se manter *alfa*, e *beta* pequenos, em 5% e 20% respectivamente

Em TH o termo Rejeitar é mais correto!

- Use os termos: **H_0 não pode ser Rejeitada** **X** **Rejeita-se H_0**

- Muito embora as pessoas digam "**Aceitar a Hipótese Nula H_0**" isso é **errado** do ponto de vista estatístico. Podem haver outras hipóteses além da hipótese nula estabelecida.

Em TH o termo Rejeitar é mais correto!

- **Rejeitar H_a** significa que não há evidência de que o acusado seja culpado. **Aceitar H_a** é rejeitar H_0 em favor de H_a. Existe suficiente evidência a favor de H_a.

- **Rejeitar H_a não siginifica Aceitar H_0**. Apenas não existe evidência para suportar H_a.O júri não está dizendo que o acusado é inocente. Apenas não encontrou evidências para condená-lo.

Testes Paramétricos x Testes Não Paramétricos

$H_o : \theta = \theta_o$

$H_1 : \theta \neq \theta_o$ (teste bilateral), ou

$H_1 : \theta > \theta_o$ (teste unilateral), ou

$H_1 : \theta < \theta_o$ (teste unilateral),

Paramétricos

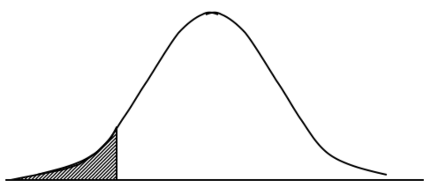
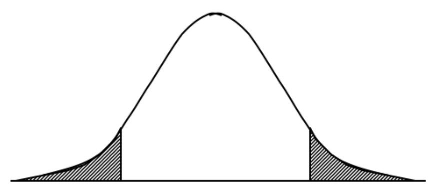
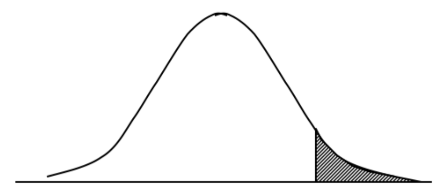
Não Paramétricos

H_o: Dados Normais H_1: Dados não normais

A Hipótese Nula sempre estabelece que o parâmetro iguala o valor especificado na Hipótese Alternativa.

H_1 é a hipótese a ser investigada (*research hypothesis)*

Testes Unilaterais e Bilaterais

	One-Tail Test (left tail)	Two-Tail Test	One-Tail Test (right tail)
Ex.:	$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$
			

Testes Unilaterais são também chamados Direcionais
(**directional test** or **one-tailed test**)

Alfa e Beta estão associados a erros

- Erro do Tipo I ≡ Rejeitar H_o sendo H_o verdadeira, com a seguinte probabilidade de ocorrência:

  P (Erro I) = P (rejeitar H_o | H_o é verdadeira) = α

  P_value

- Erro do Tipo II ≡ Não rejeitar H_o sendo H_o falsa, com a seguinte probabilidade.

  P (Erro II) = P (não rejeitar H_o / H_o é falsa) = β

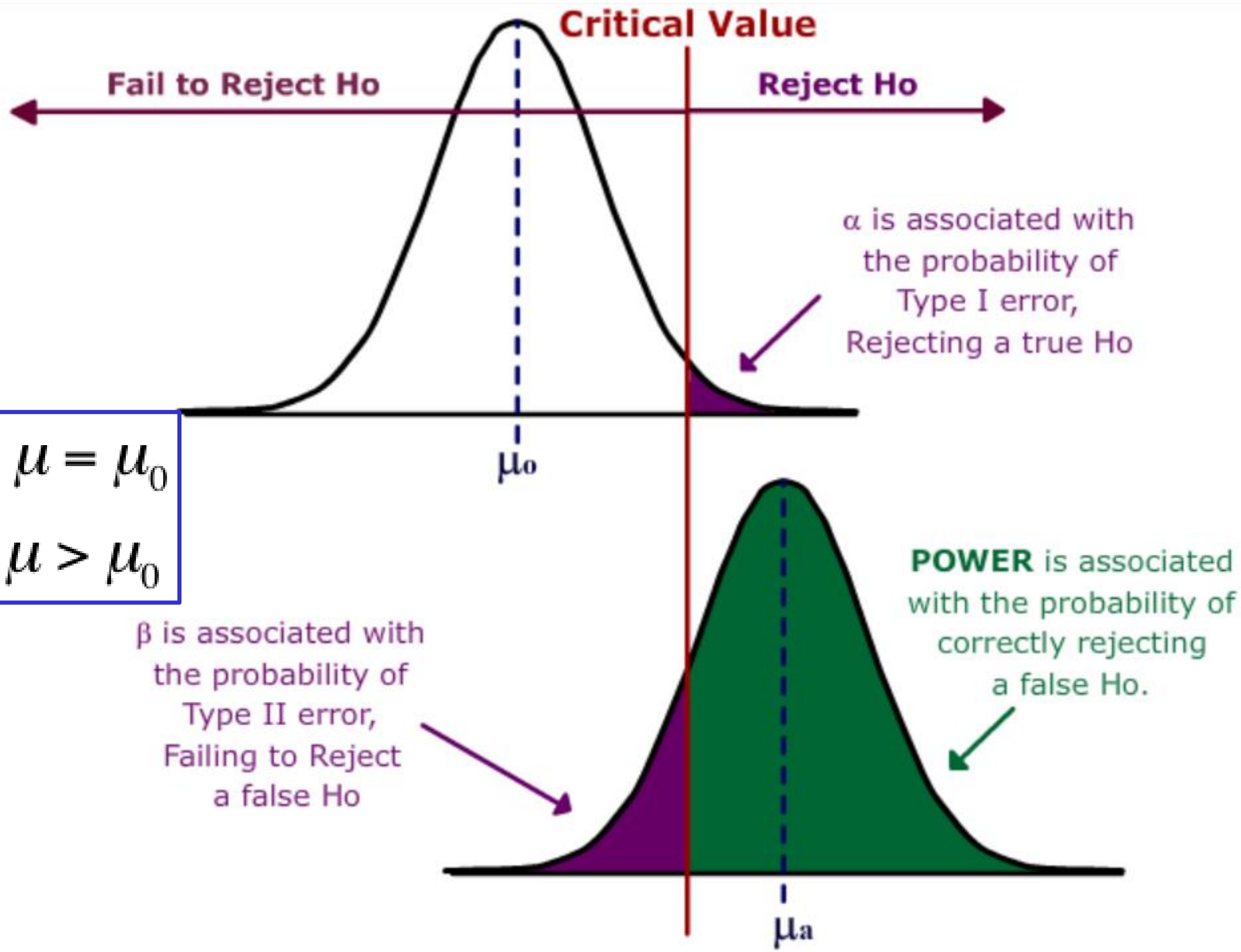
As duas probabilidades são inversamente relacionadas para um mesmo tamanho de amostra. Aumentando uma, a outra diminui.

Aumentando-se o tamanho da amostra as duas probabilidades diminuem simultaneamente.

Alfa é o P-Value definido no estudo: Geralmente 5% ou 1%

Teste de hipóteses graficamente

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$



O algoritmo básico de implementação de Teste de Hipóteses

Primeiro passo:	Definir $H_o : \theta = \theta_0$ e H_1
Segundo passo:	Escolher a Estatística de teste adequada, $\hat{\theta}$
Terceiro passo:	Escolher α e estabelecer RC
Quarto passo:	Calcular $\hat{\theta}$ com base em uma amostra de tamanho n extraída da população.
Quinto passo:	Rejeitar H_o caso $\hat{\theta} \in RC$. Não rejeitar H_o em caso contrário.

No Minitab: Análise do p-value !

Algoritmo Básico: exemplo 1/4

Tentando encorajar as pessoas a deixar o carro na garagem, a prefeitura de uma cidade informa que o tempo de se estacionar um carro pela manhã no centro da cidade é de **15 minutos ou mais** em média.

Uma pessoa desconfia que tal informação é falsa e resolve investigar. Ele pretende provar que o tempo médio é **menor que 15** minutos e, para tanto, observa o tempo que 10 carros levam para estacionar.

Nesse caso tem-se

Primeiro Passo: Definir H_0 e H_1

$$\mathrm{H}_0 : \mu = 15$$

$$\mathrm{H}_a : \mu < 15$$

Algoritmo Básico: exemplo 2/4

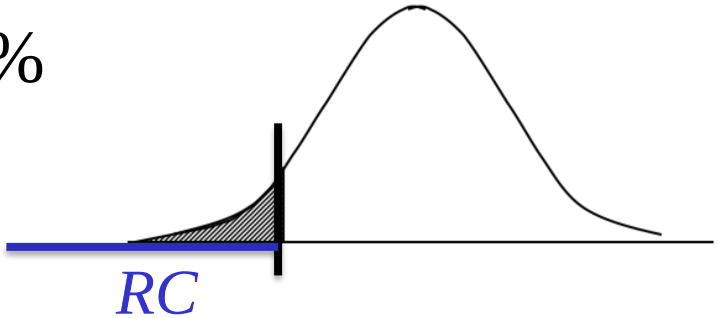
Segundo Passo: Escolha da Estatística de Teste Adequada

Pelos resultados do Teorema Central do Limite, com desvio padrão desconhecido, a estatística t=student deve ser empregada.

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

Terceiro Passo: Escolha de Alfa e da Região Crítica

α=5%



$t_{\alpha,n-1}$=$t_{5\%,9}$=-1,833

No lugar de RC o Minitab apresenta Lower Bounds /Upper Bounds.

Para teste unilateral a Direita:

$$95\%LB = \bar{x} - t_{\alpha,n-1} s/\sqrt{n}$$

Para teste unilateral a Esquerda:

$$95\%UB = \bar{x} + t_{\alpha,n-1} s/\sqrt{n}$$

Para teste bilateral

$$95\%B = \bar{x} \pm t_{\alpha/2,n-1} s/\sqrt{n}$$

Algoritmo Básico: exemplo 3/4

Quarto Passo: Cálculo da estatística de teste com base na amostra

$X = \{13\ 13\ 14\ 23\ 12\ 11\ 7\ 18\ 12\ 16\}$

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

$\bar{X} = 13{,}90$

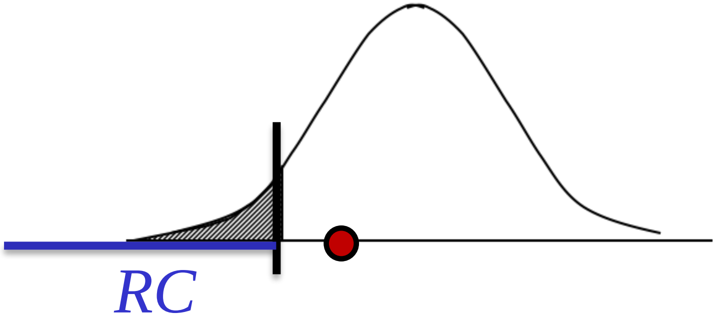
$s = 4{,}33$

$n = 10$

$$t_0 = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{13{,}9 - 15}{4{,}33/\sqrt{10}} = -0{,}80$$

Quinto Passo: Tomada de Decisão

Como t_0 **<u>Não</u>** pertence a região crítica, ***Não se Rejeita H₀***

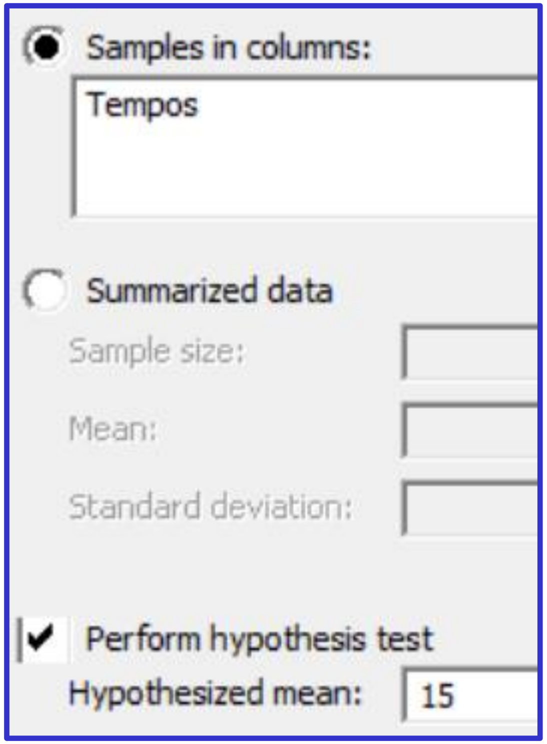


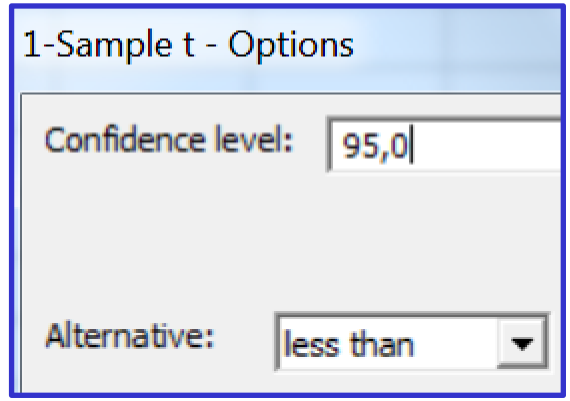
$t_{\alpha,n-1} = t_{5\%,9} = -1{,}833$

Não há evidência para se rejeitar a declaração da prefeitura

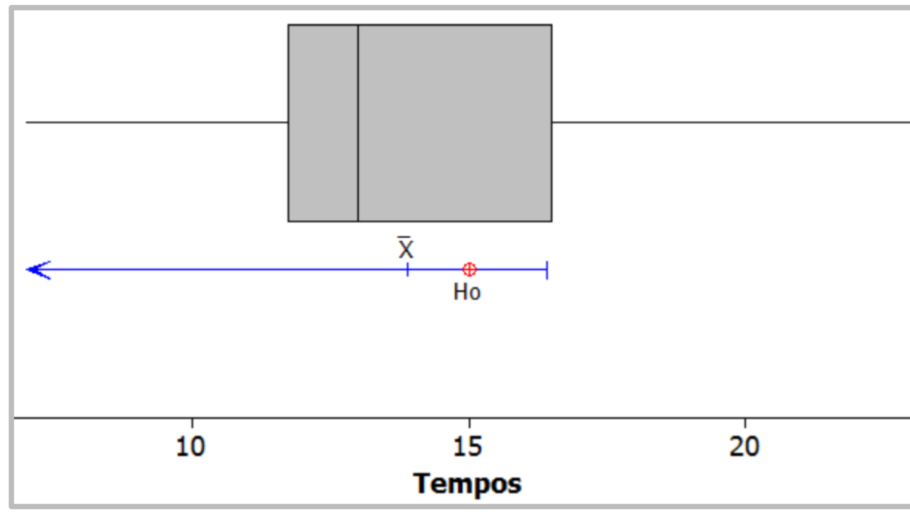
Algoritmo Básico: exemplo 4/4

Tempos
13
13
14
23
12
11
7
18
12
16





1-Sample t



Test of mu = 15 vs < 15

Variable	N	Mean	StDev	SE Mean	95% Upper Bound	T	P
Tempos	10	13,90	4,33	1,37	16,41	-0,80	**0,221**

$$95\%UB = \bar{x} + t_{\alpha,n-1} s/\sqrt{n}$$

$$SE\ Mean = StDev/\sqrt{n}$$

Não se Rejeita H_0

Interpreting the P-value

Simbologia comum para valores de P-value:

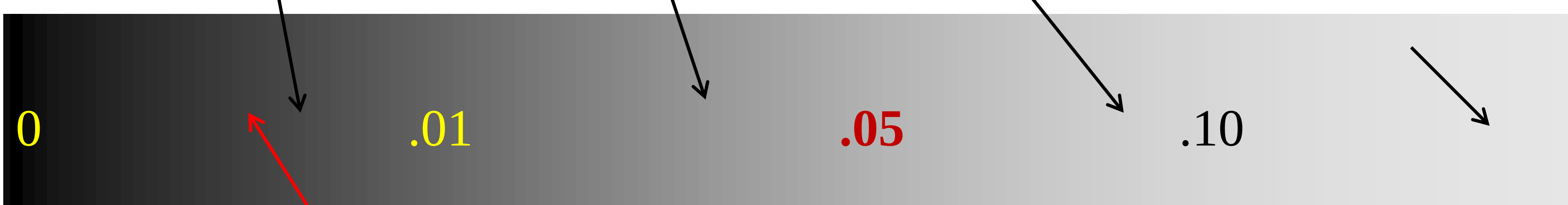
n.s	p>0,05
*	0,01<p<0,05
**	0,005<p<0,01
***	0,0001<p<0,005

Fortíssima Evidência
(Altamente Significante)

Forte Evidência
(Significante)

Fraca Evidência
(Pouco Significante)

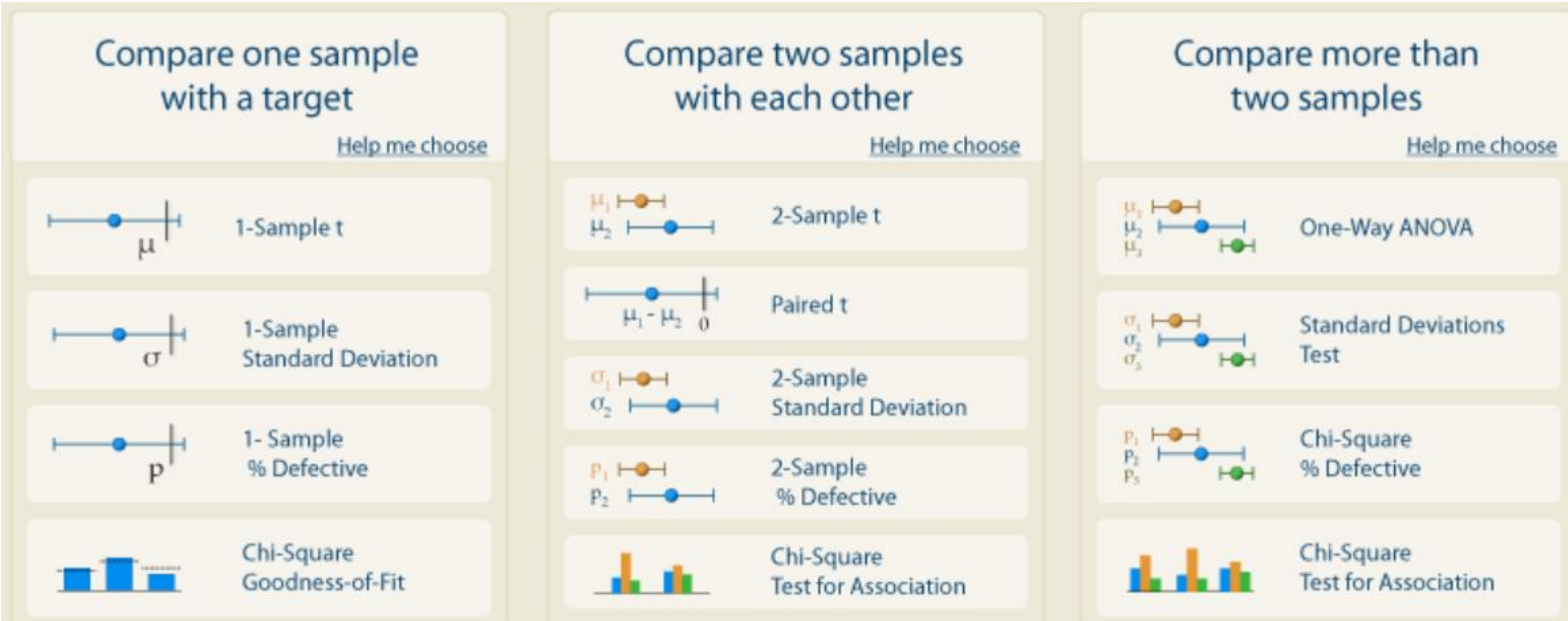
Nenhuma Evidência
(Nenhuma Significância)



p=.0069

Evidência para se Rejeitar H_0

Testes comuns no Minitab



Existem inúmeros Testes de Hipóteses na literatura nos mais diversos campos de estudo.

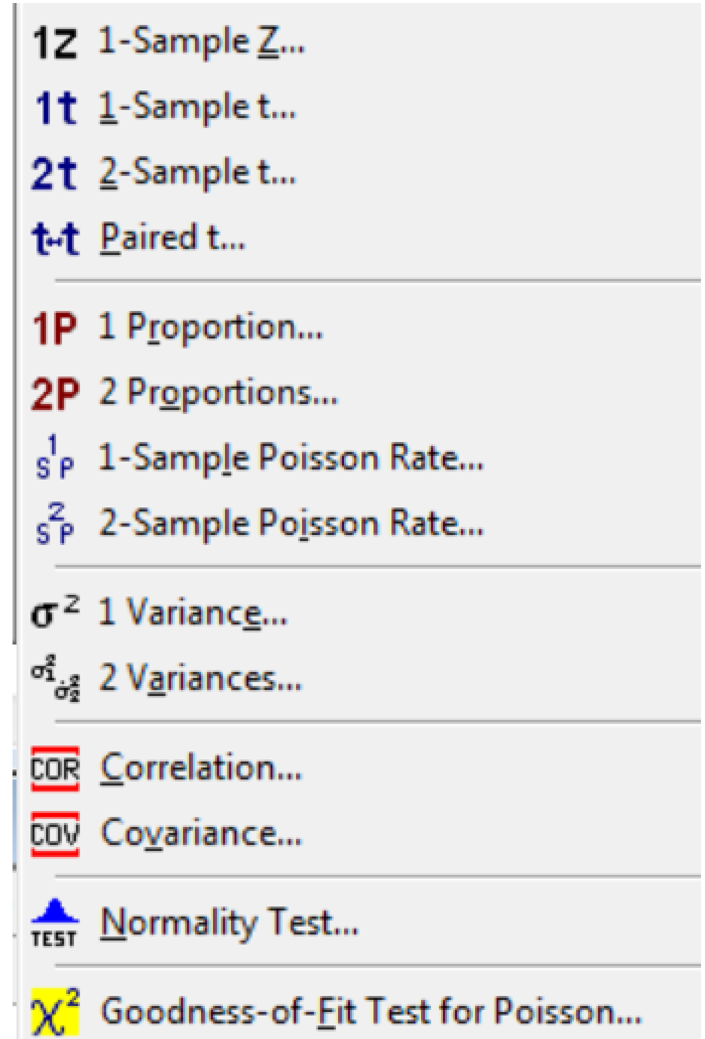
Veja a associação com vários testes

• Duas linhas de produção supostamente idênticas estão apresentando resultados diferentes. Como confirmar isso?

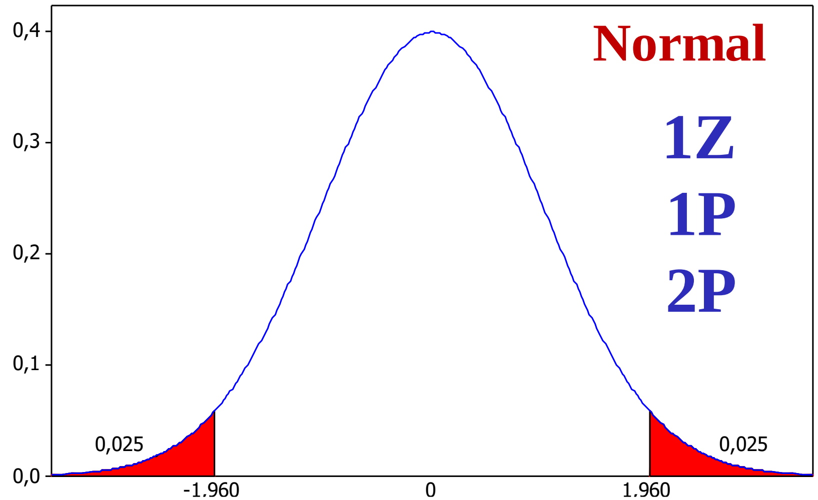
• Como saber estatisticamente se dois funcionários tem o mesmo desempenho?

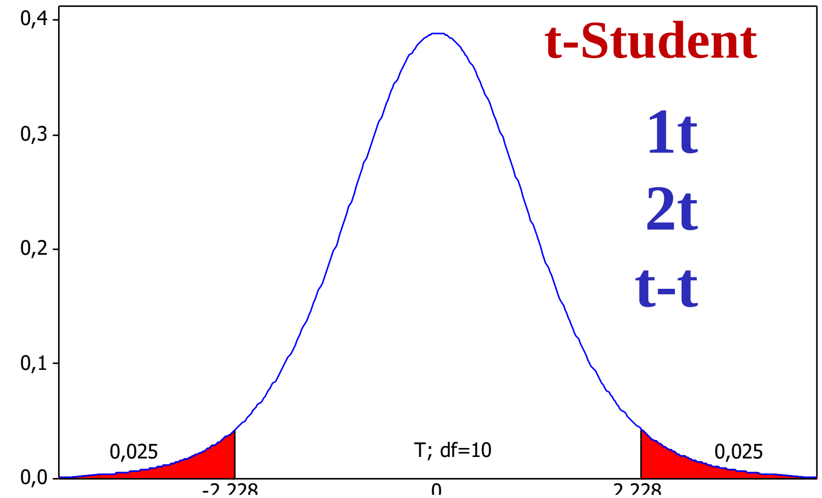
•A variabilidade de um processo é maior que outro. Temos certeza?

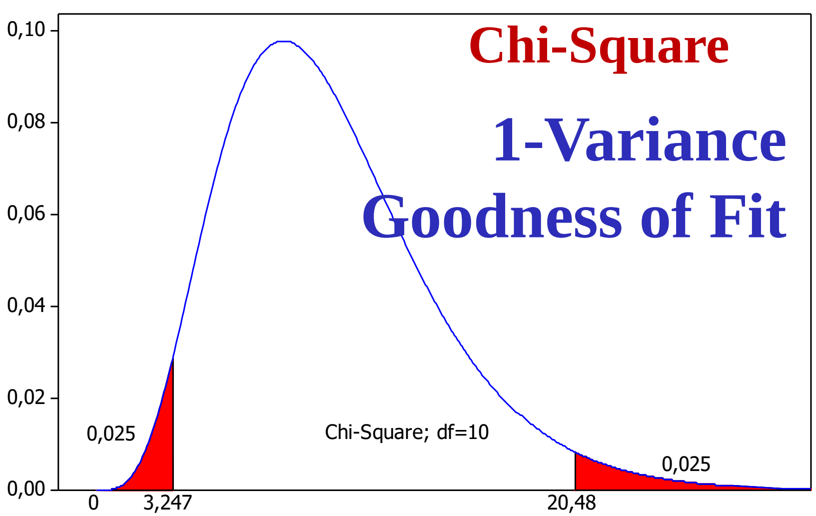
• Os dados estão normalmente distribuídos?

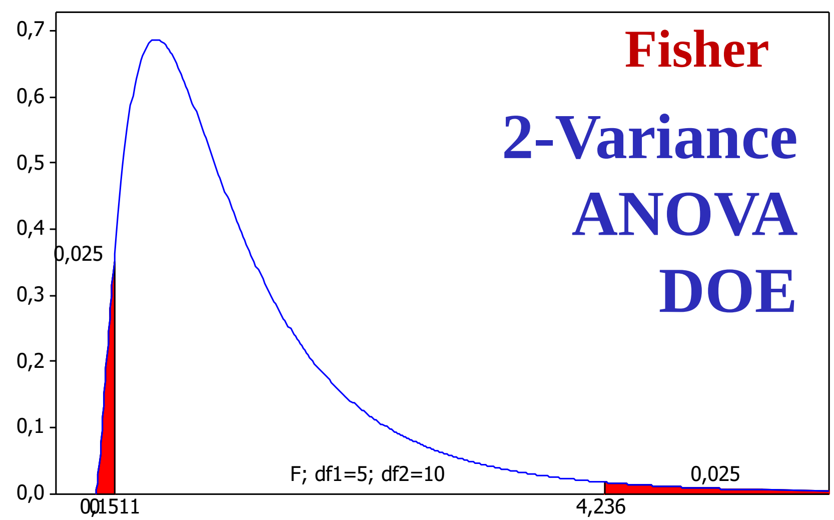


Distribuições de Probabilidade Fundamentais em Teste de Hipóteses









<1-Sample t>: exemplo 1/4

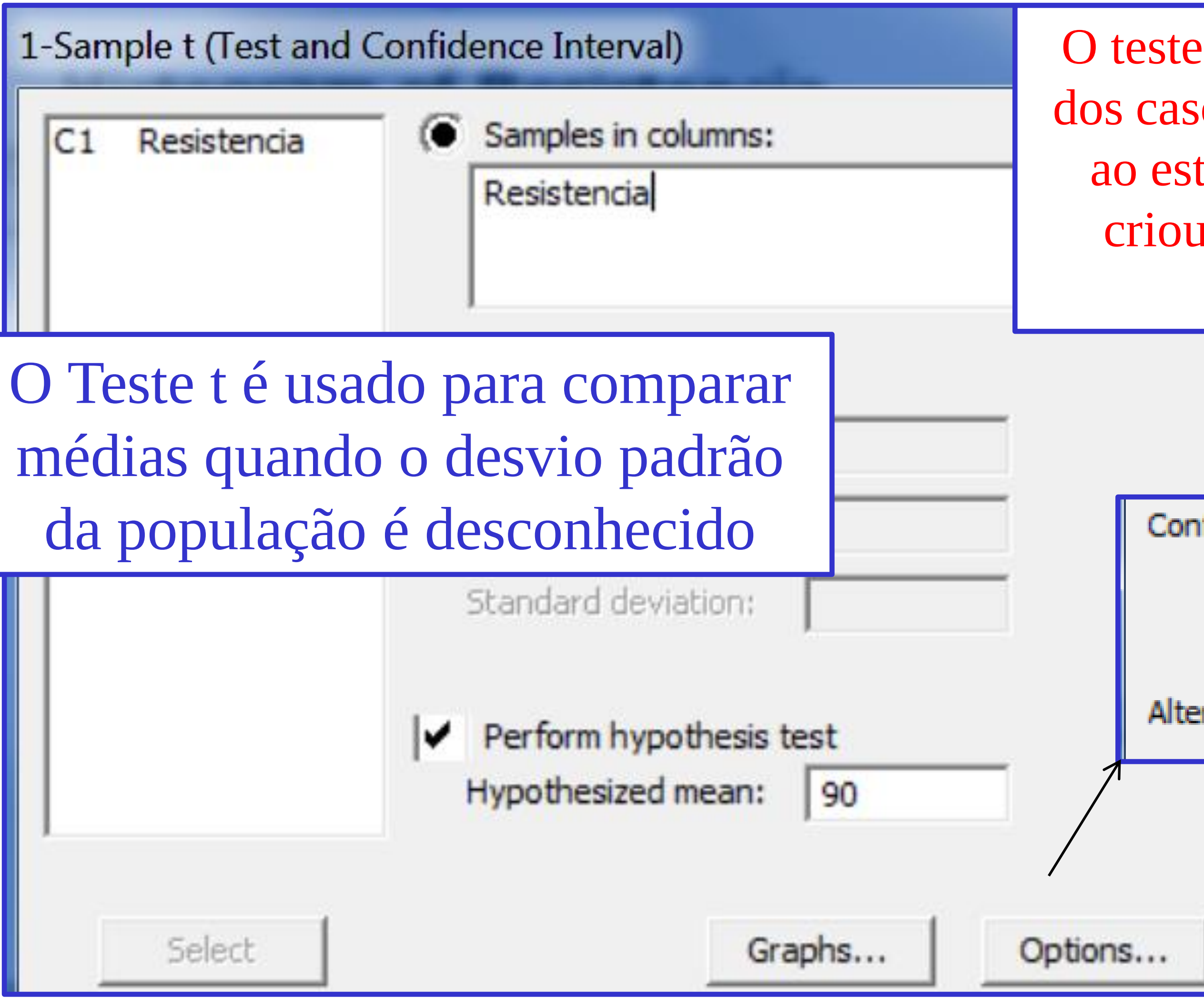
A Resistência média ao Estufamento das latas para a inspeção final deve ser **maior que 90 psi**. Tal resistência obedece a uma distribuição normal.

As medidas da Resistência para uma determinada linha/turno estão dadas na planilha Resistência.MTW

Teste a Hipótese de que as medidas da Resistência média ao Estufamento estão dentro do limite de especificação. (**Prove que as medidas são maiores que 90**)

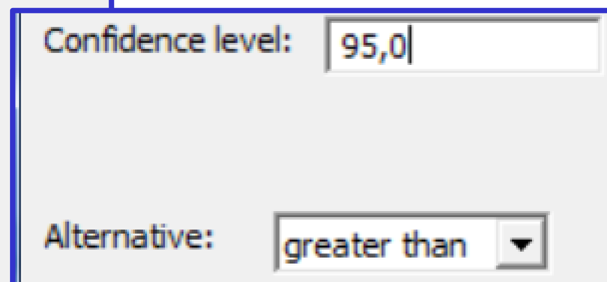
↓	C1
	Resistencia
1	91.17
2	91.02
3	90.61
4	89.45
5	90.96
6	90.22
7	93.03
8	91.29
9	91.17
10	90.89
11	91.82
12	91.96
13	91.14
14	91.63
15	90.31

<1-Sample t>: exemplo 2/4



O teste t é usado na maioria dos casos. O termo t deve-se ao estatístico **Gosset** que criou a distribuição t de Student.

O Teste t é usado para comparar médias quando o desvio padrão da população é desconhecido



H_1 é geralmente o que se deseja provar

<1-Sample t>: exemplo 3/4

Boa regra:

P_value< 0,05, rejeita-se H_o

One-Sample T: Resistencia

H_0 H_1

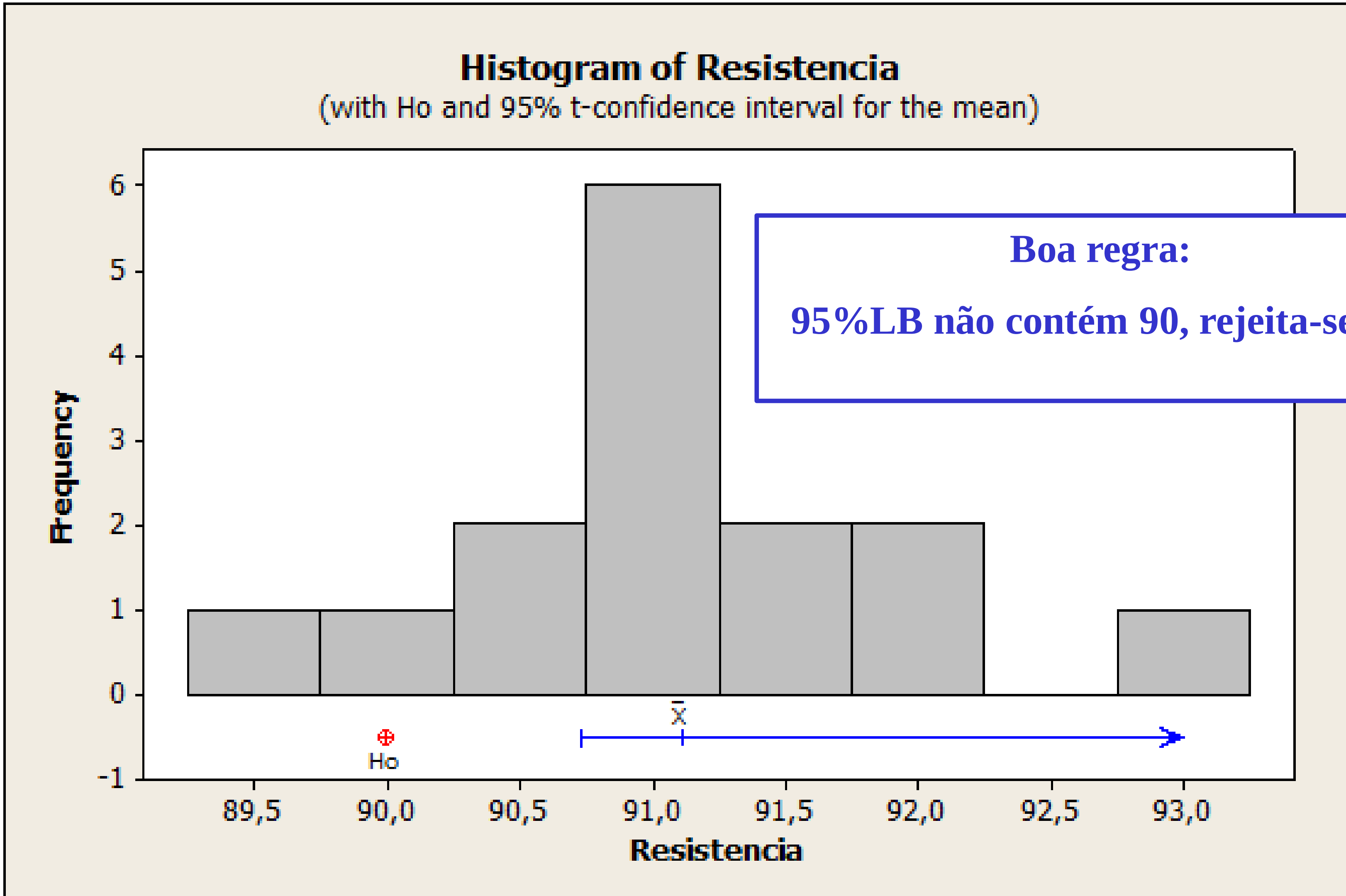
Test of mu = 90 vs > 90

Rejeita-se H_0

Variable	N	Mean	StDev	SE Mean	95% Lower Bound	T	P
Resistencia	15	91,111	0,834	0,215	90,731	5,16	0,000

$$95\%LB = \bar{x} - t_{\alpha,n-1}s/\sqrt{n}$$

<1-Sample t>: exemplo 4/4



Boa regra:

95%LB não contém 90, rejeita-se H_o

<2 Variances>: exemplo 1/4

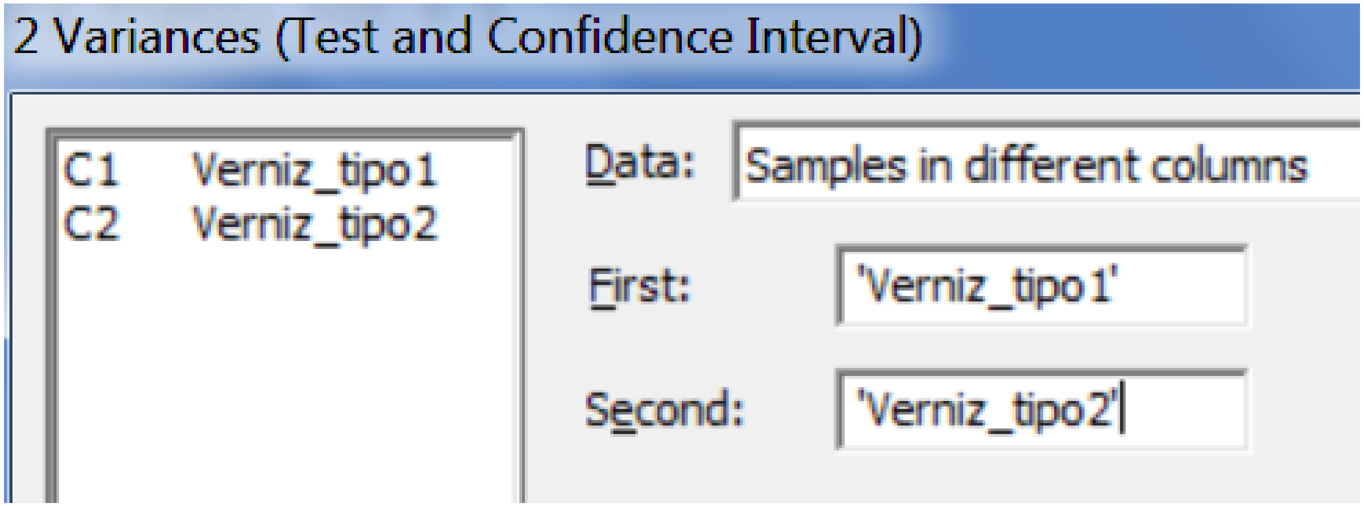
Dois tipos de Bico de Aplicação de verniz (Tipo I e Tipo II) foram avaliados. Deseja-se investigar o efeito desses dois Bicos com relação ao Peso do Verniz (em mg) medido após o processo. Tais medidas são dadas na planilha ao lado.

As variâncias são iguais? (Teste a Hipótese nula de que os dois bicos produzem um peso de Verniz com mesma variância, ou seja, mesma dispersão).

	Verniz_tipo1	Verniz_tipo2
1	111.540	112.501
2	110.615	111.621
3	110.396	112.311
4	110.702	111.897
5	111.396	112.542
6	111.586	112.537
7	110.577	111.860
8	110.863	112.160
9	110.207	112.104
10	110.043	112.519

peso_verniz.mtw

<2 Variances>: exemplo 2/4



Obs.: Teste o Procedimento
Stack Columns

Veja Options

<2 Variances>: exemplo 3/4

```
Null hypothesis          σ(Verniz_tipo1) / σ(Verniz_tipo2) = 1
Alternative hypothesis   σ(Verniz_tipo1) / σ(Verniz_tipo2) ≠ 1
Significance level       α = 0,05

Statistics
                                          95% CI for
Variable       N   StDev  Variance          StDevs
Verniz_tipo1  10   0,548     0,300   (0,392; 0,953)
Verniz_tipo2  10   0,331     0,110   (0,223; 0,611)

Ratio of standard deviations = 1,655
Ratio of variances = 2,738

95% Confidence Intervals
                                CI for
         CI for StDev          Variance
Method      Ratio               Ratio
Bonett  (0,889; 3,352)   (0,791; 11,238)
Levene  (0,665; 2,939)   (0,442;  8,637)

Tests
                      Test
Method  DF1  DF2  Statistic  P-Value
Bonett    1    –       2,76    0,097
Levene    1   18       1,51    0,236
```

Testes:

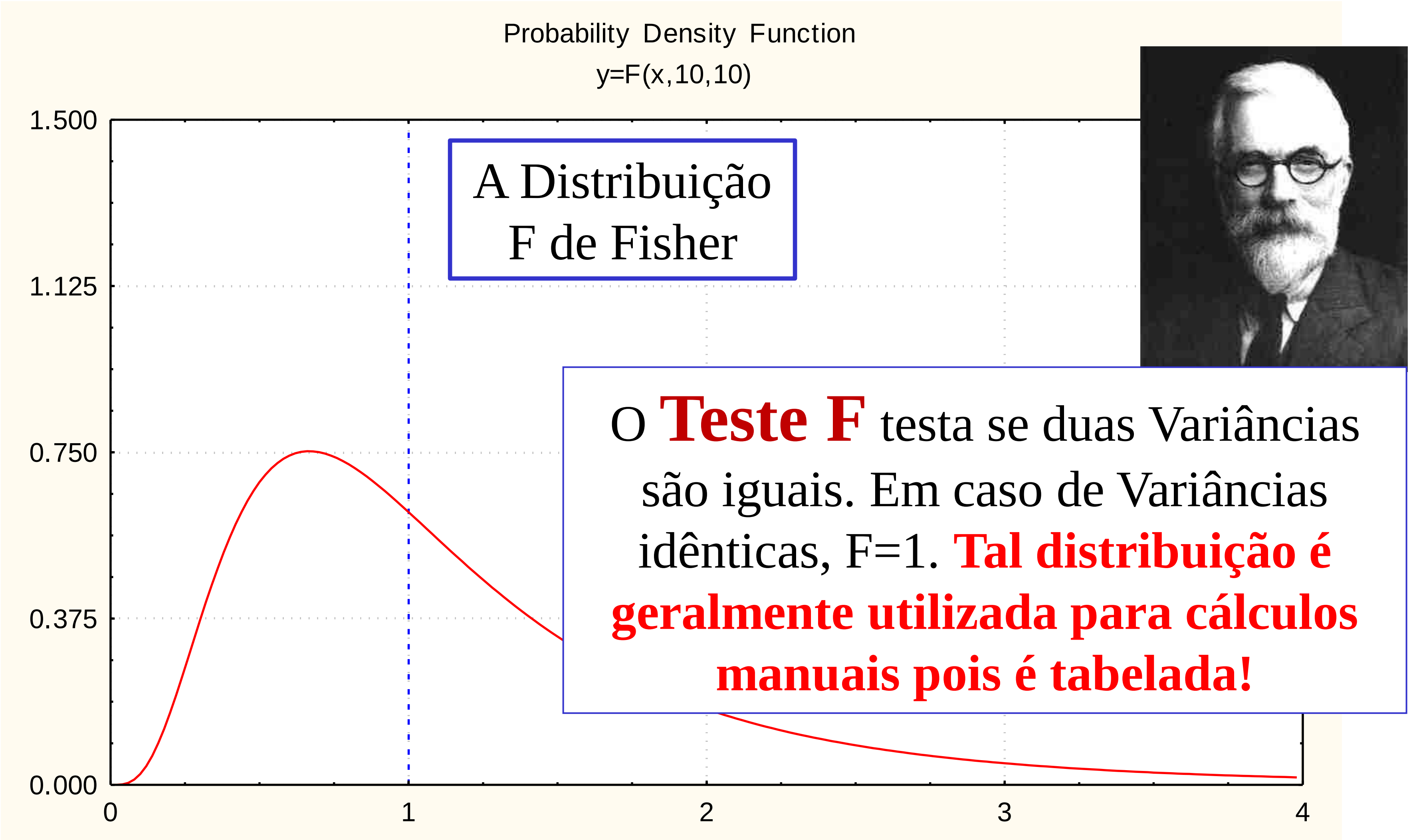
Bonett: Mais moderno e genérico

Levene: Amostra menor que 20 e valores contínuos

F: Somente para distribuições Normais

As variâncias são iguais!

<2 Variances>: exemplo 4/4

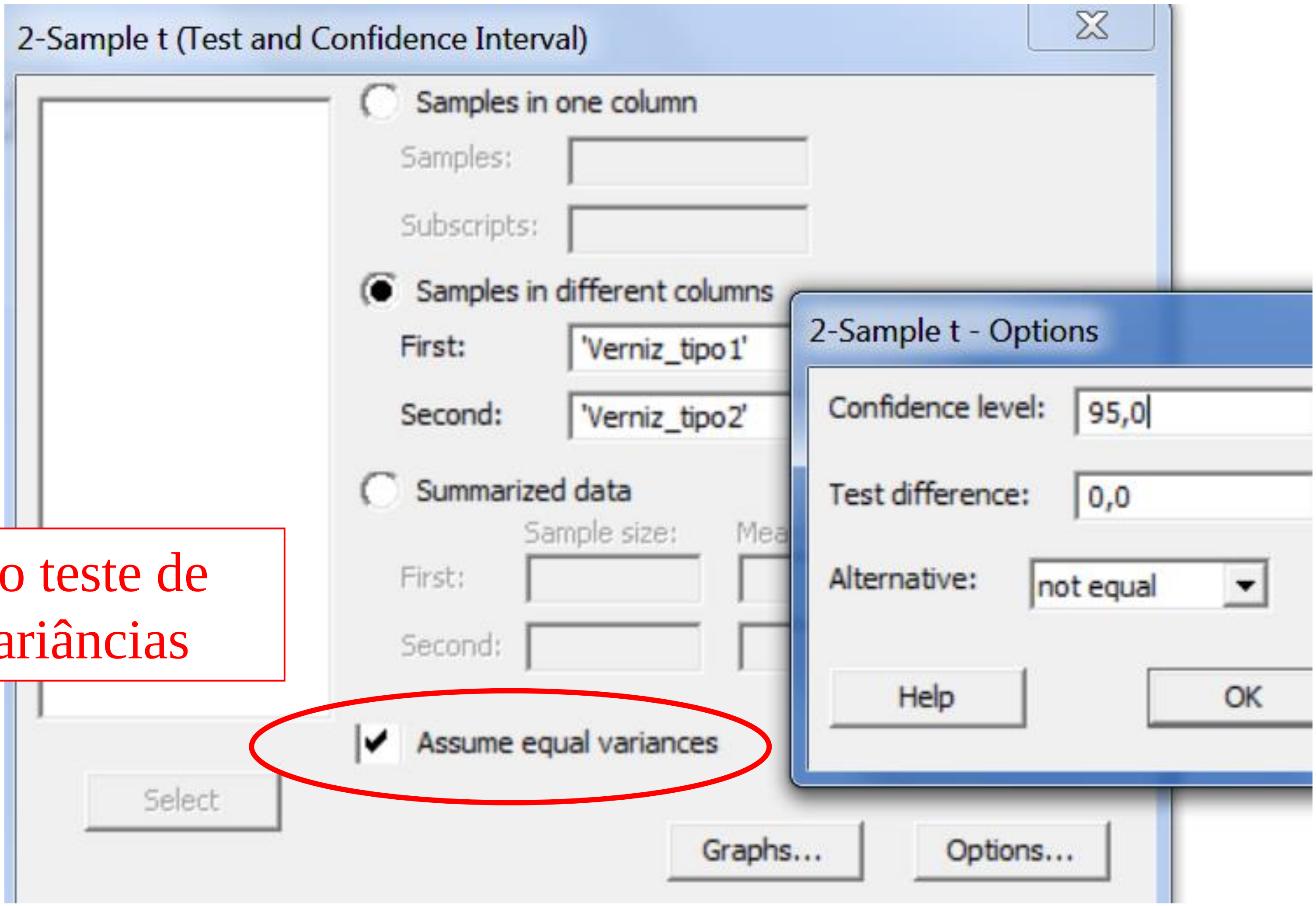


A Distribuição F de Fisher

O **Teste F** testa se duas Variâncias são iguais. Em caso de Variâncias idênticas, F=1. **Tal distribuição é geralmente utilizada para cálculos manuais pois é tabelada!**

<2 –Sample t>: exemplo 1/2

Em relação ao problema anterior, teste se as médias são diferentes.
(Peso_Verniz.MTW)



Do teste de Variâncias

<2 –Sample t>: exemplo 2/2

Two-sample T for Verniz_tipo1 vs Verniz_tipo2

	N	Mean	StDev	SE Mean
Verniz_tipo1	10	110,792	0,548	0,17
Verniz_tipo2	10	112,205	0,331	0,10

Difference = μ (Verniz_tipo1) - μ (Verniz_tipo2)
Estimate for difference: -1,413
95% CI for difference: (-1,838; -0,987)
T-Test of difference = 0 (vs $\neq$): T-Value = -6,97 P-Value = 0,000 DF = 18
Both use Pooled StDev = 0,4529

Médias diferentes