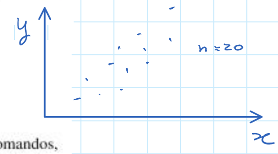
3.36. Um enólogo pretende avaliar a acidez total de um vinho. Para isso selecciona aleatoriamente 20 garrafas de vinho na adega e mede a acidez através do método clássico e de um dispositivo de titulação automática. Alguns resultados das análises, em g/l, foram:



	garrafa	1	2	3	⋯	18	19	20
(x)	método clássico	4.8	3.4	2.5	⋯	2.9	5.4	2.1
(y)	titulação automática	6.1	5.1	2.1	⋯	4.5	3.9	1.5

Os dados foram introduzidos no *software* R. Abaixo apresentam-se resultados de comandos, alguns inadequados. Responda às seguintes questões utilizando os resultados apresentados abaixo.

a) De acordo com a legislação em vigor um vinho de mesa deverá ter uma acidez total superior a 3.5 g/l. Com base nos resultados das análises efectuadas pelo método clássico, o enólogo poderá concluir que o seu vinho cumpre os requisitos de acidez impostos pela legislação? Explicite e valide os pressupostos necessários à resolução do problema.

b) Com base nos valores obtidos poder-se-á concluir que os dois métodos de análise da acidez total do vinho têm resultados significativamente diferentes? Explicite e valide os pressupostos necessários à resolução do problema.

X v.A. ACIDEZ MEDIDA PELO MÉTODO CLÁSSICO (g/l)

Y v.A. " " " " AUTOMÁTICO (g/l)

SEJAM $\mu_1 = E[X]$ E $\mu_2 = E[Y]$.

CLARAMENTE, AS AMOSTRAS SÃO EMPARELHADAS. POIS HÁ 20 GARRAFAS AO TODO, E CADA GARRAFA É SUJEITA AOS 2 MÉTODOS.

a) PRETENDE-SE DISCUTIR SE $E[X] > 3.5\, g/l$.

1º) DADO QUE A AMOSTRA É PEQUENA $(n=20)$, É NECESSÁRIO PODER ADMITIR QUE A DISTRIBUIÇÃO DE X É NORMAL, ISTO É, $X \cap N(\mu_1, \sigma_1)$.

ANEXO:

```
> classico
 [1]  4.8 3.4 2.5 3.8 4.3 3.6 3.5 3.5 4.0 3.6 6.3 3.0 3.1 3.7 2.8
 [16] 5.1 4.0 2.9 5.4 2.1
> automatico
 [1]  6.1 5.1 2.1 4.9 6.6 4.0 4.5 1.0 4.7 3.5 8.2 3.9 3.6 4.5 3.5
 [16] 6.3 4.7 4.5 3.9 1.5

> var(classico)        > var(automatico)        > var(classico-automatico)
[1] 1.040105            [1] 2.887868              [1] 1.326605

> shapiro.test(classico)                  > shapiro.test(automatico)
  Shapiro-Wilk normality test               Shapiro-Wilk normality test
data:  classico                           data:  automatico
W = 0.951, p-value = 0.3827               W = 0.9625, p-value = 0.5959
```

TESTE DE NORMALIDADE PARA X.

FORMULÁRIO:

regra de decisão:
– rejeitar H_0 se *p-value* $\leq \alpha$

O Teste de Shapiro Wilk
H_0 : X tem distribuição normal
H_1: X não tem distribuição normal

DADO QUE NO TESTE A $H_0: X \cap N(\mu_1, \sigma_1)$, O p-value $p = 0.3827 > \alpha$, CONSIDERANDO UM NÍVEL DE SIGNIFICÂNCIA $\alpha = 0.05$, ENTÃO ADMITE-SE H_0.

2º) VAI ENTÃO REALIZAR-SE O TESTE

μ	σ desconhecido $\boxed{\text{distribuição normal}}$	H_0: $\mu = \mu_0$ / $\mu \le \mu_0$ / $\mu \ge \mu_0$	H_1: $\mu \neq \mu_0$ / $\boxed{\mu > \mu_0}$ / $\mu < \mu_0$	$T = \dfrac{\overline{X} - \mu_0}{S_x/\sqrt{n}}$ c/ $(n-1)$g.l.	$T < -t_{\alpha/2}$ ou $T > t_{\alpha/2}$ / $\boxed{T > t_\alpha}$ / $T < -t_\alpha$

VERIFICADO EM 1º)

CRITÉRIO IMPOSTO PELA LEGISLAÇÃO

ASSIM, O TESTE SERÁ

1. $\alpha = 0.05$; $H_0: \mu \leq 3.5$; $H_1: \mu > 3.5$

2. $T = \dfrac{\overline{X} - \mu_0}{S_x/\sqrt{n}}$ SOB H_0, $T \cap t_{(n-1)}$

3. REGIÃO CRÍTICA $T > t_\alpha$

$t_{\alpha;(n-1)} = t_{0.05;(19)} = 1.7291$

$\uparrow$ $n=20$ TABELA

4. $T_{CALC} = 1.184$

```
> t.test(classico,mu=3.5,alternative="greater")
        One Sample t-test
data:  classico
t = 1.184, df = 19, p-value = 0.1255
alternative hypothesis: true mean is greater than 3.5
95 percent confidence interval:
 3.375677      Inf
sample estimates:
mean of x
     3.77
```

NOTA: $T_{CALC} = \dfrac{3.77 - 3.5}{\sqrt{1.04}/\sqrt{20}}$

5. DADO QUE T_{CALC} NÃO PERTENCE À REGIÃO CRÍTICA, NÃO SE REJEITA H_0 E ADMITE-SE COM $\alpha = 0.05$ QUE A ACIDEZ MÉDIA PODERÁ SER INFERIOR A 3.5 g/l.

b) Com base nos valores obtidos poder-se-á concluir que os dois métodos de análise da acidez total do vinho têm resultados significativamente diferentes? Explicite e valide os pressupostos necessários à resolução do problema.

VAMOS VER COMO RESPONDER À QUESTÃO DE 3 FORMAS ALTERNATIVAS DIFERENTES.

ALTERNATIVA 1 : USANDO UM INTERVALO DE CONFIANÇA PARA $\mu_1 - \mu_2$:

μ_D diferença de médias	amostras emparelhadas; σ_D desconhec. e pop. das diferenças normal	$\overline{d} - t_{\alpha/2;\,(n-1)}\, s_D/\sqrt{n} < \mu_D < \overline{d} + t_{\alpha/2;\,(n-1)}\, s_D/\sqrt{n}$

EM QUE D É A V.A. DEFINIDA POR $D = X - Y$

PRESSUPOSTOS PARA A CONSTRUÇÃO DO I.C. PARA $\mu_D = \mu_1 - \mu_2$:

i) AMOSTRAS EMPARELHADAS ✓

ii) POP. DAS DIFERENÇAS NORMAL:

FORMULÁRIO: regra de decisão:

ii) POP. DAS DIFERENÇAS NORMAL:

```
> shapiro.test(classico-automatico)
        Shapiro-Wilk normality test
data:  classico - automatico
W = 0.9163, p-value = 0.08413
```

FORMULÁRIO:

regra de decisão:
– rejeitar H_0 se *p-value* $\leq \alpha$

O Teste de Shapiro Wilk
H_0 : X tem distribuição normal
H_1: X não tem distribuição normal

COMO p-VALUE = 0.08413 > α = 0.05, ENTÃO NÃO SE REJEITA H_0 E ADMITE-SE QUE $D \cap N(\mu_D, \sigma_D)$.

INTERVALO A 95% DE CONFIANÇA PARA $\mu_D = \mu_1 - \mu_2$

O OUTPUT DE R DÁ O INTERVALO DE CONFIANÇA A (POR OMISSÃO) 95% DE CONFIANÇA:

I.C. PARA $\mu_1 - \mu_2$

AMOSTRAS EMPARELHADAS

```
> t.test(classico,automatico,paired=TRUE)
        Paired t-test
data:  classico and automatico
t = -2.2714, df = 19, p-value = 0.03493
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.12405128 -0.04594872
sample estimates:
mean of the differences
                 -0.585
```

$= \bar{x} - \bar{y}$

ASSIM, DADO QUE ZERO NÃO PERTENCE AO INTERVALO, NÃO SE ADMITE, COM 95% DE CONFIANÇA, QUE $\mu_1 - \mu_2 = 0$, OU SEJA QUE $\mu_1 = \mu_2$.

PORTANTO, CONCLUI-SE QUE OS DOIS MÉTODOS DE ANÁLISE DA ACIDEZ DO VINHO TEM RESUTADOS SIGNIFICATIVAMENTE DIFERENTES.

ALTERNATIVA 2: TESTE À HIPOTESE $H_0: \mu_1 = \mu_2$ (5 PASSOS)

COMO AS AMOSTRAS SÃO EMPARELHADAS, O TESTE É:

μ_D diferença de duas médias	amostras emparelhadas; pop.das diferenças normal e σ_D desconhec.	$\boxed{\mu_D = 0}$ $\mu_D \leq 0$ $\mu_D \geq 0$	$\boxed{\mu_D \neq 0}$ $\mu_D > 0$ $\mu_D < 0$	$T = \frac{\overline{D}}{S_D/\sqrt{n}}$ c/ $(n-1)$ g.l.	$\boxed{T < -t_{\alpha/2} \text{ ou } T > t_{\alpha/2}}$ $T > t_\alpha$ $T < -t_\alpha$

EM QUE D É A V.A. DEFINIDA POR $D = X - Y$

ASSIM $\mu_D = \mu_1 - \mu_2$, E "$\mu_D = 0$" É O MESMO QUE "$\mu_1 = \mu_2$"

PRESSUPOSTOS DO TESTE:

i) AMOSTRAS EMPARELHADAS ✓

ii) POP. DAS DIFERENÇAS NORMAL:

```
> shapiro.test(classico-automatico)
        Shapiro-Wilk normality test
data:  classico - automatico
W = 0.9163, p-value = 0.08413
```

FORMULÁRIO: **regra de decisão:**
– rejeitar H_0 se *p-value* $\leq \alpha$

O Teste de Shapiro Wilk
H_0 : X tem distribuição normal
H_1: X não tem distribuição normal

COMO p-VALUE $= 0.08413 > \alpha = 0.05$, ENTÃO NÃO SE REJEITA H_0 E ADMITE-SE QUE $D \frown N(\mu_D, \sigma_D)$.

TESTE

1. $\alpha = 0.05$, $\underbrace{H_0 : \mu_1 = \mu_2}_{\text{ou } H_0 : \mu_D = 0}$; $\underbrace{H_1 : \mu_1 \neq \mu_2}_{\text{ou } H_1 : \mu_D \neq 0}$

2. ESTATÍSTICA DE TESTE: $T = \dfrac{\overline{D}}{S_D/\sqrt{n}}$

SOB H_0, $T \frown t_{(n-1)}$

3. REGIÃO CRÍTICA BILATERAL. $t_{\alpha/2;(n-1)} = t_{0.025;(19)} = 2.093$

TABELA t

ASSIM, A REGIÃO CRÍTICA É:

$]-\infty, -2.093[\cup]2.093, +\infty[$

4. $T_{CALC} = -2.2714$

ANEXO (VER ACIMA):

```
> t.test(classico,automatico,paired=TRUE)
        Paired t-test
data:  classico and automatico
t = -2.2714, df = 19, p-value = 0.03493
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.12405128 -0.04594872
sample estimates:
mean of the differences
                 -0.585
```

AMOSTRAS EMPARELHADAS

5. DADO QUE T_{calc} ESTÁ NA REGIÃO CRÍTICA, REJEITA-SE H_0 COM NÍVEL DE SIGNIFICÂNCIA $\alpha = 0.05$, E PORTANTO CONCLUI-SE QUE OS DOIS MÉTODOS DE ANÁLISE DA ACIDEZ DO VINHO TEM RESUTADOS SIGNIFICATIVAMENTE DIFERENTES.

ALTERNATIVA 3: COMO ACIMA (TESTE À HIPÓTESE "$H_0 : \mu_1 = \mu_2$") MAS USANDO O "p-VALUE" EM VEZ DE

FAZER TODOS OS PASSOS:

COMO AS AMOSTRAS SÃO EMPARELHADAS, O TESTE É:

μ_D diferença de duas médias	amostras emparelhadas; pop.das diferenças normal e σ_D desconhec.	$\boxed{\mu_D = 0}$	$\boxed{\mu_D \neq 0}$	$T = \frac{\overline{D}}{S_D/\sqrt{n}}$	$\boxed{T < -t_{\alpha/2} \text{ ou } T > t_{\alpha/2}}$
		$\mu_D \leq 0$	$\mu_D > 0$	c/ $(n-1)$ g.l.	$T > t_\alpha$
		$\mu_D \geq 0$	$\mu_D < 0$		$T < -t_\alpha$

EM QUE D É A V.A. DEFINIDA POR $D = X - Y$

ASSIM $\mu_D = \mu_1 - \mu_2$, E "$\mu_D = 0$" É O MESMO QUE "$\mu_1 = \mu_2$"

PRESSUPOSTOS DO TESTE:

i) AMOSTRAS EMPARELHADAS ✓

ii) POP. DAS DIFERENÇAS NORMAL:

```
> shapiro.test(classico-automatico)
        Shapiro-Wilk normality test
data:  classico - automatico
W = 0.9163, p-value = 0.08413
```

FORMULÁRIO:

regra de decisão:
– rejeitar H_0 se *p-value* $\leq \alpha$

O Teste de Shapiro Wilk
H_0 : X tem distribuição normal
H_1: X não tem distribuição normal

COMO p-VALUE = 0.08413 > α = 0.05, ENTÃO NÃO SE REJEITA H_0 E ADMITE-SE QUE $D \frown N(\mu_D, \sigma_D)$.

TESTE

1. $\alpha = 0.05$, $\underbrace{H_0: \mu_1 = \mu_2}_{\text{ou } H_0:\ \mu_D = 0}$; $\underbrace{H_1: \mu_1 \neq \mu_2}_{\text{ou } H_1:\ \mu_D \neq 0}$

2. ESTATÍSTICA DE TESTE: $T = \frac{\overline{D}}{S_D/\sqrt{n}}$

SOB H_0, $T \frown t_{(n-1)}$

3-5. ANEXO (VER ACIMA):

```
> t.test(classico,automatico,paired=TRUE)
        Paired t-test
data:  classico and automatico
t = -2.2714, df = 19, p-value = 0.03493
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.12405128 -0.04594872
sample estimates:
mean of the differences
               -0.585
```

COMO O VALOR DE PROVA DO TESTE É $0.03493 < \alpha = 0.05$, REJEITA-SE H_0 COM NÍVEL DE SIGNIFICÂNCIA $\alpha = 0.05$, E PORTANTO CONCLUI-SE QUE OS DOIS MÉTODOS DE ANÁLISE DA ACIDEZ DO VINHO TEM RESUTADOS SIGNIFICATIVAMENTE

DIFERENTES.