Ex. 3.41

5 de janeiro de 2021 17:29

3.41. (*Exame 25.01.2019*) Realizou-se um estudo de investigação em modelos animais (ratinhos), onde se pretende conhecer a importância de algumas proteínas do sistema imune na infecção com malária na gravidez. Quantificou-se a expressão relativa dos genes associados a uma proteína, A, em placentas de 6 fêmeas não infectadas (NI) e 6 fêmeas infectadas (INF), tendo-se obtido os seguintes indicadores:

	Média	Variância
NI	1.520	0.612
INF	0.347	0.131

O cientista suspeita que a expressão relativa dos genes da proteína A pode sofrer alterações, em média, com a infecção.

Os resultados da experiência foram analisados com recurso ao *software* R e estão apresentados no Anexo Utilize-os, sempre que possível, para responder de forma completa às questões apresentadas a seguir.

a) Utilizando um intervalo de confiança adequado (consultar o Anexo) esclareça a dúvida do cientista. Responda de forma completa à questão, justificando convenientemente a sua resposta.

b) Um parâmetro de interesse neste estudo é a variabilidade da expressão relativa dos genes quando há infecção.

i) Indique o estimador associado ao estudo desse parâmetro?

ii) Com base nos dados recolhidos poder-se-á admitir que o valor daquele parâmetro é inferior a 0.3? Justifique convenientemente.

X v.a EXPRESSÃO PARA INFECTADO (INF)

Y " " " NÃO INFECTADO (NI)

$\bar{x} = 1.520$ $s_x^2 = 0.612$

$\bar{y} = 0.347$ $s_y^2 = 0.131$

a) $\mu_1 = E[X]$; $\mu_2 = E[Y]$

O CIENTISTA QUESTIONA SE $\mu_1 = \mu_2$?

2 ABORDAGENS ALTERNATIVAS
- I.C. PARA $\mu_1 - \mu_2$
- TESTE A $H_0: \mu_1 = \mu_2$

$\mu_1 - \mu_2$	amostras independentes, pop. normais e σ_1 e σ_2 desconhec. mas supostos iguais	$\overline{x}_1 - \overline{x}_2 - t_{\alpha/2;\,(n_1+n_2-2)}\, s_p \sqrt{(1/n_1)+(1/n_2)} < \mu_1 - \mu_2 < \overline{x}_1 - \overline{x}_2 + t_{\alpha/2;\,(n_1+n_2-2)}\, s_p \sqrt{(1/n_1)+(1/n_2)}$; $s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$

		H_0	H_1		
$\mu_1 - \mu_2$	amostras independentes,	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$T = \frac{\overline{X}_1 - \overline{X}_2}{S_p \sqrt{(1/n_1)+(1/n_2)}}$	$T < -t_{\alpha/2}$ ou $T > t_{\alpha/2}$
	pop. normais e σ_1 e σ_2 desconh.	$\mu_1 \leq \mu_2$	$\mu_1 > \mu_2$	$S_p = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$	$T > t_\alpha$
	mas supostos iguais	$\mu_1 \geq \mu_2$	$\mu_1 < \mu_2$	c/ (n_1+n_2-2) g.l.	$T < -t_\alpha$
		1.	1.	2./4.	3.

PRESSUPOSTOS:

i) POP. NORMAIS, I.E, $X \cap N(\mu_1, \sigma_1)$ E $Y \cap N(\mu_2, \sigma_2)$

DADO QUE NO OUTPUT DO TESTE DE SHAPIRO-WILK PARA X E PARA Y, P-VALUE $> \alpha = 0.05$ ENTÃO ADMITE-SE QUE

$X \cap N(\mu_1, \sigma_1)$ E $Y \cap N(\mu_2, \sigma_2)$

ii) $\sigma_1^2 = \sigma_2^2$

		H_0	H_1		
$\frac{\sigma_1^2}{\sigma_2^2}$	distribuições	$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$		$F < f_{1-\alpha/2}$ ou $F > f_{\alpha/2}$
	normais	$\sigma_1^2 \leq \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$F = \frac{S_1^2}{S_2^2}$	$F > f_\alpha$
	independ.	$\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$	c/ (n_1-1, n_2-1) g.l.	$F < f_{1-\alpha}$

TESTAR $H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$

DADO QUE P-VALUE DO TESTE É $0.1159 > \alpha = 0.05$ ENTÃO ADMITE-SE QUE $\sigma_1^2 = \sigma_2^2$

ANEXO

```
> var.test(A_NI, A_INF)
        F test to compare two variances
data:  A_NI and A_INF
F = 4.6731, num df = 5, denom df = 5, p-value = 0.1159
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.6539058 33.3954743
sample estimates:
ratio of variances
         4.67306
```

ii) > 0.05

I.C. PARA σ_1^2 / σ_2^2

```
> shapiro.test(A_NI)
 Shapiro-Wilk normality test
```

```
> shapiro.test(A_INF)
 Shapiro-Wilk normality test
```

```
> shapiro.test(A_NI)
 Shapiro-Wilk normality test
data:  A_NI
W = 0.92543, p-value = 0.5453
```

```
> shapiro.test(A_INF)
  Shapiro-Wilk normality test
data:  A_INF
W = 0.79852, p-value = 0.05699
```

```
> shapiro.test(placenta$A_NI-placenta$A_INF)

        Shapiro-Wilk normality test

data:  placenta$A_NI - placenta$A_INF
W = 0.9378, p-value = 0.6414
```

AMOSTRAS INDEPENDENTES ✓

FALSE

```
> t.test(A_NI,A_INF, paired=F, var.equal=T,alternative="two.sided")
        Two Sample t-test
data:  A_NI and A_INF
t = 3.334, df = 10, p-value = 0.007566
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
0.3891588   ????
sample estimates:
mean of x     mean of y
1.5199600     0.3467328
```

$H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

T_{CALC}

I.C. A 95% PARA $\mu_1 - \mu_2$

TRUE

```
> t.test(A_NI, A_INF, paired=T, var.equal=T,alternative="two.sided")
        Paired t-test
data:  A_NI and A_INF        t = 2.66, df = 5, p-value = 0.04488
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
0.03945195    ?????
sample estimates:
mean of the differences
         1.173227
```

X AMOSTRAS EMPARELHADAS

Passos a seguir na construção de um **teste estatístico**:

* 1. Identificar o(s) parâmetro(s); especificar $\mathbf{H}_0$ **e** $\mathbf{H}_1$ e o nível de significância α.

* 2. Escolher uma variável aleatória – estatística de teste, que sob $\mathbf{H}_0$ terá distribuição conhecida (pelo menos aproximadamente).

* 3. Definir a **região de rejeição** ou **região crítica – RC** (conjunto de valores da estatística que são menos "plausíveis" caso $\mathbf{H}_0$ seja verdadeira, portanto levam a rejeitar $\mathbf{H}_0$).

4. Calcular **o valor** da estatística de teste, para a amostra observada.

5. **Se o valor calculado** $\in$ **RC** $\longrightarrow$ **rejeita-se** $\mathbf{H}_0$
Se o valor calculado $\notin$ **RC** $\longrightarrow$ **não se rejeita** $\mathbf{H}_0$

1. $\alpha = 0.05$; $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

2. SOB H_0 ,
$$T = \frac{\overline{X}_1 - \overline{X}_2}{S_p \sqrt{(1/n_1) + (1/n_2)}}$$
$$S_p = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$$
c/ $(n_1 + n_2 - 2)$ g.l.

$\frown t_{(n_1 + n_2 - 2)}$, with $n_1 = 6$, $n_2 = 6$

3. REGIÃO CRÍTICA:

$T < -t_{\alpha/2}$ ou $T > t_{\alpha/2}$

$t_{\alpha/2} = t_{0.05/2\,;(10)} = t_{0.025;(10)} = 2.228$

$]-\infty, -2.228[\cup]2.228, +\infty[$ TABELA t

	α							
n	0.4	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	0.32492	1.00000	3.07768	6.31375	12.70620	31.82052	63.65674	318.30884
2	0.28868	0.81650	1.88562	2.91999	4.30265	6.96456	9.92484	22.32712
3	0.27667	0.76489	1.63774	2.35336	3.18245	4.54070	5.84091	10.21453
4	0.27072	0.74070	1.53321	2.13185	2.77645	3.74695	4.60409	7.17318
5	0.26718	0.72669	1.47588	2.01505	2.57058	3.36493	4.03214	5.89343
6	0.26483	0.71756	1.43976	1.94318	2.44691	3.14267	3.70743	5.20763
7	0.26317	0.71114	1.41492	1.89458	2.36462	2.99795	3.49948	4.78529
8	0.26192	0.70639	1.39682	1.85955	2.30600	2.89646	3.35539	4.50079
9	0.26096	0.70272	1.38303	1.83311	2.26216	2.82144	3.24984	4.29681
10	0.26018	0.69981	1.37218	1.81246	2.22814	2.76377	3.16927	4.14370

4. $T_{CALC} = 3.334$ (OUTPUT DO R)

5. COMO $T_{CALC} = 3.334 \in$ REGIÃO CRÍTICA
ENTÃO REJEITA-SE H_0 E

CONCLUI-SE QUE $\mu_1 \neq \mu_2$ COM NÍVEL

DE SIGNIFICÂNCIA $\alpha = 0.05$.

$\underbrace{\text{OU } [3.-5.]}_{\text{USANDO p-VALUE}}$: COMO P-VALUE DO TESTE É 0.0075,
MUITO BAIXO $(0.0075 < \alpha = 0.05)$,
REJEITA-SE H_0 E CONCLUI-SE
QUE $\mu_1 \neq \mu_2$.

$\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$ e $\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \frown \mathcal{N}(0,1)$ ESTIMADOR DO $\mu = E[X]$

$S^2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{[illegible]}$ e $\frac{(n-1)S^2}{[illegible]} \frown \chi^2_{[illegible]}$ " DE $\sigma^2 = VAR[X]$

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \quad \text{e} \quad \frac{X-\mu}{\sigma/\sqrt{n}} \frown \mathcal{N}(0,1)$$

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} \quad \text{e} \quad \frac{(n-1)S^2}{\sigma^2} \frown \chi^2_{(n-1)}$$

ESTIMADOR DE $\mu = E[\;]$

" DE $\sigma^2 = VAR[X]$

b) i) O ESTIMADOR DA VARIÂNCIA É A V.A. $S^2 = \dfrac{\sum (x_i - \bar{x})^2}{n-1}$

ii) $\underline{s_1^2 = 0.131}$; $\sigma_1^2 = VAR[X]$; $\boxed{\sigma_0^2 = 0.3}$

		H_0	H_1		
σ^2	distribuição normal	$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ c/ (n-1) g.l.	$\chi^2 < \chi^2_{1-\alpha/2}$ ou $\chi^2 > \chi^2_{\alpha/2}$
		$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$		$\chi^2 > \chi^2_\alpha$
		$\boxed{\sigma^2 \geq \sigma_0^2}$	$\sigma^2 < \sigma_0^2$		$\chi^2 < \chi^2_{1-\alpha}$
		BENEFÍCIO DA DÚVIDA	1.	2./4.	3.

1. $\alpha = 0.05$; $H_0: \sigma_1^2 \geq 0.3$ vs $\underbrace{H_1: \sigma_1^2 < 0.3}_{}$ Hip. Alternativa

2. SOB H_0, $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} \frown \chi^2_{(n-1)}$

3. REG. CRÍTICA: $\chi^2 < \chi^2_{1-\alpha}$ $\qquad \chi^2_{1-\alpha} = \chi^2_{0.95;(5)} = 1.145$

$\chi^2 < 1.145$ $\qquad$ TABELA χ^2

	α														
n	0.999	0.995	0.99	0.975	0.95	0.9	0.75	0.5	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	0.000	0.000	0.000	0.001	0.004	0.016	0.102	0.455	1.323	2.706	3.841	5.024	6.635	7.879	10.828
2	0.002	0.010	0.020	0.051	0.103	0.211	0.575	1.386	2.773	4.605	5.991	7.378	9.210	10.597	13.816
3	0.024	0.072	0.115	0.216	0.352	0.584	1.213	2.366	4.108	6.251	7.815	9.348	11.345	12.838	16.266
4	0.091	0.207	0.297	0.484	0.711	1.064	1.923	3.357	5.385	7.779	9.488	11.143	13.277	14.860	18.467
5	0.210	0.412	0.554	0.831	$\boxed{1.145}$	1.610	2.675	4.351	6.626	9.236	11.070	12.833	15.086	16.750	20.515

4. $\chi^2_{CALC} = \dfrac{(6-1) \times 0.131}{0.3} = 2.183$

5*0.131/0.3=2.1833

5. Como $\chi^2_{CALC} = 2.183 \notin$ R.C.,

ENTÃO ADMITE-SE H_0, OU SEJA, ADMITE-SE QUE σ_1^2 POSSA SER SUPERIOR A 0.3