Universidade Federal de Lavras
Departamento de Estatística
Prof. Daniel Furtado Ferreira
$12^{\underline{a}}$ Teoria da Decisão Estatística

1) Os dados a seguir referem-se às mensurações da glicose arterial em mM em amostras independentes de animais (ruminantes) tratados e não tratados (controle) com o medicamento Phlorizin.

Quantidades	Controle (Não tratados)	Tratados (Phlorizin)
n_i	10	14
$\bar{X}_i$	3,21	3,11
S_i^2	0,85	0,80

Aplicar o teste da hipótese $H_0 : \mu_1 - \mu_2 = 0$, considerando o valor nominal de significância ($\alpha = 5\%$), para verificar se existem diferenças entre as médias da glicose arterial na população controle e tratada com Phlorizin. Considere as variâncias populacionais iguais. Tire as conclusões de interesse. Adaptado de Bauer et al. (1995).

Dados: $t_{0,025;\nu=22} = 2{,}074$.

2) Neste mesmo trabalho Bauer et al. (1995) estudando o efeito do phlorizin no fluxo do sangue arterial obtiveram os seguintes resultados em l/h.

Quantidades	Controle (Não tratados)	Tratados (Phlorizin)
n_i	10	14
$\bar{X}_i$	94	120
S_i^2	4	36

Aplicar o teste da hipótese $H_0 : \mu_1 - \mu_2 = 0$, considerando o valor nominal de significância ($\alpha = 5\%$), para verificar se existem diferenças entre as médias do fluxo de sangue arterial na população controle e tratada com Phlorizin. Considere as variâncias populacionais heterogêneas. Tire as conclusões de interesse.

Dados: $t_{0,025;\nu=22} = 2{,}074$ e $t_{0,025;\nu=17} = 2{,}110$.

3) Utilizando os dados de glicose, exercício 1, e fluxo sanguíneo, exercício 2, aplicar o teste para a hipótese $H_0 : \sigma_1^2 = \sigma_2^2$ em ambos os casos, sendo que o índice 1, representa o controle e o índice 2, o grupo tratado. Utilize um coeficiente de confiança de 95%.

Dados: $F_{0,025;\nu_1=9,\nu_2=13} = 3{,}312$ e $F_{0,025;\nu_1=13,\nu_2=9} = 3{,}831$.

Resolução

1) A hipótese de interesse é dada por:

$$H_0: \mu_1 - \mu_2 = 0 \qquad \text{vs} \qquad H_1: \mu_1 - \mu_2 \neq 0.$$

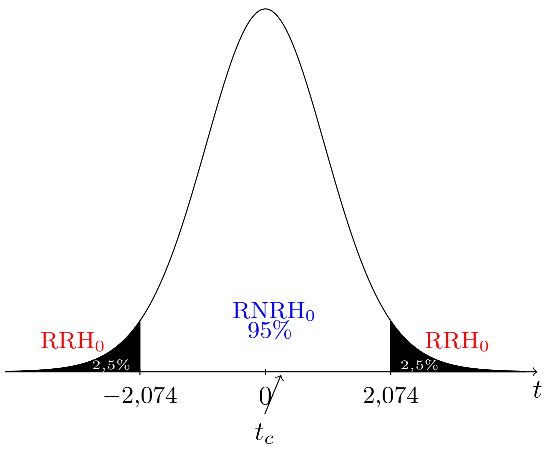
Como neste caso as variâncias são consideradas homogêneas, a variância comum S_p^2 é dada por:

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{9 \times 0{,}85 + 13 \times 0{,}80}{22} = 0{,}8205.$$

A estatística do teste é:

$$t_c = \frac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{S_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \frac{3{,}21 - 3{,}11 - 0}{\sqrt{0{,}8205\left(\dfrac{1}{10} + \dfrac{1}{14}\right)}} = 0{,}2666.$$

Os graus de liberdade para a distribuição da estatística, no caso de variâncias homogêneas, são $\nu = n_1 + n_2 - 2 = 22$. Assim, a região crítica (região de rejeição da hipótese nula), sabendo que $t_{0{,}025;\nu=22} = 2{,}074$, é dada por:



Como o valor de t_c pertence **a região de não rejeição da hipótese**, pelo teste t, com 95% de confiança, a hipótese nula **não deve ser** rejeitada, ou seja, concluímos que as médias de glicose das populações tratada e não tratada são as mesmas, o que indica que o medicamento **não** tem o efeito de alterar a média de glicose dos animais.

2) A hipótese de interesse, para o fluxo sanguíneo, é dada por:

$$H_0: \mu_1 - \mu_2 = 0 \qquad \text{vs} \qquad H_1: \mu_1 - \mu_2 \neq 0.$$

Como neste caso as variâncias são consideradas heterogêneas, e não devemos e nem podemos estimar uma variância comum S_p^2. A estatística do teste é:

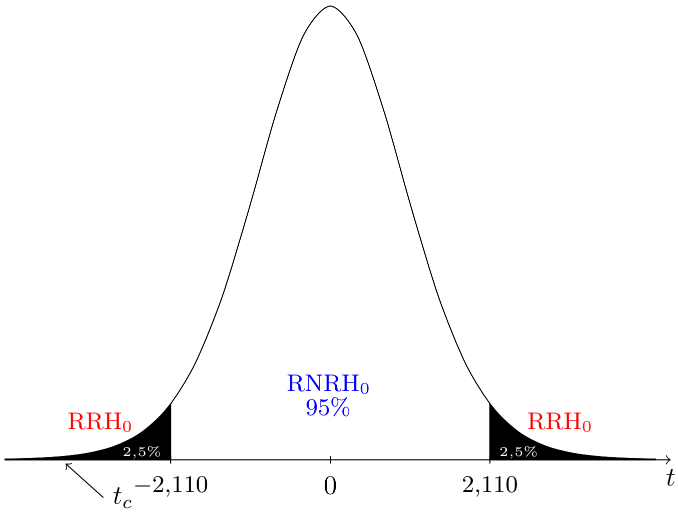
$$t_c = \frac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} = \frac{94 - 120 - 0}{\sqrt{\dfrac{4}{10} + \dfrac{36}{14}}} = -15{,}083.$$

Os graus de liberdade para a distribuição da estatística, no caso de variâncias heterogêneas, são

$$\nu \cong \frac{\left(\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1-1}+\dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2-1}}$$

$$\cong \frac{\left(\dfrac{4}{10}+\dfrac{36}{14}\right)^2}{\dfrac{\left(\dfrac{4}{10}\right)^2}{10-1}+\dfrac{\left(\dfrac{36}{14}\right)^2}{14-1}} = 16{,}77 \approx 17.$$

Assim, a região crítica (região de rejeição da hipótese nula), sabendo que $t_{0,025;\nu=17} = 2{,}110$, é dada por:



Como o valor de t_c pertence **a região de rejeição da hipótese**, pelo teste t, com 95% de confiança, a hipótese nula **deve ser rejeitada**, ou seja, concluímos que a média do fluxo sanguíneo da população tratada é maior do que a da população não tratada, o que indica que o medicamento tem o efeito de aumentar a média do fluxo sanguíneo dos animais.

Obs. Como sabemos que o efeito, detectado como significativo, foi de aumentar a média e não de diminuir? A resposta para isso é simples: a população não tratada foi a número 1 e a diferença amostral deu negativa, o que pode ser visto pelo valor negativo da estatística, e as médias são diferentes (hipótese nula rejeitada), então *juntando as peças*, concluímos (inferimos) que a média da população tratada é maior.

3) A seguir estão as resoluções para cada caso separadamente.

(a) A hipótese de interesse, para a glicose, é dada por:

$$H_0: \sigma_1^2 = \sigma_2^2 \qquad \text{vs} \qquad H_1: \sigma_1^2 \neq \sigma_2^2.$$

A estatística do teste é:

$$F_c = \frac{S^2_{\text{Maior}}}{S^2_{\text{Menor}}} = \frac{0{,}85}{0{,}80} = 1{,}0625.$$

Usamos o **artifício** de colocar sempre a maior variância amostral no numerador para não termos que preocupar com o valor tabelado $F_{1-\alpha/2}$, uma vez que para $\alpha < 0{,}5$, esse valor é inferior a unidade e o valor calculado é no mínimo igual a unidade. Esses valores não são tabulados nos livros, o que não é um problema, pois temos programas que os calculam diretamente. Mas esse artifício é conveniente para facilitar o processo. Apenas isso.

Os graus de liberdade para a distribuição da estatística são $\nu_1 = 9$ e $\nu_2 = 13$. O valor tabelado de F é portanto $F_{0,025;\nu_1=9,\nu_2=13} = 3{,}312$.

Como o valor de F_c pertence **a região de não rejeição da hipótese**, pois $F_c < F_{0,025;\nu_1=9,\nu_2=13}$, pelo teste F, com 95% de confiança, a hipótese nula **não deve ser rejeitada**, ou seja, concluímos que a variância da população tratada não difere da variância da variância da população não tratada para a glicose, ou em outras palavras, concluímos que as variâncias das duas populações são homogêneas.

(b) A hipótese de interesse, para o fluxo sanguíneo, é dada por:

$$H_0: \sigma_1^2 = \sigma_2^2 \qquad \text{vs} \qquad H_1: \sigma_1^2 \neq \sigma_2^2.$$

A estatística do teste é:

$$F_c = \frac{S^2_{\text{Maior}}}{S^2_{\text{Menor}}} = \frac{36}{4} = 9{,}0.$$

Os graus de liberdade para a distribuição da estatística são $\nu_1 = 13$ e $\nu_2 = 9$. O valor tabelado de F é portanto $F_{0{,}025;\nu_1=13,\nu_2=9} = 3{,}831$.

Como o valor de F_c pertence **a região de rejeição da hipótese**, pois $F_c > F_{0{,}025;\nu_1=13,\nu_2=9}$, pelo teste F, com 95% de confiança, a hipótese nula **deve ser rejeitada**, ou seja, concluímos que a variância da população tratada difere da variância da variância da população não tratada para o fluxo sanguíneo, ou em outras palavras, concluímos que as variâncias das duas populações são heterogêneas.

PS. Espero que todos façam uma boa prova e que não me encontrem de novo na sala de aula da GEX112 no próximo ano. Boa sorte para vocês todos. Lembrem-se de que a estrada do conhecimento é longa e sinuosa, mas o importante é estar nela.