ME - 731 Análise Multivariada
Segundo semestre de 2009
Lista de Exercícios II
Entrega: Todos os exercícios em 22/09/2009

Obs1: Para todas as questões referentes à análise de dados, considere o arquivo relativo aos dados da Iris, disponível na página do curso e no programa R sob o nome de *iris*. Você pode utilizar qualquer pacote estatístico e/ou linguagem de programação para resolver tais exercícios, desde que você a(s) mencione em seu trabalho.

Obs2: Não é necessário digitar a resolução da lista, os exercícios podem ser entregues feitos à mão.

Legenda do banco de dados da Iris.

Sepal.Length - comprimento da sépala, Sepal.Width - largura da sépala, Petal.Width - largura da pétala, Petal.Length - comprimento da pétala.

Exercícios

1. Considere $\boldsymbol{X}_j \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}), j = 1, ..., n$, uma amostra aleatória em que $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ são ambos desconhecidos. Obtenha a estatística do teste da razão de verossimilhanças (t.r.v.), e sua respectiva distribuição assintótica sob H_0, para testar $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ vs $H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$, em que $\boldsymbol{\Sigma}_{0(p \times p)}$ é uma matriz conhecida.

2. Utilize o resultado da Questão 1 para testar a referida hipótese considerando ($\alpha = 5\%$), nos seguintes casos:

   a) $\boldsymbol{\Sigma}_0 = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$

   b) $\boldsymbol{\Sigma}_0 = \begin{bmatrix} \sigma^2 & \sigma_{11} \\ \sigma_{11} & \sigma^2 \end{bmatrix}$

   Utilize as variáveis *Petal.Length* e *Petal.Width* e o grupo *iris setosa*.

   O que significa cada hipótese? Para cada item anterior, apresente a matriz de covariâncias estimada, o valor da estatística do teste, o nível descritivo e a respectiva conclusão.

3. Considere $\boldsymbol{X}_{ij} \sim N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), i = 1, 2. j = 1, ..., n_i$, amostras aleatórias de duas populações independentes, em que $(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), i = 1, 2$ são ambos desconhecidos. Obtenha a estatística do t.r.v., e sua respectiva distribuição assintótica sob H_0, para testar $H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ vs $H_1 : \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$.

4. Utilize o resultado da Questão 3 para testar a referida hipótese considerando ($\alpha = 5\%$), tendo como população 1 o grupo *iris setosa* e a população 2 o grupo *iris versicolor*. Considere as variáveis *Petal.Length* e *Petal.Width*. Apresente as matrizes de covariâncias estimadas de cada grupo e a ponderada, o valor da estatística do teste, o nível descritivo e a respectiva conclusão.

5. Considere $\boldsymbol{X}_{ij} \sim N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), i = 1, 2. j = 1, ..., n_i$, amostras aleatórias de duas populações independentes, em que $(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), i = 1, 2$ são ambos desconhecidos. Proponha uma estatística de teste, baseada na forma quadrática

$$T^2 = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1} \left(\overline{\boldsymbol{X}}_1 - \overline{\boldsymbol{X}}_2 - \boldsymbol{\Delta}\right)' \widehat{\boldsymbol{\Sigma}}^{-1} \left(\overline{\boldsymbol{X}}_1 - \overline{\boldsymbol{X}}_2 - \boldsymbol{\Delta}\right),$$

para testar as hipóteses $H_0 : \boldsymbol{R}\left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\right) = \boldsymbol{\Delta}$ vs $\boldsymbol{R}\left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\right) \neq \boldsymbol{\Delta}$, em que $\boldsymbol{R}_{(c \times p)}$, $c \leq p$ e $\boldsymbol{\Delta}_{(c \times 1)}$ são matrizes conhecidas de posto coluna completo. Encontre sua distribuição exata e mostre como calcular o valor crítico e o nível descritivo associado.

6. Com relação à Questão 5), obtenha a distribuição aproximada e mostre como calcular o valor crítico e o nível descritivo associado, baseado na respectiva distribuição assintótica.

7. Utilize o teste F, para a igualdade de vetores de médias, apresentando em classe para testar as hipóteses $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ vs $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, no conjunto de dados da Iris considerando os grupos setosa e versicolor como 1 e 2, respectivamente ($\alpha = 5\%$). Considere que $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$. Utilize as variáveis *Petal.Length* e *Petal.Width*. Apresente os vetores de médias estimados de cada grupo, as matrizes de covariâncias estimadas de cada grupo e a ponderada, o valor da estatística do teste, o nível descritivo e a respectiva conclusão. Caso você tenha rejeitado H_0, identifique em qual(is) variável(is) reside a diferença, através de alguma metodologia adequada.

8. Utilize o teste de Box, para a igualdade de matrizes de covariâncias, apresentando em classe para testar as hipóteses $H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_3$ vs H_1 : pelo menos uma diferença, no conjunto de dados da Iris considerando os grupos setosa, versicolor e virginica como 1, 2 e 3, respectivamente. Utilize as variáveis *Petal.Length* e *Petal.Width*. Apresente as matrizes de covariâncias estimadas de cada grupo e a ponderada, o valor da estatística do teste, o nível descritivo e a respectiva conclusão. Apresente os vetores de médias

estimados de cada grupo, o valor da estatística do teste, o nível descritivo e a respectiva conclusão.

9. Utilize o teste de Wilks, para igualdade de vetores de médias, apresentando em classe para testar as hipóteses $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{\mu}_3$ vs H_1 : pelo menos uma diferença, no conjunto de dados da Iris considerando os grupos setos, versicolor e virginica como 1, 2 e 3, respectivamente.

10. Com relação ao modelo linear multivariado homocedástico, considere-o em sua forma vetorial tal como visto em sala. Responda os itens:

    a) Prove que o e.m.v de $\boldsymbol{\beta}$ é dado por $\widehat{\boldsymbol{\beta}} = \left[(\boldsymbol{X}^*)'(\boldsymbol{\Sigma}^*)^{-1}\boldsymbol{X}^*\right]^{-1}(\boldsymbol{X}^*)'(\boldsymbol{\Sigma}^*)^{-1}\boldsymbol{Y}^*$.

    b) Encontre a distribuição exata de $\widehat{\boldsymbol{\beta}}$.

    c) Proponha uma estatística do teste, com sua respectiva distribuição (exata ou aproximada) para testar $\boldsymbol{R\beta} = \boldsymbol{b}$.

    d) Prove que o e.m.v de $\boldsymbol{B}$ é dado por $\widehat{\boldsymbol{B}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Y}$.

    e) Prove que o e.m.v de $\boldsymbol{\Sigma}$ é dado por $\widehat{\boldsymbol{\Sigma}} = \frac{1}{n}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{B}}\right)'\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{B}}\right)$.