3ª Lista de Exercícios - SME0822 Análise Multivariada - Profª Cibele Russo
Data: 10/10/2014.

Exercício 1. Seja $\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_n$ uma amostra aleatória de vetores de uma distribuição $N_p(\underset{\sim}{\mu}, \Sigma)$. Prove que $\widehat{\Sigma}$ (o estimador de máxima verossimilhança de Σ) é viesado para Σ e S (a matriz de covariâncias amostrais) é não-viesado para Σ.

Exercício 2. Se $\underset{\sim}{Y}_1, \ldots, \underset{\sim}{Y}_n$ são independentes seguindo distribuições $N_p(\underset{\sim}{\mu}_i, \Sigma_i)$, $i = 1, \ldots, n$, prove que $\sum_{i=1}^n a_i \underset{\sim}{Y}_i \sim N_p\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \Sigma_i\right)$, onde $\underset{\sim}{a}$ é um vetor de reais.

Exercício 3. Sejam $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$. Defina $Q = \sum_{i=1}^n (y_i - \bar{y})^2$.

(a) Prove que $\operatorname{cov}(\bar{y}, y_1 - \bar{y}) = 0$ e conclua que $\bar{y}$ é estatísticamente independente de Q.

(b) Qual a distribuição de Q/σ^2?

Exercício 4. Em cada uma das densidades bivariadas a seguir apresente o vetor de médias e a matriz de covariâncias

(a) $f(y_1, y_2) = \dfrac{1}{2\pi} \exp\left\{-\dfrac{1}{2}\left[(y_1-1)^2 + (y_2-2)^2\right]\right\}$

(b) $f(y_1, y_2) = \dfrac{1}{2.4} \exp\left\{-\dfrac{y_1^2/4 - 8.8 y_1 y_2 + y_2^2}{0.72}\right\}$

(c) $f(y_1, y_2) = \dfrac{1}{2\pi} \exp\left\{-\dfrac{1}{2}\left(y_1^2 + y_2^2 + 4y_1 - 6y_2 + 13\right)\right\}$

(d) $f(y_1, y_2) = \dfrac{1}{2\pi} \exp\left\{\dfrac{1}{2}\left(2y_1^2 + y_2^2 + 2y_1 y_2 - 22y_1 - 14y_2 + 65\right)\right\}$

(e) Há itens em que as variáveis são independentes? Quais?

(f) Represente graficamente as densidades e curvas de nível de cada item. Podem ser utilizados recursos computacionais (como por exemplo o R).

Exercício 5. Seja $\underset{\sim}{Y} \sim N_3(\underset{\sim}{0}, \Sigma)$, com

$$\Sigma = \begin{bmatrix} 1 & 0.8 & -0.4 \\ 0.8 & 1 & -0.56 \\ -0.4 & -0.56 & 1 \end{bmatrix}.$$

(a) Determine a distribuição condicional de y_1 e y_3 dado y_2.

(b) Calcule o coeficiente de correlação parcial entre y_1 e y_3 dado y_2.

Exercício 6. (a) Calcule a estatística T^2 para o teste $H_0 : \underset{\sim}{\mu}' = [7 \;\; 11]$, usando os dados:

$$X = \begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix}.$$

(b) Especifique a distribuição de T^2 para a situação (a).

(c) Usando (a) e (b), teste H_0 ao nível de significância $\alpha = 0.05$. Qual conclusão você obtém?

Exercício 7. (Mínimos Quadrados Ponderados). Seja

$$\underset{\sim}{Y}_{(n\times 1)} = \underset{\sim}{Z}_{(n\times(r+1))}\underline{\beta}_{((r+1)\times 1)} + \underset{\sim}{\epsilon}_{(n\times 1)}$$

onde $E[\underset{\sim}{\epsilon}] = \underset{\sim}{0}$ mas $E[\underset{\sim}{\epsilon}\underset{\sim}{\epsilon}'] = \sigma^2\underset{\sim}{V}$, com $\underset{\sim}{V}_{n\times n}$ conhecida e positiva definida. Para $\underset{\sim}{V}$ de posto completo, mostre que o estimador de mínimos quadrados ponderados é

$$\underline{\widehat{\beta}}_w = (\underset{\sim}{Z}'\underset{\sim}{V}\underset{\sim}{Z})^{-1}\underset{\sim}{Z}'\underset{\sim}{V}^{-1}\underset{\sim}{Y}.$$

Se σ^2 não for conhecido, um estimador não viesado é

$$(n-r-1)^{-1} \times (\underset{\sim}{Y} - \underset{\sim}{Z}\underline{\widehat{\beta}}_w)'\underset{\sim}{V}^{-1}(\underset{\sim}{Y} - \underset{\sim}{Z}\underline{\widehat{\beta}}_w).$$

Dica: $\underset{\sim}{V}^{-1/2}\underset{\sim}{Y} = (\underset{\sim}{V}^{-1/2}\underset{\sim}{Z})\underline{\beta} + \underset{\sim}{V}^{-1/2}\underset{\sim}{\epsilon}$ é a forma clássica do modelo de regressão linear $\underset{\sim}{Y}^* = \underset{\sim}{Z}^*\underline{\beta} + \underset{\sim}{\epsilon}^*$, com $E[\underset{\sim}{\epsilon}^*] = \underset{\sim}{0}$ e $E[\underset{\sim}{\epsilon}^*\underset{\sim}{\epsilon}^*] = \sigma^2\underset{\sim}{I}$. Assim, $\underline{\widehat{\beta}}_w = \underline{\widehat{\beta^*}} = (\underset{\sim}{Z}^*\underset{\sim}{Z}^*)^{-1}\underset{\sim}{Z}^{*\prime}\underset{\sim}{Y}^*$.

Exercício 8. Use o estimador de mínimos quadrados ponderados do exercício anterior para derivar uma expressão para a estimativa da inclinação β no modelo $Y_j = \beta_j z_j + \epsilon_j,\ j = 1, \ldots, n$, onde (a) $\mathrm{Var}[\epsilon_j] = \sigma^2$, (b) $\mathrm{Var}[\epsilon_j] = \sigma^2 z_j$, e (c) $\mathrm{Var}[\epsilon_j] = \sigma^2 z_j^2$. Comente de que maneira em que as variância desiguais para os erros influecia na escolha ótima para β_w.

Referências bibliográficas

- Johnson, R. A. and Wichern, D. W. (2007) Applied Multivariate Statistical Analysis. 5th edition. Prentice-Hall
- Coteia WIKI http://wiki.icmc.usp.br/index.php/SME-263.