Análise Multivariada II
Lista II

Breno Cauã Rodrigues da Silva

Questão 1. Análise A Posteriori

Pode-se estimar os vetores $\boldsymbol{\tau_k}$ para os tratamentos A, B, C e D de forma que

$$\begin{aligned}
\widehat{\boldsymbol{\tau}}_{\boldsymbol{A}} &= \left(\overline{\boldsymbol{X}}_A - \overline{\boldsymbol{X}}\right)^\intercal = \begin{pmatrix}-274.95 & -0.59\end{pmatrix}^\intercal \\
\widehat{\boldsymbol{\tau}}_{\boldsymbol{B}} &= \left(\overline{\boldsymbol{X}}_B - \overline{\boldsymbol{X}}\right)^\intercal = \begin{pmatrix}-202.75 & 0.37\end{pmatrix}^\intercal \\
\widehat{\boldsymbol{\tau}}_{\boldsymbol{C}} &= \left(\overline{\boldsymbol{X}}_C - \overline{\boldsymbol{X}}\right)^\intercal = \begin{pmatrix}179.45 & 0.03\end{pmatrix}^\intercal \\
\widehat{\boldsymbol{\tau}}_{\boldsymbol{D}} &= \left(\overline{\boldsymbol{X}}_D - \overline{\boldsymbol{X}}\right)^\intercal = \begin{pmatrix}298.25 & 0.20\end{pmatrix}^\intercal
\end{aligned}$$

Usando a Soma de Quadrados do Resíduo de Within ($\mathbf{W}$) expressa por

$$\begin{aligned}
\mathbf{W} &= \sum_{k=1}^{g}\sum_{i=1}^{n_k}(\mathbf{x}_{ki} - \overline{\mathbf{x}}_k)(\mathbf{x}_{ki} - \overline{\mathbf{x}}_k)^\intercal = n_1\mathbf{S}_1 + n_2\mathbf{S}_2 + \ldots + n_g\mathbf{S}_g \\
\mathbf{W} &\approx \begin{bmatrix}29058.55 & 10.26 \\ 10.26 & 0.32\end{bmatrix},
\end{aligned}$$

com $n - g = 20 - 4 = 16$ graus de liberdade. Onde $n = n_1 + n_2 + \ldots + n_g$ é o tamanho total da amostra e $g = 4$ o número de tratamentos em comparação. Temos que a diferença entre a produtividade de dois tratamentos quaisquer pode ser obtida por

$$\widehat{\boldsymbol{\tau}}_{k_1} - \widehat{\boldsymbol{\tau}}_{j_1},$$

com $\widehat{Var}\left[\widehat{\boldsymbol{\tau}}_{k_1} - \widehat{\boldsymbol{\tau}}_{j_1}\right] = \left(\frac{1}{n_k} + \frac{1}{n_j}\right)\frac{\mathbf{W}_{11}}{n-g}$. De posse da estatística $\widehat{\boldsymbol{\tau}}_{k_1} - \widehat{\boldsymbol{\tau}}_{j_1}$ e de sua variância, podemos montar o intervalo de $100(1-\alpha)\%$ de confiança, expresso por

$$(\widehat{\boldsymbol{\tau}}_{k_1} - \widehat{\boldsymbol{\tau}}_{j_1}) \pm t_{n-g\ ;\ \alpha/pg(p-1)}\sqrt{\widehat{Var}\left[\widehat{\boldsymbol{\tau}}_{k_1} - \widehat{\boldsymbol{\tau}}_{j_1}\right]}.$$

Comparação entre a ***Produtividade*** A e B

Temos que

$$\widehat{\boldsymbol{\tau}}_{A_1} - \widehat{\boldsymbol{\tau}}_{B_1} = -274.95 - (-202.75) = -72.20,$$

e $\widehat{Var}\left[\widehat{\boldsymbol{\tau}}_{A_1} - \widehat{\boldsymbol{\tau}}_{B_1}\right] = \left(\frac{1}{5} + \frac{1}{5}\right)\frac{29058.55}{20-4} = 726.46$, logo, $\sqrt{726.46} \approx 26.95$. Fixando um $\alpha = 0.05$, temos que o quantil de acordo com a correção de Bonferroni $t_{20-4\ ;\ 0.05/2\cdot4\cdot(4-1)} = t_{16\ ;\ 0.002} \approx 3.34$. Então, o intervalo de confiança fica

$$-72.20 \pm 3.34 \times 26.95 \approx [-162,21\ ;\ 17.81].$$

Como o intervalo contém o 0, diz-se que não há diferença entre a Produtividade do Tratamento A e do Tratamento B, ao nível de significância de 5%.

Comparação entre a ***Produtividade*** A e C

Temos que

$$\widehat{\boldsymbol{\tau}}_{A_1} - \widehat{\boldsymbol{\tau}}_{C_1} = -274.95 - 179.45 = -454.4,$$

e $\widehat{Var}\left[\widehat{\boldsymbol{\tau}}_{A_1} - \widehat{\boldsymbol{\tau}}_{C_1}\right] = \left(\frac{1}{5} + \frac{1}{5}\right)\frac{29058.55}{20-4} = 726.46$, logo, $\sqrt{726.46} \approx 26.95$. Fixando um $\alpha = 0.05$, temos que o quantil de acordo com a correção de Bonferroni $t_{20-4\ ;\ 0.05/2\cdot 4\cdot(4-1)} = t_{16\ ;\ 0.002} \approx 3.34$. Então, o intervalo de confiança fica

$$-454.4 \pm 3.34 \times 26.95 \approx [-364.39\ ;\ -544.41].$$

Como o intervalo não contém o 0, diz-se que há diferença entre a Produtividade do Tratamento A e do Tratamento C, ao nível de significância de 5%.

Comparação entre a ***Produtividade*** A e D

Temos que

$$\widehat{\boldsymbol{\tau}}_{A_1} - \widehat{\boldsymbol{\tau}}_{D_1} = -274.95 - 289.25 = -564.2,$$

e $\widehat{Var}\left[\widehat{\boldsymbol{\tau}}_{A_1} - \widehat{\boldsymbol{\tau}}_{D_1}\right] = \left(\frac{1}{5} + \frac{1}{5}\right)\frac{29058.55}{20-4} = 726.46$, logo, $\sqrt{726.46} \approx 26.95$. Fixando um $\alpha = 0.05$, temos que o quantil de acordo com a correção de Bonferroni $t_{20-4\ ;\ 0.05/2\cdot 4\cdot(4-1)} = t_{16\ ;\ 0.002} \approx 3.34$. Então, o intervalo de confiança fica

$$-564.2 \pm 3.34 \times 26.95 \approx [-654.21\ ;\ -474.19].$$

Como o intervalo não contém o 0, diz-se que há diferença entre a Produtividade do Tratamento A e do Tratamento D, ao nível de significância de 5%.

Comparação entre a ***Produtividade*** B e C

Temos que

$$\widehat{\boldsymbol{\tau}}_{B_1} - \widehat{\boldsymbol{\tau}}_{C_1} = -202.25 - 179.45 = -381.7,$$

e $\widehat{Var}\left[\widehat{\boldsymbol{\tau}}_{B_1} - \widehat{\boldsymbol{\tau}}_{C_1}\right] = \left(\frac{1}{5} + \frac{1}{5}\right)\frac{29058.55}{20-4} = 726.46$, logo, $\sqrt{726.46} \approx 26.95$. Fixando um $\alpha = 0.05$, temos que o quantil de acordo com a correção de Bonferroni $t_{20-4\ ;\ 0.05/2\cdot 4\cdot(4-1)} = t_{16\ ;\ 0.002} \approx 3.34$. Então, o intervalo de confiança fica

$$-381.7 \pm 3.34 \times 26.95 \approx [-471,71\ ;\ -291,69].$$

Como o intervalo não contém o 0, diz-se que há diferença entre a Produtividade do Tratamento A e do Tratamento D, ao nível de significância de 5%.

Comparação entre a ***Produtividade*** B e D

Temos que

$$\widehat{\boldsymbol{\tau}}_{B_1} - \widehat{\boldsymbol{\tau}}_{D_1} = -202.25 - 298.75 = -501,$$

e $\widehat{Var}\left[\widehat{\boldsymbol{\tau}}_{B_1} - \widehat{\boldsymbol{\tau}}_{D_1}\right] = \left(\frac{1}{5} + \frac{1}{5}\right)\frac{29058.55}{20-4} = 726.46$, logo, $\sqrt{726.46} \approx 26.95$. Fixando um $\alpha = 0.05$, temos que o quantil de acordo com a correção de Bonferroni $t_{20-4\ ;\ 0.05/2\cdot 4\cdot(4-1)} = t_{16\ ;\ 0.002} \approx 3.34$. Então, o intervalo de confiança fica

$$-501 \pm 3.34 \times 26.95 \approx [-591.01\ ;\ -410.99].$$

Como o intervalo não contém o 0, diz-se que há diferença entre a Produtividade do Tratamento A e do Tratamento D, ao nível de significância de 5%.

Comparação entre a ***Produtividade*** C e D

Temos que

$$\widehat{\boldsymbol{\tau}}_{C_1} - \widehat{\boldsymbol{\tau}}_{D_1} = 179.45 - 298.25 = -118.8,$$

e $\widehat{Var}\left[\widehat{\boldsymbol{\tau}}_{C_1} - \widehat{\boldsymbol{\tau}}_{D_1}\right] = \left(\frac{1}{5} + \frac{1}{5}\right)\frac{29058.55}{20-4} = 726.46$, logo, $\sqrt{726.46} \approx 26.95$. Fixando um $\alpha = 0.05$, temos que o quantil de acordo com a correção de Bonferroni $t_{20-4\ ;\ 0.05/2\cdot 4\cdot(4-1)} = t_{16\ ;\ 0.002} \approx 3.34$. Então, o intervalo de confiança fica

$$-118.8 \pm 3.34 \times 26.95 \approx [-208.81\ ;\ -28.79].$$

Como o intervalo não contém o 0, diz-se que há diferença entre a Produtividade do Tratamento A e do Tratamento D, ao nível de significância de 5%.

Questão 2. MANOVA