

Exercise 6.4 (in the following ' is used for the transpose matrix)

Consider 3 normal populations corresponding to mean values

$$\boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \quad \boldsymbol{\mu}_3 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

and the common dispersion matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

1. Verify that the matrix corresponding to the variation between group means is

$$\mathbf{B} = \sum_{i=1}^3 (\boldsymbol{\mu}_i - \bar{\boldsymbol{\mu}})(\boldsymbol{\mu}_i - \bar{\boldsymbol{\mu}})' = \begin{bmatrix} 14 & 4 \\ 4 & 2 \end{bmatrix}$$

Here $\bar{\boldsymbol{\mu}}$ is the average of the three mean values

Consider a random variable \mathbf{X} following one of the three distributions and introduce the linear combination

$$Y = \mathbf{d}' \mathbf{X} = (a \ b) \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = aX_1 + bX_2.$$

For $i=1, 2, 3$ the mean and variance of Y are equal to

$$\nu_i = E(Y|i) = \mathbf{d}' \boldsymbol{\mu}_i$$

$$\sigma_Y^2 = V(Y|i) = \mathbf{d}' \boldsymbol{\Sigma} \mathbf{d}$$

2. Verify for $\bar{\nu}$ equal to the average of the ν_i -values that (3-1) \times (the empirical variance of the ν_i -values) is

$$\sum_{i=1}^3 (\nu_i - \bar{\nu})^2 = \mathbf{d}' \mathbf{B} \mathbf{d}$$

Given the vectors \mathbf{d}_1 and \mathbf{d}_2 consider the ratio between (3-1) \times (the empirical variance of the ν_i -values) and the variance of Y

$$\frac{\sum_{i=1}^3 (\nu_i - \bar{\nu})^2}{\sigma_Y^2} = \frac{\mathbf{d}'_k \mathbf{B} \mathbf{d}_k}{\mathbf{d}'_k \boldsymbol{\Sigma} \mathbf{d}_k}, \quad k = 1, 2.$$

3. Determine the vectors \mathbf{d}_k so that the ratio is maximized for $k=1$ and maximized under the constraint $\mathbf{d}'_1 \boldsymbol{\Sigma} \mathbf{d}_2 = 0$ for $k=2$.

4. Same question as above, but now for the case

$$\boldsymbol{\Sigma} = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}$$