Homework 6 (Due 5/22)

Question 1 Consider a sample $\vec{x}_1, \ldots, \vec{x}_{10}$ from $\mathcal{N}_2(\vec{\mu}, \mathbf{\Sigma})$. The summary statistics are

$$\overline{\boldsymbol{x}} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad \boldsymbol{S} = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}.$$

(a) Find one-at-a-time 95% confidence intervals for μ_1 and μ_2;
(b) Find simultaneous $\geq 95\%$ confidence intervals for $\mu_1, \mu_2, \mu_1 + \mu_2, \mu_1 - \mu_2$ based on T^2;
(c) Find simultaneous $\geq 95\%$ confidence intervals for μ_1 and μ_2 based on Bonferroni correction.

Question 2 Let $\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim \mathcal{N}_3(\vec{\mu}, \mathbf{\Sigma})$ with $\vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$. Let $\vec{x}_1, \ldots, \vec{x}_{10}$ be an observed sample from the above population, for which the sample mean and sample covariance matrix are

$$\bar{\vec{x}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{S}_x = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix},$$

respectively.

1. Find $\geq 95\%$ simultaneous confidence intervals for $\mu_1 - \mu_2$ and $\mu_1 - \mu_3$ with the T^2 based on the sample $\vec{x}_1, \ldots, \vec{x}_{10}$;
2. Suppose that the original sample is transformed to the 2-variate sample $\vec{y}_1, \ldots, \vec{y}_{10}$ by the linear transformation $Y_1 = X_1 - X_2$ and $Y_2 = X_1 - X_3$. Find $\geq 95\%$ simultaneous confidence intervals for $\mu_1 - \mu_2$ and $\mu_1 - \mu_3$ with the T^2 based on the new sample;
3. Find $\geq 95\%$ simultaneous confidence intervals for $\mu_1 - \mu_2$ and $\mu_1 - \mu_3$ with Bonferroni correction;
4. Compare the above results.

Question 3 We have introduced in class the Bonferroni-corrected two-sample test: $H_0 : \vec{\mu}_1 - \vec{\mu}_2 = \vec{\delta}_0$ is rejected if

$$\max_{1 \leq j \leq p} \left| \frac{(\overline{x}_{1j} - \overline{x}_{2j}) - \delta_{0j}}{s_{pooled,j}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| \geq t_{n_1+n_2-2}\left(\frac{\alpha}{2p}\right).$$

Prove the following Type I error control:

$$\mathbb{P}_{null}\left(\max_{1 \leq j \leq p} \left| \frac{(\overline{X}_{1j} - \overline{X}_{2j}) - \delta_{0j}}{S_{pooled,j}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| \geq t_{n_1+n_2-2}\left(\frac{\alpha}{2p}\right) \right) \leq \alpha.$$

Question 4 Consider a sample of size $n_1 = 14$ from $\mathcal{N}_2(\vec{\mu}_1, \mathbf{\Sigma}_1)$ and a sample of size $n_2 = 14$ from $\mathcal{N}_2(\vec{\mu}_2, \mathbf{\Sigma}_2)$. Assume $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$. The summary statistics for these two samples are

$$\bar{\vec{x}}_1 = \begin{bmatrix} 8 \\ 5 \end{bmatrix}, \quad \boldsymbol{S}_1 = \begin{bmatrix} 3 & -1 \\ -1 & 6 \end{bmatrix}, \quad \bar{\vec{x}}_2 = \begin{bmatrix} 10 \\ 4 \end{bmatrix}, \quad \boldsymbol{S}_2 = \begin{bmatrix} 3 & -1 \\ -1 & 4 \end{bmatrix}.$$

1. Test $H_0 : \vec{\mu}_1 = \vec{\mu}_2$ at the level of $\alpha = .05$ with Hotelling's T^2;
2. Find $\geq 95\%$ simultaneous confidence intervals for $\mu_{1j} - \mu_{2j},\ j = 1, \ldots, p$ with Bonferroni corrected t-tests;
3. Test $H_0 : \vec{\mu}_1 = \vec{\mu}_2$ at the level of at most $\alpha = .05$ with Bonferroni corrected t-tests.

Question 5 Suppose we have the following two samples:

$$\text{Sample 1: } \vec{x}_{11}, \ldots, \vec{x}_{1n_1} \text{ from } \mathcal{N}_p(\vec{\mu}_1, \mathbf{\Sigma}_1),$$

$$\text{Sample 2: } \vec{x}_{21}, \ldots, \vec{x}_{2n_2} \text{ from } \mathcal{N}_p(\vec{\mu}_2, \mathbf{\Sigma}_2).$$

Two new samples are defined through the linear transformations $\vec{y}_{lj} = \boldsymbol{C}\vec{x}_{lj} + \vec{d}$ for all $l = 1, 2$ and $j = 1, 2, \ldots, n_l$, where $\boldsymbol{C}$ is a $p \times p$ nonsingular matrix and $\vec{d}$ is a $p \times 1$ vector. Based on Samples 1 and 2, the T^2-statistic for testing $\vec{\mu}_1 = \vec{\mu}_2$ is denoted as T_x^2. On the other hand, based on the two samples $\vec{y}_{11}, \ldots, \vec{y}_{1n_1}$ and $\vec{y}_{21}, \ldots, \vec{y}_{2n_2}$, the T^2-statistic for testing the equality of vector means is denoted as T_y^2. Show that $T_x^2 = T_y^2$.

Question 6 Consider two independent samples from 3-variate multivariate normal populations:

$$\text{Population 1 with } \vec{\mu}_1 = \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \end{bmatrix} \text{: sample size } n_1 = 10,\ \bar{\vec{x}}_1 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix},\ \boldsymbol{S}_1 = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 2 \end{bmatrix};$$
$$\text{Population 2 with } \vec{\mu}_2 = \begin{bmatrix} \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} \text{: sample size } n_2 = 10,\ \bar{\vec{x}}_2 = \begin{bmatrix} 4 \\ 6 \\ 6 \end{bmatrix},\ \boldsymbol{S}_2 = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 4 \end{bmatrix}.$$

Furthermore, assume the population covariance matrices of the two populations are the same. We aim to test

$$H_0 : \mu_{11} - \mu_{21} = \mu_{12} - \mu_{22} = \mu_{13} - \mu_{23}.$$

1. Test H_0 with $\alpha = 0.05$ by Hotelling's T^2;
2. Test H_0 with $\alpha \leq 0.05$ by Bonferroni correction.

Hint: Find $\boldsymbol{C}$ such that the null can be written as $H_0 : \boldsymbol{C}\vec{\mu}_1 = \boldsymbol{C}\vec{\mu}_2$.