Homework 5 (Due 5/8)

Question 1 Let $\vec{X}_1, \ldots, \vec{X}_n$ be a random sample from $\mathcal{N}_p(\vec{\mu}, \mathbf{\Sigma})$. Show that

1. $\overline{\vec{X}} \sim \mathcal{N}_p(\vec{\mu}, \frac{1}{n}\mathbf{\Sigma})$;
2. $\mathbb{E}(\boldsymbol{S}) = \mathbf{\Sigma}$.

Question 2 Consider a sample of size $n = 10$ from a 2-variate population. The summery statistics are

$$\overline{\vec{x}} = \begin{bmatrix} 1.5 \\ 1.6 \end{bmatrix}, \quad \boldsymbol{S} = \begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix}.$$

1. Test $\mu_1 = 0$ with $\alpha = .05$;
2. Test $\mu_2 = 0$ with $\alpha = .05$;
3. Test $\mu_1 = \mu_2 = 0$ with $\alpha = .05$;
4. Plot the 95% confidence region for $\vec{\mu}$.

Question 3 Let

$$\vec{X}_1, \ldots, \vec{X}_{25}$$

be a random sample from $\mathcal{N}_2(\vec{\mu}, \mathbf{\Sigma})$, where $\vec{\mu} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{\Sigma} = \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix}$. Denote the sample mean and sample covariance by $\overline{\vec{X}}$ and $\boldsymbol{S}$, respectively.

1. Derive the distribution of $\overline{\vec{X}}$.
2. Derive the distribution of $\left(\overline{\vec{X}} - \vec{\mu}\right)^\top \boldsymbol{S}^{-1} \left(\overline{\vec{X}} - \vec{\mu}\right)$. Furthermore, generate a histogram of this random variable by generating an independent sample of $(\overline{\vec{X}}, \boldsymbol{S})$, and compare the histogram with the distribution you have derived.

Question 4 Consider a sample of size $n = 10$ from a 3-variate population. The summary statistics for this sample are

$$\overline{\boldsymbol{x}} = \begin{bmatrix} 5 \\ 4 \\ 6 \end{bmatrix}, \quad \boldsymbol{S} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Test $H_0 : \mu_1 = \mu_2 = \mu_3$ with $\alpha = .05$.

Question 5 For a sample $\vec{x}_1, \ldots, \vec{x}_{10}$ from $\mathcal{N}_4(\vec{\mu}, \mathbf{\Sigma})$, the sample mean and sample covariance matrix are

$$\bar{\vec{x}} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{S}_x = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix},$$

respectively.
(a) Test $\mu_1 - \mu_3 = \mu_2 - \mu_4 = 0$ with $\alpha = .05$ by Hotelling's T^2.
(b) Test $\mu_1 - \mu_2 = \mu_2 - \mu_3 = \mu_3 - \mu_4$ with $\alpha = .05$ by Hotelling's T^2.

Question 6 Consider a sample $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ from $\mathcal{N}_2(\vec{\mu}, \mathbf{\Sigma})$. The sample mean and sample covariance are

$$\bar{\vec{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{S} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix},$$

respectively.
(a) Find a 95% confidence interval of $\mu_1 - \mu_2$.
(b) Assume $s_{12} > 0$, and someone ignores this positive correlation and takes the wrong sample covariance

$$\widetilde{\boldsymbol{S}} = \begin{bmatrix} s_{11} & 0 \\ 0 & s_{22} \end{bmatrix}.$$

Will this person derive a wider or narrower 95% confidence interval of $\mu_1 - \mu_2$ than the correct one? Explain.