Homework 3

Deadline: May 18, 2022

1. Let $\mathbf{x}_\alpha$ be distributed according to $N(c_\alpha \boldsymbol{\gamma}, \boldsymbol{\Sigma})$, for $\alpha = 1, \ldots, N$, where $\sum_{c=1}^N c_\alpha^2 > 0$.
   (a) Show that the distribution of
   $$\mathbf{g} = \frac{1}{\sum_{\alpha=1}^N c_\alpha^2} \sum_{c=1}^N c_\alpha \mathbf{x}_\alpha$$
   is $\mathcal{N}\big(\boldsymbol{\gamma}, 1/\big(\sum_{\alpha=1}^N c_\alpha^2\big)\big)$.
   (b) Show that
   $$\mathbf{E} = \sum_{\alpha=1}^N (\mathbf{x}_\alpha - c_\alpha \mathbf{g})(\mathbf{x}_\alpha - c_\alpha \mathbf{g})^\top$$
   is independently distributed as $\sum_{\alpha=1}^{N-1} \mathbf{z}_\alpha \mathbf{z}_\alpha^\top$, where $\mathbf{z}_1, \ldots, \mathbf{z}_N$ are independent, each with distribution $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$.

2. Suppose we have N observations $\mathbf{x}_1, \ldots, \mathbf{x}_N$ which are independently distributed according to $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Define the the sample mean and the sample covariance matrix as
   $$\bar{\mathbf{x}} = \sum_{\alpha=1}^N \mathbf{x}_\alpha \quad \text{and} \quad \mathbf{S} = \frac{1}{N-1} \sum_{\alpha=1}^N (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top$$
   respectively.
   (a) Show that $\bar{\mathbf{x}}$ is efficient for estimating $\boldsymbol{\mu}$.
   (b) Show that $\bar{\mathbf{x}}$ and $\mathbf{S}$ have efficiency
   $$\left(\frac{N-1}{N}\right)^{\frac{p(p+1)}{2}}$$
   for estimating $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

3. Let $T^2 = N\bar{\mathbf{x}}^\top \mathbf{S}^{-1} \bar{\mathbf{x}}$, where $\bar{\mathbf{x}}$ and $\mathbf{S}$ are the mean vector and covariance matrix of a sample of N from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Show that T^2 is distributed the same when $\boldsymbol{\mu}$ is replaced by $\boldsymbol{\lambda} = [\tau, 0, \ldots, 0]^\top$, where $\tau^2 = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$, and $\boldsymbol{\Sigma}$ is replaced by $\mathbf{I}$.

4. Let $\mathbf{x}_\alpha$ be distributed according to $\mathcal{N}_p(\boldsymbol{\mu} + (z_\alpha - \bar{z})\boldsymbol{\beta}, \boldsymbol{\Sigma})$ for $\alpha = 1, \ldots, N$, where $\bar{\mathbf{z}} = \frac{1}{N}\sum_{\alpha=1}^N \mathbf{z}_\alpha$. Let
   $$\mathbf{b} = \frac{1}{\sum_{\alpha=1}^N (z_\alpha - \bar{z})^2} \sum_{\alpha=1}^N (z_\alpha - \bar{z}) \mathbf{x}_\alpha, \quad (N-2)\mathbf{S} = \sum_{\alpha=1}^N (\mathbf{x} - \bar{\mathbf{x}} - (z_\alpha - \bar{z})\mathbf{b})(\mathbf{x} - \bar{\mathbf{x}} - (z_\alpha - \bar{z})\mathbf{b})^\top$$
   and
   $$T^2 = \sum_{\alpha=1}^N (z_\alpha - \bar{z})^2 \mathbf{b}^\top \mathbf{S}^{-1} \mathbf{b}.$$
   Find c such that cT^2 is distributed according to the F-distribution and find the parameters of this F-distribution.

5. Let

$$\mathbf{X} = \begin{bmatrix} 1.9 & 0.7 \\ 0.8 & -1.6 \\ 1.1 & -0.2 \\ 0.1 & -1.2 \\ -0.1 & -0.1 \\ 4.4 & 3.4 \\ 5.5 & 3.7 \\ 1.6 & 0.8 \\ 4.6 & 0 \\ 3.4 & 2 \end{bmatrix} \in \mathbb{R}^{10\times 2}$$

and denote α-th row of $\mathbf{X}$ be $\mathbf{x}_\alpha^\top$. We suppose that each $\mathbf{x}_\alpha$ is an observation from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = [\mu_1, \mu_2]^\top$.

(a) Give a confidence region for $\boldsymbol{\mu}$ with confidence coefficient 0.95.

(b) Test the hypothesis that both μ_1 and μ_2 are non-negative at significance level 0.01.

6. Let $\mathbf{x}_\alpha^{(i)}$ be observations from $\mathcal{N}(\boldsymbol{\mu}^{(i)}, \boldsymbol{\Sigma}_i)$ for $\alpha = 1, \dots, N_i$, $i = 1, 2$. Find the likelihood ratio criterion for testing the hypothesis $\boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)}$.

7. Let $\left\{\mathbf{x}_\alpha^{(i)}\right\}$ for $\alpha = 1, \dots, N_i$, $i = 1, \dots, q$ be independent samples from $\mathcal{N}(\boldsymbol{\mu}^{(i)}, \boldsymbol{\Sigma}_i)$ for $i = 1, 2$, respectively. We suppose $N_1 < N_2$ and define

$$\mathbf{y}_\alpha = \mathbf{x}_\alpha^{(1)} - \sqrt{\frac{N_1}{N_2}}\mathbf{x}_\alpha^{(2)} + \frac{1}{\sqrt{N_1 N_2}} \sum_{\beta=1}^{N_1} \mathbf{x}_\beta^{(2)} - \frac{1}{N_2} \sum_{\gamma=1}^{N_2} \mathbf{x}_\gamma^{(2)},$$

for $\alpha = 1, \dots, N_1$. Then we have

$$\bar{\mathbf{y}} = \frac{1}{N_1} \sum_{\alpha=1}^{N_1} \mathbf{y}_\alpha = \bar{\mathbf{x}}_\alpha^{(1)} - \bar{\mathbf{x}}_\alpha^{(2)}$$

and

$$\mathrm{Cov}(\mathbf{y}_\alpha, \mathbf{y}_{\alpha'}) = \begin{cases} \boldsymbol{\Sigma}_1 + \frac{N_1}{N_2}\boldsymbol{\Sigma}_2, & \alpha = \alpha', \\ \mathbf{0}, & \text{otherwise.} \end{cases}$$