Homework-3

1. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be a random sample of size n from a p-dimensional normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Show that

$$T^2 = n(\bar{\boldsymbol{X}} - \boldsymbol{\mu})^T \mathbf{S}^{-1}(\bar{\boldsymbol{X}} - \boldsymbol{\mu})) \sim \frac{(n-1)p}{n-p} F_{p,n-p}$$

.

2. Show that for $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, non-random matrices $\mathbf{A}_{q\times p}$ and $\mathbf{B}_{r\times p}$, $\mathbf{Y} = \mathbf{AX}$ and $\mathbf{Z} = \mathbf{BX}$ are independent if and only if $\mathbf{A\Sigma B}^T = 0$.

3. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{38}$ be a random sample of size $n = 32$ from a bivariate normal distribution, i.e., $\mathbf{X}_I \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. We obtain the sample mean $\bar{\mathbf{X}} = (0.564, 0.603)^T$ and the sample covariance $\mathbf{S} = \begin{pmatrix} 0.0144 & 0.0117 \\ 0.0117 & 0.0146 \end{pmatrix}$. It is desire to construct a confidence region of size α for $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$.

   a) Find the MLE of $\boldsymbol{\Sigma}$.

   b) Evaluate the expression for 95% elliptical confidence region for $\boldsymbol{\mu}$. Denote this region by R_1.

   c) Is $\boldsymbol{\mu}_0 = (0.60, 0.58)^T$ fall in R_1?

   d) Evaluate the simultaneous confidence intervals for μ_1 and μ_2. Denote this region by R_2.

   e) Conduct a hypothesis test for $Ho : \boldsymbol{\mu} = (0.60, 0.58)^T$ versus $H_a : \boldsymbol{\mu} \neq (0.60, 0.58)^T$. Report the details of the test statistics, the distribution of it under the null hypothesis, and the p-value.

4. *Bonferroni method for multiple comparison.* An alternative to the simultaneous confidence interval is the Bonferroni method for multiple comparison. In problem 3, since there are only two parameters μ_1 and μ_2) involved. Notice that Bonferroni method has some advantageous compared to the simultaneous confidence intervals.

   In general, for m linear combinations $\mathbf{a}_1^T\boldsymbol{\mu}, \mathbf{a}_2^T\boldsymbol{\mu}, \ldots, \mathbf{a}_m^T\boldsymbol{\mu}$, assume C_i is the confidence interval for the ith linear combination, $\mathbf{a}_i^T\boldsymbol{\mu}$ with confidence level α_i. Then, we have

$$\begin{aligned} P(\mathbf{a}_i^T\boldsymbol{\mu} \in C_i \text{for all} i &= 1 - P(\text{at least for one } i, \mathbf{a}_1^T\boldsymbol{\mu} \notin C_i) \\ &= \geq 1 - \sum_{i=1}^{m} P(\mathbf{a}_i^T\boldsymbol{\mu} \notin C_i) \\ &= 1 - (\alpha_1 + \alpha_2 + \cdots + \alpha_m) \end{aligned} \tag{1}$$

   a) Show that the interval $C_i(\alpha_i) = \mathbf{a}_i^T\bar{\mathbf{x}} \pm t_{n-1}(\alpha_i/2)\sqrt{\mathbf{a}_i^T\mathbf{S}\mathbf{a}_i/n}$ is a confidence interval of $\mathbf{a}_i^T\boldsymbol{\mu}$ with confidence level α_i.

b) Verify the inequality (1), for $m = 3$.

c) verify

$$P(\mathbf{a}_i^T \boldsymbol{\mu} \in C_i(\alpha/m) \text{for all } i) \geq 1 - \alpha.$$

This gives a Bonferroni type simultaneous confidence intervals $C_i(\alpha/m)$ for level α.

d) Using the data in problem 3, evaluate the simultaneous confidence intervals for μ_1 andμ_2 using the Bonferroni method. Denote this region by R_3

e) Sketch a graph for the regions R_1, R_2, R_3 in $x - y$ coordinates. Which one do you prefer and why?

5. Consider the monthly returns of three stocks (Coke, IBM, and Caterpillar with ticker symbols KO, IBM, and CAT, respectively) and 3 market indexes (value-weighted (VW), equal-weighted (EW), and S&P composite index) from January 1961 to December 2015. The returns include dividends. The data are in the file *"three-stocks.txt"*.

   a) Are the monthly returns of three stocks jointly normally distributed? Why?

   b) Consider jointly the returns of the three stocks and three indexes. Test the null hypothesis $Ho : \boldsymbol{\mu} = 0$ versus $H_a : \boldsymbol{\mu} \neq 0$, where $\boldsymbol{\mu}$ denotes the mean vector of the monthly simple returns. Perform the test and make your conclusion.

   c) Provide justifications for the reference distribution used in the testing of part

   d) Transform the returns to log returns. Construct simultaneous T^2, Bonferroni, marginal, and asymptotic chi-square confidence intervals for the means of the log returns.

   e) Now consider again the monthly *log* returns of three stocks and three indexes of part (d). Divide the sample into two parts non-overlapping, part one, from 1 to 330, and part 2 from 331 to 660 in time order.

   I. Let $\boldsymbol{\Sigma}_i$ be the covariance matrix of the sub-period $i = 1, 2$. Test $H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ versus $H_a : \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$. Make your conclusion.

   II. Let $\boldsymbol{\mu}_i$ be the vector of mean returns in the sub-period $i = 1, 2$. Based on the result of Part(II), test $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ versus $H_a : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$. Make your conclusion.

6. **Reading assignments:** Chapters 5 to 7 of the textbook.