Jared Yu 914640019

STA 135 HW 5

1. Let $\vec{X}_1, \cdots, \vec{X}_n$ be a random sample from $\mathcal{N}_p(\vec{\mu}, \mathbf{\Sigma})$. It will next be shown that,
   1. $\bar{\vec{X}} \sim N_P\left(\vec{\mu}, \frac{1}{n}\mathbf{\Sigma}\right)$;
   2. $E(\boldsymbol{S}) = \mathbf{\Sigma}$.

   To show (1), the following will be done:

   Let

$$\bar{\vec{X}} = \frac{1}{n}\vec{X}_1 + \frac{1}{n}\vec{X}_2 + \cdots + \frac{1}{n}\vec{X}_n,$$

where $\vec{X}_1, \cdots, \vec{X}_n \sim N_p(\vec{\mu}, \mathbf{\Sigma})$ for $i = 1, \cdots, n$. Then below is $E\left(\bar{\vec{X}}\right)$:

$$E\left(\bar{\vec{X}}\right) = E\left(\frac{1}{n}\vec{X}_1 + \frac{1}{n}\vec{X}_2 + \cdots + \frac{1}{n}\vec{X}_n\right) = E\left(\frac{1}{n}\vec{X}_1\right) + E\left(\frac{1}{n}\vec{X}_2\right) + \cdots + E\left(\frac{1}{n}\vec{X}_n\right)$$

$$= \frac{1}{n}E(\vec{X}_1) + \frac{1}{n}E(\vec{X}_2) + \cdots + \frac{1}{n}E(\vec{X}_n) = \frac{1}{n}\vec{\mu} + \frac{1}{n}\vec{\mu} + \cdots + \frac{1}{n}\vec{\mu} = \vec{\mu}$$

Next $Cov\left(\bar{\vec{X}}\right)$ will be shown:

$$Cov\left(\bar{\vec{X}}\right) = E\left\{\left(\bar{\vec{X}} - \vec{\mu}\right)\left(\bar{\vec{X}} - \vec{\mu}\right)^T\right\} = E\left\{\left(\frac{1}{n}\Sigma_{j=1}^n(\vec{X}_j - \vec{\mu})\right)\left(\frac{1}{n}\Sigma_{l=1}^n(\vec{X}_l - \vec{\mu})\right)^T\right\}$$

$$= E\left\{\frac{1}{n^2}\Sigma_{j=1}^n\Sigma_{l=1}^n(\vec{X}_j - \vec{\mu})(\vec{X}_l - \vec{\mu})^T\right\} = \frac{1}{n^2}\left\{\Sigma_{j=1}^n\Sigma_{l=1}^n E\left[(\vec{X}_j - \vec{\mu})(\vec{X}_l - \vec{\mu})^T\right]\right\}$$

It will be noted first that $E\left[(\vec{X}_j - \vec{\mu})(\vec{X}_l - \vec{\mu})^T\right]$ is also $Cov(\vec{X}_j, \vec{X}_l)$. Since it is a random sample, the $\vec{X}_j$ and $\vec{X}_l$ are independent for $j \neq l$. Therefore, the covariance for them is zero.

$$\frac{1}{n^2}\left\{\Sigma_{j=1}^n\Sigma_{l=1}^n E\left[(\vec{X}_j - \vec{\mu})(\vec{X}_l - \vec{\mu})^T\right]\right\} = \frac{1}{n^2}\left\{\Sigma_{j=1}^n E\left[(\vec{X}_j - \vec{\mu})(\vec{X}_j - \vec{\mu})^T\right]\right\}$$

Then, let $\mathbf{\Sigma} = E\left[(\vec{X}_j - \vec{\mu})(\vec{X}_j - \vec{\mu})^T\right]$ denote the population covariance matrix. So, it follows then that:

$$\frac{1}{n^2}\left\{\Sigma_{j=1}^n E\left[(\vec{X}_j - \vec{\mu})(\vec{X}_j - \vec{\mu})^T\right]\right\} = \frac{1}{n^2}\Sigma_{j=1}^n\mathbf{\Sigma} = \frac{1}{n}\mathbf{\Sigma}$$

From this it follows that $\bar{\vec{X}} \sim N_P\left(\vec{\mu}, \frac{1}{n}\mathbf{\Sigma}\right)$, where $\bar{\vec{X}}$ has mean $\vec{\mu}$ and variance-covariance matrix $\frac{1}{n}\mathbf{\Sigma}$. The reason is that $\vec{X}_1, \cdots, \vec{X}_n \sim N_p(\vec{\mu}, \mathbf{\Sigma})$, and so the linear combination $\bar{\vec{X}}$ also follows a normal distribution $\bar{\vec{X}} \sim N_P\left(\vec{\mu}, \frac{1}{n}\mathbf{\Sigma}\right)$.

To show (2), the following will be done:

First it is given that $\boldsymbol{S}_n$, the biased sampled variance-covariance matrix is as follows,

$$\boldsymbol{S}_n = \frac{1}{n}\Sigma_{j=1}^{n}\left(\vec{X}_j - \bar{\vec{X}}\right)\left(\vec{X}_j - \bar{\vec{X}}\right)^T.$$

Then it can also be written as follows:

$$\boldsymbol{S}_n = \frac{1}{n}\Sigma_{j=1}^{n}\left(\vec{X}_j - \bar{\vec{X}}\right)\left(\vec{X}_j - \bar{\vec{X}}\right)^T = \frac{1}{n}\left\{\Sigma_{j=1}^{n}\left(\vec{X}_j - \bar{\vec{X}}\right)\vec{X}_j^T + \left(\Sigma_{j=1}^{n}\left(\vec{X}_j - \bar{\vec{X}}\right)\right)\left(-\bar{\vec{X}}\right)^T\right\}$$

$$= \frac{1}{n}\left\{\Sigma_{j=1}^{n}\left(\vec{X}_j\vec{X}_j^T - \bar{\vec{X}}\vec{X}_j^T\right) + \vec{0}\left(-\bar{\vec{X}}\right)^T\right\} = \frac{1}{n}\left\{\Sigma_{j=1}^{n}\vec{X}_j\vec{X}_j^T - n\bar{\vec{X}}\bar{\vec{X}}^T\right\}$$

$\because \Sigma_{j=1}^{n}\left(\vec{X}_j - \bar{\vec{X}}\right) = \Sigma_{j=1}^{n}\vec{X}_j - n \times \frac{1}{n}\Sigma_{i=1}^{n}\vec{X}_i = \vec{0}$ and therefore $n\bar{\vec{X}}^T = \Sigma_{j=1}^{n}\vec{X}_j^T$

Then taking the expectation,

$$E\left[\frac{1}{n}\left\{\Sigma_{j=1}^{n}\vec{X}_j\vec{X}_j^T - n\bar{\vec{X}}\bar{\vec{X}}^T\right\}\right] = \frac{1}{n}E\left(\Sigma_{j=1}^{n}\vec{X}_j\vec{X}_j^T - n\bar{\vec{X}}\bar{\vec{X}}^T\right)$$

$$= \frac{1}{n}\left\{\Sigma_{j=1}^{n}E\left(\vec{X}_j\vec{X}_j^T\right) - nE\left(\bar{\vec{X}}\bar{\vec{X}}^T\right)\right\}.$$

It must be shown first that (1) $E\left(\vec{X}_j\vec{X}_j^T\right) = \boldsymbol{\Sigma} + \vec{\mu}\vec{\mu}^T$ and (2) $E\left(\bar{\vec{X}}\bar{\vec{X}}^T\right) = \frac{1}{n}\boldsymbol{\Sigma} + \vec{\mu}\vec{\mu}^T$.

$$Var\left(\vec{X}_j\right) = E\left(\vec{X}_j\vec{X}_j^T\right) - E\left(\vec{X}_j\right)E\left(\vec{X}_j^T\right) = Cov\left(\vec{X}_j\right)$$

$$E\left(\vec{X}_j\vec{X}_j^T\right) = Cov\left(\vec{X}_j\right) + E\left(\vec{X}_j\right)E\left(\vec{X}_j^T\right)$$

$$E\left(\vec{X}_j\vec{X}_j^T\right) = \boldsymbol{\Sigma} + \vec{\mu}\vec{\mu}^T \tag{1}$$

$$Var\left(\bar{\vec{X}}\right) = E\left(\bar{\vec{X}}\bar{\vec{X}}^T\right) - E\left(\bar{\vec{X}}\right)E\left(\bar{\vec{X}}\right)^T = Cov\left(\bar{\vec{X}}\right)$$

$$E\left(\bar{\vec{X}}\bar{\vec{X}}^T\right) = Cov\left(\bar{\vec{X}}\right) + E\left(\bar{\vec{X}}\right)E\left(\bar{\vec{X}}\right)^T$$

$$E\left(\bar{\vec{X}}\bar{\vec{X}}^T\right) = \frac{1}{n}\boldsymbol{\Sigma} + \vec{\mu}\vec{\mu}^T \tag{2}$$

Now returning to the expectation from behavior:

$$\frac{1}{n}\left\{\Sigma_{j=1}^{n}E\left(\vec{X}_j\vec{X}_j^T\right) - nE\left(\bar{\vec{X}}\bar{\vec{X}}^T\right)\right\} = \frac{1}{n}\left\{n(\boldsymbol{\Sigma} + \vec{\mu}\vec{\mu}^T) - n\left(\frac{1}{n}\boldsymbol{\Sigma} + \vec{\mu}\vec{\mu}^T\right)\right\}$$

$$= \frac{1}{n}\{n\boldsymbol{\Sigma} + n\vec{\mu}\vec{\mu}^T - \boldsymbol{\Sigma} - n\vec{\mu}\vec{\mu}^T\} = \frac{(n-1)}{n}\boldsymbol{\Sigma}$$

Then the unbiased sample variance-covariance matrix $\boldsymbol{S}$ can be written as follows,

$$\boldsymbol{S} = \left(\frac{n}{n-1}\right)\boldsymbol{S}_n = \frac{1}{n-1}\Sigma_{j=1}^{n}\left(\vec{X}_j - \bar{\vec{X}}\right)\left(\vec{X}_j - \bar{\vec{X}}\right)^T.$$

It has been shown that $E(\boldsymbol{S}_n) = \frac{(n-1)}{n}\boldsymbol{\Sigma}$ and so it follows that $E(\boldsymbol{S}) = \boldsymbol{\Sigma}$. ■

2. Given a sample size of $n = 10$, from a 2-variate population, the following are the summary statistics,

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

The following is a test for $\mu_1 = 0$ with $\alpha = 0.05$:

1) It can be said that $\vec{X}$ is the vector of samples, where, $\vec{X} = \begin{bmatrix}\vec{X}_1^T\\ \vec{X}_2^T\\ \vdots\\ \vec{X}_{10}^T\end{bmatrix}$. Then the distribution of $\vec{X}$ is as follows,

$$\vec{X} \sim N_2(\vec{\mu}, \boldsymbol{\Sigma}).$$

Then the $\vec{X}_i$ for $i = 1,2,\cdots,n = 10$ are represented as $\vec{X}_i = \begin{bmatrix}X_{i,1}\\X_{i,2}\end{bmatrix}$ where the jth column is the jth variable of the 2-variate sample.

Assume $X_{1,1}, X_{2,1}, \cdots, X_{n,1} \sim N(\mu_1, \sigma_{1,1})$ where $X_{i,j}$ is the ith sample in the jth column. The sample mean is $\bar{x}_1 = 1.5$ and the sample variance is $s_{1,1}^2 = 4$. Then the hypothesis test is $H_0{:}\,\mu_1 = 0$ vs $H_1{:}\,\mu_1 \neq 0$ at $\alpha = 0.05$. The test statistic is,

$$\frac{\bar{x}_1 - \mu_1}{s_{1,1}/\sqrt{n}} \sim t_{n-1}.$$

NOTE: $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ and $t = \frac{Z}{\sqrt{\frac{\chi^2_{n-1}}{n-1}}} = \frac{\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{\frac{(n-1)s^2}{\sigma^2}}{n-1}}} = \frac{\bar{x}-\mu}{s/\sqrt{n}}$

Under H_0 is true,

$$t^* = \frac{1.5 - 0}{2/\sqrt{10}} = 2.371708.$$

The critical value of a t distribution with $d.f. = 9$ is ± 2.262157. The rejection rule is to reject the null hypothesis if $|t^*| > 2.262157$. Since the test statistic is larger than the critical value, the decision is to reject the null hypothesis that $\mu_1 = 0$ in favor of the alternative that $\mu_1 \neq 0$ at confidence level $\alpha = 0.05$.

2) Assume $X_{1,2}, X_{2,2}, \cdots, X_{n,2} \sim N(\mu_2, \sigma_{2,2})$ where $X_{i,j}$ is the ith sample in the jth column. The sample mean is $\bar{x}_2 = 1.6$ and the sample variance is $s_{2,2}^2 = 4$. Then the hypothesis test is $H_0: \mu_2 = 0$ vs $H_1: \mu_2 \neq 0$ at $\alpha = 0.05$. The test statistic is,

$$\frac{\bar{x}_2 - \mu_2}{s_{2,2}/\sqrt{n}} \sim t_{n-1}.$$

NOTE: $Z = \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ and $t = \frac{Z}{\sqrt{\frac{\chi_{n-1}^2}{n-1}}} = \frac{\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{\frac{(n-1)s^2}{\sigma^2}}{n-1}}} = \frac{\bar{x}-\mu}{s/\sqrt{n}}$

Under H_0 is true,

$$t^* = \frac{1.6 - 0}{2/\sqrt{10}} = 2.529822.$$

The critical value of a t distribution with $d.f. = 9$ is ± 2.262157. The rejection rule is to reject the null hypothesis if $|t^*| > 2.262157$. Since the test statistic is larger than the critical value, the decision is to reject the null hypothesis that $\mu_2 = 0$ in favor of the alternative that $\mu_2 \neq 0$ at confidence level $\alpha = 0.05$.

3) Assume $\vec{X}_1, \vec{X}_2, \cdots, \vec{X}_{10}$ is a random sample from a $N_2(\vec{\mu}, \mathbf{\Sigma})$ population. The hypothesis test is $H_0: \mu_1 = \mu_2 = 0$ vs $H_1: \mu_1, \mu_2 \neq 0$ at $\alpha = 0.05$. Then the test statistic T^2 is,

$$T^{2^*} = n(\bar{x} - \mu_0)^T \boldsymbol{S}^{-1}(\bar{x} - \mu_0) \sim \frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha).$$

Under H_0 is true,

$$T^{2^*} = 10\left(\begin{bmatrix}1.5\\1.6\end{bmatrix}\right)^T \begin{pmatrix} \frac{4}{15} & \frac{1}{15} \\ \frac{1}{15} & \frac{4}{15} \end{pmatrix} \left(\begin{bmatrix}1.5\\1.6\end{bmatrix}\right) = ([15 \quad 16]) \begin{pmatrix} \frac{4}{15} & \frac{1}{15} \\ \frac{1}{15} & \frac{4}{15} \end{pmatrix} \left(\begin{bmatrix}1.5\\1.6\end{bmatrix}\right)$$
$$= \left(\left[\left(4 + \frac{16}{15}\right) \quad \left(1 + \frac{64}{15}\right)\right]\right)\left(\begin{bmatrix}1.5\\1.6\end{bmatrix}\right) = \left(\begin{bmatrix}\frac{76}{15} & \frac{79}{15}\end{bmatrix}\right)\left(\begin{bmatrix}1.5\\1.6\end{bmatrix}\right) = 16.02667.$$

NOTE: $\boldsymbol{S}^{-1} = \frac{1}{16-1}\begin{pmatrix}4 & 1\\1 & 4\end{pmatrix} = \begin{pmatrix} \frac{4}{15} & \frac{1}{15} \\ \frac{1}{15} & \frac{4}{15} \end{pmatrix}$.

Also,

$$\frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha) = \frac{(10-1)2}{(10-2)} F_{2,10-2}(0.05) = \frac{9}{4} F_{2,10-2}(0.05) = 10.032268$$

The rejection rule is to reject the null hypothesis if $\left|T^{2^*}\right| > 10.032268$. Since the test statistic is larger than the critical value, the decision is to reject the null hypothesis that $\mu_1 = \mu_2 = 0$ in favor of the alternative that $\mu_1, \mu_2 \neq 0$ at confidence level $\alpha = 0.05$.

4) To plot the 95% confidence region for $\vec{\mu}$, it will require that an ellipse is plotted with center $\vec{x}$ and the axes of the confidence ellipsoid are $\pm\sqrt{\lambda_i}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)}e_i$ where $\boldsymbol{S}e_i = \lambda_i e_i, i = 1,2$. First the $\lambda_i's$ will be shown:

$$|\boldsymbol{S} - \lambda \boldsymbol{I}| = 0$$

$$\begin{vmatrix} 4-\lambda & -1 \\ -1 & 4-\lambda \end{vmatrix} = 0$$

$$(4-\lambda)^2 - (-1)^2 = 0$$

$$(4-\lambda-1)\big(4-\lambda-(-1)\big) = 0$$

$$(3-\lambda)(5-\lambda) = 0$$

$$\lambda_1 = 5, \lambda_2 = 3$$

In the case of $\lambda_1 = 5$,

$$\begin{vmatrix} 4-5 & -1 \\ -1 & 4-5 \end{vmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{vmatrix} -1 & -1 \\ -1 & -1 \end{vmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$R_1 - R_2 \rightarrow R_2, \qquad \begin{vmatrix} -1 & -1 \\ 0 & 0 \end{vmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$-1R_1 \rightarrow R_1, \qquad \begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$x + y = 0$$

$$0 + 0 = 0$$

$$\frac{x}{y} = -1$$

$$\boldsymbol{x} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$e_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

In the case of $\lambda_2 = 3$,

$$\begin{vmatrix} 4-3 & -1 \\ -1 & 4-3 \end{vmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$R_1 + R_2 \to R_2, \qquad \begin{vmatrix} 1 & -1 \\ 0 & 0 \end{vmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$x - y = 0$$

$$0 + 0 = 0$$

$$\frac{x}{y} = 1$$

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$e_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Below is a plot of the 95% confidence region for $\vec{\mu}$.

3. Given that $\vec{X}_1, \cdots, \vec{X}_{25}$ is a random sample from $N_2(\vec{\mu}, \mathbf{\Sigma})$, where $\vec{\mu} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{\Sigma} = \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix}$. The sample mean and sample covariance are $\vec{\bar{X}}$ and $\boldsymbol{S}$, respectively.

   1. The following will show the distribution of $\vec{\bar{X}}$:

$$\vec{\bar{X}} = \frac{1}{25}\vec{X}_1 + \frac{1}{25}\vec{X}_2 + \cdots + \frac{1}{25}\vec{X}_{25}$$

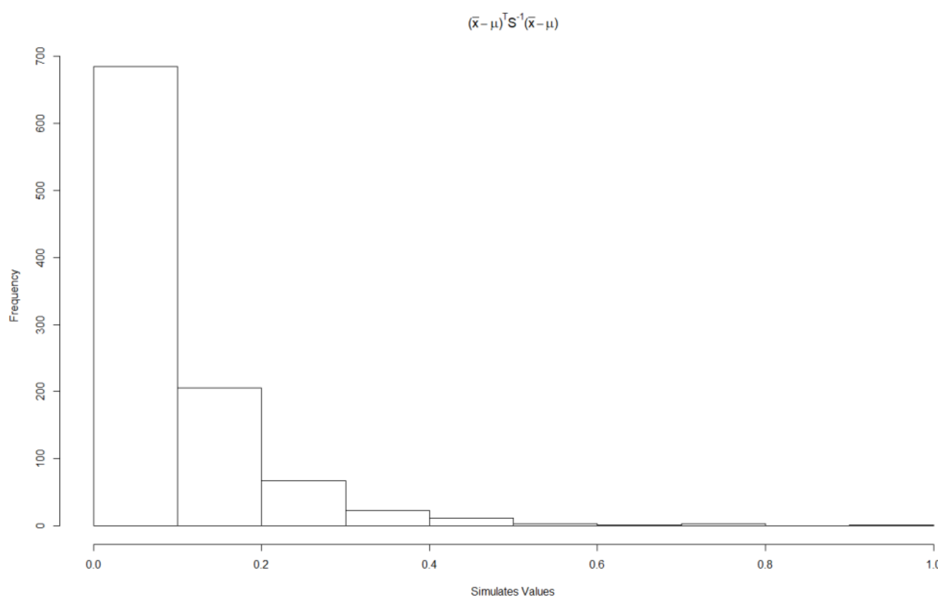
      The assumption for question 1 stated that, let $\vec{X}_1, \cdots, \vec{X}_n$ be a random sample from $\mathcal{N}_p(\vec{\mu}, \mathbf{\Sigma})$. From this it was determined that $\vec{\bar{X}} \sim N_P\left(\vec{\mu}, \frac{1}{n}\mathbf{\Sigma}\right)$. In this case, $n = 25$ and $p = 2$. So it can be said that $\vec{\bar{X}} \sim N_2\left(\vec{\mu}, \frac{1}{25}\mathbf{\Sigma}\right)$ where $\vec{\mu} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{\Sigma} = \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix}$.

   2. Assuming that $\vec{X}_1, \cdots, \vec{X}_p \sim N_p(\vec{\mu}, \mathbf{\Sigma})$, the hypothesis test for μ is $H_0: \vec{\mu} = \vec{\mu}_0$ vs $H_1: \vec{\mu} \neq \vec{\mu}_0$. Then under the null hypothesis, the test statistic is

$$T^2 = n\left(\vec{\bar{X}} - \vec{\mu}_0\right)^T \boldsymbol{S}^{-1}\left(\vec{\bar{X}} - \vec{\mu}_0\right) \sim \frac{(n-1)p}{(n-p)} F_{p,n-p},$$

      where n is the sample size and p is the number of variables in each observation. Then it follows that the distribution $\left(\vec{\bar{X}} - \vec{\mu}\right)^T \boldsymbol{S}^{-1}\left(\vec{\bar{X}} - \vec{\mu}\right)$ will follow a $\frac{(n-1)p}{(n-p)n} F_{p,n-p}$.

      Below is a histogram of the above random variable, it involves 1,000 simulations. The histogram has a highly skewed distribution towards the right, like the appearance of the F-distribution. This makes sense, as the random variable is supposed to follow something proportional to the F-distribution.



4. Given a sample size of $n = 10$, from a 3-variate population, the following are the summary statistics,

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

The goal is to test $H_0: \mu_1 = \mu_2 = \mu_3$ with $\alpha = 0.05$. This can be done using the following hypothesis test:

$$H_0: \vec{C}\vec{\mu} = \vec{0} \text{ vs. } H_1: \vec{C}\vec{\mu} \neq \vec{0}$$

where $\vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$ and $\vec{C} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$. Under the null hypothesis, the test statistic for the above test is as follows,

$$T^2 = n\left(\vec{C}\bar{\boldsymbol{x}}\right)^T \left(\vec{C}\vec{S}\vec{C}^T\right)^{-1} \vec{C}\bar{\boldsymbol{x}}.$$

The critical value for the test is,

$$\frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha).$$

Using R, $T^2 = 7.142857$, which is smaller than the critical value of 10.03268, therefore we fail to reject the null hypothesis that $\vec{C}\vec{\mu} = \vec{0}$, or $\mu_1 = \mu_2 = \mu_3$. It is worth noting that in the critical value, $p = 2$ since the value inside is transformed rather than being the original x.

5. Given a sample $\vec{x}_1, \cdots, \vec{x}_{10}$ from $N_4(\vec{\mu}, \boldsymbol{\Sigma})$, the sample mean, and sample covariance matrix are

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

respectively.

a) The goal is to test $H_0: \mu_1 - \mu_3 = \mu_2 - \mu_4 = 0$ with $\alpha = 0.05$. This can be done using the following hypothesis test:

$$H_0: \vec{C}\vec{\mu} = \vec{0} \text{ vs. } H_1: \vec{C}\vec{\mu} \neq \vec{0}$$

where $\vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix}$ and $\vec{C} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$. Under the null hypothesis, the test statistic for the above test is as follows,

$$T^2 = n\left(\vec{C}\bar{\boldsymbol{x}}\right)^T \left(\vec{C}\vec{S}\vec{C}^T\right)^{-1} \vec{C}\bar{\boldsymbol{x}}.$$

The critical value for the test is,

$$\frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha).$$

Using R, $T^2 = 1.714286$, which is smaller than the critical value of 10.03268, therefore we fail to reject the null hypothesis that $\vec{C}\vec{\mu} = \vec{0}$, or $\mu_1 - \mu_3 = \mu_2 - \mu_4 = 0$.

b) The goal is to test $H_0: \mu_1 - \mu_2 = \mu_2 - \mu_3 = \mu_3 - \mu_4$ with $\alpha = 0.05$. This can be done using the following hypothesis test:

$$H_0: \vec{C}\vec{\mu} = \vec{0} \text{ vs. } H_1: \vec{C}\vec{\mu} \neq \vec{0}$$

where $\vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix}$ and $\vec{C} = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix}$. Under the null hypothesis, the test statistic for the above test is as follows,

$$T^2 = n\left(\vec{C}\bar{\boldsymbol{x}}\right)^T \left(\vec{C}\vec{S}\vec{C}^T\right)^{-1} \vec{C}\bar{\boldsymbol{x}}.$$

The critical value for the test is,

$$\frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha).$$

Using R, $T^2 = 4$, which is smaller than the critical value of 10.03268, therefore we fail to reject the null hypothesis that $\vec{C}\vec{\mu} = \vec{0}$, or $\mu_1 - \mu_2 = \mu_2 - \mu_3 = \mu_3 - \mu_4$.

6. Given a sample $\vec{x}_1, \cdots, \vec{x}_n$ from $N_2(\vec{\mu}, \boldsymbol{\Sigma})$, the sample mean, and sample covariance matrix are

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

respectively.

a) a) A 95% confidence interval for $\mu_1 - \mu_2$ requires a paired comparison. Using a linear combination of $\vec{X}$, where $Z = a_1X_1 + a_2X_2$, $a_1 = 1, a_2 = -1$. This can be rewritten as $\vec{a}^T\vec{X}$. Then it also follows that, $\mu_z = E(Z) = \vec{a}^T\vec{\mu}$ and $\sigma_Z^2 = Var(Z) = \vec{a}^T\Sigma\vec{a}$. Then the sample mean and variance of the observed values are $\bar{z} = \vec{a}^T\bar{\vec{x}}$ and $s_Z^2 = \vec{a}^T\vec{S}\vec{a}$. Then for $\vec{a}$ fixed and σ_Z^2 unknown, a $100(1-\alpha)\%$ confidence interval for $\mu_Z = \vec{a}^T\vec{\mu}$ is based on student's t-ratio

$$t = \frac{\bar{z} - \mu_Z}{s_Z/\sqrt{n}} = \frac{\sqrt{n}\left(\vec{a}^T\bar{\vec{x}} - \vec{a}^T\vec{\mu}\right)}{\sqrt{\vec{a}^T\vec{S}\vec{a}}}$$

which leads to

$$\bar{z} - t_{n-1}\left(\frac{\alpha}{2}\right)\frac{s_z}{\sqrt{n}} \leq \mu_Z \leq \bar{z} + t_{n-1}\left(\frac{\alpha}{2}\right)\frac{s_z}{\sqrt{n}}$$

or

$$\vec{a}^T\bar{\vec{x}} - t_{n-1}\left(\frac{\alpha}{2}\right)\frac{\sqrt{\vec{a}^T\vec{S}\vec{a}}}{\sqrt{n}} \le \vec{a}^T\vec{\mu} \le \vec{a}^T\bar{\vec{x}} + t_{n-1}\left(\frac{\alpha}{2}\right)\frac{\sqrt{\vec{a}^T\vec{S}\vec{a}}}{\sqrt{n}}$$

So, the following can be plugged in:

$$\sqrt{\vec{a}^T\vec{S}\vec{a}} = \sqrt{\begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}s_{11} & s_{12}\\ s_{21} & s_{22}\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix}} = \sqrt{\begin{bmatrix}s_{11}-s_{21} & s_{12}-s_{22}\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix}} = \sqrt{s_{11}-s_{21}-s_{12}+s_{22}}$$
$$= \sqrt{s_{11}-2s_{12}+s_{22}}$$

$$\vec{a}^T\bar{\vec{x}} = \begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}\bar{x}_1\\ \bar{x}_2\end{bmatrix} = \bar{x}_1 - \bar{x}_2$$

$$\vec{a}^T\vec{\mu} = \begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} = \mu_1 - \mu_2$$

To get the following 95% confidence interval for $\mu_1 - \mu_2$:

$$\bar{x}_1 - \bar{x}_2 - t_{n-1}(0.025)\frac{\sqrt{s_{11}-2s_{12}+s_{22}}}{\sqrt{n}} \le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + t_{n-1}(0.025)\frac{\sqrt{s_{11}-2s_{12}+s_{22}}}{\sqrt{n}}$$

b) If $s_{12} > 0$ and that the wrong sample covariance,

$$\tilde{\boldsymbol{S}} = \begin{bmatrix}s_{11} & 0\\ 0 & s_{22}\end{bmatrix},$$

is taken, then the confidence interval will become wider. The reason is that if $s_{12} > 0$, then the effect is that there is no longer a $-2s_{12}$ component to the $\frac{\sqrt{\vec{a}^T\vec{S}\vec{a}}}{\sqrt{n}}$ portion of the confidence interval. Therefore, since the actual value is positive, there would have been a narrower confidence interval.