Homework 4 Solutions

1. We compute the joint density

$$p(x,y) \propto \exp\left(-\frac{1}{2}\begin{pmatrix}(x-\mu_x)^\top & (y-\mu_y)^\top\end{pmatrix}\begin{pmatrix}\Sigma_{xx} & \Sigma_{xy}\\ \Sigma_{yx} & \Sigma_{yy}\end{pmatrix}^{-1}\begin{pmatrix}x-\mu_x\\ y-\mu_y\end{pmatrix}\right).$$

We use the following inversion formula (https://en.wikipedia.org/wiki/Block_matrix#Inversion):

$$\begin{pmatrix}\Sigma_{xx} & \Sigma_{xy}\\ \Sigma_{yx} & \Sigma_{yy}\end{pmatrix}^{-1} = \begin{pmatrix}(\Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1} & -(\Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}\Sigma_{xy}\Sigma_{yy}^{-1}\\ -\Sigma_{yy}^{-1}\Sigma_{yx}(\Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1} & \cdots\end{pmatrix},$$

where the last block is not important for our purpose. From this, we can see that

$$\begin{aligned}
&\begin{pmatrix}(x-\mu_x)^\top & (y-\mu_y)^\top\end{pmatrix}\begin{pmatrix}\Sigma_{xx} & \Sigma_{xy}\\ \Sigma_{yx} & \Sigma_{yy}\end{pmatrix}^{-1}\begin{pmatrix}x-\mu_x\\ y-\mu_y\end{pmatrix}\\
&\quad= (x-\mu_x)^\top(\Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}(x-\mu_x)\\
&\qquad- (x-\mu_x)^\top(\Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}\Sigma_{xy}\Sigma_{yy}^{-1}(y-\mu_y)\\
&\qquad- (y-\mu_y)^\top\Sigma_{yy}^{-1}\Sigma_{yx}(\Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}(x-\mu_x)\\
&\qquad+ \text{ terms depending only on } y\\
&\quad= (x-\mu_x-\Sigma_{xy}\Sigma_{yy}^{-1}(y-\mu_y))^\top(\Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}(x-\mu_x-\Sigma_{xy}\Sigma_{yy}^{-1}(y-\mu_y))\\
&\qquad+ \text{ terms depending only on } y.
\end{aligned}$$

Therefore, the conditional density $p(x\,|\,y) = \frac{p(x,y)}{p(y)}$ depends on x only through the term

$$\exp\left(-\frac{(x-\mu_x-\Sigma_{xy}\Sigma_{yy}^{-1}(y-\mu_y))^\top(\Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1}(x-\mu_x-\Sigma_{xy}\Sigma_{yy}^{-1}(y-\mu_y))}{2}\right),$$

which means that

$$X\,|\,Y=y \sim N(\mu_x+\Sigma_{xy}\Sigma_{yy}^{-1}(y-\mu_y), \Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}).$$

Hence, we can say that

$$X\,|\,Y \sim N(\mu_x+\Sigma_{xy}\Sigma_{yy}^{-1}(Y-\mu_y), \Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}).$$

2-(a). We know that $(X_1, \bar{X})$ has to be jointly normal. We have $\mathbb{E}(X_1) = \mathbb{E}(\bar{X}) = \mu$, $\mathrm{Var}(\bar{X}) = \frac{1}{n}$, and $\mathrm{Cov}(X_1, \bar{X}) = \frac{1}{n}$. Hence,

$$\begin{pmatrix}X_1\\ \bar{X}\end{pmatrix} \sim N\left(\begin{pmatrix}\mu\\ \mu\end{pmatrix}, \begin{pmatrix}1 & \frac{1}{n}\\ \frac{1}{n} & \frac{1}{n}\end{pmatrix}\right).$$

From Q1, we have

$$\mathbb{E}(X_1\,|\,\bar{X}) = \mu + \frac{1}{n}n(\bar{X}-\mu) = \bar{X}.$$

2-(b). By symmetry, $\mathbb{E}(X_2 \mid \bar{X})$ has to be $\bar{X}$ as well. Hence,

$$\mathbb{E}\left(\frac{X_1+X_2}{2} \,\middle|\, \bar{X}\right) = \bar{X}.$$

2-(c). Similarly, $\mathbb{E}(X_3 \mid \bar{X}) = \bar{X}$. Hence,

$$\mathbb{E}\left(\frac{X_1+X_2+X_3}{3} \,\middle|\, \bar{X}\right) = \bar{X}.$$

2-(d). From this, we can deduce that the conditional expectation of the average of any subset of $\{X_1, \ldots, X_n\}$ given the $\bar{X}$ has to be $\bar{X}$.

3. Recall that

$$\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Hence, we have an α-level test by rejecting the null if

$$\sum_{i=1}^n (X_i - \bar{X})^2 > \chi^2_{n-1,1-\alpha}.$$

4. Let z_α be the α-th quantile of $N(0,1)$. Then, we have

$$\mathbb{P}(p(X) \le \alpha) = \mathbb{P}\left(\Phi\left(-\frac{\sqrt{n}\bar{X}}{\sigma}\right) \le \alpha\right) = \mathbb{P}\left(-\frac{\sqrt{n}\bar{X}}{\sigma} \le z_\alpha\right).$$

Under the null hypothesis, we have $\mu \le 0$ and thus

$$\mathbb{P}\left(-\frac{\sqrt{n}\bar{X}}{\sigma} \le z_\alpha\right) \le \mathbb{P}\left(-\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \le z_\alpha\right) = \mathbb{P}(N(0,1) \le z_\alpha) = \alpha.$$

Hence, under the null hypothesis, we have $\mathbb{P}(p(X) \le \alpha) \le \alpha$. When $\mu < 0$, we can see that the above inequality is strict. Hence, the distribution of p-value is not uniform.

5-(a). We have

$$\mathbb{E}(M) = \sum_{j=1}^m \mathbb{E}\mathbb{I}\left\{\mu_j \notin [\hat{\mu}_{j,\text{left}}, \hat{\mu}_{j,\text{right}}]\right\} = \sum_{j=1}^m \mathbb{P}(\mu_j \notin [\hat{\mu}_{j,\text{left}}, \hat{\mu}_{j,\text{right}}]) = m\alpha.$$

5-(b). For $j \in \{1, \ldots, m\}$, we know that

$$\mathbb{P}(\mu_j \in [\bar{X}_j - z_{1-\frac{\alpha}{2m}}, \bar{X}_j + z_{1-\frac{\alpha}{2m}}]) = 1 - \frac{\alpha}{m},$$

where $\bar{X}_j = \frac{1}{n}\sum_{i=1}^n X_{i,j}$ is the mean of the j-th coordinates of $X_1, \ldots, X_n$. Hence, letting $\hat{\mu}_{j,\text{left}} = \bar{X}_j - z_{1-\frac{\alpha}{2m}}$ and $\hat{\mu}_{j,\text{right}} = \bar{X}_j + z_{1-\frac{\alpha}{2m}}$, we have

$$\mathbb{P}(\mu_j \notin [\hat{\mu}_{j,\text{left}}, \hat{\mu}_{j,\text{right}}] \text{ for some } j \in \{1, \ldots, m\}) \leq \sum_{j=1}^m \mathbb{P}(\mu_j \notin [\hat{\mu}_{j,\text{left}}, \hat{\mu}_{j,\text{right}}]) = \alpha.$$

Hence, we have

$$\mathbb{P}(\mu_j \in [\hat{\mu}_{j,\text{left}}, \hat{\mu}_{j,\text{right}}] \;\; \forall j \in \{1, \ldots, m\}) \geq 1 - \alpha.$$