Homework 4

1. Suppose $\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}\right)$. Find the conditional distribution of $X|Y$. Hint: use the formula $p(x|y) = \frac{p(x,y)}{p(y)}$.

2. Consider i.i.d. observations $X_1, ..., X_n \sim N(\mu, 1)$.

   (a) Compute $\mathbb{E}(X_1|\bar{X})$. Hint: use the above problem, and find the conditional distribution of X_1 given $\bar{X}$ first.

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

   (c) What about $\mathbb{E}\left(\frac{X_1+X_2+X_3}{3}\middle|\bar{X}\right)$?

   (d) Discuss your finding.

3. Consider i.i.d. observations $X_1, ..., X_n \sim N(\mu, \sigma^2)$ with both μ, σ^2 unknown. Construct an α-level test for $H_0 : \sigma^2 \leq 1$ vs $H_1 : \sigma^2 > 1$ using chi-squared distribution.

4. Consider i.i.d. observations $X_1, ..., X_n \sim N(\mu, \sigma^2)$ with μ unknown and σ^2 known. For the problem $H_0 : \mu \leq 0$ vs $H_1 : \mu > 0$, show that the p-value $p(X) = \Phi\left(-\frac{\sqrt{n}\bar{X}}{\sigma}\right)$ satisfies $\mathbb{P}\left(p(X) \leq \alpha\right) \leq \alpha$ for all $\mu \leq 0$ under the null hypothesis. Is the distribution of p-value uniform under the null?

5. Suppose you have m different parameters $\mu_1, ..., \mu_m$, and you construct m $(1-\alpha)$-confidence intervals, denoted by $[\hat{\mu}_{j,\text{left}}, \hat{\mu}_{j,\text{right}}]$ for $j = 1, ..., m$. It is guaranteed that each confidence interval has the $(1-\alpha)$-coverage in the sense that $\mathbb{P}\left(\mu_j \in [\hat{\mu}_{j,\text{left}}, \hat{\mu}_{j,\text{right}}]\right) = 1 - \alpha$.

   (a) The number of parameters that are not covered by the confidence intervals is $M = \sum_{j=1}^m \mathbb{I}\{\mu_j \notin [\hat{\mu}_{j,\text{left}}, \hat{\mu}_{j,\text{right}}]\}$. What is the expected value of M?

   (b) You hope all the parameters $\mu_1, ..., \mu_m$ can be covered by the m confidence intervals simultaneously. Mathematically, you want to construct intervals $[\hat{\mu}_{j,\text{left}}, \hat{\mu}_{j,\text{right}}]$ that satsify

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

   Consider a concrete setting with i.i.d. observations $X_1, ..., X_n \sim N(\mu, I_m)$, where $\mu = (\mu_1, ..., \mu_m)^T \in \mathbb{R}^m$. Apply the idea of Bonferroni correction and construct confidence intervals that satisfy the above requirement.