SM-4331 Exercise 3

1. Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathrm{N}(\mu, \sigma^2)$. Prove that the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ is distributed according to $\bar{X} \sim \mathrm{N}(\mu, \sigma^2/n)$. *Hint: Find the expectation and variance of $\bar{X}$, and use the linearity property of normal distributions.*

2. Suppose that we plan to take a random sample of size n from a normal distribution with mean μ and standard deviation $\sigma = 2$.
   (a) Suppose $\mu = 4$ and $n = 20$.
      i. What is the probability that the mean $\bar{X}$ of the sample is greater than 5?
      ii. What is the probability that $\bar{X}$ is smaller than 3?
      iii. What $\mathrm{P}(|\bar{X} - \mu| \leq 1)$ in this case?
   (b) How large should n be in order that $\mathrm{P}(|\bar{X} - \mu| \leq 0.5) \geq 0.95$ for every possibly value of μ?
   (c) It is claimed that the true value of μ is 5 in a population. A random sample of size $n = 100$ is collected from this population, and the mean for this sample is $\bar{X} = 5.8$. Based on the result in (b), what would you conclude from this value of $\bar{X}$?

3. (a) If Z is a random variable with a standard normal distribution, what is $\mathrm{P}(Z^2 < 3.841)$?
   (b) Suppose that X_1 and X_2 are independent $\mathrm{N}(0, 4)$ random variables. Compute $\mathrm{P}(X_1^2 < 36.84 - X_2^2)$.
   (c) Suppose that $X_1, X_2, X_3 \overset{\text{iid}}{\sim} \mathrm{N}(0, 1)$, while Y independently follows a χ^2_5 distribution. Compute $\mathrm{P}(X_1^2 + X_2^2 < 7.236Y - X_3^2)$.

4. Let X_i, $i = 1, 2, 3$ be independent with $\mathrm{N}(i, i^2)$ distributions. For each of the following situations, use the X_is to construct a statistic with the indicated distribution:
   (a) χ^2-distribution with 3 degrees of freedom;
   (b) t-distribution with 2 degrees of freedom; and
   (c) F-distribution with 1 and 2 degrees of freedom.

5. Imagine rolling an r-sided die n number of times independently. Define the indicator function

$$\mathbb{1}_{[k=i]}(k) = \begin{cases} 1 & \text{if roll } k \text{ is equal to } i \\ 0 & \text{otherwise} \end{cases}$$

   Suppose further that $\mathrm{P}(\mathbb{1}_{[k=i]}(k) = 1) = p_i$.
   (a) What is $\mathrm{E}\left[\mathbb{1}_{[k=i]}(k)\right]$ and $\mathrm{Var}\left[\mathbb{1}_{[k=i]}(k)\right]$?
   (b) Calculate $\mathrm{E}\left[\mathbb{1}_{[k=i]}(k)\,\mathbb{1}_{[l=j]}(l)\right]$ when $k \neq l$.
   (c) Argue that $\mathrm{E}\left[\mathbb{1}_{[k=i]}(k)\,\mathbb{1}_{[l=j]}(l)\right] = 0$ when $k = l$.
   (d) Let X_i be the number of rolls that result in side i facing up. Write down the equation relating X_i and the indicator functions above. What possible values can X_i take?

(e) Determine $\mathrm{E}(X_i)$.

(f) Consider two random variables X_i and X_j defined as per (d). From your answers to (a), (b) and (c), calculate $\mathrm{E}(X_i X_j)$.

(g) Now calculate the covariance between X_i and X_j.

6. Let $\{X_1, \ldots, X_n\}$ be a random sample from a $\mathrm{N}(\mu, \sigma^2)$ population.

   (a) Let $M = \sum_{i=1}^{n}(X_i - \bar{X})^2$, where $\bar{X}$ is the sample mean. Work out the distribution of M/σ^2.

   (b) Let $\alpha = 0.05$. Using the χ^2 probability tables, determine the values of $\chi^2_{14}(\alpha/2)$ and $\chi^2_{14}(1-\alpha/2)$, i.e. the top and bottom $\alpha/2$ point of the χ^2_{14} distribution where $\mathrm{P}\left(Y > \chi^2_k(a)\right) = a$ when $Y \sim \chi^2_k$.

   (c) Suppose $n = 15$ and the sample variance is $s^2 = 24.5$. What is a 95% confidence interval for σ^2?

7. Let $\{Y_{ij}\}$ be sample from $\mathrm{N}(\mu_j, \sigma^2)$, $i = 1, \ldots, n_j$ and $j = 1, \ldots, m$. In total there are $n = \sum_{j=1}^{m} n_j$ samples. Further, let $S = \sum_{i=1}^{n_j}\sum_{j=1}^{m}(Y_{ij} - \bar{Y})^2$, where $\bar{Y} = \frac{1}{n}\sum_{i=1}\sum_{j=1}^{m} Y_{ij}$.

   (a) Define the sample group means to be $\bar{Y}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} Y_{ij}$. Add and subtract the sample group mean $\bar{Y}_j$ into the squared sum in S to show that

   $$\sum_{i=1}^{n_j}\sum_{j=1}^{m}(Y_{ij} - \bar{Y})^2 = \sum_{i=1}^{n_j}\sum_{j=1}^{m}(Y_{ij} - \bar{Y}_j)^2 + \sum_{j=1}^{m} n_j(\bar{Y}_j - \bar{Y})^2$$

   (b) What is the distribution of $\bar{Y}$ and $\bar{Y}_j$?

   (c) Assuming that $\mu_j = \mu$, for all $j = 1, \ldots, m$ and using your answer to (b), determine then the following distributions

      i. $\frac{1}{\sigma^2}\sum_{i=1}^{n_j}\sum_{j=1}^{m}(Y_{ij} - \mu)^2$

      ii. $\frac{n}{\sigma^2}(\bar{Y} - \mu)^2$

      iii. $\frac{1}{\sigma^2}\sum_{i=1}^{n_j}\sum_{j=1}^{m}(Y_{ij} - \bar{Y})^2$

      iv. $\frac{1}{\sigma^2}\sum_{j=1}^{m} n_j(\bar{Y}_j - \mu)^2$

      v. $\frac{1}{\sigma^2}\sum_{i=1}^{n_j}\sum_{j=1}^{m}(Y_{ij} - \bar{Y}_j)^2$

   Hint: Use the sum of squares decomposition with $\bar{Y}$ and $\bar{Y}_j$, and then use the properties of χ^2-distributions.