

SM-4331 Exercise 3

- Let $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$. Prove that the sample mean $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ is distributed according to $\bar{X} \sim N(\mu, \sigma^2/n)$. *Hint: Find the expectation and variance of \bar{X} , and use the linearity property of normal distributions.*
- Suppose that we plan to take a random sample of size n from a normal distribution with mean μ and standard deviation $\sigma = 2$.
 - Suppose $\mu = 4$ and $n = 20$.
 - What is the probability that the mean \bar{X} of the sample is greater than 5?
 - What is the probability that \bar{X} is smaller than 3?
 - What $P(|\bar{X} - \mu| \leq 1)$ in this case?
 - How large should n be in order that $P(|\bar{X} - \mu| \leq 0.5) \geq 0.95$ for every possible value of μ ?
 - 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} ?
- If Z is a random variable with a standard normal distribution, what is $P(Z^2 < 3.841)$?
 - Suppose that X_1 and X_2 are independent $N(0, 4)$ random variables. Compute $P(X_1^2 < 36.84 - X_2^2)$.
 - Suppose that $X_1, X_2, X_3 \stackrel{\text{iid}}{\sim} N(0, 1)$, while Y independently follows a χ_5^2 distribution. Compute $P(X_1^2 + X_2^2 < 7.236Y - X_3^2)$.
- Let $X_i, i = 1, 2, 3$ be independent with $N(i, i^2)$ distributions. For each of the following situations, use the X_i s to construct a statistic with the indicated distribution:
 - χ^2 -distribution with 3 degrees of freedom;
 - t -distribution with 2 degrees of freedom; and
 - F -distribution with 1 and 2 degrees of freedom.
- 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 $P(\mathbb{1}_{[k=i]}(k) = 1) = p_i$.

- What is $E[\mathbb{1}_{[k=i]}(k)]$ and $\text{Var}[\mathbb{1}_{[k=i]}(k)]$?
- Calculate $E[\mathbb{1}_{[k=i]}(k) \mathbb{1}_{[l=j]}(l)]$ when $k \neq l$.
- Argue that $E[\mathbb{1}_{[k=i]}(k) \mathbb{1}_{[l=j]}(l)] = 0$ when $k = l$.
- 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 $E(X_i)$.
- (f) Consider two random variables X_i and X_j defined as per (d). From your answers to (a), (b) and (c), calculate $E(X_i X_j)$.
- (g) Now calculate the covariance between X_i and X_j .
6. Let $\{X_1, \dots, X_n\}$ be a random sample from a $N(\mu, \sigma^2)$ 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 $\chi_{14}^2(\alpha/2)$ and $\chi_{14}^2(1-\alpha/2)$, i.e. the top and bottom $\alpha/2$ point of the χ_{14}^2 distribution where $P(Y < \chi_k^2(a)) = a$ when $Y \sim \chi_k^2$.
- (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 $N(\mu_j, \sigma^2)$, $i = 1, \dots, n_j$ and $j = 1, \dots, m$. 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 $\bar{Y} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m Y_{ij}$.
- (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 $j = 1, \dots, m$ and using your answer to (b), determine then the following distributions
- $\frac{1}{\sigma^2} \sum_{i=1}^{n_j} \sum_{j=1}^m (Y_{ij} - \mu)^2$
 - $\frac{n}{\sigma^2} (\bar{Y} - \mu)^2$
 - $\frac{1}{\sigma^2} \sum_{i=1}^{n_j} \sum_{j=1}^m (Y_{ij} - \bar{Y})^2$
 - $\frac{1}{\sigma^2} \sum_{j=1}^m n_j (\bar{Y}_j - \mu)^2$
 - $\frac{1}{\sigma^2} \sum_{i=1}^{n_j} \sum_{j=1}^m (Y_{ij} - \bar{Y}_j)^2$

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