1. (50 pts.) True or False

   (a) When choosing from a set of possible point estimators, a reasonable approach is to choose the estimator with the smallest mean squared error.

   (b) We call $\hat{\Theta}$ an unbiased estimator of θ if $\text{Var}(\hat{\Theta}) = \theta$.

   (c) If $\hat{\Theta}$ is unbiased estimator of θ, then the mean squared error of $\hat{\Theta}$ is identical to the variance of $\hat{\Theta}$.

   (d) The mean squared error of $\hat{\Theta}$ is defined by $\text{E}(\hat{\Theta} - \mu_{\hat{\Theta}}^2)$, where $\mu_{\hat{\Theta}}$ is the expected value of $\hat{\Theta}$.

   (e) The standard error of a statistic is the variance of the statistic.

   (f) If $X_1, \ldots, X_n$ are independent identically distributed (i.i.d.) random variables, then the sample mean $\overline{X}_n = \sum_{i=1}^{n} X_i/n$ follows a normal distribution for any n.

   (g) All estimators are statistics, but some statistics are not estimators.

   (h) Applying central limit theorem, normality of the random samples needs to be assumed.

   (i) Suppose that T follows a t distribution with degrees of freedom 10 and Z follows a standard normal distribution. $\text{P}(T > 2) \geq \text{P}(Z > 2)$.

   (j) If X_1, X_2, ..., X_n follow a normal distribution with mean μ and variance σ^2, it is also called a random sample.

   (k) Suppose that we obtain a 95% confidence of the mean μ to be (65.5, 68.4). We know that the unknown mean satisfies $65.5 \leq \mu \leq 68.4$.

   (l) Suppose that we obtain a 95% confidence of the mean μ to be (65.5, 68.4). We know that $\text{P}(65.5 \leq \mu \leq 68.4) = 0.95$.

   (m) Suppose that we obtain a 95% confidence of the proportion p to be (0.5, 0.6). We know that if we repeat the experiment many times, in the long run only 95% of the intervals would include the unknown proportion p, but we don't know whether (0.5, 0.6) includes p or not.

(n) The life in hours of a battery is known to be approximately normally distributed, with standard deviation $\sigma = 2$ hours. A claim is made that the battery life exceeds 50 hours. Applying the testing-hypothesis approach to make a conclusion, reasonable statistical hypotheses are
H_0: $\overline{X} = 50$, H_1: $\overline{X} > 50$, where $\overline{X}$ is the sample mean.

(o) Suppose that voters, choosing between a Republican and a Democratic candidate, give the Republican $p \times 100\%$ of the votes. We take a random sample of all voters. Consider the statistical hypotheses:
H_0: $p \leq 0.5$, H_1:$p > 0.5$. Type I error means that we claim that the Republican will receive no more than half of the votes given that Republican will lose.

(p) Suppose that voters, choosing between a Republican and a Democratic candidate, give the Republican $p \times 100\%$ of the votes. We take a random sample of all voters. Consider the statistical hypotheses:
H_0: $p \leq 0.5$, H_1: $p > 0.5$. Suppose that we take a sample of size n and make a decision that we reject the null hypothesis. That means that the testing-hypothesis approach has helped us to see the true: "Republican will win".

(q) Consider the two-sided test for the ratio of two variances:
H_0: $\sigma_1^2/\sigma_2^2 = 1$, H_1: $\sigma_1^2/\sigma_2^2 \neq 1$. If the testing-hypothesis results show that we fail to reject H_0, then it is true that $\sigma_1^2 = \sigma_2^2$.

(r) Consider a one-sided test for the unknown parameter θ:
H_0: $\theta = 1$, H_1: $\theta > 1$. Suppose that the sampling distribution of $\hat{\theta}$ is normal. If the $\hat{\theta}$, the estimate of θ, is 0.9, we will always fail to reject H_0 if the significance level α is less than 0.5.

(s) The usual 95% confidence of the variance σ^2 is
$(S^2 - c\ \mathrm{se}(S^2), S^2 - c\ \mathrm{se}(S^2))$, where S^2 is the sample variance, $\mathrm{se}(S^2)$ is the standard error of S^2, and c is some constant.

(t) Let θ denotes the population mean μ or the difference of two means $\mu_1 - \mu_2$. If the standard error of the corresponding estimator $\hat{\Theta}$ is known, then the usual 95% confidence of the population parameter θ is
$(\hat{\theta} - c\ \mathrm{se}(\hat{\Theta}), \hat{\theta} + c\ \mathrm{se}(\hat{\Theta}))$, where $\mathrm{se}(\hat{\Theta})$ is the standard error of $\hat{\Theta}$, $\hat{\theta}$ is the point estimate, and c is some constant.

2. (8 pts.) Consider a bag containing two white and four black balls, and two balls are drawn at random without replacement from the bag. Let X and Y be random variables representing the results of the 1st and 2nd drawings, respectively. Let 0 correspond to draw a black ball and 1 correspond to draw a white ball. Find $f_X(x), f_Y(y), f_{X,Y}(x,y)$ and $f_{Y|X}(y|x)$.

3. (8 pts.) (Let the three-dimensional random variable (X_1, X_2, X_3) have the density function $f_{X_1,X_2,X_3}(x_1, x_2, x_3) = cx_1x_2x_3$ for $0 < x_1, x_2, x_3 < 1$. Find c, the expected values $\mathrm{E}(3X_1 + 2X_2 + 6X_3), \mathrm{E}(X_1X_2X_3)$, and $\mathrm{E}(X_1X_2)$.

4. (10 pts.) Graph the pairs of points below. Find the least squares regression line and the standard error of estimate. How many values of Y are within 1 standard deviation or standard error from the regression line?

X	9	6	8	5
Y	5	3	5	3

5. (10 pts.) The sample variance is given by the formula:

$$S_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2,$$

where X_i denotes the i^{th} observation in the random sample and $\bar{X}$ denotes the sample mean.

   (a) Show that S_X^2 is a biased estimator of the population variance σ^2,

   (b) Construct an unbiased estimator of σ^2, and then show that the unbiased estimator is consistent.

6. (6 pts.) State the Central Limit Theorem.

7. (8 pts) Tom and Joe like to throw darts. Tom throws 100 times and hits the target 54 times; Joe throws 100 times and hits the target 49 times. Find a 95 percent confidence interval for p_1, p_2 where p_1 and p_2 represent the true proportions of hits in Tom's and Joe's tosses, respectively. Note that $z_{0.05} = 1.645, z_{0.025} = 1.96$, where z_α is the value such that $\mathrm{P}(Z > z_\alpha) = \alpha$ and Z follows a standard normal distribution.