Homework Set 1

All questions are from the end-of-chapter exercises. The question numbers refer to those in the book. The due time is **March 13$^{\text{th}}$**.

2.6 The table below gives the joint probability distribution between employment status and college graduation among those either employed or looking for work (unemployed) in the working age U.S. population for 2008.

	Unemployed (Y=0)	Employed (Y=1)	Total
Non-college grads (X=0)	0.037	0.622	0.659
College grads (X=1)	0.009	0.332	0.341
Total	0.046	0.954	1.000

a. Compute $E(Y)$.

b. The unemployment rate is the fraction of the labor force that is unemployed. Show that the unemployment rate is given by $1 - E(Y)$.

c. Calculate the unemployment rate for (i) college graduates and (ii) non-college graduates

d. A randomly selected member of this population reports being unemployed. What is the probability that this worker is a college graduate? A non-college graduates?

e. Are educational achievement and employment status independent? Explain.

2.10 Compute the following probabilities:

a. If Y is distributed $N(1, 4)$, find $\Pr(Y \leq 3)$.

b. If Y is distributed $N(3, 9)$, find $\Pr(Y > 0)$.

c. If Y is distributed $N(50, 25)$, find $\Pr(40 \leq Y \leq 52)$.

d. If Y is distributed $N(5, 2)$, find $\Pr(6 \leq Y \leq 8)$.

2.13 X is a Bernoulli random variable with $\Pr(X = 1) = 0.99$, Y is distributed $N(0, 1)$, W is distributed $N(0, 100)$, and X, Y, and W are independent. Let $S = XY + (1 - X)W$. (That is, $S = Y$ when $X = 1$, and $S = W$ when $X = 0$)

a. Show that $E(Y^2) = 1$ and $E(W^2) = 100$

b. show that $E(Y^3) = 0$ and $E(W^3) = 0$. (Hint: What is the skewness for a symmetric distribution?)

c. Show that $E(Y^4) = 3$ and $E(W^4) = 3 \times 100^2$. (Hint: Use the fact that the kurtosis is 3 for a normal distribution.)

d. Derive $E(S), E(S^2), E(S^3)$, and $E(S^4)$. (Hint: Use the law of iterated expectations conditioning on $X = 0$ and $X = 1$)

e. Derive the skewness and kurtosis for S.

2.23 This exercise provides an example of a pair of random variables X and Y for which the conditional mean of Y given X depends on X but $\text{corr}(X, Y) = 0$.

Let X and Z be two independently distributed standard normal random variables, and let $Y = X^2 + Z$.

a. Show that $E(Y|X) = X^2$

b. Show that $\mu_Y = 1$.

c. Show that $E(XY) = 0$. (Hint: Use the fact that the odd moments of a standard normal random variable are all zero.)

d. Show that $\text{cov}(X, Y) = 0$ and thus $\text{corr}(X, Y) = 0$

2.26 Suppose that $Y_1, Y_2, \ldots, Y_n$ are random variables with a common mean μ_Y, a common variance σ_Y^2, and the same correlation ρ (so that the correlation between Y_i and Y_j is equal to ρ for all pairs i and j , where $i \neq j$.)

a. Show that $\text{cov}(Y_i, Y_j) = \rho\sigma_Y^2$ for $i \neq j$.

b. Suppose that $n = 2$. Show that $E(\overline{Y}) = \mu_Y$ and $\text{var}(\overline{Y}) = \frac{1}{2}\sigma_Y^2 + \frac{1}{2}\rho\sigma_Y^2$.

c. For $n \geq 2$, show that $E(\overline{Y}) = \mu_Y$ and $\text{var}(\overline{Y}) = \sigma_Y^2/n + [(n-1)/n]\rho\sigma_Y^2$.

d. When n is very large, show that $\text{var}(\overline{Y}) \approx \rho\sigma_Y^2$.

3.3 In a survey of 400 likely voters,215 responded that they would vote for the incumbent and 185 responded that they would vote for the challenger. Let p denote the fraction of all likely voters who preferred to incumbent at the time of the survey, and let $\hat{p}$ be the fraction of survey respondents who preferred the incumbent.

a. Use the survey results to estimate p.

b. Use the estimator of the variance of $\hat{p}$, $\hat{p}(1-\hat{p})/n$, to calculate the standard error of your estimator.

c. What is the p-value for the test $H_0 : p = 0.5$ v.s. $H_1 : p \neq 0.5$?

d. What is the p-value for the test $H_0 : p = 0.5$ v.s. $H_1 : p > 0.5$?

e. Why do the results from (**c**) and (**d**) differ?

f. Did the survey contain statistically significant evidence that the incumbent was ahead of the challenger at the time of the survey? Explain.

3.9 Suppose that a lightbulb manufacturing plant produces bulbs with a mean life of 2000 hours and a standard deviation of 200 hours. An inventor claims to have developed an improved process that produces bulbs with a longer mean life and the same standard deviation. The plant manager randomly selects 100 bulbs produced by the process. She says that she will believe the inventor's claim if the sample mean life of the bulbs is greater than 2100 hours; otherwise, she will conclude that the new process is no better than the old process. Let μ denote the mean of the new process. Consider the null and alternative hypothesis $H_0 : \mu = 2000$ v.s. $H_1 : \mu > 2000$.

a. What is the size of the plant manager's testing procedure?

b. Suppose the new process is in fact better and has a mean bulb life of 2150 hours. What is the power of the plant manager's testing procedure?

c. What testing procedure should the plant manager use if she wants the size of her test to be 5%?

3.11 Consider the estimator $\tilde{Y}$, defined in Equation (3.1). Show that (a) $E(\tilde{Y}) = \mu_Y$ and (b) $\text{var}(\tilde{Y}) = 1.25\sigma_Y^2/n$.

Equation (3.1) is

$$\tilde{Y} = \frac{1}{n}(\frac{1}{2}Y_1 + \frac{3}{2}Y_2 + \frac{1}{2}Y_3 + \frac{3}{2}Y_4 + \cdots + \frac{1}{2}Y_{n-1} + \frac{3}{2}Y_n)$$

where n is assumed to be even for convenience.