ECON 307: Spring 2016
Assignment 1: Statistics Review
Questions 1 through 6: Due January 26th

1. SW, Chapter 2, Exercise 2.1: Let \( Y \) denote the number of "heads" that occur when two coins are tossed.

(a) Derive the probability distribution of \( Y \).
(b) Derive the cumulative probability distribution of \( Y \).
(c) Derive the mean and variance of \( Y \).

2. SW, Chapter 2, Exercise 2.6: The following table gives the joint probability distribution between employment status and college graduation among those either employed or looking for work (unemployed) in the working US population for 2008.

<table>
<thead>
<tr>
<th></th>
<th>Unemployed (( Y = 0 ))</th>
<th>Employed (( Y = 1 ))</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-college grads (( X = 0 ))</td>
<td>0.037</td>
<td>0.622</td>
<td>0.659</td>
</tr>
<tr>
<td>College grads (( X = 1 ))</td>
<td>0.009</td>
<td>0.332</td>
<td>0.341</td>
</tr>
<tr>
<td>Total</td>
<td>0.046</td>
<td>0.954</td>
<td>1.000</td>
</tr>
</tbody>
</table>

(a) Computer \( E(Y) \)
(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 \( E(Y|X = 1) \) and \( E(Y|X = 0) \).
(d) Calculate the unemployment rate (i) college graduates and (ii) non-college graduates.
(e) A randomly selected member of this population reports being unemployed. What is the probability that this worker is a college graduate? A non-college graduate?
(f) Are educational achievement and employment status independent? Explain.

3. SW, Chapter 2, Exercise, 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 \( N(5, 2) \), find \( Pr(6 \leq Y \leq 8) \).
4. SW, Chapter 2, Exercise 2.18: In any year, the weather can inflict storm damage to a home. From year to year, the damage is random. Let $Y$ denote the dollar value of damage in any given year. Suppose that in 95% of the years, $Y = 0$, but in 5% of the years, $Y = 20,000$.

(a) What are the mean and standard deviation of the damage in any year?

(b) Consider an insurance pool of 100 people whose homes are sufficiently dispersed so that, in any year, the damage to different homes can be viewed as independently distributed random variables. Let $\bar{Y}$ denote the average damage to these 100 homes in a year. (i) What is the expected value of the average damage $\bar{Y}$? (ii) What is the probability that $\bar{Y}$ exceeds $2000$?

5. SW, Chapter 3, Exercise 3.1: In a population $\mu_Y = 100$ and $\sigma_Y^2 = 43$. Use the central limit theorem to answer the following questions:

(a) In a random sample of size $n = 100$, find $\Pr(\bar{Y} \leq 101)$.

(b) In a random sample of size $n = 165$, find $\Pr(\bar{Y} \geq 98)$.

(c) In a random sample of size $n = 64$, find $\Pr(101 \leq \bar{Y} \leq 103)$.

6. SW, Chapter 3, Exercise 3.6: Let $Y_1, \ldots, Y_n$ be $i.i.d.$ draws from a distribution with mean $\mu$. A test of $H_0 : \mu = 5$ vs. $H_0 : \mu \neq 5$ using the usual t-statistic yields a $p$-value of 0.03.

(a) Does the 95% confidence interval contain $\mu = 5$? Explain.

(b) Can you determine if $\mu = 6$ is contained in the 95% confidence interval? Explain.

7. SW, Chapter 3, Exercise 3.7: In a given population, 11% of the likely voters are African American. A survey using a simple random sample of 600 landline telephone numbers finds 8% African Americans. Is there evidence that the survey is biased? Explain.

8. SW, Chapter 3, Empirical Exercise E3.1: On the text website, find the data file CPS92_08. It contains data on full-time, full-year workers, age 25-34, with a high school diploma or BA/BS as their highest degree. A detailed description is given in CPS92_08_Description. Answer parts (a) through (g) in the exercise. Write your answers on your do-file and turn in your do-file and log file.


(b) In 2008, the value of the Consumer Price Index (CPI) was 215.2. In 1992, the value of the CPI was 140.3. Repeat (a) but use AHE measured in real 2008 dollars ($2008); that is, adjust the 1992 data for the price inflation that occurred between 1992 and 2008.
(c) If you were interested in the change in workers’ purchasing power from 1992 to 2008, would you use the results from (a) or from (b)? Explain.

(d) Use the 2008 data to construct a 95% confidence interval for the mean of AHE for high school graduates. Construct a 95% confidence interval for the mean of AHE for workers with a college degree. Construct a 95% confidence interval for the difference between the two means.

(e) Repeat (d) using the 1992 data expressed in $2008.

(f) Did real wages of high school graduates increase from 1992 to 2008? Explain. Did real wages of college graduates increase? Did the gap between earnings of college and high school graduates increase? explain, using appropriate estimates, confidence intervals, and test statistics.

(g) Table 3.1 presents information on the gender gap for college graduates. Prepare a similar table for high school graduates using the 1992 and 2008 data. Are there any notable differences between the results for high school and college graduates?