Solution: Problem Set #1

**Exercise 1.** Think of an example involving five possible quantitative outcomes of a discrete random variable and attach a probability to each one of these outcomes. Display the outcomes, probability distribution, and cumulative probability distribution in a table. Sketch both the probability distribution and the cumulative probability distribution. Derive the mean and variance of this random variable.

Answers will vary by student. The generated table should be similar to Table 2.1 in the text, and figures should resemble Figures 2.1 and 2.2 in the text.

**Exercise 2.** Suppose $X$ is a Bernoulli random variable with $P(X=1)=p$.

a. Show $E(X^3) = p$

$E(X^3) = 0^3 * (1-p) + 1^3 * p = p$

b. Show $E(X^k) = p$ for $k > 0$

$E(X^k) = 0^k * (1-p) + 1^k * p = p$

c. Suppose that $p=0.3$. Compute the mean, variance, skewness, and kurtosis of $X$.

(Hint: You might find it helpful to use the formulas given below)

$E(x - \mu)^3 = E(X^3) - 3[E(X^2)][E(X)] + 2[E(X)]^3$

$E(x - \mu)^4 = E(X^4) - 4[E(X)][E(X^3)] + 6[E(X)][E(X^2)] - 3[E(X)]^4$

$E(X) = 0.3$

$Var(X) = E(X^2) - [E(X)]^2 = 0.3 - 0.3^2 = 0.21$

Thus, $\sigma_X = \sqrt{Var(X)} = \sqrt{0.21} = 0.45826$

To compute the skewness, using the formula

$E(x - \mu)^3 = E(X^3) - 3[E(X^2)][E(X)] + 2[E(X)]^3 = 0.3 - 3*0.3*0.3 + 2*0.3^3 = 0.084$

$Skewness=0.084/0.45826^3 = 0.87286$

$E(x - \mu)^4 = E(X^4) - 4[E(X)][E(X^3)] + 6[E(X)][E(X^2)] - 3[E(X)]^4$

$= 0.3 - 4*0.3*0.3 + 6*0.3^2*0.3 - 3*0.3^4 = 0.0777$

$Kurtosis=0.0777/0.45826^4 = 0.7619$

**Exercise 3.** $Y$ is a random variable with $\mu_Y = 0$, $\sigma_Y = 1$, skewness = 0, kurtosis = 100. Sketch a hypothetical probability distribution of Y. Explain why n random variables drawn from this distribution might have some large outliers.

The probability distribution looks like figure 2.3b from the handout, but with more mass concentrated in the tails. Because the distribution is symmetric around $\mu_Y = 0$, $Pr(Y > c) = Pr(Y < -c)$, because there is substantial mass in the tails of the distribution, $Pr(Y > c)$ remains significantly greater than zero even for large values of c.

**Exercise 4.** In September, Seattle’s daily high temperature has a mean of $70^\circ F$ and a standard deviation of $7^\circ F$. What is the mean, standard deviation, and variance in $^\circ C$?

Let $X$ denote temperature in $^\circ F$ and $Y$ denote temperature in $^\circ C$. Recall that $Y = 0$ when $X=32$ and $Y=100$ when $X=212$; this implies that $Y = (100/180) * (X - 32) = -17.78 + (5/9)X$
\[ \mu_X = 70 \, ^\circ F \Rightarrow \mu_Y = -17.78 + (5/9) \times 70 = 21.109 \, ^\circ C \]
\[ \sigma_X = 7^\circ F \Rightarrow \sigma_Y = (5/9) \times 7 = \frac{35}{9} = 3.89^\circ C \]

**Exercise 5.** Stock & Watson (2003) 2.3

(a) Compute \( E(Y) \)

The table shows that \( \Pr(X = 0, Y = 0) = 0.045, \Pr(X = 0, Y = 1) = 0.709, \Pr(X = 1, Y = 0) = 0.005, \Pr(X = 1, Y = 1) = 0.241 \).
\[
\begin{align*}
\Pr(X = 0) &= 0.754, \Pr(X = 1) = 0.246, \Pr(Y = 0) = 0.05, \Pr(Y = 1) = 0.95 \\
E(Y) &= 0 \times 0.05 + 1 \times 0.95 = 0.95
\end{align*}
\]

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

Unemployment Rate = \#Unemployed / \#LaborForce = \Pr(Y = 0) = 0.05 = 1 - E(Y)

(c) Calculate \( E(Y \mid X = 1) \) and \( E(Y \mid X = 0) \).

\[
\begin{align*}
\Pr(Y \mid X = 1) &= \frac{\Pr(Y = 0, X = 1)}{\Pr(X = 1)} = \frac{0.005}{0.246} = 0.020325 \\
\Pr(Y \mid X = 1) &= \frac{\Pr(Y = 1, X = 1)}{\Pr(X = 1)} = \frac{0.241}{0.246} = 0.97967
\end{align*}
\]
\[
E(Y \mid X = 1) = 0 \times 0.020325 + 1 \times 0.97967 = 0.97967
\]
\[
\begin{align*}
\Pr(Y \mid X = 0) &= \frac{\Pr(Y = 0, X = 0)}{\Pr(X = 0)} = \frac{0.045}{0.754} = 0.059682 \\
\Pr(Y \mid X = 0) &= 1 - \Pr(Y = 0 \mid X = 0) = 1 - 0.059682 = 0.94032 \\
E(Y \mid X = 0) &= 0 \times 0.059682 + 1 \times 0.94032 = 0.94032
\end{align*}
\]

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

The unemployment rate for college graduates is \( \Pr(Y = 0 \mid X = 1) = 0.02 \)

The unemployment rate for non-college graduates is \( \Pr(Y = 0 \mid X = 0) = 0.06 \)

(e) A randomly selected number for this population reports being unemployed. What is the probability that the work is a college graduate? A non-college graduate?

The probability that the work is a college graduate
\[
\begin{align*}
=\Pr(X = 1 \mid Y = 0) &= \Pr(X = 1, Y = 0) / \Pr(Y = 0) = 0.005 / 0.05 = 0.1 \\
\text{The probability that the work is a non-college graduate} \\
=\Pr(X = 0 \mid Y = 0) &= \Pr(X = 0, Y = 0) / \Pr(Y = 0) = 0.045 / 0.05 = 0.9
\end{align*}
\]

(f) Are educational achievement and employment status independent? Explain.

Educational achievement and employment status are not independent because they do not satisfy that, for all values of x and y, \( \Pr(Y = y \mid X = x) = \Pr(Y = y) \), for example, \( \Pr(Y = 1 \mid X = 1) \neq \Pr(Y = 1) \)

**Exercise 6.** “Any member of political party A who decides to quit his party and join party B will raise the average IQ score in both parties.” Is it possible? Explain.
It is possible if the person who leaves party A in order to join party B has an IQ which is below the average of party A and above the average of party B.