**Week #9: Review of Chapters 4, 5, and 6**

**Problem 1.** Suppose weights of the checked baggage of airline passengers follow a nearly normal distribution with a mean of 45 pounds and a standard deviation of 3.2 pounds. Most airlines charge a fee for baggage that weights in excess of 50 pounds. What percent of airline passengers incur this fee?

\[ X = \text{weight of checked baggage} \]

\[ X \sim N(\mu=45, \sigma=3.2) \]

1. **Picture**
   
   \[ P(X > 50) \]

2. **Z-score**
   
   \[ Z = \frac{\text{obs-mean}}{\text{std dev}} \]

   \[ Z = \frac{50 - 45}{3.2} = \frac{5}{3.2} = 1.56 \]

3. **Cumulative Probability**
   
   CP: \( P(X < 50) = P(Z < 1.56) = 0.9406 \)

4. **Answer**
   
   \[ P(X > 50) = P(Z > 1.56) = 1 - P(Z < 1.56) = 1 - 0.9406 = 0.0594 \]

   \[ 5.94\% \]
Problem 2. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with a mean of 55 inches and a standard deviation of 6 inches. Calculate the following using the Empirical Rule:

a. Draw the appropriate Empirical Rule diagram for this scenario.
b. What is the probability that a randomly selected 10 year old is 55 inches or taller?
c. What is the probability that a randomly selected 10 year old is 49 inches or taller?
d. What is the probability that a randomly selected 10 year old is 43 inches or shorter?
e. What is the probability that a randomly selected 10 year old is between 43 and 61 inches tall?
Problem 2. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with a mean of 55 inches and a standard deviation of 6 inches. Calculate the following using the Empirical Rule:

a. Draw the appropriate Empirical Rule diagram for this scenario.
b. What is the probability that a randomly selected 10 year old is 55 inches or taller?
c. What is the probability that a randomly selected 10 year old is 49 inches or taller?
d. What is the probability that a randomly selected 10 year old is 43 inches or shorter?
e. What is the probability that a randomly selected 10 year old is between 43 and 61 inches tall?

b. \( P(X > 55) = 0.34 + 0.135 + 0.0235 + 0.0015 = 0.50 \)

c. \( P(X > 49) = 0.34 + 0.34 + 0.135 + 0.0235 + 0.0015 = 0.84 \)

d. \( P(X < 43) = 0.0015 + 0.0235 = 0.025 \)

e. \( P(43 < X < 61) = 0.135 + 0.34 + 0.34 = 0.815 \)
Problem 3. Let’s continue looking at the scenario presented in question 2. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with a mean of 55 inches and a standard deviation of 6 inches. Calculate the following using the z-score method:

a. What is the probability that a randomly selected 10 year old is shorter than 48 inches?

\[ P(X < 48) = P(Z < \frac{48 - 55}{6}) = P(Z < -1.17) = 0.1210 \]

b. What is the probability that a randomly selected 10 year old is taller than 58 inches?

\[ Z = \frac{58 - 55}{6} = \frac{3}{6} = 0.5 
\]

c. What is the probability that a randomly selected 10 year old is between 60 and 65 inches tall?

\[ P(60 < X < 65) = P(1.67 < Z < 1.17) \]

\[ P(Z < 1.17) - P(Z < 1.67) = 0.1210 - 0.0475 = 0.0735 \]

d. If the tallest 10% of the class is considered "very tall," what is the height cutoff for being very tall?

\[ P(X > \text{cutoff}) = 0.10 \]

\[ Z = \frac{X - \mu}{\sigma} = \frac{X - 55}{6} \]

\[ Z = 1.28 \]

\[ X = 55 + 1.28(6) = 65.68 \text{ inches} \]

e. The height requirement for *Batman the Ride* at Six Flags Magic Mountain is 54 inches. What percent of 10 year olds cannot go on this ride?

\[ P(X < 54) = P(Z < \frac{54 - 55}{6}) = P(Z < -0.17) = 0.4576 \]

f. Suppose there are four 10 year olds. What is the chance that at least two of them will be able to ride *Batman the Ride*?

\[ \text{At least 2 out of 4} \]

\[ \binom{4}{2} \times (0.1210)^2 \times (0.8790)^2 + \binom{4}{3} \times (0.1210)^3 \times (0.8790) + (0.1210)^4 = 0.8922 \]
Problem 3. Let’s continue looking at the scenario presented in question 2. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with a mean of 55 inches and a standard deviation of 6 inches. Calculate the following using the z-score method:

a. What is the probability that a randomly selected 10 year old is shorter than 48 inches?

b. What is the probability that a randomly selected 10 year old is taller than 58 inches?

c. What is the probability that a randomly selected 10 year old is between 60 and 65 inches tall?

d. If the tallest 10% of the class is considered “very tall,” what is the height cutoff for being very tall?

e. The height requirement for *Batman the Ride* at Six Flags Magic Mountain is 54 inches. What percent of 10 year olds cannot go on this ride?

f. Suppose there are four 10 year olds. What is the chance that at least two of them will be able to ride *Batman the Ride*?
Problem 3. Let’s continue looking at the scenario presented in question 2. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with a mean of 55 inches and a standard deviation of 6 inches. Calculate the following using the z-score method:

a. What is the probability that a randomly selected 10 year old is shorter than 48 inches?
b. What is the probability that a randomly selected 10 year old is taller than 58 inches?
c. What is the probability that a randomly selected 10 year old is between 60 and 65 inches tall?
d. If the tallest 10% of the class is considered "very tall," what is the height cutoff for being very tall?
e. The height requirement for *Batman the Ride* at Six Flags Magic Mountain is 54 inches. What percent of 10 year olds cannot go on this ride?
f. Suppose there are four 10 year olds. What is the chance that at least two of them will be able to ride *Batman the Ride*?

\[ Z_A = \frac{60 - 55}{6} = \frac{5}{6} = 0.83 \]

\[ Z_B = \frac{65 - 55}{6} = \frac{10}{6} = 1.67 \]

\[ \text{CP}_A: P(X < 60) = P(Z < 0.83) = 0.7967 \]

\[ \text{CP}_B: P(X < 65) = P(Z < 1.67) = 0.9525 \]

\[ \text{WTK: } P(60 < X < 65) = P(0.83 < Z < 1.67) = P(Z < 1.67) - P(Z < 0.83) = 0.9525 - 0.7967 = 0.1558 \]
Problem 3. Let’s continue looking at the scenario presented in question 2. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with a mean of 55 inches and a standard deviation of 6 inches. Calculate the following using the z-score method:

a. What is the probability that a randomly selected 10 year old is shorter than 48 inches?
b. What is the probability that a randomly selected 10 year old is taller than 58 inches?
c. What is the probability that a randomly selected 10 year old is between 60 and 65 inches tall?
d. If the tallest 10% of the class is considered ‘very tall,’ what is the height cutoff for being very tall?
e. The height requirement for Batman the Ride at Six Flags Magic Mountain is 54 inches. What percent of 10 year olds cannot go on this ride?
f. Suppose there are four 10 year olds. What is the chance that at least two of them will be able to ride Batman the Ride?

Looking for
\[ P(Z < 1.28) = 0.90 \]

\[ Z = \frac{\text{obs} - \text{mean}}{\text{std dev}} \]

\[ 1.28 = \frac{\text{obs} - 55}{6} \]

\[ \text{obs} = (1.28)(6) + 55 \]

\[ = 7.68 + 55 \]

\[ = 62.68 \text{ inches} \]
Problem 3. Let’s continue looking at the scenario presented in question 2. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with a mean of 55 inches and a standard deviation of 6 inches. Calculate the following using the z-score method:

a. What is the probability that a randomly selected 10 year old is shorter than 48 inches?

b. What is the probability that a randomly selected 10 year old is taller than 58 inches?

c. What is the probability that a randomly selected 10 year old is between 60 and 65 inches tall?

d. If the tallest 10% of the class is considered "very tall," what is the height cutoff for being very tall?

e. The height requirement for Batman the Ride at Six Flags Magic Mountain is 54 inches. What percent of 10 year olds cannot go on this ride?

f. Suppose there are four 10 year olds. What is the chance that at least two of them will be able to ride Batman the Ride?

\[ Z = \frac{54 - 55}{6} = \frac{-1}{6} = -0.17 \]

CP: \( P(X < 54) = P(Z < -0.17) = 0.4325 \)

43.25% of 10 year olds are too short.
Problem 3. Let’s continue looking at the scenario presented in question 2. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with a mean of 55 inches and a standard deviation of 6 inches. Calculate the following using the z-score method:

a. What is the probability that a randomly selected 10 year old is shorter than 48 inches?

b. What is the probability that a randomly selected 10 year old is taller than 58 inches?

c. What is the probability that a randomly selected 10 year old is between 60 and 65 inches tall?

d. If the tallest 10% of the class is considered "very tall," what is the height cutoff for being very tall?

e. The height requirement for Batman the Ride at Six Flags Magic Mountain is 54 inches. What percent of 10 year olds cannot go on this ride?

f. Suppose there are four 10 year olds. What is the chance that at least two of them will be able to ride Batman the Ride?

\[ f. \ P(\text{too short}) = 0.4325 \]
\[ P(\text{tall enough}) = 1 - 0.4325 = 0.5675 \]
\[ X = \# \ of \ kids \ who \ are \ tall \ enough \]
\[ X \sim \text{Binomial} \ (n = 4, \ p = 0.5675) \]
\[ P(\text{at least 2 are tall enough}) = P(X \geq 2) \]
\[ P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) \]
\[ P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \]

- \( k = 2 \)  
- \( k = 3 \)  
- \( k = 4 \)
Problem 3. Let’s continue looking at the scenario presented in question 2. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with a mean of 55 inches and a standard deviation of 6 inches. Calculate the following using the z-score method:

a. What is the probability that a randomly selected 10 year old is shorter than 48 inches?
b. What is the probability that a randomly selected 10 year old is taller than 58 inches?
c. What is the probability that a randomly selected 10 year old is between 60 and 65 inches tall?
d. If the tallest 10% of the class is considered "very tall,” what is the height cutoff for being very tall?
e. The height requirement for Batman the Ride at Six Flags Magic Mountain is 54 inches. What percent of 10 year olds cannot go on this ride?
f. Suppose there are four 10 year olds. What is the chance that at least two of them will be able to ride Batman the Ride?

\[
\begin{align*}
\text{f (continued)} \\
P(X=2) &= \binom{4}{2} (0.5675)^2 (0.4325)^2 \\
&= 0.3615 \\
P(X=3) &= \binom{4}{3} (0.5675)^3 (0.4325)^1 \\
&= 0.3162 \\
P(X=4) &= \binom{4}{4} (0.5675)^4 (0.4325)^0 \\
&= 0.1037 \\
P(X=2) &= 0.3615 + 0.3162 + 0.1037 \\
&= 0.7814
\end{align*}
\]
Problem 4. A student is interested in knowing what proportion of TAMU students are from out of state. She takes a random sample of 43 TAMU students and determines that 7 of them are from out of state.

a. Construct a 95% confidence interval for the proportion of out of state student in the entire TAMU student body.

b. Interpret your confidence interval from part a.

c. Based on the sample above, what sample size would be required so that the margin of error of a 95% confidence interval would be at most 0.08.

d. Assuming we hadn’t taken the sample listed above, what sample size would be required so that the margin of error of a 95% confidence interval would be at most 0.08.

\[ \hat{p} = \frac{7}{43} = 0.1628 \]
\[ n = 43 \]
\[ z^* = 1.960 \]

\[ 0.1628 \pm 1.960 \sqrt{\frac{0.1628(1-0.1628)}{43}} \]

\[ 0.1628 \pm 0.1103 \]

\[ 0.1628 - 0.1103 = 0.0525 \]
\[ 0.1628 + 0.1103 = 0.2731 \]

95% CI: (0.0525, 0.2731)

b. We are 95% confident that the true proportion of all TAMU students who are from out of state is between 0.0525 and 0.2731.
Problem 4. A student is interested in knowing what proportion of TAMU students are from out of state. She takes a random sample of 43 TAMU students and determines that 7 of them are from out of state.

a. Construct a 95% confidence interval for the proportion of out of state student in the entire TAMU student body.

b. Interpret your confidence interval from part a.

c. Based on the sample above, what sample size would be required so that the margin of error of a 95% confidence interval would be at most 0.08.

d. Assuming we hadn’t taken the sample listed above, what sample size would be required so that the margin of error of a 95% confidence interval would be at most 0.08.

\[ n \geq \frac{(z^*)^2 \hat{p}(1-\hat{p})}{m^2} \]

\[ z^* = 1.960 \]

\[ m = 0.08 \]

\[ \hat{p} = 0.1628 \]

\[ n \geq \frac{(1.960)^2(0.1628)(0.8372)}{(0.08)^2} \]

\[ n \geq 81.81 \]

\[ n = 82 \]
Problem 4. A student is interested in knowing what proportion of TAMU students are from out of state. She takes a random sample of 43 TAMU students and determines that 7 of them are from out of state.

a. Construct a 95% confidence interval for the proportion of out of state student in the entire TAMU student body.

b. Interpret your confidence interval from part a.

c. Based on the sample above, what sample size would be required so that the margin of error of a 95% confidence interval would be at most 0.08.

d. Assuming we hadn’t taken the sample listed above, what sample size would be required so that the margin of error of a 95% confidence interval would be at most 0.08.

\[ n \geq \frac{(z^*)^2 \bar{p}(1-\bar{p})}{m^2} \]

\[ n = \frac{(1.960)^2(0.50)(0.50)}{0.08^2} \]

\[ n = 150.0025 \]

\[ n = 151 \]
Problem 5. The ASPCA claims that 75% of all households own either a cat or a dog. Vennessa believes that in her community, less than 75% of households own a cat or a dog. She takes a random sample of 100 people who live in her county (population = 126,739) and finds out that 68 of the households own either a cat or a dog. Using this data, conduct the appropriate hypothesis test using a 0.05 level of significance.

a. What are the hypotheses?
b. What is the significance level?
c. What is the value of the test statistic?
d. What is the p-value?
e. What is the correct decision?
f. What is the appropriate conclusion/interpretation?

\[ H_0: \hat{p} = 0.75 \]
\[ H_A: \hat{p} < 0.75 \]

\[ \alpha = 0.05 \]

\[ TS = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \]

\[ \hat{p} = \frac{68}{100} = 0.68 \]
\[ p_0 = 0.75 \]
\[ n = 100 \]

\[ TS = \frac{0.68 - 0.75}{\sqrt{\frac{(0.75)(0.25)}{100}}} = \frac{-0.07}{0.04330} = -1.62 \]
Problem 5. The ASPCA claims that 75% of all households own either a cat or a dog. Venessa believes that in her community, less than 75% of households own a cat or a dog. She takes a random sample of 100 people who live in her county (population = 126,739) and finds out that 62 of the households own either a cat or a dog. Using this data, conduct the appropriate hypothesis test using a 0.05 level of significance.

a. What are the hypotheses?
b. What is the significance level?
c. What is the value of the test statistic?
d. What is the p-value?
e. What is the correct decision?
f. What is the appropriate conclusion/interpretation?

d. \[ p\text{-value} = P(\hat{p} \leq 0.68 \mid p = 0.75) \]
   \[ = P(z < -1.62) \]
   \[ = 0.0526 \]

E. 0.0526 > 0.05
p-value > \( \alpha \)

Fail to Reject \( H_0 \)
Problem 5. The ASPCA claims that 75% of all households own either a cat or a dog. Vennessa believes that in her community, less than 75% of households own a cat or a dog. She takes a random sample of 100 people who live in her county (population = 126,739) and finds out that 62 of the households own either a cat or a dog. Using this data, conduct the appropriate hypothesis test using a 0.05 level of significance.

a. What are the hypotheses?
b. What is the significance level?
c. What is the value of the test statistic?
d. What is the p-value?
e. What is the correct decision?
f. What is the appropriate conclusion/interpretation?

f. The data does not provide statistically significant evidence that the true proportion of people in Vennessa’s community who own a cat or a dog is less than 0.75.
Problem 6. The average cholesterol level in the general US population is 189 mg/dL. A researcher wants to see if the average cholesterol for men in the US is different from 189 mg/dL. She takes a sample of 81 American males and finds a sample mean of 194 mg/dL and a sample standard deviation of 10.4. Create a 96% confidence interval for the true average cholesterol level of the general US male population.

a. What is the 96% confidence interval?
b. What is the correct interpretation of the confidence interval?
c. Are the assumptions met? Explain.
d. Based on the confidence interval, what can you say about the researcher’s question?

\[ \bar{x} = 194 \]
\[ s = 10.4 \]
\[ n = 81 \]
\[ \bar{x} \pm z^* \left( \frac{s}{\sqrt{n}} \right) \]

\[ \bar{x} = 194 \]
\[ s = 10.4 \]
\[ n = 81 \]
\[ z^* = 2.054 \]

\[ \bar{x} \pm 2.054 \left( \frac{10.4}{\sqrt{81}} \right) \]

\[ 194 \pm 2.054 \left( \frac{10.4}{9} \right) \]

\[ 194 \pm 2.054 \times 1.1556 \]

\[ 194 \pm 2.389 \]

\[ 194 - 2.389 = 191.611 \]
\[ 194 + 2.389 = 197.389 \]

\[ 96\% \text{ C.I.} = (191.611, 197.389) \]

b. We are 96\% confident that the true average cholesterol level of all US men is between 190.804 mg/dL and 197.196 mg/dL.
Problem 6. The average cholesterol level in the general US population is 189 mg/dL. A researcher wants to see if the average cholesterol for men in the US is different from 189 mg/dL. She takes a sample of 81 American males and finds a sample mean of 194 mg/dL and a sample standard deviation of 10.4. Create a 96% confidence interval for the true average cholesterol level of the general US male population.

a. What is the 96% confidence interval?
b. What is the correct interpretation of the confidence interval?
c. Are the assumptions met? Explain.
d. Based on the confidence interval, what can you say about the researcher's question?

C. ① Independent
random ? unknown
n < 10% of pop’n? ✓

② pop’n Dist’n
we don’t know what the pop’n
dist’n looks like

-D n is large (81 ≥ 30), sampling
dist’n is approx normal

96% CI: (190.804, 197.196)

-D 189 isn’t in the interval, this
supports the idea that the pop’n
mean for males is different from 189
mg/dL
Problem 7. The average cholesterol level in the general US population is 189 mg/dL. A researcher wants to see if the average cholesterol for men in the US is different from 189 mg/dL. She takes a sample of 81 American males and finds a sample mean of 194 mg/dL and a sample standard deviation of 10.4. **Conduct a hypothesis test at the 0.01 significance level to test the researcher’s claim.**

a. What are the hypotheses?
b. What is the significance level?
c. What is the value of the test statistic?
d. What is the p-value?
e. What is the correct decision?
f. What is the appropriate conclusion/interpretation?
g. Are the assumptions met? Explain.

\[ a. H_0: \mu = 189 \]
\[ H_A: \mu \neq 189 \]

\[ b. \alpha = 0.01 \]

\[ c. T_S = \frac{\bar{X} - M_o}{S/\sqrt{n}} \]
\[ = \frac{194 - 189}{10.4/\sqrt{81}} \]
\[ = \frac{5}{1.1557} \]
\[ = 4.33 \]

\[ \text{null value: } M_o = 189 \]

\[ \bar{X} = 194 \]
\[ M_o = 189 \]
\[ S = 10.4 \]
\[ n = 81 \]
Problem 7. The average cholesterol level in the general US population is 189 mg/dL. A researcher wants to see if the average cholesterol for men in the US is different from 189 mg/dL. She takes a sample of 81 American males and finds a sample mean of 194 mg/dL and a sample standard deviation of 10.4. **Conduct a hypothesis test at the 0.01 significance level to test the researcher’s claim.**

a. What are the hypotheses?
b. What is the significance level?
c. What is the value of the test statistic?
d. What is the p-value?
e. What is the correct decision?
f. What is the appropriate conclusion/interpretation?
g. Are the assumptions met? Explain.

d. \( P(\mid z \mid > 4.33) = P(z < -4.33) + P(z > 4.33) \)

\[ = 2 \times P(z < -4.33) \]

\[ = 2 \times [<0.0002] \]

\[ < 0.0004 \]

p-value < 0.0004

e. p-value < 0.0004 < 0.01

\[ P-value < 0.01 \]

\[ P-value < \alpha \]

Reject \( H_0 \)
Problem 7. The average cholesterol level in the general US population is 189 mg/dL. A researcher wants to see if the average cholesterol for men in the US is different from 189 mg/dL. She takes a sample of 81 American males and finds a sample mean of 194 mg/dL and a sample standard deviation of 10.4. **Conduct a hypothesis test at the 0.01 significance level to test the researcher’s claim.**

a. What are the hypotheses?
b. What is the significance level?
c. What is the value of the test statistic?
d. What is the p-value?
e. What is the correct decision?
f. What is the appropriate conclusion/interpretation?
g. Are the assumptions met? Explain.

f. The data does provide statistically significant evidence that the true average cholesterol level of all US men is different from 189. Based on our sample, we believe it is greater than 189.
Problem 7. The average cholesterol level in the general US population is 189 mg/dL. A researcher wants to see if the average cholesterol for men in the US is different from 189 mg/dL. She takes a sample of 81 American males and finds a sample mean of 194 mg/dL and a sample standard deviation of 10.4. **Conduct a hypothesis test at the 0.01 significance level to test the researcher’s claim.**

a. What are the hypotheses?
b. What is the significance level?
c. What is the value of the test statistic?
d. What is the p-value?
e. What is the correct decision?
f. What is the appropriate conclusion/interpretation?
g. Are the assumptions met? Explain.

9. (i) Independent?
   - Random? Unknown
   - n < 10% of pop'n? ✓
   - all males in US

(ii) pop'n dist'n?
   - don't know what the pop'n dist'n looks like
   - Dn is large (n ≥ 30), sampling dist'n is approx normal