Chapter 7: Hypothesis Testing - Solutions

7.1 Introduction to Hypothesis Testing

The problem with applying the techniques learned in Chapter 5 is that typically, the population mean (\( \mu \)) and standard deviation (\( \sigma \)) are not known. In chapter 7, we will be looking at the situation when a simple random sample is taken from a large population with unknown mean and standard deviation. However, the sample mean and sample standard deviation can be calculated.

Setup:

i) A sample is taken from a population with unknown mean and standard deviation.

ii) A claim is made about the true population mean \( \mu \).

iii) The claim is tested using the method of hypothesis testing, to be explained below.

Example: Suppose that an automobile manufacturer advertises that its new hybrid car has a mean gas mileage of 50 miles per gallon. You take a simple random sample of \( n = 30 \) hybrid vehicles and test their gas mileage. You find that in this sample, the average is \( \bar{x} = 47 \) miles per gallon with a standard deviation of \( s = 5.5 \) miles per gallon. Does this indicate that the advertiser’s statement is true or false?

Step 1:
Rewrite the claim about the population mathematically. This is called the null hypothesis, denoted

\[ H_0 : \mu = 50 \]

Assume, to start, that the null hypothesis is true. In the end, you may reject it in favor of:

Step 2:
The alternative (complement) to the null hypothesis is called the alternative hypothesis, denoted

\[ H_a : \mu \neq 50 \]

It is not possible to determine with absolute certainty whether or not \( H_0 \) is true, unless the entire population is tested, which is rarely feasible.
Step 3: Determine (or use the given) probability threshold for making the decision to reject $H_0$ when it is true. This probability is called the level of significance, denoted $\alpha$, often chosen to be 0.05 (or, for example 0.01 or 0.10).

Step 4: Calculate the mean and standard deviation for the simple random sample.

Step 5: Calculate the $P$-value, which is the probability of obtaining a sample mean that is as extreme or more extreme than the one obtained from the data (details given in 7.2 and 7.3).

Step 6: If $P \leq \alpha$, make the decision to reject $H_0$. Otherwise, fail to reject $H_0$.

(Note: Failing to reject $H_0$ is not the same as accepting it – $H_0$ was assumed to be true from the start. Failing to reject $H_0$ simply means that there is not enough statistical evidence to reject it.)

This table shows all of the possible scenarios that can occur in hypothesis testing:

<table>
<thead>
<tr>
<th></th>
<th>$H_0$ is true</th>
<th>$H_0$ is false</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fail to reject</strong> $H_0$</td>
<td>Correct decision</td>
<td>Type II error</td>
</tr>
<tr>
<td><strong>Reject</strong> $H_0$</td>
<td>Type I error</td>
<td>Correct decision</td>
</tr>
</tbody>
</table>

In summary, the inputs are:

- $H_0$ : null hypothesis
- $H_a$ : alternative hypothesis
- $\alpha$ : significance level
- $\bar{x}$ : sample mean
- $s$ : sample standard deviation
- $P$ : $P$-value

The output is a single decision to reject $H_0$ or not, based on the relationship between $P$, which you calculate, and the $\alpha$, which is typically given or determined by you.
7.2 Hypothesis Testing for the Mean (Large Samples)

In this section, we describe the complete procedure of hypothesis testing when the sample size \( n \geq 30 \).

Example 1: Suppose that an automobile manufacturer advertises that its new hybrid car has a mean gas mileage of 50 miles per gallon. You take a simple random sample of \( n = 30 \) hybrid vehicles and test their gas mileage. You find that in this sample, the average is \( \bar{x} = 47 \) miles per gallon with a standard deviation of \( s = 5.5 \) miles per gallon. Is there enough evidence to support the advertised claim using \( \alpha = 0.05 \)?

Inputs:

1. \( H_0 : \mu = 50 \)
2. \( H_a : \mu \neq 50 \)
3. \( \alpha = 0.05 \)
4. \( \bar{x} = 47 \)
   \( s = 5.5 \)

Step 5:
Calculating the \( P \)-value if \( n \geq 30 \):

i) Find the \( z \)-score

\[
z = \frac{\bar{x} - \mu}{s/\sqrt{n}}
\]

ii) Find the cumulative area corresponding to the \( z \)-score.

iii) The \( P \)-value depends on the null hypothesis (draw a picture):

If \( H_0 : \mu = k \), \( P = 2 \cdot \text{cumulative area} \) “two-tailed test”
If \( H_0 : \mu \geq k \), \( P = \text{cumulative area} \) “left-tailed test”
If \( H_0 : \mu \leq k \), \( P = 1 - \text{cumulative area} \) “right-tailed test”

Step 6:
If \( P \leq \alpha \), make the decision to reject \( H_0 \). Otherwise, fail to reject \( H_0 \).
Example 1: (continued)

The z-score is

\[ z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{47 - 50}{5.5/\sqrt{30}} \approx -3.0 \]

The cumulative area for this z-score is 0.0013. This is a two-tailed test: \( P = 2(0.0013) = 0.0026 \).

Since \( P = 0.0026 < 0.05 = \alpha \), we make the decision to reject the null hypothesis. This means that there is not enough statistical evidence to support the advertiser’s claim of an average of 50 miles per gallon.

Example 2: A pit crew claims that its mean pit stop time (for 4 new tires and fuel) during an auto race is less than 13 seconds. A random sample of 32 pit stop times has a sample mean of 12.9 seconds and a standard deviation of 0.19 second. Is there enough evidence to support the claim of the pit crew at \( \alpha = 0.01 \)?

1. \( H_0 : \mu \geq 13 \)
2. \( H_a : \mu < 13 \)
3. \( \alpha = 0.01 \)
4. \( \bar{x} = 12.9 \)
   \( s = 0.19 \)
5. \( z = \frac{12.9 - 13}{0.19/\sqrt{32}} \approx -2.98 \)
   \( P = 0.0014 \)

Note: the null hypothesis should have the equality symbol, not the alternate hypothesis.

Decision and Interpretation: \( P = 0.0014 < 0.01 = \alpha \). Therefore, we reject the null hypothesis. There is not enough evidence to reject the claim of the pit crew.
Practice Problem 1: A manufacturer of sprinkler systems designed for fire protection claims that the average activating temperature is at least 135°F. To test this claim, you randomly select a sample of 32 systems and find the mean activation temperature to be 133°F with a standard deviation of 3.3°F. At $\alpha = 0.10$, do you have enough evidence to reject the manufacturer’s claim?

1. $H_0: \mu \geq 135$
2. $H_a: \mu < 135$
3. $\alpha = 0.10$
4. $\bar{x} = 133$
   $s = 3.3$
5. $z = \frac{133 - 135}{3.3/\sqrt{32}} \approx -3.43$
   $P = 0.0003$

Decision and Interpretation: $P = 0.0003 < 0.10 = \alpha$. Therefore, we reject the null hypothesis.
7.3 Hypothesis Testing for the Mean (Small Samples)

In this section, we describe the complete procedure of hypothesis testing when the sample size \( n < 30 \) and the population is assumed to be approximately normally distributed. The z-table will not be used for small samples, instead the t-table will be used.

Example 3: A used car dealer says that the mean price of a 2008 Honda CR-V is at least $20,500. You suspect this claim is incorrect and find that a random sample of 14 similar vehicles has a mean price of $19,850 and a standard deviation of $1084. Is there enough evidence to reject the dealer’s claim at \( \alpha = 0.05 \)? Assume the population is normally distributed.

Inputs:

1. \( H_0 : \mu \geq 20,500 \) (left-tailed)
2. \( H_a : \mu < 20,500 \)
3. \( \alpha = 0.05 \)
4. \( \bar{x} = 19,850 \)
   \( s = 1,084 \)

New Input: \( n - 1 = 14 - 1 = 13 \) the degrees of freedom

Step 5:
Calculating the \( P \)-value is not possible in this case with our tables, as we do not use the z-table, but the t-table. Instead,

i) Determine the critical value, \( t_0 \) for \( t \) using the table and: \( n - 1, \alpha, H_0 \):  
   \( t_0 \) is negative for a left-tailed test
   \( t_0 \) is positive for a right-tailed test
   \( t_0 \) is negative and positive for a two-tailed test

   This is a left-tailed test, so \( t_0 = -1.771 \).

ii) Find the \( t \)-score for the sample

\[
 t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{19,850 - 20,500}{1,084\sqrt{14}} \approx -2.244
\]

Step 6:
The rejection region is drawn based on the critical value \( t_0 \). If the calculated \( t \) is in the rejection region (draw a picture), make the decision to reject \( H_0 \). Otherwise, fail to reject \( H_0 \).

\( t = -2.244 \) is in the rejection region, therefore we reject \( H_0 \): there is enough evidence at the 5% significance level to reject this claim.
Example 4: An industrial company claims that the mean pH level of the water in a nearby river is 6.8. You randomly select 19 water samples and measure their pH. The mean and standard deviation are 6.7 and 0.24, respectively. Is there enough evidence to reject the company’s claim at a significance level of 0.05? Assume the population is normally distributed.

Inputs:

1. \( H_0 : \mu = 6.8 \) (two-tailed)
2. \( H_a : \mu \neq 6.8 \)
3. \( \alpha = 0.05 \)
4. \( \bar{x} = 6.7 \)
   \( s = 0.24 \)
   \( n-1 = 18 \)
5. \( t_0 = 2.10, -2.10 \) (rejection region: \( t \geq 2.10 \) and \( t \leq -2.10 \))

\[ t \approx \frac{6.7 - 6.8}{0.24/\sqrt{19}} \approx -1.82 \]

Decision and Interpretation: The t-value \(-1.82\) is not in the rejection region, thus, there is not enough statistical evidence to reject the null hypothesis, that the mean pH of the water is equal to 6.8. We fail to reject \( H_0 \).

Practice Problem 2: The Internal Revenue Service claims that the mean wait time for callers during a recent tax filing season was at most 7 minutes. A random sample of 11 callers has a mean wait time of 8.7 minutes and a standard deviation of 2.7 minutes. Is there enough evidence to reject the claim at a significance level of 0.10?

Inputs:

1. \( H_0 : \mu \leq 7 \) (right-tailed)
2. \( H_a : \mu > 7 \)
3. \( \alpha = 0.10 \)
4. \( \bar{x} = 8.7 \)
   \( s = 2.7 \)
   \( n-1 = 10 \)
5. \( t_0 = 1.37 \) (rejection region: \( t \geq 1.372 \))

\[ t \approx \frac{8.7 - 7}{2.7/\sqrt{11}} \approx 2.09 \]

Decision and Interpretation: The t-value 2.09 is in the rejection region, thus, there is enough statistical evidence to reject the null hypothesis, that the mean waiting time is 7 minutes.