Chapter 8: Hypothesis Testing of a Single Population Parameter

THE LANGUAGE OF STATISTICAL DECISION MAKING

DEFINITIONS:
The **population** is the entire group of objects or individuals under study, about which information is desired.

A **sample** is a part of the population that is actually used to get information.

**Statistical inference** is the process of drawing conclusions about the population based on information from a sample of that population.

DEFINITIONS:
The **null hypothesis**, denoted by $H_0$, is the conventional belief—the status quo, or prevailing viewpoint, about a population.

The **alternative hypothesis**, denoted by $H_1$, is an alternative to the null hypothesis – the statement that there is an effect, a difference, a change in the population.

**Tip:** Having trouble determining the alternative hypothesis? Ask yourself “Why is the research being conducted?” The answer is generally the alternative hypothesis, which is why the alternative hypothesis is often referred to as the research hypothesis.
Aspirin Cuts Cancer Risk  Problem 1

According to the American Cancer Society, the lifetime risk of developing colon cancer is 1 in 16. A study suggests that taking an aspirin every other day for 20 years can cut your risk of colon cancer nearly in half. However, the benefits may not kick in until at least a decade of use.

(a) Write the null and the alternative hypotheses for this setting.

(b) Is this a one-sided to the right, one-sided to the left, or two-sided alternative hypothesis? *Hint:* look at the alternative hypothesis.

**Solution 1**

(a) \( H_0: \) Taking an aspirin every other day for 20 years will not change the risk of getting colon cancer, which is 1 in 16.

\( H_1: \) Taking an aspirin every other day for 20 years will reduce the risk of getting colon cancer, so the risk will be less than 1 in 16.

(b) The alternative hypothesis is one-sided to the left.

Average Life Span  Problem 2

Suppose you work for a company that produces cooking pots with an average life span of seven years. To gain a competitive advantage, you suggest using a new material that claims to extend the life span of the pots. You want to test the hypothesis that the average life span of the cooking pots made with this new material increases.

(a) Write the null and the alternative hypotheses for this setting.

(b) Is this a one-sided to the right, one-sided to the left, or two-sided alternative hypothesis? *Hint:* look at the alternative hypothesis.

**Solution 2**

(a) \( H_0: \) The average life span of the new cooking pots is 7 years.

\( H_1: \) The average life span of the new cooking pots is greater than 7 years.

(b) The alternative hypothesis is one-sided to the right.

Poll Results  Problem 3

Based on a previous poll, the percentage of people who said they plan to vote for the Democratic candidate was 50%. The presidential candidates will have daily televised commercials and a final political debate during the week before the election. You want to test the hypothesis that the population proportion of people who say they plan to vote for the Democratic candidate has changed.

(a) Write the null and the alternative hypotheses for this setting.

(b) Is this a one-sided to the right, one-sided to the left, or two-sided alternative hypothesis? *Hint:* look at the alternative hypothesis.

**Solution 3**

(a) \( H_0: \) The percentage of the Democratic votes in the upcoming election will be 50%.

\( H_1: \) The percentage of the Democratic votes in the upcoming election will be different from 50%.

(b) The alternative hypothesis is two-sided.
Let’s Do It! Practice Writing The Null and Alternative.

a. In the past, the mean running time for a certain type of flashlight battery has been 9 hours. The manufacturer has introduced a change in the production method and wants to perform a hypothesis test to determine whether the mean running time has changed as a result.

   \[ H_0: \]

   \[ H_1: \]

b. The maximum acceptable level of a certain toxic chemical in vegetables has been set at 0.6 parts per million (ppm). A consumer health group measured the level of the chemical in a random sample of tomatoes obtained from one producer to determine whether the mean level of the chemical in these tomatoes exceeds the recommended limit.

   \[ H_0: \]

   \[ H_1: \]

c. A manufacturer claims that the mean amount of juice in its 16 ounce bottles is 16 ounces. A consumer advocacy group wants to perform a hypothesis test to determine whether the mean amount is actually less than this.

   \[ H_0: \]

   \[ H_1: \]

d. A psychologist has designed a test to measure stress levels in adults. She has determined that nationwide the mean score on her test is 27. A hypothesis test is to be conducted to determine whether the mean score for trial lawyers exceeds the national mean score.

   \[ H_0: \]

   \[ H_1: \]
What Errors Could We Make?

**DEFINITION:**

Rejecting the null hypothesis $H_0$ when in fact it is true, is called **Type I error**.

Failing to reject the null hypothesis $H_0$ when in fact it is false is called a **Type II error**.

---

<table>
<thead>
<tr>
<th>The True Hypothesis</th>
<th>Null $H_0$ is true</th>
<th>Alternative $H_1$ is true</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your Decision Based on the Data</td>
<td>Null $H_0$ is supported</td>
<td>No error</td>
</tr>
<tr>
<td></td>
<td>Alternative $H_1$ is supported</td>
<td>Type I error</td>
</tr>
</tbody>
</table>

**Remember that** ... A Type I error can *only* be made if the null hypothesis is true.

A Type II error can *only* be made if the alternative hypothesis is true.

**Let’s Do It!  Rain, Rain, Go Away!**

You plan to walk to a party this evening. Are you going to carry an umbrella with you? You don’t want to get wet if it should rain. So you wish to test the following hypotheses:

- $H_0$: Tonight it is going to rain.
- $H_1$: Tonight it is not going to rain.

(a) Describe the two types of error that you could make when deciding between these two hypotheses.

(b) What are the consequences of making each type of error?
Let’s Do It!

a. A manufacturer claims that the mean amount of juice in its 16 ounces bottles is 16.1 ounces. A consumer advocacy group wants to perform a hypothesis test to determine whether the mean amount is actually less than this. The hypotheses are:

\[ H_0: \mu = 16.1 \text{ ounces} \]
\[ H_1: \mu < 16.1 \text{ ounces} \]

Where \( \mu \) is the mean amount of juice in the manufacturer's 16 ounces bottles. Suppose that the results of the sampling lead to rejection of the null hypothesis. Classify that conclusion as a Type I error, a Type II error, or a correct decision, if in fact the mean amount of juice, \( \mu \), is less than 16.1 ounces.

b. In the past, the mean running time for a certain type of flashlight battery has been 9.9 hours. The manufacturer has introduced a change in the production method and wants to perform a hypothesis test to determine whether the mean running time has increased as a result. The hypotheses are:

\[ H_0: \mu = 9.9 \text{ hours} \]
\[ H_1: \mu > 9.9 \text{ hours} \]

Where \( \mu \) is the mean running time of the new batteries. Suppose that the results of the sampling lead to non-rejection of the null hypothesis. Classify that conclusion as a Type I error, a Type II error, or a correct decision, if in fact the mean running time has increased.

**DEFINITION:**

A **critical region** is the set of values for which you would reject the null hypothesis \( H_0 \). Such values are contradictory to the null hypothesis and favor the alternative hypothesis \( H_1 \).

A **non-critical region** is the set of values for which you would Not Reject the null hypothesis \( H_0 \).

The **cut-off value** or **critical value** is the value which marks the starting point of the set of values that comprise the rejection region.
Elements of a Test of Hypothesis

1. **Null hypothesis** $H_0$: A theory about the values of one or more population parameters. The theory generally represents the status quo, which we adopt until it is proven false. By convention, the theory is stated as $H_0: \mu = \text{Value}$.

2. **Alternative (research) hypothesis** $H_1$: A theory that contradicts the null hypothesis. The theory generally represents that which we will accept only when sufficient evidence exists to establish its truth.

3. **Test statistic**: A sample statistic used to decide whether to support the research hypothesis $H_1$.

4. **Critical (A.K.A. Rejection) region** $\alpha$: The numerical values of the test statistic for which the null hypothesis $H_0$ will be rejected and the research hypothesis $H_1$ is supported. The rejection region is chosen so that the probability is that it will contain the test statistic when the null hypothesis is true, thereby leading to a Type I error $\alpha$. The value of is usually chosen to be small (e.g., .01, .05, or .10) and is referred to as the **level of significance** $\alpha$ of the test.

5. **Conclusion**:
   
a. If the numerical value of the test statistic falls into the critical region, we reject the null hypothesis and support the alternative hypothesis. We know that the hypothesis-testing process will lead to this conclusion incorrectly (a Type I error) only $\alpha$% of the time when $H_0$ is true.

   b. If the test statistic does not fall into the critical region, we do not reject $H_0$. Thus, we reserve judgment about which hypothesis is true. We do not conclude that the null hypothesis is true because we do not (in general) know the probability $\beta$ that our test procedure will lead to an incorrect acceptance of $H_0$ (a Type II error).
Large Sample Test of Hypothesis about a Population Mean $\mu$.

Steps for Selecting the Null and Alternative Hypotheses

1. Select the *alternative hypothesis* as that which the sampling experiment is intended to establish. The alternative hypothesis will assume one of three forms:
   a. One tailed, upper tailed (right -sided) *Example*: $H_1: \mu > 2400$
   b. One tailed, lower tailed (left-sided) *Example*: $H_1: \mu < 2400$
   c. Two tailed *Example*: $H_1: \mu \neq 2400$

The rejection region for a *two-tailed test* differs from that for a one-tailed test. When we are trying to detect departure from the null hypothesis in *either* direction, we must establish a rejection region in both tails of the sampling distribution of the test statistic.

Figures 8.4a and 8.4b show the one-tailed rejection regions for lower- and upper-tailed tests, respectively. The two-tailed rejection region is illustrated in Figure 8.4c. Note that a rejection region is established in each tail of the sampling distribution for a two-tailed test.

The rejection regions corresponding to typical values selected for $\alpha$ are shown in Table 8.2 for one- and two-tailed tests. Note that the smaller you select, the more evidence (the larger $z$) you will need before you can reject $H_0$ and support $H_1$.

**TABLE 8.2** Rejection Regions for Common Values of $\alpha$

<table>
<thead>
<tr>
<th>Alternative Hypotheses</th>
<th>Lower Tailed</th>
<th>Upper Tailed</th>
<th>Two Tailed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha = .10$</td>
<td>$z &lt; -1.28$</td>
<td>$z &gt; 1.28$</td>
<td>$z &lt; -1.645$ or $z &gt; 1.645$</td>
</tr>
<tr>
<td>$\alpha = .05$</td>
<td>$z &lt; -1.645$</td>
<td>$z &gt; 1.645$</td>
<td>$z &lt; -1.96$ or $z &gt; 1.96$</td>
</tr>
<tr>
<td>$\alpha = .01$</td>
<td>$z &lt; -2.33$</td>
<td>$z &gt; 2.33$</td>
<td>$z &lt; -2.575$ or $z &gt; 2.575$</td>
</tr>
</tbody>
</table>
Example
Finding the Critical Region for $\alpha = 0.01$ (Right-Tailed Test) using the standard normal distribution $N(0,1)$.

Solution
The critical value is for $\alpha = 0.01$ is 2.33.

Let’s Do It! Critical Values.

a. Finding the Critical Value for $\alpha = 0.05$ (Left-Tailed Test) using The standard normal distribution $N(0,1)$.

b. Finding the Critical Value for $\alpha = 0.05$ (Two-Tailed Test) using The standard normal distribution $N(0,1)$.

The Decision Tool (THE TEST STATISTIC) In order to perform the hypothesis test, we first determine the value of the observed statistic of our sample. Many hypotheses are tested using a statistical test based on the following general formula:

\[
\text{Test value} = \frac{(\text{observed value}) - (\text{expected value})}{\text{standard error}}
\]

There are two cases:

- **If the sample used is large**, then the distribution used for testing the hypothesis is $Z$. The standardized observed test statistic is given by:

\[
Z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}
\]

- **If the sample is small ($n < 30$) and the population is normally distributed**, then the distribution used for testing is $T$. The standardized observed test statistic is:

\[
T = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \quad (\text{We discussed this case in the next section})
\]
Hypothesis Testing

In order to perform the hypothesis test, many hypotheses are tested using a statistical test based on the following general formula:

\[
\text{Test Statistic } T = \frac{(\text{Observed average from sample}) - (\text{Expected average from null hypothesis})}{\text{Standard Error}}
\]

In particular, when testing the population mean, our test value is the standardized Test statistic

\[
\frac{\bar{x} - \mu_0}{s / \sqrt{n}}
\]

Hypothesis-Testing Problems (Traditional Method)

**Step 1:** State the null and the Alternative and identify the claim.

**Step 2:** Find the Test Statistic \( \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \) (standardized sample mean)

**Step 3:** If the test statistic \( \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \) falls in the critical region \( \alpha \), this implies that there is an evidence to reject the null and support the alternative. If \( \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \) is in the non-critical region, then the test statistic fell outside the rejection region of the null and we have no evidence to support the alternative.

**Step 4:** Summarize the results. **Note:** Never accept the null. We reserve the judgment about the hypothesis is true.

**Let’s Do It!**

A sample mean, sample standard deviation, and sample size are given. Use the one-mean t-test to perform the required hypothesis test about the mean, \( \mu \), of the population from which the sample was drawn. Use the critical value approach.

a. \( \bar{x} = 24.4, s = 9.2, n = 35, H_0: \mu = 26, H_1: \mu < 26, \alpha = 0.05 \)

b. \( \bar{x} = 22, s = 14,200, n = 47, H_0: \mu = 30,000, H_1: \mu \neq 30,000, \alpha = 0.001 \)
**Example:** A researcher thinks that the average salary of assistant professors is more than $42,000. A sample of 30 assistant professors has a mean salary of $43,260. At \( \alpha = 0.05 \), test the claim that assistant professors earn more than $42,000 a year. The standard deviation of the sample is $5230.

**Step1: State the hypotheses and identify the claim.**

\[ H_0 : \mu = 42,000 \]

\[ H_1 : \mu > 42,000 \quad \text{This is the researcher’s claim} \]

**Step2: Compute the standardized test statistic**

\[ Z = \frac{43,260 - 42,000}{5230/\sqrt{30}} = 1.32 \] (So, our sample average of 43,260 is 1.32 standard deviations from 42,000)

**Step3: Find the critical region and critical value:**

Since this is a right tailed test, the critical region on the right and starts at the critical value C.V. \( \text{C.V.} = \text{InvNorm (0.95, 0, 1)} = 1.65 \)

**Step4: Make the decision.**

Note that the test statistic +1.32 is not in the critical region, the decision is failing to reject the null and therefore, there is no evidence supporting the alternative.

**Step5: Summarize the results**

There is not enough evidence to support the claim that assistant professors earn more on average than $42,000 a year.
Let’s Do It!
A supermarket sells rotisserie chicken at a fixed price per chicken rather than by the weight of the chicken. The store advertises that the average weight of their chickens is 4.6 pounds. A random sample of 30 of the store’s chickens yielded the weights (in pounds) shown below.

4.4  4.7  4.6  4.4  4.5  4.3  4.6  4.5  4.6  4.9  
4.6  4.8  4.3  4.4  4.7  4.5  4.2  4.3  4.1  4.0  
4.5  4.6  4.2  4.4  4.7  4.8  5.0  4.2  4.1  4.5

Test whether the population’s mean weight of the chickens is less than 4.6 pounds. Use $\alpha = 0.05$

Hypothesis:

$H_0$: H_{1}: 

Test Statistic:

Critical region and critical value (draw a picture):

Decision:

Result:

Let’s Do It!
A fast food outlet reports that the mean waiting time in line is 3.5 minutes. A random sample of 60 customers has a mean of 3.6 minutes with a standard deviation of 0.6 minute. If $\alpha = 0.01$, test that the mean waiting time is different from 3.5 minutes.

Hypothesis:

$H_0$: H_{1}: 

Test Statistic:

Critical region and critical value (draw a picture):

Decision:

Result:
Small Sample Test of Hypothesis about a Population Mean $\mu$

**The Decision Tool (THE TEST STATISTIC)** In order to perform the hypothesis test, we first determine the value of the observed statistic of our sample. Many hypotheses are tested using a statistical test based on the following general formula:

$$\text{Test value} = \frac{\text{(observed value)} - \text{(expected value)}}{\text{standard error}}$$

There are two cases:

- If the sample used is large, then the distribution used for testing the hypothesis is $Z$. The standardized observed test statistic is given by:
  $$Z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$ (We discussed this case in the previous section)

- If the sample is small ($n<30$) and the population is normally distributed, then the distribution used for testing is $T$. The standardized observed test statistic is:
  $$T = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

**Example**

Find the critical $t$ value for $\alpha = 0.05$ with d.f. = 16 for a right-tailed $t$ test.

**Solution**

Refer to t-table VI Page 775 in your textbook:

Find the 0.05 column in the top row and 16 in the left-hand column. Where the row and column meet, the appropriate critical value is found; it is +1.746
Let’s Do It!

Find the critical t value for $\alpha = 0.01$ with d.f. = 22 for a left-tailed test.

Find the critical values for $\alpha = 0.10$ with d.f. = 18 for a two-tailed t test.
Example  Most water-treatment facilities monitor the quality of their drinking water on an hourly basis. One variable monitored is pH, which measures the degree of alkalinity or acidity in the water. A pH below 7.0 is acidic, one above 7.0 is alkaline, and a pH of 7.0 is neutral. One water-treatment plant has a target pH of 8.5. (Most try to maintain a slightly alkaline level.) The average and standard deviation of 1 hour’s test results, based on 17 water samples at this plant, are 8.42 and 0.16 respectively. Does this sample provide sufficient evidence that the mean pH level in the water differs from 8.5?

Hypothesis:

\[ H_0: \mu = 8.5 \text{ (Mean pH level is 8.5.)} \]
\[ H_a: \mu \neq 8.5 \text{ (Mean pH level differs from 8.5.)} \]

Test Statistic

\[
t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{8.42 - 8.50}{0.16/\sqrt{17}} = \frac{-0.08}{0.039} = -2.05
\]

Critical region and Critical Values:

Decision: Since the calculated value of \( t \) does not fall into the rejection region we fail to reject \( H_0 \).

Conclusion: There is not enough evidence to conclude that the mean pH level in the water is different from 8.5.
An educator claims that the average salary of substitute teachers in school districts in Allegheny County, Pennsylvania, is less than $60 per day. A random sample of eight school districts is selected, and the daily salaries (in dollars) are shown. Is there enough evidence to support the educator’s claim at $\alpha = 0.10$?

| Daily Salaries (in dollars) | 60 | 56 | 60 | 55 | 70 | 55 | 60 | 55 |

Hypothesis

$H_0$: 

$H_1$: 

Test Statistic:

Critical Region and Critical Values (draw the picture)

Decision: Reject $H_0$ 

Fail to Reject $H_0$

Conclusion:

The U.S. average for state and local taxes for a family of four is $4172. A random sample of 20 families in a northeastern state indicates that they paid an annual amount of $4560 with a standard deviation of $1590. At $\alpha = 0.05$, is there sufficient evidence to conclude that they pay more than the national average of $4172? 

Hypothesis

$H_0$: 

$H_1$: 

Test Statistic:

Critical Region and Critical Values (draw the picture)

Decision: Reject $H_0$ 

Fail to Reject $H_0$

Conclusion:
A report by the Gallup Poll stated that on average a woman visits her physician 5.8 times a year. A researcher randomly selects 20 women and obtained these data. At $\alpha = 0.05$ can it be concluded that the average is different from 5.8 visits per year?

\[
\begin{array}{cccccccc}
3 & 2 & 1 & 3 & 7 & 2 & 9 & 4 \\
8 & 0 & 5 & 6 & 4 & 2 & 1 & 3 & 4 & 1
\end{array}
\]

Hypothesis

$H_0$: 

$H_1$: 

Test Statistic:

Critical Region and Critical Values (draw the picture)

Decision: Reject $H_0$ Fail to Reject $H_0$

Conclusion:
Let’s Do It!
State the null and alternative hypothesis that would be used to test the following statements—these statements are the researcher’s claim, to be stated as the alternative hypothesis. All hypotheses should be expressed in terms of $p$, the population proportion of interest.

(a) More than half of the pregnancies in this country are not planned.

(b) Less than 3% of children vaccinated for chicken pox still contract the disease

(c) The proportion of people in the United States who are lactose intolerant is different from 0.25.
Example Have More Residents Quit Smoking? The proportion of Michigan adult smokers in 1992 was said to be 25.5%. The Michigan Department of Public Health conducted a survey in 1993 to assess whether the smoking level (that is, the proportion of Michigan adults who smoke) had decreased from the previous level of 25.5%. The survey estimated smoking level of 25% was based on a random sample of adults of size \( n = 2400 \).

(a) Express this situation in terms of testing hypotheses about \( p \), the true proportion of adult smokers in Michigan.

\[
H_0: p = 0.255 \quad \text{versus} \quad H_1: p < 0.255
\]

(b) Compute the test Statistic: The observed \( z \)-test statistic is computed as

\[
Z = \frac{0.25 - 0.255}{\sqrt{0.255 \cdot (1 - 0.255)/2400}} \quad \Rightarrow \quad Z = \frac{0.25 - 0.255}{0.0089} = -0.56.
\]

(c) The Critical value is -1.65.

(d) Conclusion: There is insufficient evidence that the smoking levels has decreased from 1992.

Let's Do It! The business college computing center wants to determine the proportion of business students who have laptop computers. If the proportion exceeds 25%, then the lab will scale back a proposed enlargement of its facilities. Suppose 200 business students were randomly sampled and 65 have laptops. Use \( \alpha = 0.01 \)

Check the validity of the test assumptions:

Hypothesis

\[
H_0: \quad p = 0.25 \quad \text{versus} \quad H_1: \quad p > 0.25
\]

Test Statistic:

Critical Region and Critical Values (draw the picture)

Decision: Reject \( H_0 \) \quad Fail to Reject \( H_0 \)

Conclusion:
Let's Do It! A company reports that 80% of its employees participate in the company’s stock purchase plan. A random sample of 50 employees was asked the question, "Do you participate in the stock purchase plan?" The answers are shown below.

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

Perform the appropriate test of hypothesis to investigate your suspicion that fewer than 80% of the company's employees participate in the plan. Use $\alpha = .05$.

Check the validity of the test assumptions:

Hypothesis $H_0$: $H_1$:

Test Statistic:

Critical Region and Critical Values (draw the picture)

Decision: Reject $H_0$  Fail to Reject $H_0$

Conclusion:

Let's Do It! A poll of 1000 adult Americans reveals that 48% of the voters surveyed prefer the Democratic candidate for the presidency. At the 0.05 level of significance, do the data provide sufficient evidence to conclude that the percentage of voters who prefer the Democrat is not 50%?

Check the validity of the test assumptions:

Hypothesis $H_0$: $H_1$:

Test Statistic:

Critical Region and Critical Values (draw the picture)

Decision: Reject $H_0$  Fail to Reject $H_0$

Conclusion: