Chapter 9
Hypothesis Testing

“It’s not my fault when you consider that my three husbands have had twenty wives.”
– Ava Gardner, on being asked about her failed marriages to Mickey Rooney, Artie Shaw, and Frank Sinatra

Men are generally more careful of the breed of their horses and dogs than of their children
- William Penn

---

Example: NHL

The NHL currently has a rigorous (or so they claim) drug testing policy for all players. The NHL follows the testing “innocent until proven guilty approach”

$H_0$: The player does not take steroids
$H_a$: The player does take steroids

---

Type I and Type II Errors

- Since hypothesis tests are based on sample data, we must allow for the possibility of errors.
- A **Type I error** is rejecting $H_0$ when it is true.
  - $\alpha$ (alpha) is the probability of rejecting the truth
  - $\alpha$ is called the **level of significance**. The person conducting the hypothesis test specifies the maximum allowable probability of making a Type I error
- A **Type II error** is failing to reject $H_0$ when it is false.
  - $\beta$ (beta) is the probability of failing to reject what is false
Example: NHL Drug Testing

H₀: The player does not take steroids
H₁: The player does take steroids

<table>
<thead>
<tr>
<th>Population Condition</th>
<th>H₀ True</th>
<th>H₁ True</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail to Reject H₀</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reject H₀</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NHL Example

Given: a certain player has recently taken steroids specifically for muscular enhancement.

What type of error does this player hope will be made when he is tested for steroids?

What type of error does the NHL want to avoid?

Example: Omaha EMS

Omaha provides one of the most comprehensive emergency medical services in the world. Operating in a multiple hospital system with approximately 10 mobile medical units, the service goal is to respond to medical emergencies with a mean time of 6 minutes or less.

The director of medical services wants to formulate a hypothesis test (95% confidence) that could use a sample of emergency response times to determine whether or not the service goal of 6 minutes or less is being achieved.

One-Tailed Test about a Population Mean

Let n = 40, ̅ = 7.25 minutes, s = 3.2 minutes
Example: Omaha EMS

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Conclusion and Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$: $\mu \leq __$</td>
<td>The emergency service is meeting the response goal; no follow-up action is necessary.</td>
</tr>
<tr>
<td>$H_a$: $\mu &gt; __$</td>
<td>The emergency service is not meeting the response goal; appropriate follow-up action is necessary.</td>
</tr>
</tbody>
</table>

Where $\mu$ = mean response time for the population of medical emergency requests.

Example: Omaha EMS

<table>
<thead>
<tr>
<th>Conclusion</th>
<th>Population Condition</th>
<th>$H_0$ True $\mu \leq 6$</th>
<th>$H_a$ True $\mu &gt; 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail to Reject $H_0$ (Conclude $\mu \leq 6$)</td>
<td>Correct Conclusion</td>
<td>Type II Error</td>
<td></td>
</tr>
<tr>
<td>Reject $H_0$ (Conclude $\mu &gt; 6$)</td>
<td>Type I Error</td>
<td>Correct Conclusion</td>
<td></td>
</tr>
</tbody>
</table>

The 7 Steps of Hypothesis Testing

1. Determine the appropriate hypotheses.
2. Select the test statistic and its distribution.
3. Specify the level of significance $\alpha$ for the test.
4. Develop the decision rule for rejecting $H_0$.
5. Collect the sample data and compute the value of the test statistic.
6. Statistical Decision
   - Compute the $p$-value based on the test statistic and compare it to $\alpha$, to determine whether or not to reject $H_0$.
7. Interpretation.
One-Tailed Tests About a Population Mean:
Large-Sample Case (n ≥ 30)

1. Hypotheses
   H₀: \( \mu \leq \mu_0 \)
   H₁: \( \mu > \mu_0 \)
2. Test Statistic:
   If \( \sigma \) Known, \( z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \)
   If \( \sigma \) Unknown, \( z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \)
3. Specify \( \alpha \)
   \( \alpha = .05 \)
4. Rejection Rule:
   Reject \( H_0 \) if \( P \)-value \( \leq \alpha \)

Example: Omaha EMS

One-Tailed Test about a Population Mean

Let \( \alpha = P \) (Type I Error) = .05
Sampling distribution of \( \bar{x} \)
(assuming \( H_0 \) is True and \( \mu = 6 \))

\( z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \)

\( \bar{x} = 7.25 \) minutes, \( s = 3.2 \) minutes

\( z = \frac{7.25 - 6}{3.2/\sqrt{40}} \)

\( z = 2.47 \)
The Use of $p$-Values

- The $p$-value is the probability of obtaining a sample result that is at least as unlikely as what is observed.
- The $p$-value can be used to make the decision in a hypothesis test by noting that:
  - if the $p$-value is less than the level of significance $\alpha$, the value of the test statistic is in the rejection region.
  - if the $p$-value is greater than or equal to $\alpha$, the value of the test statistic is not in the rejection region.
- Reject $H_0$ if the $p$-value < $\alpha$.

Example: Omaha EMS

- Using the $p$-value to Test the Hypothesis

Recall that $z = 2.47$ for $\bar{x} = 7.25$, $\alpha = .05$

Find $p$-value using the $z$ table. $P$-value = .0068

Since $p$-value < $\alpha$, that is .0068 < .05, we reject $H_0$ and $\bar{x}$ is greater than 6 ($\mu_0$)

Example: Omaha EMS

- Conclusion: We are 95% confident that Omaha EMS is not meeting the response goal of 6 minutes; appropriate action should be taken to improve service.
Tests About a Population Mean:
Small-Sample Case (n < 30)

Test Statistic where \( \sigma \) is unknown

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

This test statistic has a \( t \) distribution with \( n - 1 \) degrees of freedom.

- \( H_0: \mu \leq \mu_0 \) One Tailed
- \( H_0: \mu \geq \mu_0 \) One Tailed
- \( H_0: \mu = \mu_0 \) Two Tailed

\( p \)-Values and the \( t \) Distribution

- The format of the \( t \) distribution table provided in most statistics textbooks does not have sufficient detail to determine the exact \( p \) -value for a hypothesis test.
- However, we can still use the \( t \) distribution table to identify a range for the \( p \) -value.
- An advantage of computer software packages is that the computer output will provide the \( p \) -value for the \( t \)-distribution.

NVIDIA Example

NVIDIA has announced the latest turning architecture processor that computes 16 trillion floating-point operations per second, 500 trillion tensor operations per second, 10 GigaRays per second. Internally to the company, the design specification required that the population average speed for all processors be greater than 12 GigaRays. From the Quality Control department perspective, is the design specification met with 99% confidence? (the population is normally distributed)

A random sample of 21 chips gives a sample average of 12.27 GigaRays with a sample standard deviation of 0.80 GigaRays.

Research hypothesis
Using the 7 Steps of Hypothesis Testing

1. \( H_0: \mu \leq \_ \) \( H_A: \mu > \_ \)
2. \( t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \)
3. \( \alpha = .01 \)
4. Rejection Rule:
   - Reject \( H_0 \) if \( P\text{-value} < \alpha \)
   - and \( \bar{x} > 12 \)

Using the p-value approach

5. \( t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \)
6. Statistical Decision
   - \_ < 1.55 < \_
   - \_ < \text{p-value} < \_
   - \( p\text{-value} = .06882 \) (Using Excel)

   - Reject \( H_0 \) if \( p\text{-value} < \alpha \)

   - Since .06882 < .05, fail to reject \( H_0 \)

Conclusion

7. At 99% confidence, the data is not sufficient to prove that the average speed for all processors is greater than 12 GigaRays before failure.
**NVIDIA Example**

- Loper Technologies (LT), as a major client, also receives exactly the same sample information from Intel. From LT’s perspective, can the design specification be refuted with 99% confidence? (use at least an average speed of 12 GigaRays)
- A random sample of 21 chips gives a sample average of 12.27 GigaRays with a sample standard deviation of 0.03 GigaRays.

*Testing a claim*

---

**Using the 7 Steps of Hypothesis Testing**

1. \( H_0: \mu \geq 12.00 \) \( H_1: \mu < 12.00 \)
2. \( t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \)
3. \( \alpha = .01 \)
4. Rejection Rule: \[ \text{Reject } H_0 \text{ if } P\text{-value} < \alpha \text{ and } \bar{x} < 12.00 \]

---

**Using the 7 Steps of Hypothesis Testing**

5. **Statistical Decision**
   - Fail to reject \( H_0 \)
   - \( \bar{x} \) is not less than 12.00
6. At 99% confidence, the data is not sufficient to refute that the average speed for all processors is at least 12 GigaRays before failure.
   - OR-
   - At 99% confidence, the data is not sufficient to prove that the average speed for all processors is less than 12 GigaRays before failure.
Two-Tailed Tests About a Population Mean:
Large-Sample Case (n ≥ 30)
- Hypotheses:
  - \( H_0: \mu = \mu_0 \)
  - \( H_1: \mu \neq \mu_0 \)
- Test Statistic:
  - \( \sigma \) Known
    - \[ z = \frac{\bar{x} - \mu_0}{\sigma} \]
  - \( \sigma \) Unknown
    - \[ z = \frac{\bar{x} - \mu_0}{s} \]
- Rejection Rule:
  - Region of Rejection
  - Reject \( H_0 \) if \( P \)-value < \( \alpha \)
  - \( \bar{x} \) may fall on the either side of \( \mu_0 \)

Two-Tailed Tests About a Population Mean:
Small-Sample Case (n < 30)
- Hypotheses:
  - \( H_0: \mu = \mu_0 \)
  - \( H_1: \mu \neq \mu_0 \)
- Test Statistic:
  - \( \sigma \) Unknown
    - \[ t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \]
- Rejection Rule:
  - Region of Rejection
  - Reject \( H_0 \) if \( P \)-value < \( \alpha \)
  - \( \bar{x} \) may fall on the either side of \( \mu_0 \)
Loper Electronics

Loper Openair U-acoustic Devices (LOUD) manufactures flat panel speakers. The average resistance per speaker must be exactly 8 ohms. The distributor of LOUD speakers will reject a shipment of speakers if the average resistance is not exactly 8 ohms. A sample of 25 speakers yields an average resistance of 8.15 ohms with a standard deviation of .5 ohms.

Should the distributor keep or return the shipment? Let $\alpha = .05$.

Using the 7 Steps of Hypothesis Testing

1. $H_0: \mu = 8$ $H_1: \mu \neq 8$

2. $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$

3. $\alpha = .05$

4. Rejection Rule:

   Reject $H_0$ if $P-value \leq \alpha$

   $\bar{x}$ may fall on either side of $\mu = 8$

5. $t = \frac{8.15 - 8.00}{.5/\sqrt{25}} = 1.5$

6. Statistical Decision

   $1.318 < 1.5 \leq 1.711$

   $.10 < P-value < .20$

   Use Excel function T.DIST.2T to obtain true value.

   $P-value = .146656$

   Fail to reject $H_0$
Using the 7 Steps of Hypothesis Testing

7. The data is not sufficient to prove that the population average resistance of all the speakers is not 8 ohms.
   - Keep the shipment

A Summary of Forms for Null and Alternative Hypotheses about a Population Proportion

- The equality part of the hypotheses always appears in the null hypothesis.
- In general, a hypothesis test about the value of a population proportion \( p \) must take one of the following three forms (where \( p_0 \) is the hypothesized value of the population proportion).

\[
\begin{align*}
H_0: & \quad p \geq p_0 \\
H_a: & \quad p < p_0 \\
H_0: & \quad p = p_0 \\
H_a: & \quad p \neq p_0
\end{align*}
\]

Tests About a Population Proportion:

Large-Sample Case (\( np \geq 5 \) and \( n(1 - p) \geq 5 \))

- Test Statistic: \( z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} \)
- \( H_0: \ p \leq p_0 \) One-Tailed
- \( H_0: \ p \geq p_0 \) One-Tailed
- \( H_0: \ p = p_0 \) Two-Tailed
Example: UNK Parking

- UNK is considering building a parking garage for students. The administration believes that more than 60% of the students drive cars to campus. If in a random sample of 250 students, 165 indicate that they drive a car to school is the administration's position supported? Let $\alpha = 0.05$.

Using the 7 Steps of Hypothesis Testing

1. $H_0: \rho \leq 0.60$  \hspace{1cm} $H_a: \rho > 0.60$
2. $z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}}$
3. $\alpha = 0.05$
4. Rejection Rule: $P-value < \alpha$

$\rho$ is greater than 0.60.

$\bar{p} = \frac{165}{250} = 0.66$

Using the 7 Steps of Hypothesis Testing

5. $\bar{p} = \frac{165}{250}$, $\sigma_{\bar{p}} = \sqrt{\frac{0.6(1-0.6)}{250}} = 0.03098$

\[ z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \frac{0.66 - 0.60}{0.03098} = 1.936 \]

6. Statistical Decision

P-value = 0.0262

0.0262 < 0.05, Reject $H_0$, Accept $H_a$

7. The administration is 95% confident that more than 60% of the students drive cars to campus.
Chapter 10

Hypothesis Tests About the Difference Between the Means of Two Populations: Independent Samples

Hypothesis Forms:

- $H_0: \mu_1 - \mu_2 \leq 0$
- $H_0: \mu_1 - \mu_2 \geq 0$
- $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 > 0$
$H_a: \mu_1 - \mu_2 < 0$
$H_a: \mu_1 - \mu_2 \neq 0$

Test Statistic:

- Large-Sample Case (population variance is unknown)
  
  \[ z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2) \sigma}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \]

Example: Par, Inc.

Par, Inc. is a manufacturer of golf equipment. Par has developed a new golf ball that has been designed to provide “extra distance.” In a test of driving distance using a mechanical driving device, a sample of Par golf balls was compared with a sample of golf balls made by Rap, Ltd., a competitor. The sample data is below.

<table>
<thead>
<tr>
<th></th>
<th>Sample #1</th>
<th>Sample #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>$n_1 = 120$ balls</td>
<td>$n_2 = 80$ balls</td>
</tr>
<tr>
<td>Mean</td>
<td>$x_{1 \bar{}} = 225$ yards</td>
<td>$x_{2 \bar{}} = 218$ yards</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>$s_1 = 15$ yards</td>
<td>$s_2 = 20$ yards</td>
</tr>
</tbody>
</table>
Hypothesis Tests About the Difference Between the Means of Two Populations: Large-Sample Case

Can we conclude, using a .01 level of significance, that the mean driving distance of Par, Inc. golf balls is greater than the mean driving distance of Rap, Ltd. golf balls?

\[ \mu_1 = \text{mean distance for the population of Par, Inc. golf balls} \]
\[ \mu_2 = \text{mean distance for the population of Rap, Ltd. golf balls} \]

1. Hypotheses:
   - \( H_0: \mu_1 - \mu_2 \leq 0 \)
   - \( H_a: \mu_1 - \mu_2 > 0 \)

\[ z = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{225 - 218}{\sqrt{\frac{15}{120} + \frac{20}{80}}} = \frac{7}{2.62} = 2.67 \]

Example: Par, Inc.

4. Rejection Rule:
   - Reject Ho if P-value < \( \alpha \)
   - \( \bar{x}_1 \) = greater than 0.0

5. Calculation
   \[ z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{7}{2.62} = 2.67 \]

Example: Par, Inc.

6. Statistical Decision
   - P-value = .0038
   - .0038 < .01
   - Reject Ho, accept Ha

7. Conclusion: We are at least 99% confident that the mean driving distance of Par, Inc. golf balls is greater than the mean driving distance of Rap, Ltd. golf balls.
Hypothesis Testing about \( p_1 - p_2 \)

- **Hypotheses:**
  - \( H_0: p_1 - p_2 \leq 0 \)
  - \( H_a: p_1 - p_2 > 0 \)

- **Test statistic:**
  \[
  z = \left( \hat{p}_1 - \hat{p}_2 \right) \sqrt{\frac{p(1-p)}{n_1 + n_2}}
  \]

Example: The Israeli – Palestinian Conflict

- Hypothesis Tests about \( p_1 - p_2 \)
  - Can we conclude, using a .01 level of significance, that U.S. voters are less resistant to Israeli incursions into Palestinian occupied areas this week than two weeks ago?
  - \( \rho_1 \) = proportion of the voters resistant the Israeli actions this week.
  - \( \rho_2 \) = proportion of the voters resistant the Israeli actions two weeks earlier.
  - \( n_1 = 120 \)
  - \( n_2 = 80 \)
  - \( \bar{p}_1 = .10 \)
  - \( \bar{p}_2 = .22 \)

Example: The Israeli – Palestinian Conflict

- **1. Hypotheses:**
  - \( H_0: p_1 - p_2 \geq 0 \)
  - \( H_a: p_1 - p_2 < 0 \)

- **4. Rejection Rule:**
  - Reject \( H_0 \) if \( P-value \leq \alpha \)
  - \( \rho_1 - \rho_2 \) is less than 0.0
Example: The Israeli – Palestinian Conflict

5. \[
\rho = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2 - \frac{120 \times .10 + 80 \times .22}{120 + 80}}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = .148
\]

6. \[
\hat{p}_1 - \hat{p}_2 = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} = \sqrt{.148 \times .852 \left( \frac{1}{120} + \frac{1}{80} \right)} = .0512
\]

7. \[
z = \frac{\hat{p}_1 - \hat{p}_2 - (\hat{p}_1 - \hat{p}_2) \sigma}{\hat{p}_1 - \hat{p}_2} = \frac{(10 - 22) - 0}{.0512} = -2.34
\]

P-Value = .00964

6. Statistical Decision
   Reject \( H_0 \) if P-value < \( \alpha \)
   .00964 < .01,
   Reject \( H_0 \) Accept \( H_a \)

7. We are 99% confident that voters are less resistant to the Israeli incursion into occupied Palestinian areas this week than two weeks ago.