Hypothesis Testing

Presented by:
Mahendra Adhi N

Sources:
http://business.clayton.edu/arjomand/

What is a Hypothesis?

- A hypothesis is a claim (assumption) about a population parameter:
  - population mean
    - Example: The mean monthly cell phone bill of this city is \( \mu = \$42 \)
  - population proportion
    - Example: The proportion of adults in this city with cell phones is \( p = .68 \)

The Null Hypothesis, \( H_0 \)

- States the assumption (numerical) to be tested
- Example: The average number of TV sets in U.S. Homes is equal to three (\( H_0: \mu = 3 \))
- Is always about a population parameter, not about a sample statistic
- \( H_0: \mu = 3 \)
- \( H_0: \bar{x} = 3 \)

(continued)

The Alternative Hypothesis, \( H_1 \)

- Is the opposite of the null hypothesis
  - e.g., The average number of TV sets in U.S. homes is not equal to 3 (\( H_1: \mu \neq 3 \))
  - Challenges the status quo
  - Never contains the “=”, “≤” or “≥” sign
  - May or may not be supported
  - Is generally the hypothesis that the researcher is trying to support

Hypothesis Testing Process

Claim: the population mean age is 50.
(Null Hypothesis: \( H_0: \mu = 50 \))

Is \( \bar{x} = 20 \) likely if \( \mu = 50 \)?
If not likely, REJECT Null Hypothesis

Now select a random sample

Sample

Population
Reason for Rejecting $H_0$

If it is unlikely that we would get a sample mean of this value ... if in fact this were the population mean ...

... then we reject the null hypothesis that $\mu = 50$.

Level of Significance, $\alpha$
- Defines the unlikely values of the sample statistic if the null hypothesis is true
- Defines rejection region of the sampling distribution
- Is designated by $\alpha$, (level of significance)
- Typical values are .01, .05, or .10
- Is selected by the researcher at the beginning
- Provides the critical value(s) of the test

Errors in Making Decisions
- Type I Error
  - Reject a true null hypothesis
  - Considered a serious type of error
  - The probability of Type I Error is $\alpha$
  - Called level of significance of the test
  - Set by researcher in advance

Errors in Making Decisions
- Type II Error
  - Fail to reject a false null hypothesis
  - The probability of Type II Error is $\beta$

Outcomes and Probabilities
<table>
<thead>
<tr>
<th>Possible Hypothesis Test Outcomes</th>
<th>Actual Situation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision</td>
<td></td>
</tr>
<tr>
<td>Do Not Reject $H_0$</td>
<td>$H_0$ True</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>Type I Error</td>
</tr>
</tbody>
</table>

Key:
- No Error ($1 - \alpha$) (Probability)
- Type I Error ($\alpha$)
- Type II Error ($\beta$)
- No Error ($1 - \beta$)
Type I & II Error Relationship

- Type I and Type II errors can not happen at the same time
  - Type I error can only occur if $H_0$ is true
  - Type II error can only occur if $H_0$ is false

Factors Affecting Type II Error

- All else equal,
  - $\beta \uparrow$ when the difference between hypothesized parameter and its true value $\downarrow$
  - $\beta \uparrow$ when $\alpha \downarrow$
  - $\beta \uparrow$ when $\sigma \downarrow$
  - $\beta \uparrow$ when $n \downarrow$

Power of the Test

- The power of a test is the probability of rejecting a null hypothesis that is false
- i.e., $\text{Power} = P(\text{Reject } H_0 | H_1 \text{ is true})$
- Power of the test increases as the sample size increases

Two Basic Approaches to Hypothesis Testing

- There are two basic approaches to conducting a hypothesis test:
  1. p-Value Approach, and
  2. Critical Value Approach

1. p-Value Approach to One-Tailed Hypothesis Testing

- In order to accept or reject the null hypothesis the p-value is computed using the test statistic —Actual Z value.
- Reject $H_0$ if the p-value $\leq \alpha$
- Do not reject (accept) $H_0$ if the p-value $> \alpha$

2. Critical Value Approach
One-Tailed Hypothesis Testing

- Use the Z table to find the critical Z value, and
- Use the equation to find the actual Z — Z statistics.
- The rejection rule is:
  - Lower tail: Reject $H_0$ if Actual $z \leq \text{Critical} -z_{\alpha}$
  - Upper tail: Reject $H_0$ if Actual $z \geq \text{Critical} z_{\alpha}$

In other words, if the actual Z (Z statistics) is in the rejection region, then reject the null hypothesis.

Equation for finding the actual Z value:

$$ z = \frac{X - \mu}{\sigma / \sqrt{n}} $$
Hypothesis Tests for the Mean

σ Known
σ Unknown

Test of Hypothesis for the Mean (σ Known)

- Convert sample result (\( \bar{x} \)) to a \( z \) value

The decision rule is:

\[
H_0: \mu = \mu_0 \\
H_1: \mu > \mu_0
\]

(\( \mu_0 \) is chosen for this test)

Decision Rule

\[
\text{Reject } H_0 \text{ if } z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} > z_{\alpha}
\]

p-Value Approach to Testing

- \( p \)-value: Probability of obtaining a test statistic more extreme (\( \leq \) or \( \geq \)) than the observed sample value given \( H_0 \) is true
- Also called observed level of significance
- Smallest value of \( \alpha \) for which \( H_0 \) can be rejected

\[
p-value = P(Z > z_{\text{observed}} - \mu, \sigma/\sqrt{n})
\]

Example: Find Rejection Region

- Suppose that \( \alpha = .10 \) is chosen for this test
- Find the rejection region:

\[
\text{Reject } H_0 \text{ if } z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} > z_{\alpha}
\]

Alternate rule:

- \( x \)-test:

\[
p-value = P(Z > z_{\text{observed}} - \mu_0, \sigma/\sqrt{n})
\]

© Mahendra Adhi Nugroho, M.Sc, Accounting Program Study of Yogyakarta State University
For internal use only!
Example: Upper-Tail Z Test for Mean (σ Known)

A phone industry manager thinks that customer monthly cell phone bill have increased, and now average over $52 per month. The company wishes to test this claim. (Assume σ = 10 is known)

Form hypothesis test:

H₀: μ ≤ $52     the average is not over $52 per month
H₁: μ > $52     the average is greater than $52 per month (i.e., sufficient evidence exists to support the manager’s claim)

Obtain sample and compute the test statistic

Suppose a sample is taken with the following results:

n = 64, x̄ = $53.1

Using the sample results,

z = (x̄ - μ₀) / (σ / √n)

z = (53.1 - 52) / (10 / √64) = 0.88

Calculate the p-value and compare to α

p-value = P(z ≥ 0.88) = 1 - P(z ≤ 0.88)

p-value = 1 - 0.8106 = 0.1894

Do not reject H₀ since p-value = 0.1894 > α = 0.10

One-Tail Tests

- In many cases, the alternative hypothesis focuses on one particular direction

H₀: μ ≤ 3
H₁: μ > 3

This is an upper-tail test since the alternative hypothesis is focused on the upper tail above the mean of 3

H₀: μ ≥ 3
H₁: μ < 3

This is a lower-tail test since the alternative hypothesis is focused on the lower tail below the mean of 3

Upper-Tail Tests

- There is only one critical value, since the rejection area is in only one tail

H₀: μ ≤ 3
H₁: μ > 3

Z = (x̄ - μ) / (σ / √n)

Z = (x̄ - μ) / (σ / √n)

Critical value

Do not reject H₀

Reject H₀

Do not reject H₀ since Z = 0.88 < 1.28

I.e., there is not sufficient evidence that the mean bill is over $52

One-Tail Tests

In many cases, the alternative hypothesis focuses on one particular direction

H₀: μ ≤ 3
H₁: μ > 3

This is an upper-tail test since the alternative hypothesis is focused on the upper tail above the mean of 3

H₀: μ ≥ 3
H₁: μ < 3

This is a lower-tail test since the alternative hypothesis is focused on the lower tail below the mean of 3

Upper-Tail Tests

There is only one critical value, since the rejection area is in only one tail

H₀: μ ≤ 3
H₁: μ > 3

Z = (x̄ - μ) / (σ / √n)

Z = (x̄ - μ) / (σ / √n)

Critical value

Do not reject H₀

Reject H₀

Do not reject H₀ since Z = 0.88 < 1.28

I.e., there is not sufficient evidence that the mean bill is over $52

One-Tail Tests

In many cases, the alternative hypothesis focuses on one particular direction

H₀: μ ≤ 3
H₁: μ > 3

This is an upper-tail test since the alternative hypothesis is focused on the upper tail above the mean of 3

H₀: μ ≥ 3
H₁: μ < 3

This is a lower-tail test since the alternative hypothesis is focused on the lower tail below the mean of 3

Upper-Tail Tests

There is only one critical value, since the rejection area is in only one tail

H₀: μ ≤ 3
H₁: μ > 3

Z = (x̄ - μ) / (σ / √n)

Z = (x̄ - μ) / (σ / √n)

Critical value

Do not reject H₀

Reject H₀

Do not reject H₀ since Z = 0.88 < 1.28

I.e., there is not sufficient evidence that the mean bill is over $52
There is only one critical value, since the rejection area is in only one tail.

\[
\begin{align*}
H_0: & \mu \geq 3 \\
H_1: & \mu < 3
\end{align*}
\]

\[z \leq -z_\alpha \quad \text{Reject } H_0 \]
\[z > 0 \quad \text{Do not reject } H_0 \]

In some settings, the alternative hypothesis does not specify a unique direction.

\[
\begin{align*}
H_0: & \mu = 3 \\
H_1: & \mu \neq 3
\end{align*}
\]

There are two critical values, defining the two regions of rejection:

\[z < -z_\alpha/2 \quad \text{Reject } H_0 \]
\[z > +z_\alpha/2 \quad \text{Reject } H_0 \]

Do not reject } H_0 \]

Test the claim that the true mean # of TV sets in US homes is equal to 3. (Assume \(\sigma = 0.8\))

State the appropriate null and alternative hypotheses:
- \(H_0: \mu = 3\), \(H_1: \mu \neq 3\) (This is a two tailed test)

Specify the desired level of significance:
- Suppose that \(\alpha = .05\) is chosen for this test

Choose a sample size:
- Suppose a sample of size \(n = 100\) is selected

Determine the appropriate technique:
- \(\sigma\) is known so this is a \(z\) test

Set up the critical values:
- For \(\alpha = .05\) the critical \(z\) values are \(-1.96\) and \(1.96\)

Collect the data and compute the test statistic:
- Suppose the sample results are \(n = 100\), \(x = 2.84\) (\(\sigma = 0.8\) is assumed known)
- So the test statistic is:

\[z = \frac{x - \mu_0}{\sigma/\sqrt{n}} = \frac{2.84 - 3}{0.8/\sqrt{100}} = \frac{-0.16}{0.08} = -2.0\]

Is the test statistic in the rejection region?

\[z = -2.0 < -1.96\]

Reach a decision and interpret the result:

Since \(z = -2.0 < -1.96\), we reject the null hypothesis and conclude that there is sufficient evidence that the mean number of TVs in US homes is not equal to 3.
Example: p-Value

Example: How likely is it to see a sample mean of 2.84 (or something further from the mean, in either direction) if the true mean is \( \mu = 3.0 \)?

\[ z = \frac{2.84 - 3.0}{\frac{s}{\sqrt{n}}} = -2.0 \]

\[ P(z < -2.0) = 0.0228 \]

\[ P(z > 2.0) = 0.0228 \]

\[ \text{p-value} = 0.0228 + 0.0228 = 0.0456 \]

\[ t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{172.50 - 168}{15.40/\sqrt{25}} = 1.46 \]

\[ t_{24, 0.025} = \pm 2.0639 \]

\[ t = 1.46 < 2.0639 \]

\[ \text{Do not reject } H_0: \text{not sufficient evidence that true mean cost is different than$168} \]
Proportions

Sample proportion in the success category is denoted by $\hat{p}$

$$\hat{p} = \frac{x}{n} = \frac{\text{number of successes in sample}}{\text{sample size}}$$

When $nP(1 - P) > 9$, $\hat{p}$ can be approximated by a normal distribution with mean and standard deviation

$$\mu_p = p \quad \sigma_p = \sqrt{\frac{p(1-p)}{n}}$$

Example: p-Value

Compare the p-value with $\alpha$

- If p-value < $\alpha$, reject $H_0$
- If p-value ≥ $\alpha$, do not reject $H_0$

Example: p-Value

Here: p-value = .0466

$$\alpha = .05$$

Since .0466 < .05, we reject the null hypothesis

Power of the Test

Recall the possible hypothesis test outcomes:

<table>
<thead>
<tr>
<th>Actual Situation</th>
<th>Decision</th>
<th>$H_0$ True</th>
<th>$H_0$ False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do Not Reject $H_0$</td>
<td>No error</td>
<td>Type I Error</td>
<td>Type II Error</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>(1 - $\alpha$)</td>
<td>$\beta$</td>
<td>$\alpha$</td>
</tr>
</tbody>
</table>

Key: Outcome (Probability)

- $\beta$ denotes the probability of Type II Error
- $1 - \beta$ is defined as the power of the test

Power = 1 - $\beta$ = the probability that a false null hypothesis is rejected

Type II Error

Assume the population is normal and the population variance is known. Consider the test

$H_0: \mu = \mu_0$

$H_1: \mu > \mu_0$

The decision rule is:

$$\text{Reject } H_0 \text{ if } z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} > z_{\alpha}$$

If the null hypothesis is false and the true mean is $\mu^*$, then the probability of type II error is

$$\beta = P(\bar{x} < \mu^* | \mu = \mu_0) = P \left( z < \frac{\bar{x} - \mu^*}{\sigma / \sqrt{n}} \right)$$

Type II Error Example

Type II error is the probability of failing to reject a false $H_0$

Suppose we fail to reject $H_0: \mu \geq 52$ when in fact the true mean is $\mu^* = 50$

This is the true distribution of $\bar{x}$ if $\mu = 50$

This is the range of $\bar{x}$ where $H_0$ is not rejected

Type II Error Example

Suppose we do not reject $H_0: \mu \geq 52$ when in fact the true mean is $\mu^* = 50$

This is the range of $\bar{x}$ if $\mu = 50$

This is the range of $\bar{x}$ where $H_0$ is not rejected
Type II Error Example

Suppose we do not reject $H_0: \mu \geq 52$ when in fact the true mean is $\mu^* = 50$

Here, $\beta = P(x > c)$ if $\mu^* = 50$

Calculating $\beta$

Suppose $n = 64$, $\sigma = 6$, and $\alpha = .05$

$\overline{x} = \mu_0 - Z_{\alpha} \frac{\sigma}{\sqrt{n}} = 52 - 1.645 \frac{6}{\sqrt{64}} = 50.766$

So $\beta = P(x \geq 50.766)$ if $\mu^* = 50$

Power of the Test Example

If the true mean is $\mu^* = 50$,

- The probability of Type II Error = $\beta = 0.1539$
- The power of the test = $1 - \beta = 1 - 0.1539 = 0.8461$

Probability of type II error: $\beta = .1539$

Chapter Summary

- Addressed hypothesis testing methodology
- Performed $Z$ Test for the mean ($\sigma$ known)
- Discussed critical value and p-value approaches to hypothesis testing
- Performed one-tail and two-tail tests
- Performed $t$ test for the mean ($\sigma$ unknown)
- Performed $Z$ test for the proportion
- Discussed type II error and power of the test