Introduction to Hypothesis Testing

Research Question
Is the average body temperature of healthy adults different from 98.6°F?

\[
\begin{align*}
98.2 & \quad 98.5 & \quad 98.3 & \quad 98.1 \\
98.3 & \quad 98.7 & \quad 98.1 & \quad 98.4 \\
98.2 & \quad 98.4 & \quad 98.3 & \quad 98.2
\end{align*}
\]
\[\bar{x} = 98.31 \quad s = 0.17\]

Statistical Hypothesis

Null hypothesis (\(H_0\)):
Hypothesis of no difference or no relation, often has =, \(\geq\), or \(\leq\) notation when testing value of parameters.
Example:
\[H_0: \mu = 98.6°F \quad \text{or} \quad H_0: \text{Average body temperature is 98.6°F}\]

Alternative hypothesis (\(H_A\)):
[or \(H_1\) or \(H_a\)]
Usually corresponds to research hypothesis and opposite to null hypothesis, often has >, < or \(\neq\) notation in testing mean.
Example:
\[H_A: \mu \neq 98.6°F \quad \text{or} \quad H_A: \text{Average body temperature is not 98.6°F}\]

Hypotheses Statements Example

• A researcher is interested in finding out whether average hourly salary for babysitting is different from $10.00.

\[H_0: \mu = 10 \quad \text{\(H_A: \mu \neq 10\)}\]

[Two-tailed test]

• A researcher is interested in finding out whether average life time of male is higher than 77 years.

\[H_0: \mu = 77 \quad (\text{or } \mu \leq 77) \quad H_A: \mu > 77\]

[Right-tailed test]
Hypotheses Statements Example

A researcher is interested in finding out whether the average regular gasoline price is less than $3.00 in Mid-West region.

\[ H_0: \mu = 3.00 \quad \text{or} \quad \mu \geq 3.00 \]
\[ H_A: \mu < 3.00 \]
[Left-tailed test]

One Sample Z-Test for Mean (Large sample test)

One-Sided Test

I. Hypothesis

One wishes to test whether the average body temperature for healthy adults is less than 98.6°F.

\[ H_0: \mu = 98.6°F \quad \text{v.s.} \quad H_A: \mu < 98.6°F \]

This is a one-sided test, left-side test.

Evidence

What will be the key statistic (evidence) to use for testing the hypothesis about population mean?

Sample mean: \( \bar{X} \)

A random sample of 36 subjects is chosen and the sample mean is 98.46°F and sample standard deviation is 0.3°F.

II. Test Statistic

If \( H_0: \mu = 98.6°F \) is true, sampling distribution of mean \( \rightarrow \text{Normal} \) with mean \( = 98.6 \) standard deviation \( = \frac{0.3}{\sqrt{36}} = 0.05 \).

\[ \sigma_{\bar{X}} = 0.05 \]

This implies that the statistic is 2.8 standard deviations away from the mean 98.6 in \( H_0 \), and is to the left of 98.6 (or less than 98.6).
Introduction to Hypothesis Testing for One Population Mean

**Level of Significance**

*Level of significance for the test (α)*

A probability level selected by the researcher at the beginning of the analysis that defines unlikely values of sample statistic if null hypothesis is true.

\[ \text{c.v.} = \text{critical value} \quad \text{Total tail area} = \alpha \]

**III. Decision Rule**

*Critical value approach:* Compare the test statistic with the critical values defined by significance level α, usually α = 0.05.

We reject the null hypothesis, if the test statistic \( z < -z_\alpha = -z_{0.05} = -1.64 \).

*P-value approach:* The probability of obtaining a test statistic that is as extreme or more extreme than actual sample statistic value given null hypothesis is true.

It is a probability that indicates the extremeness of evidence against \( H_0 \).

**IV. Draw conclusion**

Since from either critical value approach \( z = -2.8 < -z_\alpha = -1.64 \) or p-value approach \( p\text{-value} = .003 < \alpha = .05 \), we reject null hypothesis.

Therefore we conclude that there is sufficient evidence to support the alternative hypothesis that the average body temperature is less than 98.6°F.

**Decision Rule**

*P-value approach:* Compute p-value,

if \( H_A: \mu \neq \mu_0 \), \( p\text{-value} = 2 \cdot P(Z \geq |z|) \)

if \( H_A: \mu > \mu_0 \), \( p\text{-value} = P(Z \geq z) \)

if \( H_A: \mu < \mu_0 \), \( p\text{-value} = P(Z \leq z) \)

reject \( H_0 \) if \( p\text{-value} < \alpha \)
Introduction to Hypothesis Testing for One Population Mean

Steps in Hypothesis Testing
1. State hypotheses: $H_0$ and $H_A$.
2. Choose a proper test statistic, collect data, checking the assumption and compute the value of the statistic.
3. Make decision rule based on level of significance($\alpha$).
4. Draw conclusion.
   (Reject null hypothesis if p-value < $\alpha$.)

Errors in Hypothesis Testing
Possible statistical errors:
- **Type I error**: The null hypothesis is true, but we reject it.
- **Type II error**: The null hypothesis is false, but we don’t reject it.

Can we see data and then make hypothesis?
1. Choose a test statistic, collect data, checking the assumption and compute the value of the statistic.
2. State hypotheses: $H_0$ and $H_A$.
3. Make decision rule based on level of significance($\alpha$).
4. Draw conclusion.

One Sample t-Test for Mean

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

One-sample Test with Unknown Variance $\sigma^2$

In practice, population variance is unknown most of the time. The sample standard deviation $s^2$ is used instead for $\sigma^2$. If the random sample of size $n$ is from a normal distributed population and if the null hypothesis is true, the test statistic (standardized sample mean) will have a t-distribution with degrees of freedom $n-1$.

Test Statistic: $$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

I. State Hypothesis

One-side test example:
If one wish to test whether the body temperature is less than 98.6 or not.
$H_0$: $\mu = 98.6$ v.s. $H_A$: $\mu < 98.6$

(Left-sided Test)
II. Test Statistic

If we have a random sample of size 16 from a normal population that has a mean of 98.46°F, and a sample standard deviation 0.2. The test statistic will be a \( t \)-test statistic and the value will be: (standardized score of sample mean)

\[
\text{Test Statistic: } t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{98.46 - 98.6}{0.2} = -2.8
\]

Under null hypothesis, this \( t \)-statistic has a \( t \)-distribution with degrees of freedom \( n - 1 \), that is, \( 15 = 16 - 1 \).

III. Decision Rule

Critical Value Approach:

To test the hypothesis at \( \alpha \) level 0.05, the critical value is \(-t_{\alpha} = -t_{0.05} = -1.753\).

Decision Rule: Reject null hypothesis if \( t < -1.753 \).

III. Decision Rule

Decision Rule: Reject null hypothesis if \( p \)-value < \( \alpha \).

p-value Calculation

\( p \)-value corresponding the test statistic:

For \( t \) test, unless computer program is used, \( p \)-value can only be approximated with a range because of the limitation of \( t \)-table.

\[
\text{\( p \)-value} = P(T < -2.8) < P(T < -2.602) = 0.01
\]

Since the area to the left of -2.602 is .01, the area to the left of -2.8 is definitely less than 0.01.

IV. Conclusion

Decision Rule:

If \( t < -1.753 \), we reject the null hypothesis, or if \( p \)-value < 0.05, we reject the null hypothesis.

Conclusion: Since \( t = -2.8 < -1.753 \), or say \( p \)-value < 0.01 < 0.05, we reject the null hypothesis. There is sufficient evidence to support the research hypothesis that the average body temperature is less than 98.6°F.
What if we wish to test whether the average body temperature is different from 98.6°F using t-test with the same data?

The \( p \)-value is equal to twice the \( p \)-value of the left-sided test which will be less than .02.

Decision Rule

\[ p\text{-value approach: Compute } p\text{-value,} \]

- if \( H_A : \mu \neq \mu_0 \), \( p\)-value = \( 2 \cdot P( T \geq |t|) \)
- if \( H_A : \mu > \mu_0 \), \( p\)-value = \( P( T \geq t ) \)
- if \( H_A : \mu < \mu_0 \), \( p\)-value = \( P( T \leq t ) \)

reject \( H_0 \) if \( p\)-value < \( \alpha \)

When do we use this t-test for testing the mean of a population?

- A random sample from normally distributed population with unknown variance.
- or
- When sample size is relatively large the t-score is approximately equal to z-score therefore t-test will be almost the same as z-test.

Remarks

- If the sample size is relatively large (>30) both z and t tests can be used for testing hypothesis.
- “t-test” is robust against normality.
- If the sample size is small and the sample is from a very skewed or other non-normal distribution, we can use nonparametric alternatives Signed-Rank Test.
- Many commercial packages only provide t-test since standard deviation of the population is often unknown.

Average Weight for Female Ten Year Children In US

Info. from a random sample:

\[ \bar{x} = 80 \text{ lb,} \]
\[ s = 18.05 \text{ lb.} \]

Is average weight greater than 78 lb at \( \alpha = 0.05 \) level?

\( H_0 : \) Mean is 78 lb
\( H_a : \) Mean is greater than 78 lb

Sampling Distribution

\[ \text{S.E.} = \frac{18.05}{\sqrt{10}} = 5.71 \]
\[ n = 10 \]
\[ \bar{x} \]

\[ \text{S.E.} = \frac{18.05}{\sqrt{400}} = 0.90 \]
\[ n = 400 \]

Practical Significance? 78 80
**Type I & Type II Error**

- **Type I Error**: reject the null hypothesis when it is true. The probability of a Type I Error is denoted by $\alpha$.
- **Type II Error**: accept the null hypothesis when it is false and the alternative hypothesis is true. The probability of a Type II Error is denoted by $\beta$.

$$1 - \beta : \text{Power of the test}$$

Power of the Test

The power of the statistical test is the ability of the study as designed to distinguish between the hypothesized value and some specific alternative value. That is, the power is the probability of rejecting the null hypothesis if the null hypothesis is false.

$$\text{Power} = P(\text{reject } H_0 \mid H_0 \text{ is false})$$

$$\text{Power} = 1 - \beta$$

The power is usually calculated given a simple alternative hypothesis: $H_a: \mu = 211 \text{ mg/100ml}$

$$\text{Power} = P(\text{reject } H_0 \mid H_a \text{ is true})$$

For a two-sided test, the formula is

$$n = \left[ \frac{(z_{\alpha/2} + z_\beta) \cdot \sigma}{\mu_a - \mu_0} \right]^2$$

Or, if $\delta$ is the effect size, in terms of number of standard deviations, to be detected, the sample size would be

$$n = \left[ \frac{(z_{\alpha/2} + z_\beta)}{\delta} \right]^2$$

**Power and Sample Size in Testing One Mean**

<table>
<thead>
<tr>
<th>Decision</th>
<th>Null hypothesis</th>
<th>Type I Error</th>
<th>Type II Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td>$\alpha$</td>
<td>$1 - \beta$</td>
<td></td>
</tr>
<tr>
<td>Do not reject $H_0$</td>
<td>$1 - \alpha$</td>
<td>$\beta$</td>
<td></td>
</tr>
</tbody>
</table>

**Formula**: The estimated sample size for a one-sided test with level of significance $\alpha$ and a power of $1 - \beta$ to detect a difference of $\mu_a - \mu_0$ is,

$$n = \left[ \frac{(z_{\alpha} + z_\beta) \cdot \sigma}{\mu_a - \mu_0} \right]^2$$

Or, if $\delta$ is the effect size, in terms of standard deviations, to be detected, the sample size

$$n = \left[ \frac{(z_{\alpha} + z_\beta)}{\delta} \right]^2$$
Example:

For testing hypothesis that

$$H_0: \mu = \mu_0 = 70 \quad \text{v.s.} \quad H_a: \mu \neq 70$$

(Two-sided Test)

with a level of significance of $\alpha = .05$. Find the sample size so that one can have a power of $1 - \beta = .90$ to reject the null hypothesis if the actual mean is 5 unit different from $\mu_0$. The standard deviation, $\sigma$, is approximately equal 15.

$$n = \left[ \frac{(z_{\alpha/2} + z_{\beta}) \cdot \sigma}{\mu_a - \mu_0} \right]^2$$

$$\alpha = .05$$

$1 - \beta = .90$ \Rightarrow $\beta = .10$

$$z_{\alpha/2} = z_{.025} = 1.96$$

$$z_{\beta} = z_{.1} = 1.28$$

$$n = \left[ \frac{(1.96 + 1.28) \cdot 15}{5} \right]^2 = 94.48 \approx 95$$

The sample size needed is 95.