Chapter 6 Hypothesis Testing

Key words:

• Null hypothesis H_{0} , Alternative hypothesis H_{A} , testing hypothesis, test statistic, P-value

Hypothesis Testing

 One type of statistical inference, estimation, was discussed in Chapter 5.

 The other type, hypothesis testing, is discussed in this chapter.

Definition of a hypothesis

It is a statement about one or more populations.

It is usually concerned with the parameters of the population. e.g. the hospital administrator may want to test the hypothesis that the average length of stay of patients admitted to the hospital is 5 days

Definition of Statistical hypothesis

- They are hypothesis that are stated in such a way that they may be evaluated by appropriate statistical techniques.
- There are two hypotheses involved in hypothesis testing
- Null hypothesis H₀: It is the hypothesis to be tested.
- Alternative hypothesis H_A: It is a statement of what we believe is true if our sample data cause us to reject the null hypothesis

DEFINITION

The level of significance α is a probability and, in fact, is the probability of rejecting a true null hypothesis.

Condition of Null Hypothesis

Possible Action

	True	False
Fail to reject H_0	Correct action	Type II error
Reject H_0	Type I error	Correct action

Conclusion. If H_0 is rejected, we conclude that H_A is true. If H_0 is not rejected, we conclude that H_0 may be true.

7.2 Testing a hypothesis about the mean of a population:

- We have the following steps:
- **I.Data**: determine variable, sample size (n), sample mean($\bar{\chi}$), population standard deviation or sample standard deviation (s) if is unknown
- 2. Assumptions: We have two cases:
- Case I: Population is normally or approximately normally distributed with known or unknown variance (sample size n may be small or large),
- Case 2: Population is not normal with known or unknown variance (n is large i.e. n≥30).

- 3. Hypothesis:
- we have three cases
- Case $I: H_0: \mu = \mu_0$ $H_A: \mu \neq \mu_0$
- e.g. we want to test that the population mean is different than 50
- Case II : H_0 : $\mu = \mu_0$ H_A : $\mu > \mu_0$
- e.g. we want to test that the population mean is greater than 50
- Case III : $H_{0:} \mu = \mu_0$ $H_A: \mu < \mu_0$
- e.g. we want to test that the population mean is less than 50

Testing hypothesis for the mean μ:

When the value of sample size (n):



population is normal or not normal $(n \ge 30)$



population is normal (n < 30)



σ is known



σ is not known



σ is known



σ is not known

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \qquad Z = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

$$T = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$

Test Procedures:

Test Hocedules.					
Hypotheses	H_o : $\mu = \mu_o$	H_o : $\mu \le \mu_o$	H_o : $\mu \ge \mu_o$		
	H_A : $\mu \neq \mu_o$	H_A : $\mu > \mu_o$	H_A : $\mu < \mu_o$		
Test Statistic	Calculate the value of: $Z = \frac{\overline{X} - \mu_0}{\sqrt{\overline{C}}} \sim N(0,1)$				
(T.S.)	Calculate the value of: $Z = \frac{1}{\sigma/\sqrt{n}}$				
R.R. & A.R.					
of H _o	1=a		/ \		
	$\alpha/2$ $A.R. of H0 \alpha/2$	$\int 1-\alpha \alpha$	α 1-α		
	R.R. Z α/2 Z _{1-α/2} R.R. of H _o	A.R. of Ho Z1- a ef H	R.R. 7 A.R. of Ho		
	$= -Z_{\alpha/2}$	A.R. of H_0 $Z_{1-\alpha}$ of H_0 = - Z_{α}	R.R. Z _α A.R. of H _o		
Critical	$Z_{\alpha/2}$ and $-Z_{\alpha/2}$	$Z_{1-\alpha} = -Z_{\alpha}$	Ζα		
value (s)					
Decision:	We reject H _o (and accept H _A) at the significance level α if:				
	$Z < Z_{\alpha/2}$ or	$Z > Z_{1-\alpha} = -Z_{\alpha}$	$Z < Z_{\alpha}$		
	$Z > Z_{1-\alpha/2} = -Z_{\alpha/2}$				
	Two-Sided Test	One-Sided Test	One-Sided Test		

Test Procedures:

1 CSt 1 TOCCULICS.					
Hypotheses	H_o : $\mu = \mu_o$	H_o : $\mu \le \mu_o$	H_o : $\mu \ge \mu_o$		
	$H_A \mu \neq \mu_o$	H_A : $\mu > \mu_o$	H_A : $\mu < \mu_o$		
Test Statistic (T.S.)	Calculate the value of: $t = \frac{\overline{X} - \mu_o}{S / \sqrt{n}} \sim t(n-1)$				
	$(\mathrm{df} = v = n - 1)$				
R.R. & A.R. of H _o	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	A.R. of H ₀ $t_{1} \cdot \alpha \text{of H}_{0}$ $= -t_{\alpha}$	α 1-α R.R. of Ho		
Critical value (s)	$t_{\alpha/2}$ and $-t_{\alpha/2}$	$t_{1-\alpha} = -t_{\alpha}$	t _a		
Decision:	We reject H _o (and accept H _A) at the significance level α if:				
	$t < t_{\alpha/2}$ or	$t > t_{1-\alpha} = -t_{\alpha}$	$t < t_{\alpha}$		
	$t > t_{1-\alpha/2} = -t_{\alpha/2}$ Two-Sided Test	One-Sided Test	One-Sided Test		

The Use of P - Values in Decision Definition Making::

Note: Using P- Value as a decision tool:

P-value is the smallest value of α for which we can reject the null hypothesis H_o .

Calculating P-value:

- * Calculating P-value depends on the alternative hypothesis H_A.
- * Suppose that $Z_c = \frac{\overline{X} \mu_o}{\sigma / \sqrt{n}}$ is the computed value of the test Statistic.
- * The following table illustrates how to compute P-value, and how to use P-value for testing the null hypothesis:

Alternative Hypothesis:	H_A : $\mu \neq \mu_o$	H_A : $\mu > \mu_o$	H_A : $\mu < \mu_o$
P-Value =	$2\times P(Z> z_c)$	$P(Z > Z_c)$	$P(Z \leq Z_c)$
Significance Level =	α		
Decision:	Reject H _o if P-value $< \alpha$.		

Example:

For the previous example, we have found that:

$$Z_C = \frac{\overline{X} - \mu_o}{\sigma / \sqrt{n}} = 2.02$$

The alternative hypothesis was H_A : $\mu > 70$.

$$P-Value = P(Z > Z_C)$$

= $P(Z > 2.02) = 1 - P(Z < 2.02) = 1 - 0.9783 = 0.0217$

The level of significance was $\alpha = 0.05$.

Since P-value $< \alpha$, we reject H_o.

6.Decision :

- If we reject H_0 , we can conclude that H_A is true.
- If ,however ,we do not reject H_0 , we may conclude that H_0 is true.

An Alternative Decision Rule using the p - value Definition

 The p-value is defined as the smallest value of α for which the null hypothesis can be rejected.

- If the p-value is less than or equal to α , we reject the null hypothesis ($p \le \alpha$)
- If the p-value is greater than α , we do not reject the null hypothesis (p > α)

Example 7.2.1 Page 223

- Researchers are interested in the mean age of a certain population.
- A random sample of 10 individuals drawn from the population of interest has a mean of 27.
- Assuming that the population is approximately normally distributed with variance 20,can we conclude that the mean is different from 30 years ? (α =0.05).
- If the p value is 0.0340 how can we use it in making a decision?

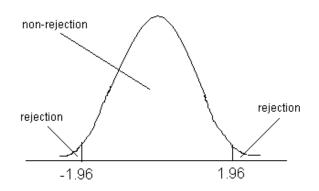
Solution

- **I-Data:** variable is age, n=10, $\bar{x} = 27$, $\sigma^2 = 20$, $\alpha = 0.05$
- 2-Assumptions: the population is approximately normally distributed with variance 20
- **3-Hypotheses:**
- $H_0: \mu=30$
- H_A : $\mu \neq 30$

4-Test Statistic:

$$\mathbf{Z} = -2.12$$

5. Decision Rule



The alternative hypothesis is

$$H_A$$
: $\mu \neq 30$

• Hence we reject H_0 if $Z > Z_{1-0.025} = Z_{0.975}$

or
$$Z < -Z_{1-0.025} = -Z_{0.975}$$

Z_{0.975}=1.96(from table D)

6.Decision:

• We reject H_0 , since -2.12 is in the rejection region.

We can conclude that µ is not equal to 30

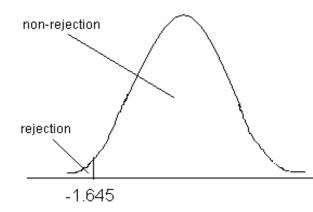
Using the p value, we note that p-value
 =0.0340< 0.05, therefore we reject H0

Example 7.2.2 page 227

- Referring to example 7.2.1. Suppose that the researchers have asked: Can we conclude that μ <30.
- I.Data.see previous example
- 2. Assumptions . see previous example
- 3. Hypotheses:
- $H_0 \mu = 30$
- H_{οA}: μ < 30

4. Test Statistic:

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{27 - 30}{\sqrt{\frac{20}{10}}} = -2.12$$



- 5. Decision Rule: Reject H_0 if $Z < -Z_{1-\alpha}$, where
- $\mathbf{Z}_{1-\alpha} = 1.645$. (from table D)
- 6. **Decision**: Reject H_0 , thus we can conclude that the population mean is smaller than 30.

Example 7.2.4 page 232

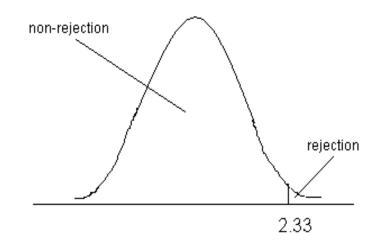
 Among I57 African-American men ,the mean systolic blood pressure was 146 mm Hg with a standard deviation of 27. We wish to know if on the basis of these data, we may conclude that the mean systolic blood pressure for a population of African-American is greater than 140. Use $\alpha = 0.01$.

Solution

- I. Data: Variable is systolic blood pressure, n=157, =146, s=27, $\alpha=0.01$.
- **2. Assumption:** population is not normal, σ^2 is unknown
- 3. Hypotheses: $H_0: \mu=140$ $H_A: \mu>140$
- 4. Test Statistic:

5. Decision Rule:

we reject H_0 if $Z>Z_{1-\alpha}$ = $Z_{0.99}$ = 2.33 (from table D)



6. Decision: We reject H₀.

Hence we may conclude that the mean systolic blood pressure for a population of African-American is greater than 140.

Exercises

Escobar performed a study to validate a translated version of the Western Ontario and McMaster University index (WOMAC) questionnaire used with spanish-speaking patient s with hip or knee osteoarthritis. For the 76 women classified with sever hip pain. The WOMAC mean function score was 70.7 with standard deviation of 14.6, we wish to know if we may conclude that the mean function score for a population of similar women subjects with sever hip pain is less than 75. Let $\alpha = 0.01$

Solution:

I.Data:

2. Assumption:

3. Hypothesis:

4. Test statistic:

5.Decision Rule

6. Decision:

Exercises

Q7.2.3:

The purpose of a study by Luglie was to investigate the oral status of a group of patients diagnosed with thalassemia major (TM). One of the outcome measure s was the decayed, missing, filled teeth index (DMFT). In a sample of 18 patients, the mean DMFT index value was 10.3 with standard deviation of 7.3. Is this sufficient evidence to allow us to conclude that the mean DMFT index is greater than 9 in a population of similar subjects?

Let $\alpha = 0.1$

Solution:

I.Data:

2. Assumption:

3. Hypothesis:

4. Test statistic:

5.Decision Rule

6. Decision:

For <u>Q7.2.3:</u>

Take the p- value = 0.22, Use the P-value to make your decision ??