Chap 9 Notes:
Hypothesis Testing

J. Harner A. Billings

Department of Statistics
West Virginia University

Stat 211 Fall 2007

Outline

Chap.9 Notes: Hypothesis Testing
Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

The Scientific Method

1. State the problem.
2. Formulate the hypothesis.
3. Design the experiment or survey.
4. Interpret the data.
5. Draw conclusions.

Hypothesis Testing - Introduction:

- Two major areas of statistics:
    1. Descriptive ($\bar{X}$, S , graphs, etc.)
    2. Inference
- Hypothesis testing is a procedure in which we use data to decide which of two hypotheses is more likely to be true.
- Hypothesis testing allows us to make an inference.

7 steps in performing a Test of Hypothesis:

1. State **null hypothesis**, H_0.
2. State **alternate hypothesis**, H_1.
3. Decide on **significance level**, α.
4. Calculate the appropriate **test statistics**.
5. Use tables to find the **"p - value"** of the test statistics.
6. Make decision.
7. State conclusions.
   (and assumptions, if any)

Hypothesis

A hypothesis is a statement about a population parameter.

- A hypothesis may be either true or false.
- In hypothesis testing, we must decide which of two mutually exclusive hypotheses is supported by the data.
- The null hypothesis is labeled as H_0.
- The alternate (or research) hypothesis is H_1 or H_A.

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

- The null hypothesis H_0 typically expresses the idea of **"no difference"**,**"no change"** or **"equality"**.
- H_0 typically contains an = ($\geq$ or $\leq$) sign.
- The alternate hypothesis H_1 expresses the idea of **"some difference"** or **"some change"** or **"inequality"**.
- H_A typically contains an inequality symbol ($<$ or $>$ or $\neq$).

Example:

a. The mean age of all college students is 21 years.

$$H_0 : \mu = 21$$

b. The mean age of all college students is not 21 years.

$$H_A : \mu \neq 21$$

(two - sided alternative)

c. mean age of all college students is less than 21 years.

$$H_A : \mu < 21$$

d. Mean age of all college students is greater than 21 years.

$$H_A : \mu > 21$$

- **c. , d.** are examples of one sided alternatives .
- H_0, H_A are specified before any data is collected.

- The null hypothesis H_0 may also state:

  H_0 : the observed results are due to chance.

- The alternate hypothesis H_A may say:

  H_A: the observed results are due to the effect of some treatment.

Example:

Randomized Controlled Experiment

$$\begin{aligned} \text{Treatment} &\quad - \quad \text{take aspirin} \\ \text{Control} &\quad - \quad \text{take placebo} \end{aligned}$$

H_0 : Any difference in rate of heart attack is due to chance

So,

$$H_0 : p_T = p_C$$

H_A : Aspirin reduces the rate of heart attack

So,

$$H_A : p_T < p_C$$

Hypothesis Test on Binomial Probability p

- We can test hypothesis involving the **binomial parameter** p.

  Example:

  Some people "sleep in" on weekends to make up for the sleep deficits during the week. The "Better Sleep Council" reports that 61% of people get more than 7 hours of sleep on Saturday night. A random sample of 350 adults found that 235 had more than 7 hours sleep last Saturday.

- Does this evidence support the Better Sleep Council's claim.
  (Use α = 0.5)

(Contd.)

- First note that this is a binomial experiment. Sleeping more than 7 hours is a "Success", and binomial parameter p represents the proportion of people who sleep more than 7 hours.
    1. $H_0 : p = 0.61$
    2. $H_1 : p = 0.61$
    3. $\alpha = 0.05$
- Before we, proceed we must check to see if:
    1. $n.p > 5 \quad \text{and}$
    2. $n(1 - p) > 5$

(Contd.)

- In our example, the null hypothesis states $H_0 : p = 0.61$, which is the value of p we use to check the two conditions above.

$$\begin{aligned} n.p = 350.(0.61) = 213.5 &> 5 \\ n.(1-p) = 350.(1-0.61) = 136.5 &> 5 \end{aligned}$$

- So both conditions are satisfied, we can proceed with our hypothesis test on p.

4. Test Statistics:

- The test statistics is calculated using statistic(s) that are computed from the data from our sample. We will also use the the hypothesized value of the population parameter - we find this value in H_0.
- For a hypothesis test on p, our test statistics has the form

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

where,

$$\begin{aligned} n &= \text{number of observations(trials)} \\ X &= \text{number of Successes} \\ \hat{p} &= \frac{X}{n} \\ p &= \text{hypothesized value in } H_0 \end{aligned}$$

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

- For our example:

$$\begin{aligned} n &= 350 \\ X &= 235 \\ \hat{p} &= \frac{X}{n} = \frac{235}{350} = 0.6714 \\ p &= 0.61 \end{aligned}$$

- Our test statistics is computed as:

$$\begin{aligned} z &= \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \\ &= \frac{0.6714 - 0.61}{\sqrt{\frac{.061(1-0.61)}{350}}} \\ &= \frac{0.0314}{\sqrt{0.0006797143}} \\ &= \frac{0.0614}{0.0260713} \\ &= 2.3551 \end{aligned}$$

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

5. Find the p - value of the test - statistics:

- The p - value of the test - statistic gives us the probability of getting a value of $\hat{p}$ as larger, or larger, as the value of $\hat{p}$ we calculated from the sample; under the assumption that the true population parameter's value is p (in H_0).
- For our example, the p - value will tell us the probability of getting a value of 0.671 for $\hat{p}$ if indeed the true value for the population parameter p is 0.61(in H_0).

(Contd.)

- we do this by finding

$$\begin{aligned} P(z \geq 2.35) &= 1 - P(z < 2.35) \\ &= 1 - 0.9906 \\ &= 0.0096 \end{aligned}$$

- So it is not likely that we would observe $\hat{p} = 0.671$ if p is really 0.61.
- This should cause us to suspect that the H_0 : p = 0.61 is false.

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

6. Make the decision:

- We will decide to reject (disbelieve)

$$H_0 \text{ **if p - value** } < \alpha$$

  Otherwise, we will not reject H_0

  - Since our p - value < α (i.e., 0.0096 < 0.05) we **reject** H_0.

7. State the conclusion - and assumptions:

- State the conclusion and any assumptions made in order to perform this hypothesis test. The conclusion should also indicate the significance level of the hypothesis test.

(Contd.)

- For our example, we would write the conclusion as follows:

  At the 5% significance level, there is evidence that more than 61% of people "sleep-in", i.e., get 7 hours or more sleep on Saturday night.

- **Note:** The conditions required to perform this test were:
  1. binomial experiment.
  2. n.p>5
  3. n.(1-p)>5

- We checked these conditions, and all conditions were satisfied.

Example:

See Example 9.6[1]

[1] refer to Page: 509 of the Text Book

Lets put all 7 steps together:

Example:

The West Virginia Dept. of revenue states that 20% of West Virginians have income below "poverty level". A.B. suspects that this percentage is lower than 20% in Monongalia County. he obtains a random sample of 400 Monongalia County residents and finds that 70 are living below the "poverty line".

- Does this data support A.B.'s suspicion that less than 20% of Monongalia County residents live below the poverty level? (Use a significance level of 0.05).

1. H_0 : p = 0.20
2. H_1 : p < 0.20
3. α = 0.05
4.

$$
\begin{aligned}
n &= 400 \\
X &= \text{number of Successes} \\
&= \text{number of residents below poverty line} \\
&= 70 \\
\hat{p} &= \frac{X}{n} = \frac{70}{400} = 0.175
\end{aligned}
$$

- The test statistics is

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

$$
\begin{aligned}
z &= \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \\
&= \frac{0.175 - 0.20}{\sqrt{\frac{0.02(1-0.20)}{400}}} \\
&= \frac{-0.025}{\sqrt{0.0004}} \\
&= \frac{-0.025}{0.02} \\
&= -1.25
\end{aligned}
$$

(Contd.)

5. Find the value of the test statistics, i.e., find the probability of getting $\hat{p}$ = 0.175 or lower, if the true proportion is p = 0.20.
   - Do this by finding

$$\begin{aligned} P(\hat{p} \leq 0.175 | p = 0.20) &= P(z \leq -1.25) \\ &= 0.1056 \end{aligned}$$

   - value of z can be obtained from Table: E of the Text Book[2].
   - **Note:** the tabulated value of z in Table: E corresponding to -1.25 should be considered as 0.1056 not as 0.0156.

[2] refer to Page: 699 of the Text Book

(Contd.)

6. Make decision:

- Reject H_0 if p - value $\leq \alpha$;
  Otherwise, do not reject H_0.
  - Since p - value $\nleq \alpha$

    i.e., p - value > α
    i.e., 0.1056 > 0.05
- **We do not reject H_0.**

7. State conclusion:

- At the 5% significance level,
  the poverty level in Monongalia County is not significantly different from the West Virginia poverty rate of 20%.

Chap.9 Notes: Hypothesis Testing
Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

Hypothesis test on μ- (Large Sample) Data from Normal Population, σ unknown:

Example

A company owns a fleet of cars with mean MPG known to be $\mu = 30$. This company will use a gasoline additive only if the additive increases gasoline MPG.

A sample of 36 cars used the additive. The sample mean was calculated to be $\bar{X}$ = 3.13 miles with standard deviation s = 7.0 miles.

- Does the additive significantly increase mean MPG?
  (Use α = 0.10)

1. $H_0 : \mu = 30$
2. $H_1 : \mu > 30$
3. $\alpha = 0.10$
4. n = 36 $\quad \bar{X}$ = 31.3 $\quad$ s = 7.0

$$T = \frac{\bar{X} - \mu_0}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{31.3 - 30.0}{\left(\frac{7}{\sqrt{36}}\right)} = \frac{1.3}{1.17} = 1.11$$

5. The p - value is the probability of getting a value of 1.11 or larger for the test statistics (since $H_A : \mu > 30$).The test statistics follows a t - distribution.
   But since n is large the CLT tells us that this t - distribution will be nearly identical to a standard normal distribution. So we can use the standard normal distribution (Table: E) to find our p - value.

(Contd.)

$$i.e. \quad P(T \geq 1.11) \simeq P(z \geq 1.11) = 0.1335$$

6. Reject H_0 if

$$\text{p - value} \leq \alpha$$

- Since $0.1335 \nleq 0.10$
  We do not reject H_0.

7. At the 10% significance level we conclude that the gasoline additive does not significantly increase MPG of the entire fleet of cars.

- **Note:** Some books call this a z - test, since we use a standard normal distribution to find the p - value.

  Assumptions:

- random sample of independent observations.

Hypothesis Test on μ - Data from Normal Population, σ known:

- The following procedure is appropriate when σ is known (unlikely) and the data comes from a Normal population.
  The sample size n may be large or small, i.e., for any sample size $n \geq 2$

  Example:

  A CNC machine (computerized numeric control) is set to produce rods 10 cm. long. The manufacturer of the CNC machine states that the standard deviation of the rods will always be 0.75 cm, no matter the length of the rods produced.

(Contd.)

A production technician believes too many "short" rods are being produced, indicating that the CNC machine needs re-calibration.

A random sample of 9 rods produced a sample mean equal to 9.34 cm.

- Does the data support the technician's belief?
  You may assume that the rod lengths are randomly distributed.
  (Use α = 0.05)

1. $H_0 : \mu = 10$
2. $H_A : \mu < 10$
2. $\alpha = 0.05$
4. **Test - statistics**

$$\begin{aligned} z &= \frac{\bar{X} - \mu_0}{\left(\frac{\sigma}{\sqrt{n}}\right)} \\ &= \frac{9.34 - 10}{\left(\frac{0.75}{\sqrt{9}}\right)} \\ &= \frac{-0.66}{0.25} \\ &= -2.64 \end{aligned}$$

5. **p- value of Test Statistics**
   Since $H_A : \mu < 10$, we must find the probability of getting a test - statistics value -2.64 or smaller,

$$i.e., \qquad P(z \leq -2.64) = 0.0041$$

6. **Decision**
   Reject H_0, if
   $$\text{p - values} \leq \alpha$$
   Since
   $$0.0041 \leq 0.05$$
   we **reject** H_0
7. **Conclude**
   At the 5% significance level,
   we conclude that the mean length of rods produced by by the CNC machine is significantly less than 10 cm.

- **Note:** based on this conclusion, we will schedule the CNC machine for re-clibration.

- **Assumptions:**

1. Random sample of Independent observations from a Normal Distribution.
2. σ is known.

Hypothesis Test on μ - (Small Sample) Data From Normal Population, σ unknown:

Example:

Automobile exhaust contains an average of 90 parts per million (ppm) of carbon monoxide. A new pollution control device is placed on ten randomly selected cars. For these 10 cars, mean carbon monoxide emission was 75 ppm, with a standard deviation of 20 ppm.

- Does the new pollution control device significantly reduce carbon monoxide emission?

  You may assume the data comes from a Normal Distribution.
  (Use α = 0.05)

1. $H_0 : \mu = 90$
2. $H_A : \mu < 90$
3. $\alpha = 0.05$
4. **Test - statistics:**

$$\begin{aligned} T &= \frac{\bar{X} - \mu_0}{\left(\frac{S}{\sqrt{n}}\right)} \\ &= \frac{75 - 90}{\left(\frac{20}{\sqrt{10}}\right)} \\ &= \frac{-15}{6.3246} \\ &= -2.37 \end{aligned}$$

5. **p - value of text statistics:**
   Since $H_A : \mu < 90$, we must find the probability of getting a value of -2.37 or smaller for test statistics, i.e.,

$$P(T \leq -2.37)$$

using a t - distribution with d.f. = n - 1 = 9.

(Contd.)

- **Note:** In this example, we can not use a standard normal distribution to find this probability since n is **NOT** large.
- Looking in the t - dist. Table F[3]along the the row for 9 d.f., we try to find -2.37, but notice that all values in the t - table are positive. But we can use symmetry of t - dist. to help us. Note that 2.37 is between 2.262 (under the column labeled 0.025) and 2.398 (under the column labeled 0.02).
  So our p - value is somewhere between 0.025 and 0.02.

[3]refer to Page: 701 of the Text Book

(Contd.)

6. **Decision:**

   Reject H_0 , if

   $$\text{p - value} \leq \alpha$$

   We know that our p- value is

   $$\text{p - value} < 0.025$$

   and since

   $$0.025 < 0.05 (\text{our } \alpha)$$

   then

   $$\text{p - value} \leq 0.05$$

   So, we reject H_0.

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

7. At the 5% significance level, the evidence indicates that the new pollution control device does significantly reduce carbon monoxide emission.

- **Note:** Most statistical computer software produces exact p - values.

- **Assumptions:**

  Random sample of independent observations from a Normal Distribution.

Hypothesis Test - Mean of Paired Differences μ_d:

1. **Small Sample Case**

   Example:

   Manufacturer of Brand A tires claims that their tires last longer than Brand B tires. One tire of each brand is mounted on the front ends 6 randomly selected cars.
   After 10,000 miles, tire wear (i.e., amount of tread reduction) was measured (thousandths of in.)

   - Does the data support the claim of the Brand A manufacturer?(Use $\alpha = 0.05$)

- In order to use the following test, the differences must be Normally Distributed.
  We will assume the differences are Normally distributed.
  We will compute the differences as d = A - B for each car. Thus, a negative value of d indicates less wear for brand A tire.

1. $H_0 : \mu_d = 0$
2. $H_0 : \mu_d < 0$
3. $\alpha = 0.05$

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

4. **Collect data; Calculate test statistics**

Car	A	B	d = A - B
1	125	133	-8
2	64	65	-1
3	94	103	-9
4	38	37	1
5	90	102	-12
6	106	115	-9

$$\bar{d} = \frac{\sum d}{n} = \frac{-38}{6} = -6.33$$

$$\begin{aligned} S_d &= \sqrt{\frac{\sum d^2 - \frac{\left(\sum d\right)^2}{n}}{n-1}} \\ &= \sqrt{\frac{372 - \frac{\left(-38\right)^2}{6}}{5}} \\ &= 5.13 \end{aligned}$$

(Contd.)

- The **test statistics** is

$$T = \frac{\bar{d} - \mu_d}{\left(\frac{S_d}{\sqrt{n}}\right)} = \frac{-6.33 - 0}{\left(\frac{5.13}{\sqrt{6}}\right)} = -3.14$$

5. **p - value**

   We want to find

$$P(T \leq -3.14)$$

   (since $H_0 : \mu_d < 0$)

- By symmetry of t- distribution, this is the same as P(T$\geq$ 3.14). Looking in t - table F along the row for d.f. = n - 1= 5 try to find 3.14.

Chap.9 Notes: Hypothesis Testing
Sect 9.1:The Null Hypothesis And The Alternative Hypothesis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

- We notice that 3.14 is between is between 2.757 (under the column labeled 0.02) and 3.365 (under the column labeled 0.01). So our p - values is between 0.01 and 0.02.

6. **Decision**

   Reject H_0 if

   $$\text{p - value} \leq \alpha$$

   Our p - value is between 0.01 and 0.02. Our p - value is clearly less than α = 0.05.
   We **reject** H_0.

7. At the 5% significance level,
the data indicates that Brand A tires wear out less quickly than Brand B tires, i.e., Brand A tires last longer than Brand B tires.

- **Assumptions:**

  Random sample of independent observations from a Normal Distribution.

Hypothesis Test - Mean of Paired Differences μ_d:

1. **Large Sample Case**

   Example:

   A random sample of 36 persons was put on an exercise program to improve strength. Before the program each person's strength was measured. After 10 weeks of exercise each person's strength was measured again.

   - Did the exercise program result in an increase in strength?

- **Note:** Paired Differences arise naturally in "Before and After" studies.

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

- **Note:** We will compute the differences as

$$d = \text{After - Before}$$

for each person. Thus a positive value for d indicates a gain in strength.

1. $H_0 : \mu_d = 0$
2. $H_0 : \mu_d > 0$
3. $\alpha = 0.05$
4. **Collect data; Calculate test statistics**

Subject	Before	After	d = A - B
1	87	92	5
2	72	74	2
⋮	⋮	⋮	⋮
36	48	45	-3

- From the sample data, we compute

$$\bar{d} = \frac{\sum d}{n} = \ldots = 2.2$$

$$S_d = \sqrt{\frac{\sum d^2 - \frac{\left(\sum d\right)^2}{n}}{n-1}} = \ldots = 4.5$$

- The test statistic is

$$\begin{aligned} T &= \frac{\bar{d} - \mu_d}{\left(\frac{S_d}{\sqrt{n}}\right)} \\ &= \frac{2.2 - 0}{\left(\frac{4.5}{\sqrt{36}}\right)} \\ &= \frac{2.2}{0.75} \\ &= 2.93 \end{aligned}$$

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

5. **p - value**
Find probability of getting 2.93 or larger for the value of test statistics, if indeed H_0 is true. So

$$\begin{aligned} P(T \geq 2.93) \simeq P(z \geq 2.93) &= 1 - P(z < 2.93) \\ &= 1 - 0.9983 \\ &= 0.0017 \end{aligned}$$

6. **Decision**

Reject H_0 if

$$\text{p - value} \leq \alpha$$

Since $0.0017 < 0.01$, we **reject** H_0.

7. **Conclude**
At the 1% significance level, there was an increase in strength due to the exercise program.

- **Note:** Since n is large, no need to assume data comes from Normal Distribution.

Hypothesis Test - Difference in Independent Population Means $\mu_X - \mu_Y$

I. **Small Sample Case** (n<20 and m<20)

Example:

A researcher wishes to asses a "new" teaching method for "slow learners". A random sample of 8 students use the new method and a random sample of 12 students use the "standard" teaching method. After 6 months, an exam is administered to each student.

- Does the data indicate that the new teaching method is preferable? (Use α = 0.05)

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

Exam Scores:

New Method: (Group 1)

$$80, 76, 80, 66, 79, 79, 81, 76$$

$$\bar{X} = 77.125, S_X = 4.853, n = 8$$

Standard Method: (Group 2)

$$79, 73, 72, 62, 76, 68$$

$$70, 86, 75, 68, 73, 66$$

$$\bar{Y} = 72.333, S_Y = 6.344, m = 12$$

1. $H_0 : \mu_X - \mu_Y = 0$
2. $H_0 : \mu_X - \mu_Y > 0$ $\qquad (\mu_X > \mu_Y)$
3. $\alpha = 0.05$
4. **Test Statistics**

$$T = \frac{\bar{X} - \bar{Y}}{S_p . \sqrt{\frac{1}{n} + \frac{1}{m}}}$$

where

$$S_p = \sqrt{\frac{(n-1).S_X^2 + (m-1).S_Y^2}{n+m-2}}$$

(Contd.)

- Here

$$
\begin{aligned}
S_p &= \sqrt{\frac{(8-1).4.853^2 + (12-1).6.344^2}{8+12-2}} \\
&= \vdots \\
&= \sqrt{33.754} \\
&= 5.81
\end{aligned}
$$

So,

$$
\begin{aligned}
T &= \frac{\bar{X} - \bar{Y}}{S_p.\sqrt{\frac{1}{n} + \frac{1}{m}}} \\
&= \frac{77.125 - 72.333}{5.81.\sqrt{\frac{1}{8} + \frac{1}{12}}} \\
&= \vdots \\
&= 1.825
\end{aligned}
$$

(Contd.)

5. **p - value**
   Find the probability of getting 1.825 or larger for value of test statistics, H_0 is true. We will examine values from t - distribution with d.f.= n+m-2 = 18 d.f. Since $H_0 : \mu_X - \mu_Y > 0$ we will try to find $P(T \geq 1.825)$.

Looking for 1.825 along the row for 18 d.f we see

$$P(T \geq 1.73) = 0.05$$

Since 1.825>1.73; Then

$$P(T \geq 1.825) < 0.05$$

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

6. **Decision**

   Reject H_0 if

   $$\text{p - value} \leq \alpha$$

   Since, we found

   $$0.0017 < 0.01$$

   in step 5, we **reject** H_0.

7. **Conclude**
   At the 5% significance level, the evidence indicates that the new teaching method produces higher mean score than the standard teaching method.

- **Note:**
  1. The data from each population must be normally distributed.
  2. The population standard deviation must be equal, i.e., $\sigma_1 = \sigma_2$

Hypothesis Test - Difference in Independent Population Means $\mu_X - \mu_Y$

I. **Large Sample Case** ($n \geq 20$ and $m \geq 20$)

Example:

A college statistics professor conjectures that student with good high school math background (2 or more math courses) perform better in college statistics course than students with a poor high school math background (1 or fewer math courses). he randomly selects 35 students with good math background, and 45 students with poor math background, and records Exam 1 scores from a college statistics course.

(Contd.)

- Test the hypothesis that the mean score of the "good background" students will be higher than the mean score of "poor math background" students.
  (Use α = 0.10)

  Summary data follows:
  Two or more math courses:

$$n = 35, \bar{X} = 84.2, S_X = 10.2$$

  One or fewer math courses:

$$m = 45, \bar{Y} = 73.1, S_Y = 14.3$$

1. $H_0 : \mu_X - \mu_Y = 0$
2. $H_0 : \mu_X - \mu_Y > 0$
3. $\alpha = 0.10$

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

4. **Test Statistics**

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$$

where $\mu_X - \mu_Y$ is the hypothesized difference in H_0.

$$\begin{aligned} T &= \frac{(84.2 - 73.1) - 0}{\sqrt{\frac{10.2^2}{35} + \frac{14.3^2}{45}}} \\ &= \frac{11.1}{\sqrt{7.51679}} \\ &= \frac{11.1}{2.74168} \\ &= 4.0486 \end{aligned}$$

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

5. **p - value**

   We must find probability of getting 4.05 or larger value for the test statistics (since, $H_0 : \mu_X - \mu_Y > 0$), under the assumption that H_0 is true:

$$\begin{aligned} P(T \geq 4.05) \simeq P(z \geq 4.05) &= 1 - P(z < 4.05) \\ &= 1 - 1 \\ &= 0 \end{aligned}$$

   So it is highly unlikely we would get a value 4.05 or larger for test statistics if H_0 is true.
   This provides evidence that H_0 is likely false.

6. **Decision**

   Reject H_0 if

$$\text{p - value} \leq \alpha$$

   Since, we found

$$0.0017 < 0.01$$

   we **reject** H_0.

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

7. **Conclusion**
   At the 10% significance level, there is evidence that students with good high school math background have a higher mean score on Exam 1 than students with poor high school math background.

- **Assumption:**
  Independent random samples from the populations.

Hypothesis Test - Two Independent Population Proportions $p_X - p_Y$

- The following procedure may be used, if

  1. $n\hat{p}_1 > 5$
  2. $n(1 - \hat{p}_1) > 5$
  3. $m\hat{p}_2 > 5$
  4. $n(1 - \hat{p}_2) > 5$

Example:

American Cancer Society wants to determine if the proportion of smokers in the population of Americans has decreased over the last decade.
In 1992, a random sample of 150 Americans showed 58 who smoked.
In 2002, a random sample of 200 Americans included 64 who smoked.

(Contd.)

- Does the data indicates that the proportion of smokers has decreased over the past decade?
  (Use α = 0.05)

1. $H_0 : p_1 - p_2 = 0$
2. $H_A : p_1 - p_2 > 0$
3. $\alpha = 0.05$
4. **Test Statistics**
   For 1992 (group 1)
   $$\hat{p}_1 = \frac{X}{n} = \frac{58}{150} = 0.3867$$
   For 2002 (Group 2)
   $$\hat{p}_2 = \frac{Y}{m} = \frac{64}{200} = 0.3200$$
   IF H_0 is true the
   $$p_1 = p_2 = p$$
   So both $\hat{p}_1$ and $\hat{p}_2$ estimate p (if H_0 is true)

Chap.9 Notes: Hypothesis Testing
Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

- Combining the data from the two samples, we can estimate p as

$$\hat{p} = \frac{X+Y}{n+m} = \frac{58+64}{150+200} = \frac{122}{350} = 0.3486$$

The **test statistics** is

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}.(1-\hat{p})\left(\frac{1}{n}+\frac{1}{m}\right)}}$$

where $p_1 - p_2$ is the hypothesized difference in proportion in H_0.

(Contd.)

$$
\begin{aligned}
Z &= \frac{(0.3867 - 0.3200) - 0)}{\sqrt{0.3486.(1 - 0.3486)\left(\frac{1}{150} + \frac{1}{200}\right)}} \\
&= \frac{0.0667}{\sqrt{(0.227078).(0.003257)}} \\
&= \frac{0.0667}{\sqrt{0.000739593}} \\
&= \frac{0.0667}{0.0271955} \\
&= 2.4526
\end{aligned}
$$

5. **p - value**
   Find P(Z$\geq$ 2.45) (since, $H_A : p_1 - p_2 > 0$)

$$
\begin{aligned}
P(Z > 2.45) &= 1 - P(Z < 2.45) \\
&= 1 - 0.9929 \\
&= 0.0071
\end{aligned}
$$

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

6. **Decision**

   Reject H_0 if

$$\text{p - value} \leq \alpha$$

   Since, we found

$$0.0071 < 0.05$$

   we **reject** H_0.

7. **Conclusion**
   At the 5% significance level, we conclude that the proportion of smokers in the American populations has decreased over the last 10 years.

- **Assumptions:**
  Independent random samples.

Two-Sided Hypothesis Tests:

- Thus far, all examples have used one-sided alternate hypothesis

$$H_A : \mu < \mu_0$$

or

$$H_A : \mu > \mu_0$$

The alternate hypothesis H_A may also have the form

$$H_A : \mu \neq \mu_0$$

This form of H_A is called a two-sided alternative hypothesis, because H_A can be true in either of two ways:

$$\text{if } \mu < \mu_0 \text{ or } \mu > \mu_0$$

then clearly $\mu \neq \mu_0$ and H_A is true.

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

- Since there are two ways for H_A to be true we must divide the significance level α by 2, allocating $\frac{\alpha}{2}$ to the possibility that $\mu < \mu_0$, and allocating $\frac{\alpha}{2}$ to the possibility that $\mu > \mu_0$.
  - For the event $\mu < \mu_0$, the p - value we use is

$$\mathrm{P(T < test\ statistics\ value)}$$

  - For the event $\mu > \mu_0$, the p - value we use is

$$\mathrm{P(T > test\ statistics\ value)}$$

- Thus, in a two-sided hypothesis test, the decision rule is
  Reject H_0, if

$$\mathrm{P(T < test\ statistics\ value)} \leq \frac{\alpha}{2}$$

  or,

$$\mathrm{P(T > test\ statistics\ value)} \leq \frac{\alpha}{2}$$

- This can be written succinctly using properties of inequalities, as
  Reject H_0, if

$$\mathrm{P}(\mathrm{T} > |\,\text{test statistics value}\,|) \leq \frac{\alpha}{2}$$

This form of our decision rule can also be written as
Reject H_0, if

$$2 \times \mathrm{P}(\mathrm{T} < \text{test statistics value}) \leq \alpha$$

this is the form of the decision rule we will use when we have a two-sided H_0.

Example:

The quality control manager at a sugar processing and packaging plant must make sure that two-pound bags of sugar actually contains two pounds of sugar. He randomly select 50 bags and weighs their contents. The sample mean is 1.962 pounds with standard deviation of 0.160 pounds.

- Dies the data indicate that the mean wean weight μ of all bags of sugar is different from 2 pounds?
  (Use α = 0.05)

1. $H_0 : \mu = 2$
1. $H_0 : \mu \neq 2$
1. $\alpha = 0.05$
4. **Test statistics**

- **Note:** n = 50 is a large sample and σ is unknown.

$$
\begin{aligned}
T &= \frac{\bar{X} - \mu_0}{\left(\frac{S}{\sqrt{n}}\right)} \\
&= \frac{1.962 - 2.000}{\left(\frac{0.160}{\sqrt{50}}\right)} \\
&= \frac{-0.038}{0.0226274} \\
&= -1.6794
\end{aligned}
$$

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

5. **p - value**
   Since we have a two-sided H_A the p - value of the test statistics is

$$
\begin{aligned}
p - value &= 2 \times \mathrm{P(T > | \text{ test statistics value } |)} \\
&= 2 \times P(T > |-1.68|) \\
&\simeq 2 \times P(Z > 1.68) \;\; \text{(since n is large)} \\
&= 2 \times [1 - P(Z < 1.68)] \\
&= 2 \times [1 - 0.9535] \\
&= 2 \times (0.0435 \\
&= 0.093
\end{aligned}
$$

(Contd.)

6. **Decision**

   Reject H_0 if

$$\text{p - value} \leq \alpha$$

   Since,

$$0.0093 \nleq 0.05 \text{ i.e, } 0.093 > 0.05$$

   we **do not reject** H_0.

7. **Conclusion**
   At the 5% significance level, the data indicates that the mean weight of two-pound bags of sugar is not significantly different from 2 pounds.

Example:

In 1990, 17.6% of young adults (age 18 - 24) were attending college. A sociologist wishes to determine if this percentage has changed. She selects a random sample of 500 young adults, of which 110 were attending college.

- Has the percentage of young adults attending college changed since 1990? (Use α = 0.05)

before we proceed, we must check **two conditions**:

$$n.p > 5 \text{ and } n(1-p) > 5$$

using p =0.176 (this p appears in H_0)

$$\begin{aligned} n.p = 500.(0.176) &= 88 > 5 \\ n.(1-p) = 500.(0.824) &= 412 > 5 \end{aligned}$$

1. $H_0 : p = 0.176$
2. $H_A : p \neq 0.176$
3. $\alpha = 0.05$
4. **Test statistics**

$$n = 500 \qquad X = 110$$

$$\hat{p} = \frac{X}{n} = \frac{110}{500} = 0.22$$

$$\begin{aligned} Z &= \frac{\hat{p} - p}{\sqrt{\frac{p.(1-p)}{n}}} \\ &= \frac{0.22 - 0.176}{\sqrt{\frac{0.176.(1-0.176)}{500}}} \\ &= \frac{0.044}{\sqrt{0.000290048}} \\ &= \frac{0.044}{0.017031} \\ &= 2.5836 \end{aligned}$$

(Contd.)

5. **p - value**
   Since H_A is two - sided, our p - value is

$$
\begin{aligned}
Z \times \mathrm{P(Z > \mid test\ statistics\ value \mid)} &= Z \times P(Z > |2.58|) \\
&= Z \times P(Z > 2.58) \\
&= Z \times [1 - P(Z \leq 2.58)] \\
&= Z \times [1 - 0.9951] \\
&= Z \times 0.0049 \\
&= 0.0098
\end{aligned}
$$

6. **Decision**

   Reject H_0 if

$$\text{p - value} \leq \alpha$$

   Since,

$$0.0098 < 0.05$$

   So, we **reject** H_0.

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

7. **Conclusion**
   At the 5% significance level, we conclude that the percentage of young adults attending college has changed since 1990.

- **Note:** We can not say that the percentage of young adults attending college has increased.

- In order to check for an increase in percentage attending college,
  the alternative hypothesis would be

$$H_A : p > 0.176$$

Acceptance/Rejection Testing:

- In some situation, we can not simply state which hypothesis has the higher probability of being true - we must either

$$\text{accept } H_0 \text{ (and reject } H_A)$$

$$\text{or, accept } H_A \text{ (and reject } H_0).$$

Example:

production manager must decide if a shipment of 10,000 brackets meets specified standards.

$$\begin{aligned} H_0 &: \text{Brackets meet standards.} \\ &\quad \text{(accept the shipment)} \\ H_A &: \text{Brackets do not meet standards.} \\ &\quad \text{(return the shipment)} \end{aligned}$$

- **Note:** If the “wrong” decision is made, the company can lose money; other “adverse” consequences.

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

Types of Errors:

		Reality	
		H_0 true (H_A false)	H_0 false (H_A true)
Our Decision	Reject H_0 (Accept H_A)	Type I Error	Correct Decision
	Do not Reject H_0 (Accept H_0)	Correct Decision	Type II Error

$$
\begin{aligned}
\alpha &= \text{P(Type I Error)} \\
&= \text{P(reject } H_0 \mid H_0 \text{ true)} \\
\beta &= \text{P(Type II Error)} \\
&= \text{P(do not reject } H_0 \mid H_0 \text{ false)}
\end{aligned}
$$

$$
\begin{aligned}
1-\beta &= \text{P(reject } H_0 | H_0 \text{ false)} \\
&= \text{P(detecting a false } H_0\text{)} \\
&= \text{Power of the test}
\end{aligned}
$$

Example: **Criminal Trial**

$$
\begin{aligned}
H_0 &: \text{Defendant is innocent} \\
H_A &: \text{Defendant is guilty}
\end{aligned}
$$

Type I Error: H_0 is true but judge decides to believe H_A, i.e., the defendant is innocent but the judge finds him/her guilty.

Type II Error: H_0 is false but judge decides to believe H_0, i.e., the defendant is guilty but the judge finds him/her innocent.

Example: Accident victim brought to emergency room. Doctor must decide between

$$H_0 : \text{Victim is alive.}$$
$$H_1 : \text{Victim is dead.}$$

Type I Error: H_0 is actually true but doctor decides that H_0 is false, i.e.,victim is alive, but doctor believes victim is dead.

Type II Error: H_0 is actually false but doctor decides that H_0 is true, i.e.,victim is dead, but doctor believes victim is alive.

Significance Level - α

- Similar to the concept of confidence level. **Significance level** is usually chosen to be 5%. (Sometimes 1% or 10%).
- **Relation to Confidence Level**

  If a C.I. has confidence level 95%, then

  1. 95% probability that μ is in the C.I.
  2. 5% probability that μ is not in the C.I.

- For hypothesis test with **significance level** 0.05.
  1. 95% chance of making a correct decision, when H_0 is true.
  2. 5% chance of **Type I error**.

- **Note:** Most books use α to represent significance level.

$$\alpha = \textbf{P(Type I error)}$$

Choice of α level

- **Recall:**

$$\alpha = \text{P(Type I Error)}$$

$$\beta = \text{P(Type II Error)}$$

Typical α - values are

$$0.001, 0.005, 0.01, 0.05, 0.10$$

- the values of β depends on α, n, and the alternate hypothesis as well as s, the (pooled) standard deviation.

- Suppose n is fixed, i.e., we can not acquire more observations.
  - If **Type I** is the worse type of error to make, choose α small. The "price" we pay is that β increases.
  - If **Type II** is the worse type of error to make, choose α large ($\alpha = 0.10$). This will make β decrease.
- If n is not fixed, we can make both α and β decrease by increasing the sample size (n). More data means more information which decreases the probability of making errors.
- In fact, if we could obtain data from everyone in a population, we could eliminate all error, making $\alpha = 0$ and $\beta = 0$.

Example:

Determine an appropriate α level for the following hypothesis test.

$$
\begin{aligned}
H_0 &: \text{The person accused of the crime is not guilty.}\\
H_1 &: \text{The accused person is guilty.}
\end{aligned}
$$

Type I Error - Reject H_0 when H_0 is really true. The accused person is not guilty but the judge decides he is guilty.

Type II Error - Fail to reject H_0 when H_0 is really false. The accused person is guilty but the judge decides he is not guilty.

Assess the consequences of a Type I Error:

1. Innocent person jailed.
2. Innocent person loses job.
3. Innocent person no longer pays taxes.
4. Waste, money, resources jailing an innocent person.
5. Guilty person still free.
6. Justice system credibility is damaged.

Assess the consequence of a Type II Error:

1. Guilty person still free.
2. No justice for victim.
3. Other criminals may emulate to the same crime.
4. Cost of continuing investigation.

- It appears that a **Type I Error** is worse in this situation; So chose α small (0.01 or less).

When the Assumptions are Violated:

- IF one (or more) assumption is not valid, the result of the hypothesis test will not be valid.
  - we must use some other procedure (Non parametric test).

Connection Between Confidence Interval and Hypothesis Testing:

- There is a relationship between C.I.s and hypothesis tests involving a two - sided alternative.
- IF the hypothesized value of μ (stated in H_0) is inside the $(1-\alpha)100\%$ confidence interval, we would fail to reject H_0, at the specified α - level.
- IF the hypothesized value of μ (stated in H_0) is inot in the $(1-\alpha)100\%$ confidence interval, we would reject H_0, at the specified α - level.

Example:

Suppose a 95%C.I. on μ goes from 43.5 to 62.7.

$$
\begin{aligned}
IF \qquad H_0 &: \mu = 58 \\
H_1 &: \mu \neq 58 \\
\alpha &= 0.05
\end{aligned}
$$

We would fail to reject H_0

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis
Sect. 9.2:Tests For A Population Proportion
Sect. 9.4: Tests For A Population Mean
Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means
Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing
Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods
Sect. 9.8:Tests For A Population Standard deviation:Issues

Hypothesis Test - Equality of Variances:

We are often interested in determining if

$$\sigma_1 = \sigma_2,$$

e.g., by testing the null hypothesis

$$H_0 : \sigma_1^2 = \sigma_2^2.$$

Example:

(Continue the example from Sec. 9.5)
use α = 0.05 (teaching methods)

1. $H_0 : \sigma_1^2 = \sigma_2^2$
2. $H_1 : \sigma_1^2 \neq \sigma_2^2$ (always)
3. $\alpha = 0.05$

Chap.9 Notes: Hypothesis Testing

Sect 9.1:The Null Hypothesis And The Alternative Hypothsis

Sect. 9.2:Tests For A Population Proportion

Sect. 9.4: Tests For A Population Mean

Sect. 9.5: Tests For Equality Of two Population Proportions And Equality Of Two Population Means

Sect. 9.6:One-Sided And Two-Sided Hypothesis Testing

Sect. 9.7: Significance Testing Versus Acceptance/Rejection Testing: Concepts And Methods

Sect. 9.8:Tests For A Population Standard deviation:Issues

(Contd.)

4. **Test statistics**

$$F^* = \frac{\text{Bigger}S^2}{\text{Smaller}S^2} = \frac{6.344^2}{4.853^2} = 1.709$$

5. **Critical Value**
   Use Table H.1 in the Text Book[4]

$$\begin{aligned} df_N &= \text{numerator degrees of freedom} \\ &= m - 1 \\ &= 12 - 1 \\ &= 11 \end{aligned}$$

[4] refer to Page: 711 of the Text Book

(Contd.)

$$\begin{aligned} df_N &= \text{denominator degrees of freedom} \\ &= n-1 \\ &= 8-1 \\ &= 7 \end{aligned}$$

6. Reject H_0, if

$$F^* > F \qquad \text{(always)}$$

$$i.e. \text{ if test statistics} > \text{critical value}$$

Since,

$$1.709 \ngtr 3.60$$

We **do not reject** H_0.

(Contd.)

7. The data indicates that there is no significant difference between the variances (and hence, the standard deviations) of the two population of students,
   at 5% significance level.

- **Assumptions:**

1. The data from each population are normally distributed.

- **Note:** If a 1% significance level (α) is used, the critical value can be found in in Table H.2 in the Text book[5].

---

[5] refer to Page: 712 of the Text Book