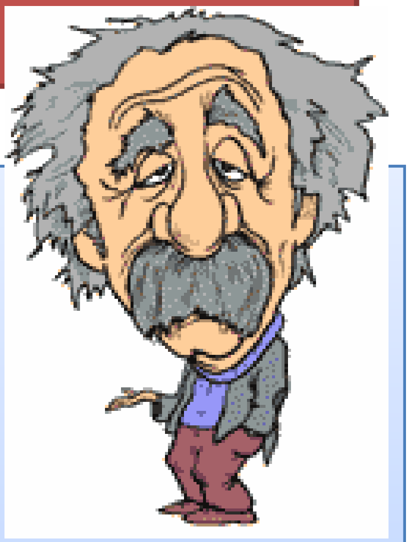
Hypothesis Testing



Out line

1. Testing of a hypothesis,
2. Tests of significance for large samples,
3. Student’s t-distribution(applications only),
4. Chi-square test of goodness of fit.

LEARNING OBJECTIVES
After reading this chapter, you should be able to:

- Identify the four steps of hypothesis testing.
- Define null hypothesis, alternative hypothesis, level of significance, test statistic, p value, and statistical significance.
- Define Type I error and Type II error, and identify the type of error that researchers control
- Calculate the one-independent sample z test and interpret the results.
- Distinguish between a one-tailed and two-tailed test,
- and explain why a Type III error is possible only with
- one-tailed tests.
- Explain what effect size measures and compute a Cohen's d for the one-independent sample z test.
- Define power and identify six factors that influence power.
- 8 Summarize the results of a one-independent sample
- z test in American Psychological Association (APA)
- format.

Definition

- **Hypothesis Test**

  is a standard procedure for testing a claim about a property of a population

Components of a Formal Hypothesis Test

- **Null Hypothesis: H_0**

❖ Statement about value of population parameter that is equal to some claimed value

$H_0: p = 0.5$ $H_0: \mu = 98.6$ $H_0: \sigma = 15$

❖ Test the Null Hypothesis directly

❖ Reject H_0 or fail to reject H_0

Alternative Hypothesis: H_1

- the statement that the parameter has a value that somehow differs from the null

- Must be true if H_0 is false

- $\neq$, $<$, $>$

Claim: Using math symbols

H_0: **Must contain equality**

H_1: **Will contain $\neq$, $<$, $>$**

- Hypothesis testing or significance testing is a method for testing a claim or hypothesis about a parameter in a population, using data measured in a sample. In this method, we test some hypothesis by determining the likelihood that a sample statistic could have been selected, if the hypothesis regarding the population parameter were true.
- FOUR STEPS TO HYPOTHESIS TESTING

  Step 1: State the hypotheses.

  Step 2: Set the criteria for a decision.

  Step 3: Compute the test statistic.

  Step 4: Make a decision.

Level of significance, or significance level refers to a criterion of judgment upon which a decision is made regarding the value stated in a null hypothesis. The criterion is based on the probability of obtaining a statistic measured in a sample if the value stated in the null hypothesis were true. In behavioral science, the criterion or level of significance is typically set at 5%. When the probability of obtaining a sample mean is less than 5% if the null hypothesis were true, then we reject the value stated in the null hypothesis.

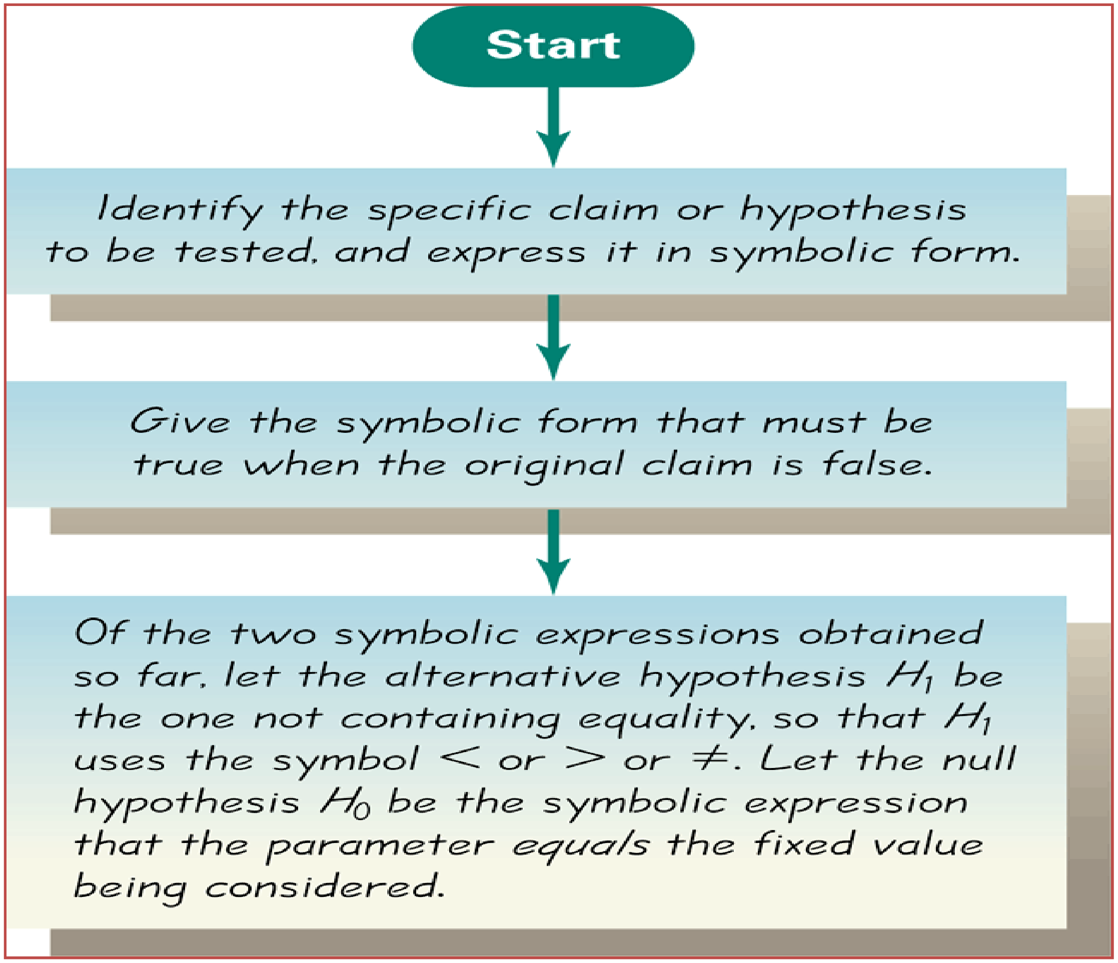
The null hypothesis (H0), stated as the null, is a statement about a population parameter, such as the population mean, that is assumed to be true. The null hypothesis is a starting point. We will test whether the value stated in the null hypothesis is likely to be true.

An alternative hypothesis (H1) Is a statement that directly contradicts a null hypothesis by stating that that the actual value of a population parameter is less than, greater than, or not equal to the value stated in the null hypothesis. Significance, or statistical significance, describes a decision made concerning a value stated in the null hypothesis. When the null hypothesis is rejected, we reach significance. When the null hypothesis is retained, we fail to reach significance.

***NOTE:** The level of significance in hypothesis testing is the criterion we use to decide whether the value stated in the null hypothesis is likely to be true.*

***NOTE:** We use the value of the test statistic to make a decision regarding the null hypothesis.*

Note about Identifying H_0 and H_1



Note about Forming Your Own Claims (Hypotheses)

- If you are conducting a study and want to use a hypothesis test to <u>support</u> your claim, the claim must be worded so that it becomes the alternative hypothesis. This means your claim must be expressed using only $\neq$, $<$, $>$

Test Statistic

- The test statistic is a value computed from the sample data, and it is used in making the decision about the rejection of the null hypothesis.

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Test statistic for proportions

Test Statistic

- The test statistic is a value computed from the sample data, and it is used in making the decision about the rejection of the null hypothesis.

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$$

Test statistic for mean

Test Statistic

- The test statistic is a value computed from the sample data, and it is used in making the decision about the rejection of the null hypothesis.

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

Test statistic for mean

Test Statistic

- The test statistic is a value computed from the sample data, and it is used in making the decision about the rejection of the null hypothesis.
- Test statistic for standard deviation

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Example: A survey of n = 880 randomly selected adult drivers showed that 56%(or $\hat{p}$ = 0.56) of those respondents admitted to running red lights. Find the value of the test statistic for the claim that the majority of all adult drivers admit to running red lights (In Section 1 we will see that there are assumptions that must be verified. For this example, assume that the required assumptions are satisfied and focus on finding the indicated test statistic.)

- **Solution:** The preceding example showed that the given claim results in the following null and alternative hypotheses: H_0: p = 0.5 and H_1: p > 0.5. Because we work under the assumption that the null hypothesis is true with p = 0.5, we get the following test statistic:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.56 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{880}}} = 3.56$$

Interpretation: We know from previous chapters that a z score of 3.56 is exceptionally large. It appears that in addition to being "more than half," the sample result of 56% is ***significantly*** more than 50%.

See Figure 7-3 where we show that the sample proportion of 0.56 (from 56%) does fall within the range of values considered to be significant because they are so far above 0.5 that they are not likely to occur by chance (assuming that the population proportion is p = 0.5).

See Figure 1 where we show that the sample proportion of 0.56 (from 56%) does fall within the range of values
considered to be significant because they are so far above 0.5 that they are not likely to occur by chance
(assuming that the population proportion is p = 0.5).

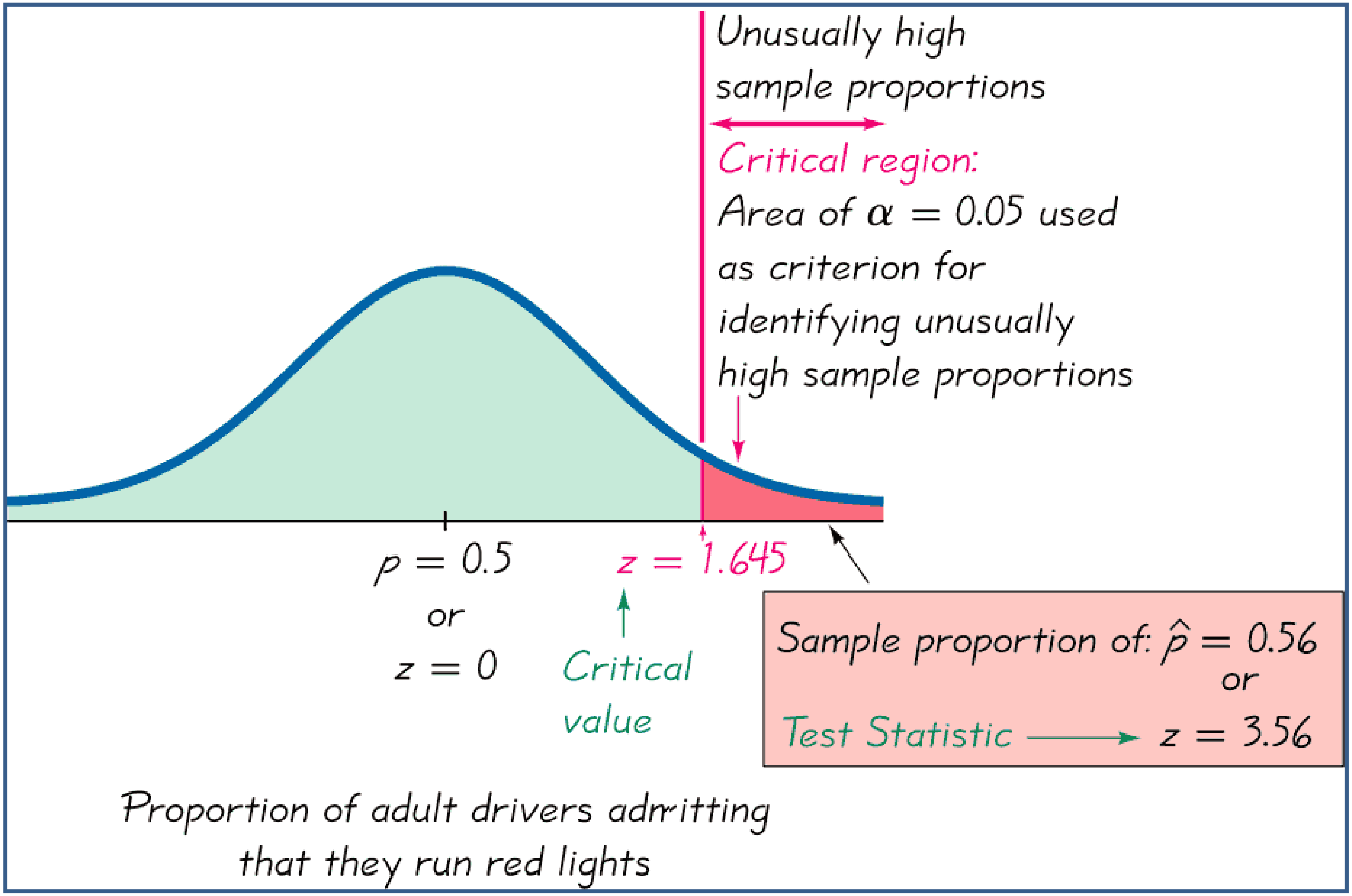
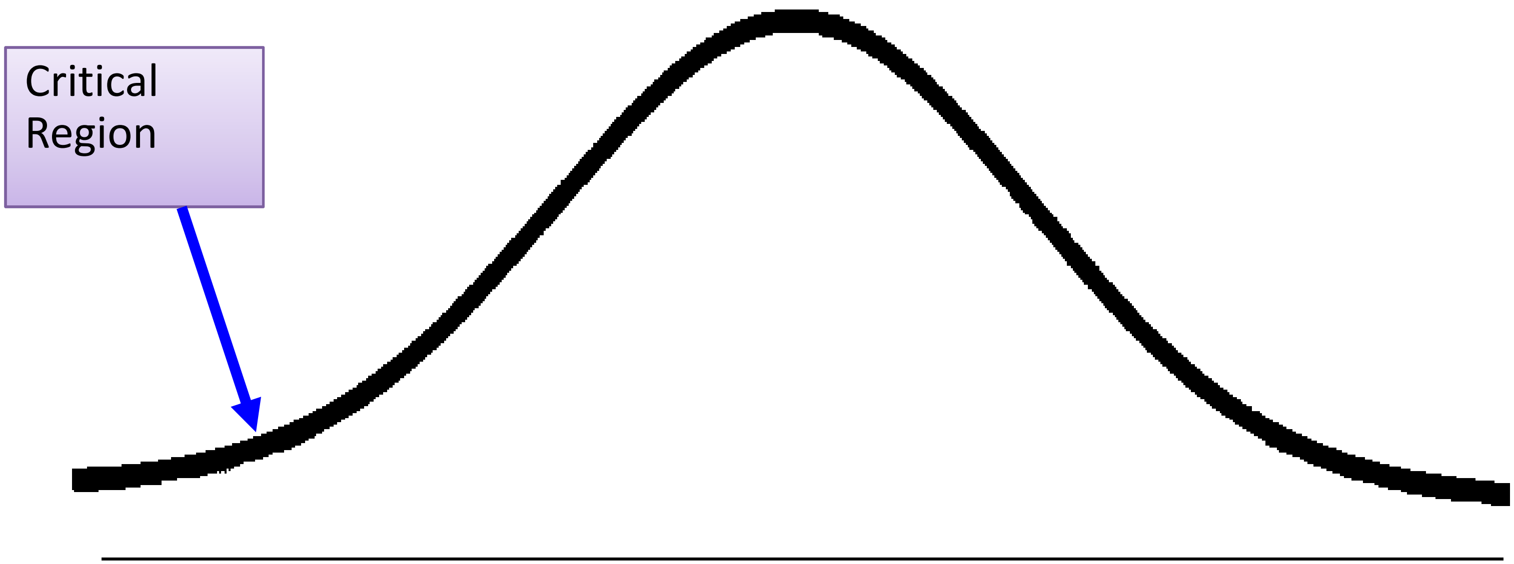

Unusually high
sample proportions
Critical region:
Area of $\alpha = 0.05$ used
as criterion for
identifying unusually
high sample proportions
$p = 0.5$
or
$z = 0$
$z = 1.645$
Critical
value
Sample proportion of: $\hat{p} = 0.56$
or
Test Statistic → $z = 3.56$
Proportion of adult drivers admitting
that they run red lights

Figure 1

Critical Region

Set of all values of the test statistic that would cause a rejection of the null hypothesis

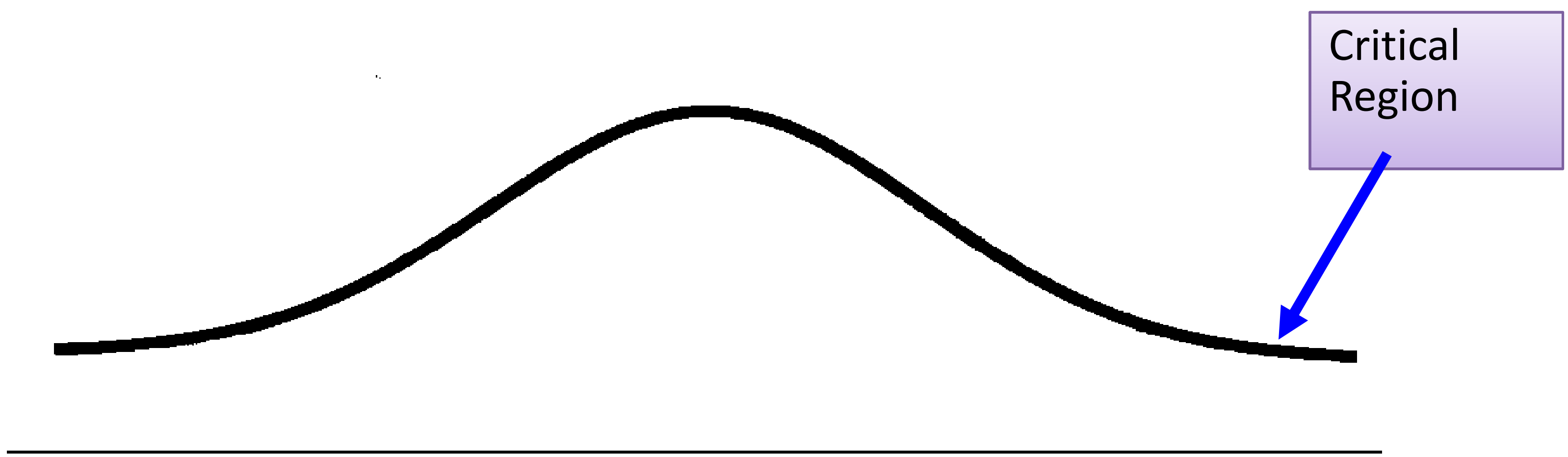
Set of all values of the test statistic that would cause a rejection of the null hypothesis



Critical Region

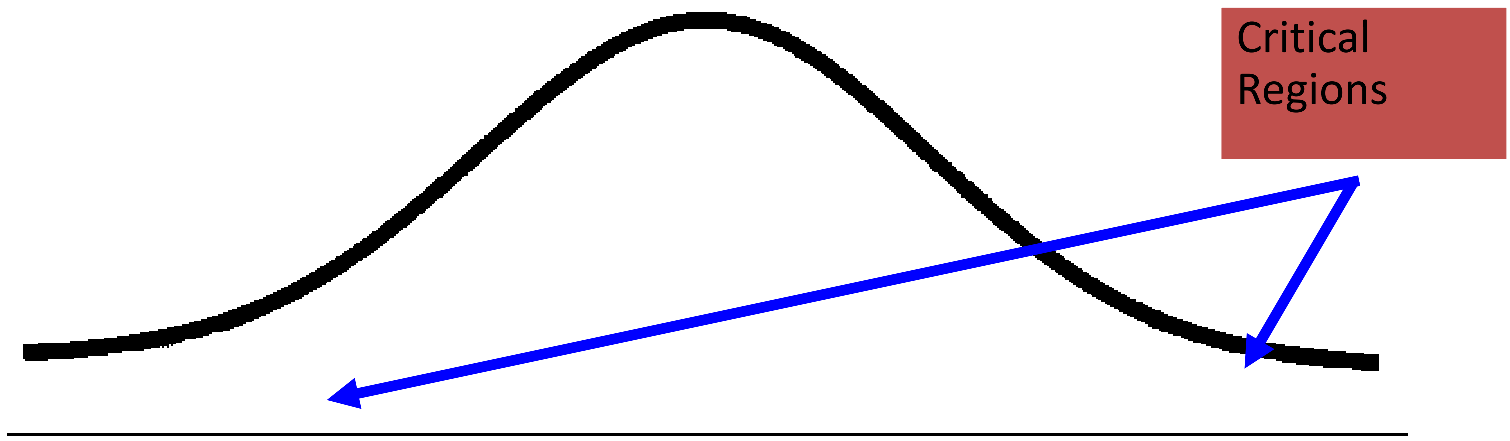
Set of all values of the test statistic that would cause a rejection of the null hypothesis

Set of all values of the test statistic that would cause a rejection of the null hypothesis



Critical Region

Set of all values of the test statistic that would cause a rejection of the null hypothesis



Example

- **A normal population has a mean of 6.8 and standard deviation of 1.5. A sample of 400 members gave a mean of 6.75. is the difference significant?**
- **Sol.** H_0 **: There is no significant difference between** $\bar{x}$ **and** μ

Given $\mu = 6.8, \quad \sigma = 1.5, \quad \bar{x} = 6.75 \quad \& \quad n = 400$

Conclusion as the calculated value of

$$|z| = \left| \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \right| = \left| \frac{6.75 - 6.8}{1.5 / \sqrt{900}} \right| = |-0.67| = 0.67$$

$$|z| < z_a = 1.96$$

at 5% level of significance, H_0 **is accepted i.e. there is no significant difference between** $\bar{x}$ **and** μ

Type I Error

- A **Type I** error is the mistake of rejecting the null hypothesis when it is true.
- The symbol α (alpha) is used to represent the probability of a type I error.

Type II Error

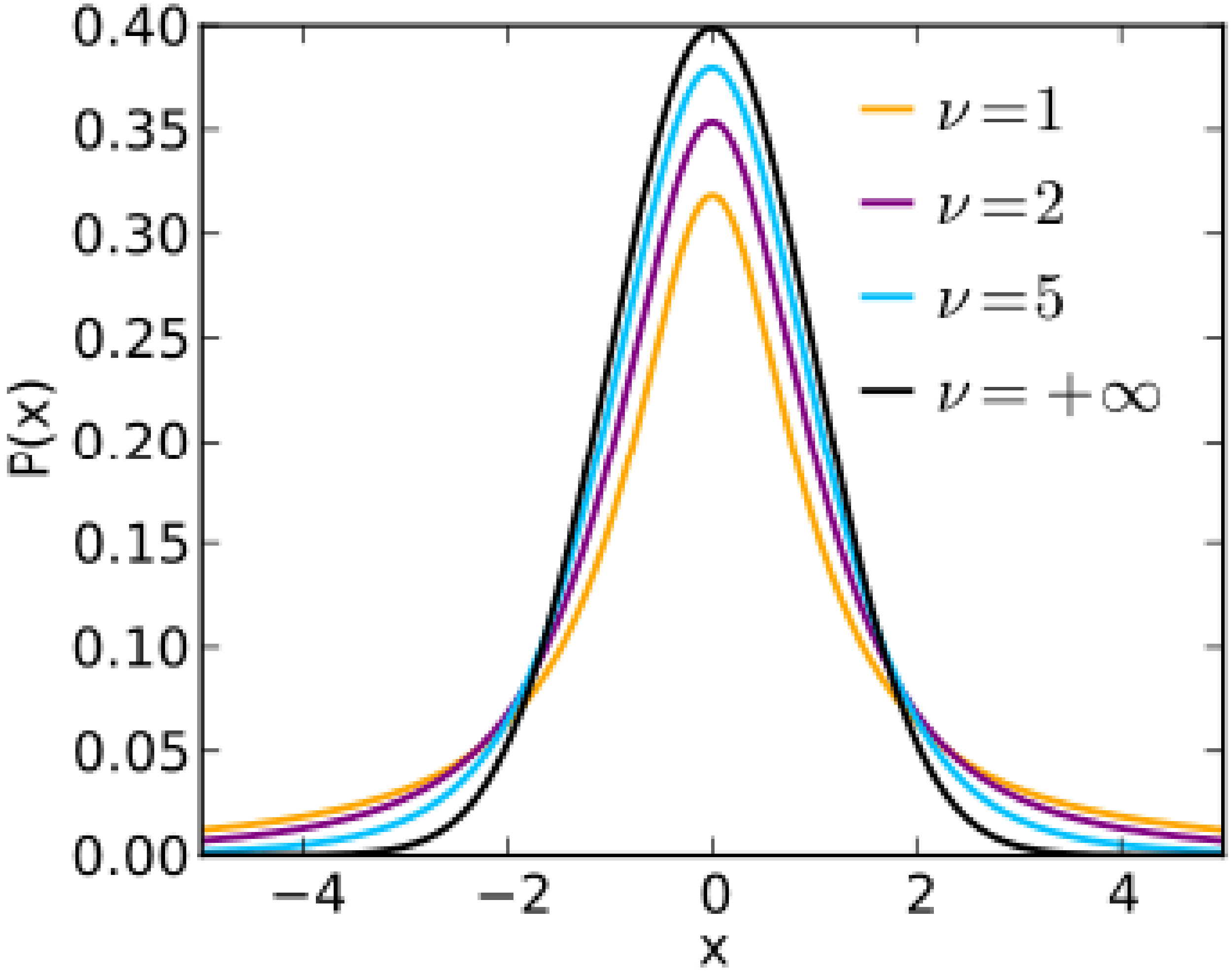
- A **Type II** error is the mistake of failing to reject the null hypothesis when it is false.

- The symbol β (beta) is used to represent the probability of a type II error.

Student's *t*-distribution

- In probability and statistics, **Student's *t*-distribution** (or simply the ***t*-distribution**) is a family of continuous probability distributions that arises when estimating the mean of a normally distributed population in situations where the sample size is small and population standard deviation is unknown. It plays a role in a number of widely used statistical analyses, including the Student's *t*-test for assessing the statistical significance of the difference between two sample means, the construction of confidence intervals for the difference between two population means, and in linear regression analysis. The Student's *t*-distribution also arises in the Bayesian analysis of data from a normal family.

Student's t Probability density function



Pearson's chi-squared test

- Pearson's chi-squared test uses a measure of goodness of fit which is the sum of differences between observed and expected outcome frequencies (that is, counts of observations), each squared and divided by the expectation:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

- Where: O_i = an observed frequency (i.e. count) for b in *i*
  E_i = an expected (theoretical) frequency for bin *i*, asserted by the null hypothesis.

For example

- **To test the hypothesis that a random sample of 100 people has been drawn from a population in which men and women are equal in frequency, the observed number of men and women would be compared to the theoretical frequencies of 50 men and 50 women. If there were 44 men in the sample and 56 women**, then

$$\chi^2 = \frac{(44-50)^2}{50} + \frac{(56-50)^2}{50} = 1.44$$

- If the null hypothesis is true (i.e., men and women are chosen with equal probability in the sample), the test statistic will be drawn from a chi-squared distribution with one degree of freedom. Though one might expect two degrees of freedom (one each for the men and women), we must take into account that the total number of men and women is constrained (100), and thus there is only one degree of freedom (2 − 1). Alternatively, if the male count is known the female count is determined, and vice-versa.
- Consultation of the chi-squared distribution for 1 degree of freedom shows that the probability of observing this difference (or a more extreme difference than this) if men and women are equally numerous in the population is approximately 0.23. This probability is higher than conventional criteria for statistical significance (.001-.05), so normally we would not reject the null hypothesis that the number of men in the population is the same as the number of women (i.e. we would consider our sample within the range of what we'd expect for a 50/50 male/female ratio.)

Table 7-1 Type I and Type II Errors

		True State of Nature	
		The null hypothesis is true.	The null hypothesis is false.
Decision	We decide to reject the null hypothesis.	**Type I error** (rejecting a true null hypothesis) α	Correct decision
	We fail to reject the null hypothesis.	Correct decision	**Type II error** (failing to reject a false null hypothesis) β

Controlling Type I and Type II Errors

- For any fixed α, an increase in the sample size ***n*** will cause a decrease in β.

- For any fixed sample size ***n*** , a decrease in α will cause an increase in β. Conversely, an increase in α will cause a decrease in β .

- To decrease both α and β, increase the sample size.

Traditional Method

Start

1 Identify the specific claim or hypothesis to be tested, and put it in symbolic form.

2 Give the symbolic form that must be true when the original claim is false.

3 Of the two symbolic expressions obtained so far, let the alternative hypothesis H_1 be the one not containing equality, so that H_1 uses the symbol $>$ or $<$ or $\neq$. Let the null hypothesis H_0 be the symbolic expression that the parameter *equals* the fixed value being considered.

4 Select the significance level α based on the seriousness of a type 1 error. Make α small if the consequences of rejecting a true H_0 are severe. The values of 0.05 and 0.01 are very common.

5 Identify the statistic that is relevant to this test and determine its sampling distribution (such as normal, t, chi-square).

6 Find the test statistic, the critical values, and the critical region. Draw a graph and include the test statistic, critical value(s), and critical region.

7 Reject H_0 if the test statistic is in the critical region. Fail to reject H_0 if the test statistic is not in the critical region.

8 Restate this previous decision in simple, nontechnical terms, and address the original claim.

Stop

Comprehensive Hypothesis Test

Problems for Practice

- **A sample of 1000 students from a university was taken and their average weight was found to be 112 pounds with a S.D. of 20 pounds. Could the mean weight of students in the population be 120 pounds?**
- **A random sample of 200 measurements from a large population gave a mean value of 50 and a S.D. of 9. Determine 95% confidence interval for the mean of population.**

Problems for Practice

- **Two independent samples of sizes 7 and 9 have the following values:**

City 1	10	12	10	13	14	11	10		
City 2	10	13	15	12	10	14	11	12	11

Assignment-1

Q1. Find the student's t test for the following variable values in a sample of eight: -4, -2, -2, 0, 2, 2, 3, 3 taking the mean of the universe to be zero.

Q2. Ten individuals are chosen at random from a normal population of students and their marks found to be 63,63,66,67,68,69,70,71,71. In the light of these data discuss the suggestion that mean mark of the population of students is 66.

Q3. A sample of 18 items has mean 24 units and standard deviation 3 units. Test the hypothesis that it is a random sample from a normal population with mean 27 units.

Q4. A sample of 20 items has mean 42 units and standard deviation 5 units. Test the hypothesis that it is a random sample from a normal population with mean 45 units.

Q5. Memory capacity of 9 students was tested before and after a course of mediation for a month. State whether the course was effected or not from the data below(in same unit)

Before	10	15	9	3	7	12	16	17	4
After	12	17	8	5	6	11	18	20	3

Q6. The following figures refer to observation in live independent samples:

Sample-I	25	35	28	34	24	20	13	32	22	38
Sample-II	40	34	22	20	31	40	30	23	36	17

Assignment-2

Q1. The following table gives the number of accidents that took place in a industry during various days a week. Test if accidents are uniformly distributed over the week.

Days	Mon	Tue	Wed	Thu	Fri	Sat
No. of accidents	14	18	12	11	15	14

Q2. The following figures refer to distribution of digits in numbers chosen at random from a telephone directory.

Digits	0	1	2	3	4	5	6	7	8	9
Frequency	1026	1107	997	966	1075	933	1107	972	964	853

Q3. The following data collected on two characters:

	smoker	Non smoker
Literate	83	57
illiterate	45	68

Assignment-3

Q1. In a experiment on the immunization of goats from anthrax, the following results were obtained. Derive your inferences on the efficiency of the vaccine:

	Died	Survived
Inoculated with vaccine	2	10
Not Inoculated	6	6

Q2. Samples of sizes 10 and 14 were taken from two normal populations with S.D. 3.5 and 5.2. The sample means were found to be 20.3 and 18.6. Test whether the means of the two populations are the same at 5% level.

- Thank you