Testing of Hypothesis

Introduction

- Testing of hypothesis is a process to test a hypothesis about the value of a population parameter on the basis of sample drawn from the population.
- Or in other words, Hypothesis testing is an act in statistics whereby an analyst tests an assumption regarding a population parameter.

Key Terms....

- Null Hypothesis
- Alternate Hypothesis
- Simple Hypothesis
- Composite Hypothesis
- Test Statistic
- Acceptance Region
- Rejection Region
- One Tailed Test & Two Tailed Test
- Level of Significance

Null Vs Alternate Hypothesis

- **Null Hypothesis:**

  It is denoted by H_0. A hypothesis which is to be tested for possible rejection under the assumption that it is true is called Null hypothesis.

- **Alternate Hypothesis:**

  Any other hypothesis which we accept when the null hypothesis H_0 is rejected is called Alternate hypothesis and is denoted by H_1.

Null Vs Alternate Hypothesis

- The **alternate hypothesis** is usually your initial hypothesis that predicts a relationship between variables.
- The **null hypothesis** is a prediction of no relationship between the variables you are interested in.
- For example, You want to test whether there is a relationship between gender and height. Based on your knowledge of human physiology, you formulate a hypothesis that men are, on average, taller than women. To test this hypothesis, you restate it as:

H_o: Men are, on average, not taller than women.
H_a: Men are, on average, taller than women.

Null Vs Alternate Hypothesis

Equal Sign is compulsory for Null Hypothesis

Null Hypothesis: H_0

$=$

$\geq$

$\leq$

Alternate Hypothesis will always contain inequality sign

Alternate Hypothesis: H_1

$\neq$

$<$

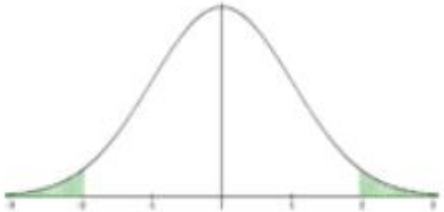
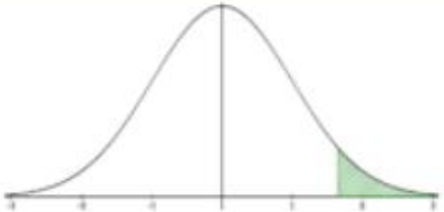
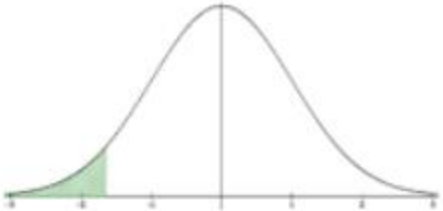
$>$

Null Vs Alternate Hypothesis

The null and alternate hypothesis will always be opposite to each other .

For example:

If our null hypothesis is $H_0 = 35$ then Alternate hypothesis may be $H_1 \neq 35$ **Or** $H_1 < 35$ **Or** $H_1 > 35$

Two-tailed Test	Right-tailed Test	Left-tailed Test
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$
		
	Claim is in the Null Hypothesis	
=	≤	≥
Is equal to	Is less than or equal to	Is greater than or equal to
Is exactly the same as	Is at most	Is at least
Has not changed from	Is not more than	Is not less than
Is the same as	Within	Is more than or equal to
	Claim is in the Alternative Hypothesis	
≠	>	<
Is not	More than	Less than
Is not equal to	Greater than	Below
Is different from	Above	Lower than
Has changed from	Higher than	Shorter than
Is not the same as	Longer than	Smaller than
	Bigger than	Decreased
	Increased	Reduced

Simple Vs Composite Hypothesis

- **Simple Hypothesis :**

A hypothesis which uniquely specifies all the parameters of a distribution is called simple hypothesis.

For Example:

Mean of a normal distribution is 50 and Variance is 5 .

Simple Vs Composite Hypothesis

- **Composite Hypothesis :**

A hypothesis which does not specify all the parameters of a distribution is called composite hypothesis.

For Example:

If we say that mean of distribution is 25 and variance is less than 6.

Simple Vs Composite Hypothesis

Simple Hypothesis	Composite Hypothesis
• $\mu=1200$	• $\mu<58$
• $\delta= 40$	• $\delta \geq 5$
• $\delta^2= 25$	• $\delta^2 \neq 64$
• $P= 75\%$	• $P \leq 29\%$

Test - Statistic

- **Test Statistic:**

  A statistic (a value obtained from the sample data) on which the decision is based whether to accept or reject the null hypothesis H_0 is called Test Statistic. Every test statistic has its own sampling distribution. The commonly used test statistic are Z , T , F etc

Acceptance vs Rejection Region

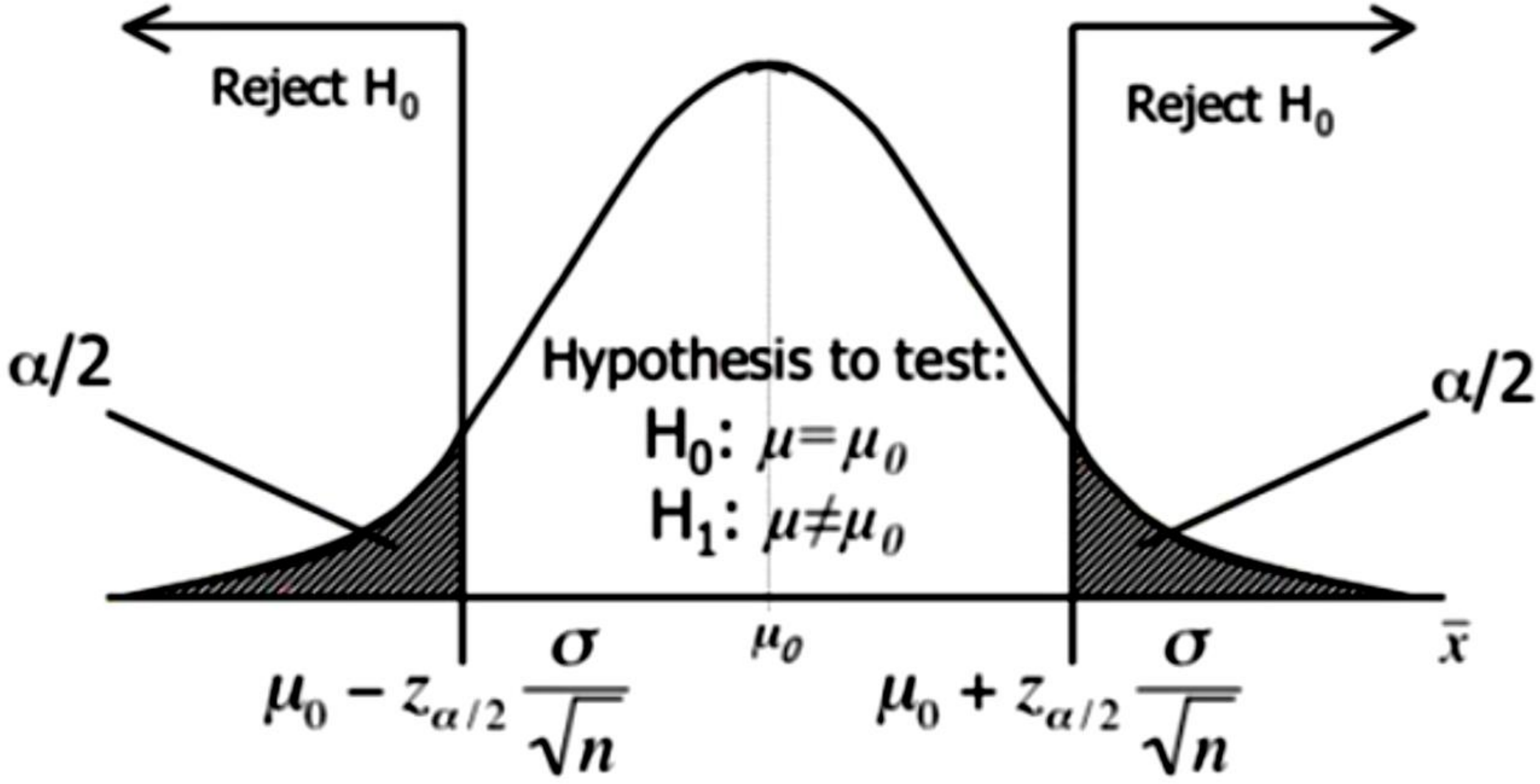
- **Acceptance Region:**

  A region in which the values of a test statistic leading to the acceptance of the null hypothesis H_0 is called Acceptance region.

- **Rejection Region:**

  A region in which the values of a test statistic leading to the rejection of the null hypothesis H_0 is called Rejection region. The rejection region is also known as critical region.

Acceptance vs Rejection Region



One Tailed vs two Tailed test

- **One Tailed Test:**

  A test for which the rejection region is located entirely at one end of the sampling distribution of a test statistic (either left end or right end) then it is called One tailed test.

  A one tailed test is used when the alternate hypothesis H_1 is formulated in the following form

  $H_1: \mu > 45$ (Right tail) $H_1: \mu < 45$ (Left tail)

One Tailed vs two Tailed test

- **Two Tailed Test:**

When the rejection region is located at both the ends of the sampling distribution of a test statistic then it is called two tailed test.

A two tailed test is used when the alternate hypothesis H_1 is formulated in the following form

$H_1: \mu \neq 70$

Level of Significance

- **Level of Significance:**

The probability of rejecting the null hypothesis H_0 when H_0 is assumed to be true is called Level of significance. It is denoted by "α". In testing of hypothesis the most commonly used values of "α" are 5% or 1% . By 5% we mean that there are 5 percent chances in 100% of rejecting a true null hypothesis H_0: or in other words we are 95% confident in making a correct decision.

Claim

When a researcher conducts a study, he is generally looking for evidence to support a claim. Therefore, the claim should be stated as the alternative hypothesis, (research hypothesis).

A claim, though, can be stated as either the null hypothesis or the alternative hypothesis;

- A **statistical test** uses the data obtained from a sample to make a decision about whether the null hypothesis should be rejected or not.

- The numerical value obtained from a statistical test is called the **test value**.

Sample mean

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \quad \ell_0$$

z test for the mean when σ is given

- **Testing hypothesis for the mean μ :**
  - **When the value of sample size (n):**

- **population is normal or not normal**
  (n ≥ 30)
  - **σ is known**
    $$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$
  - **σ is not known**
    $$Z = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$
- **population is normal**
  (n< 30)
  - **σ is known**
    $$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$
  - **σ is not known**
    $$T = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$

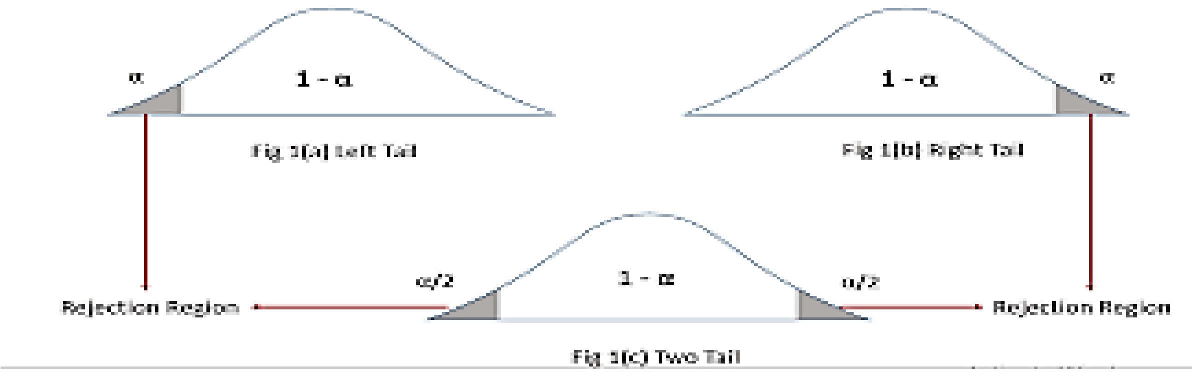
Finding the Critical Values for Specific α Values, Using Table:

Draw the figure and indicate the appropriate area.

1. If the test is left-tailed, the critical region, with an area equal to α, will be on the left side of the mean.

2. If the test is right-tailed, the critical region, with an area equal to α, will be on the right side of the mean.

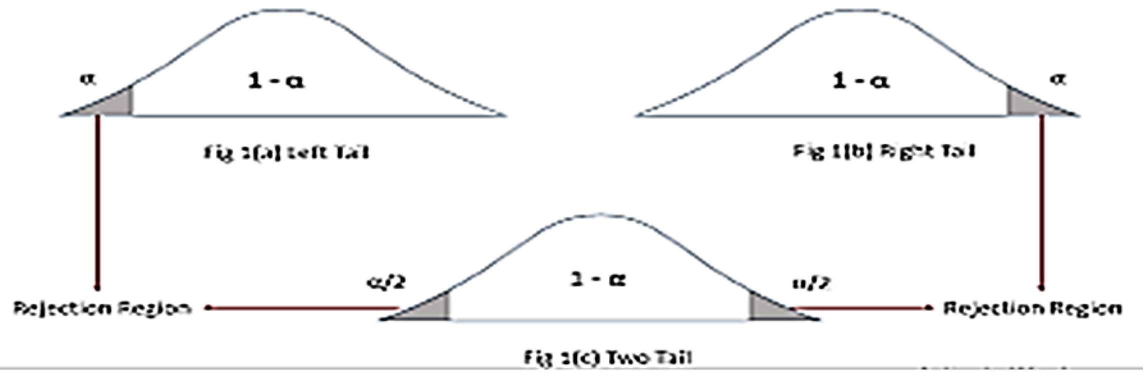
3. If the test is two-tailed, α must be divided by 2: one-half of the area will be to the right of the mean, and one-half will be to the left of the mean.



1. For a left-tailed test, use the z value that corresponds to the area equivalent to α in Table.

2. For a right-tailed test, use the z value that corresponds to the area equivalent to $1 - \alpha$.

3. For a two-tailed test, use the z value that corresponds to $\alpha/2$ for the left value. It will be negative. For the right value, use the z value that corresponds to the area equivalent to $1 - \alpha/2$. It will be positive.



Testing of Hypothesis
(Two Tail Test)

General Procedure for Testing of Hypothesis:

The procedure for testing a hypothesis about the population parameter involves the following **Six steps:**

1) Formulate the null and alternate Hypothesis.
2) Decide the level of significance.
3) Decide upon the test statistic.
4) Computation.
5) Critical Region.
6) Conclusion or Decision.

Example#01

- The average height of females in the freshman class of a certain college has been 162.5 centimeters with a standard deviation of 6.9 centimeters.
- Is there reason to believe that there has been a change in the average height if a random sample of 50 females in the present freshman class has an average height of 165.2 centimeters?
- Use a 0.05 level of significance.

Solution…

So we have $\mu=162.5$, $\delta= 6.9$,
$n=50$, $\bar{x} = 165.2$, $\alpha = 0.05$

- **Step 1:** H_0: $\mu=162.5$ **Vs** H_1: $\mu \neq 162.5$
- **Step 2:** $\alpha = 0.05$
- **Step 3: Test Statistic:** $$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

- **Step 4: Calculation:**

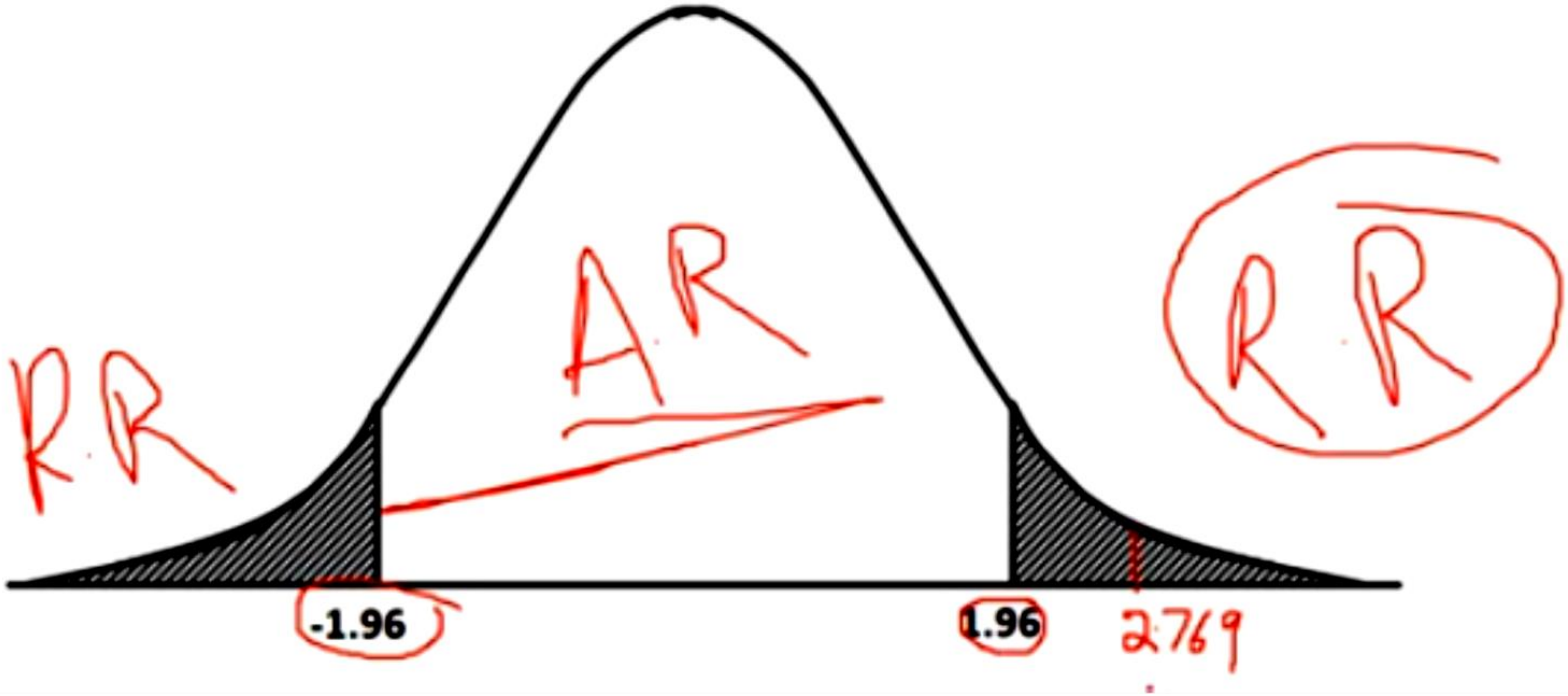
$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

$$= \frac{165.2 - 162.5}{6.9/\sqrt{50}} = \frac{2.7}{6.9/7.07}$$

$$= \frac{2.7}{0.975}$$

$$\mathbf{Z_{cal} = 2.769}$$

α	Two- Tailed Test	One-Tailed test
0.10	±1.645	+1.28(Right tail) -1.28(Left tail)
0.05	±1.96	+1.645(Right tail) -1.645(Left tail)
0.01	±2.58	+2.33(Right tail) -2.33(Left tail)


R.R
A.R
R.R
-1.96
1.96
2.769

- **Step 5: Critical Region:** We will reject H_0: if

$$Z_{cal} \leq -Z_{\alpha/2} \quad \textbf{Or} \quad Z_{cal} \geq Z_{\alpha/2}$$

$$Z_{cal} \leq -1.96 \quad \textbf{Or} \quad Z_{cal} \geq 1.96$$

- **Step 6: Decision:** Since $Z_{cal} = 2.769$ falls in critical region, so therefore we reject H_0: and accept H_1: $\mu \neq 162.5$.

OR

Since $Z_{cal} = 2.769$ is more then the critical value **(1.96)**, so we reject the null Hypothesis and accept the Alternate **H_1: $\mu \neq 162.5$**

Testing of Hypothesis
(One Tail –Test)

General Procedure for Testing of Hypothesis:

The procedure for testing a hypothesis about the population parameter involves the following **Six steps:**

1) Formulate the null and alternate Hypothesis.
2) Decide the level of significance.
3) Decide upon the test statistic.
4) Computation.
5) Critical Region.
6) Conclusion or Decision.

Example

- An oil company introduces a new fuel that they claim contains, on average, **no more than 150 milligrams** of toxic matter per liter.
- We sample 55 separate liters of this fuel and found mean of 154.1 milligrams/liter with a variance of 625.
- A consumer group wants to test this at the α = 10% significance level.

Solution…

Data: $\mu \leq 150$, $S^2 = 625$, $n = 55$, $\bar{x} = 154.1$, $\alpha = 0.10$

- **Step 1:** $H_0: \mu \leq 150$ **Vs** $H_1: \mu > 150$

- **Step 2:** $\alpha = 0.10$

- **Step 3: Test Statistic:** $z = \dfrac{\bar{X} - \mu}{s/\sqrt{n}}$

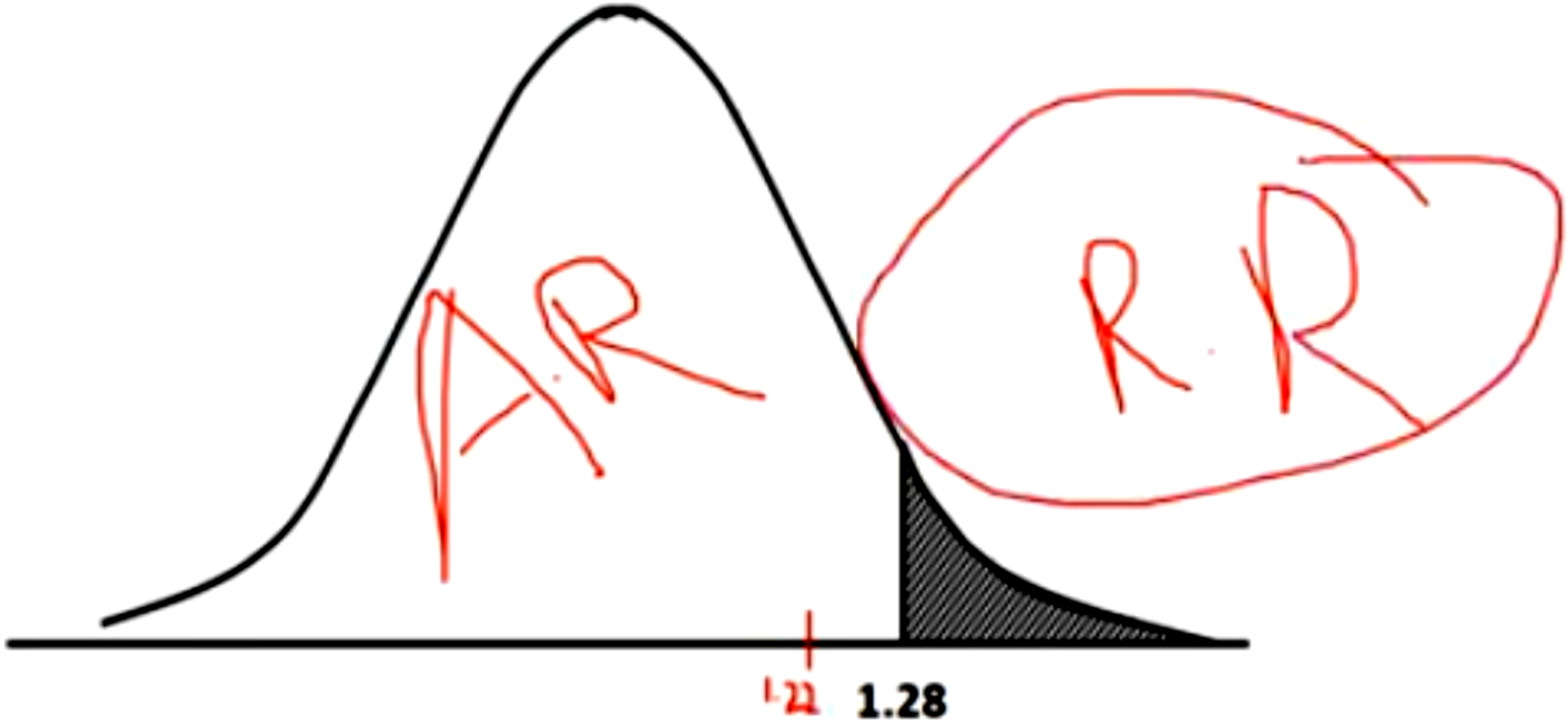
- **Step 4: Calculation:**

$$z = \frac{\overline{X} - \mu}{s/\sqrt{n}} \qquad = \frac{154.1 - 150}{25/\sqrt{55}}$$

$$= \frac{4.1}{25/7.41} = \frac{4.1}{3.37}$$

$$\mathbf{Z_{cal} = 1.22}$$

α	Two- Tailed Test	One-Tailed test
0.10	±1.645	+1.28(Right tail) -1.28(Left tail)
0.05	±1.96	+1.645(Right tail) -1.645(Left tail)
0.01	±2.58	+2.33(Right tail) -2.33(Left tail)


AR
R.R
1.28

- **Step 5: Critical Region:** We will reject H_0: if

  $$Z_{cal} \geq Z_{\alpha}$$

  $$Z_{cal} \geq 1.28$$

- **Step 6: Decision:** Since $Z_{cal} = 1.22$, does not falls in critical region, so therefore we accept **$H_0: \mu \leq 150$**

OR

Since $Z_{cal} = 1.22$ is less then the critical value **(1.28),** so we accept the null hypothesis **$H_0: \mu \leq 150$**

Practice Questions # 1

- A random sample of 100 recorded deaths in the united states during the past year showed an average life of 71.8 years, with a standard deviation of 8.9 years.
- Does this seem to indicate that the average life today is at most 70 years? Use a 0.05 level of significance.

Z-value= 1.645

P Value Method For Hypothesis Testing

Hypothesis Testing

When we perform a hypothesis test in statistics, a *p*-value helps us determine the significance of the results.

The p-value is used as an alternative to rejection points to provide the smallest level of significance at which the null hypothesis would be rejected.

The ***P*-value** (or probability value) is the probability of getting a sample statistic (such as the mean) or a more extreme sample statistic in the direction of the alternative hypothesis when the null hypothesis is true.

Hypothesis Testing

Decision rule when using a P-value

- If P-value $\leq \alpha$, reject the null hypothesis
- If P-value $> \alpha$, do not reject the null hypothesis

Solving Hypothesis-Testing Problems (*P*-Value Method)

Step 1 State the hypotheses and identify the claim.

Step 2 Compute the test value.

Step 3 Find the *P*-value.

Step 4 Make the decision.

Step 5 Summarize the results.

Example: A researcher wishes to test the claim that the average cost of tuition and fees at a two-year college is greater than \$5,550. She selects a random sample of 36 two-year colleges and finds the mean to be \$5,800. The population standard deviation is \$600. Is there evidence to support the claim at $\alpha = 0.05$? Use the *P*-value method.

Step 1: State the hypotheses and identify the claim.

Null H

$H_0 = \mu = \$5{,}500$

Alternative H

$H_1 = \mu > 5{,}500$

Claim

Step 2: Compute the test value.

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{5800 - 5500}{600/\sqrt{36}}$$

$$= \frac{250}{100} \Rightarrow \underline{\underline{2.50}}$$

Step 3: Find the *P*-value.

0.9938

$1 - 0.9938$

$P \; \boxed{= .0062}$

$\alpha = .05$

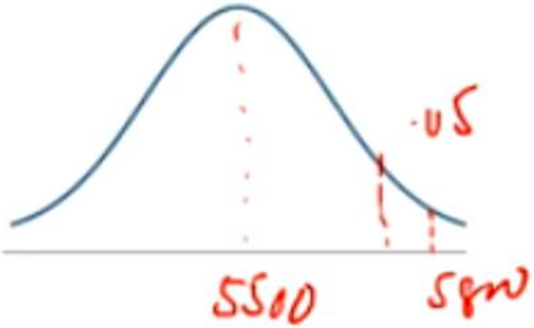
$>$

$P < \alpha$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9958	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

Step 4: Make the decision.

$P < \alpha$



Step 5: Summarize the results.

There is enough evidence to support the claim that the tuition and fees at two-year colleges are greater than $5,550.

$\bar{X}$

$H_0 = \mu = 5,500$

$H_0 = \mu > 5500$

Example 2: A researcher claims that the average wind speed in a certain city is 9 miles per hour. A sample of 36 days has an average wind speed of 9.3 miles per hour. The standard deviation of the population is 0.8 mile per hour. At $\alpha = 0.01$, is there enough evidence to reject the claim? Use the *P*-value method.

Step 1: State the hypotheses and identify the claim.

NA

$$H_0: \mu = 9$$

A H

$$H_0: \mu \neq 9$$

Step 2: Compute the test value.

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{9.3 - 9}{.8/\sqrt{36}} = \frac{0.3}{0.133} = 2.25$$

Step 3: Find the *P*-value.

0.9878

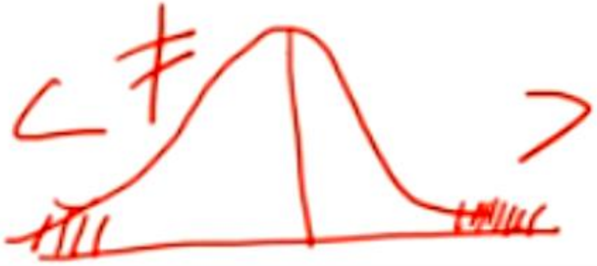
$1 - 0.9878$

$= .0122 \times 2$ ✓

$\boxed{P = 0.0244}$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9958	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

$\alpha = \boxed{0.01}$



$P > \alpha$

Step 4: Make the decision.

The decision is not to reject the null hypothesis, since the *P*-value is greater than 0.01.

Step 5: Summarize the results.

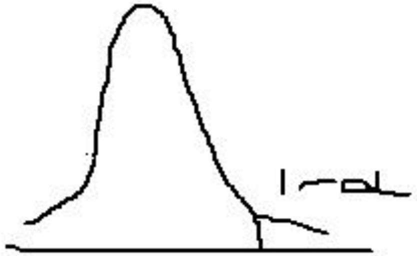
There is not enough evidence to reject the claim that the average wind speed is 9 miles per hour.

Practice Problem: Suppose the average number of Facebook friends is 150. The population standard deviation is 40.3. a random sample of 64 high school students in a particular country revealed that the average number of Facebook friends was 160. At $\alpha = 0.05$, is there sufficient evidence to conclude that the mean number of friends is greater than 150. Ans: P = 0.0239

$H_0: \mu = 150$

$H_1: \mu > 150$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$



Chi Square Test for Comparing Variance (Hypothesis Testing)

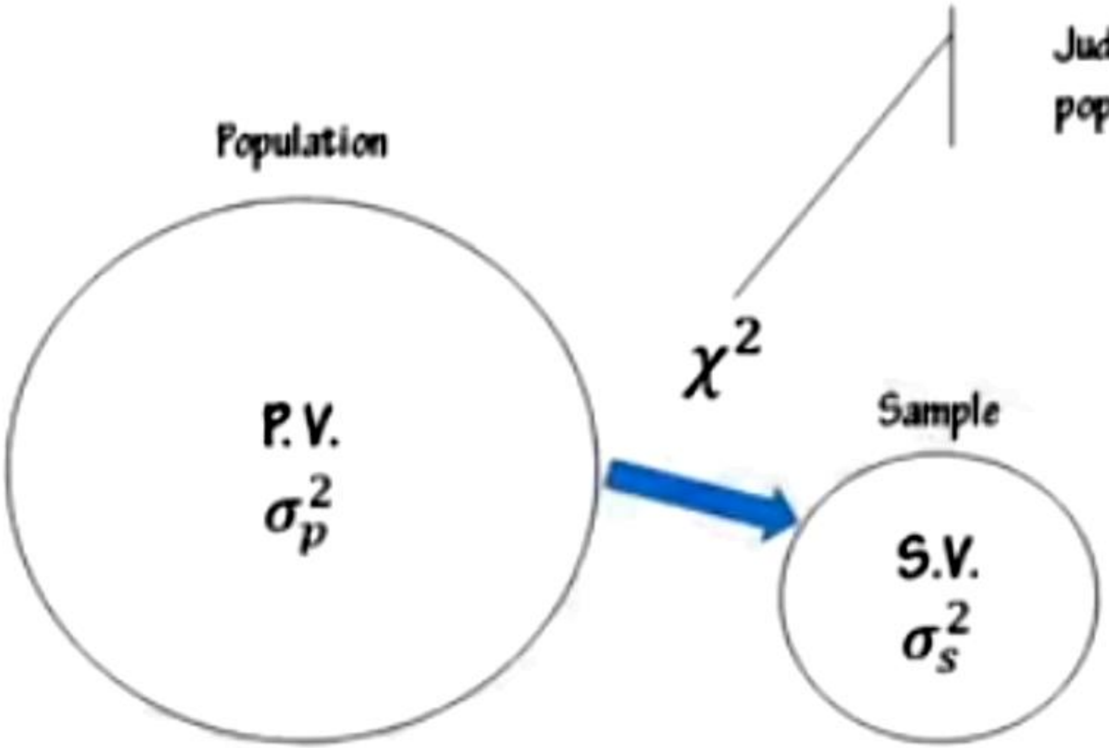
Hypothesis Testing

Chi-Square Test

$$H_0: \sigma_s^2 = \sigma_p^2$$
$$H_a: \sigma_s^2 \neq \sigma_p^2$$

$$\chi^2 = \frac{\sigma_s^2}{\sigma_p^2}(n-1)$$

Chi-Square Test



Judge, if a random sample has been drawn from a normal population with a specified variance.

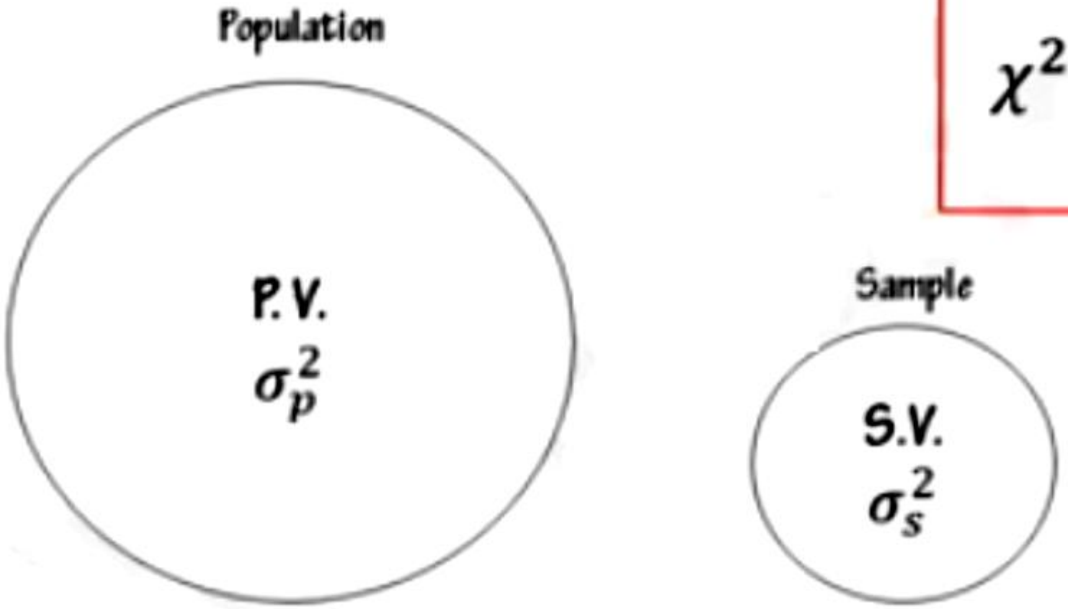
$$H_0: \sigma_s^2 = \sigma_p^2$$

Sample has been drawn from the population

$$H_a: \sigma_s^2 \neq \sigma_p^2$$

Sample has not been drawn from the population

Test Statistic



Sample variance

$$\boxed{\chi^2 = \frac{\sigma_s^2}{\sigma_p^2}(n-1)}$$

Degrees of Freedom

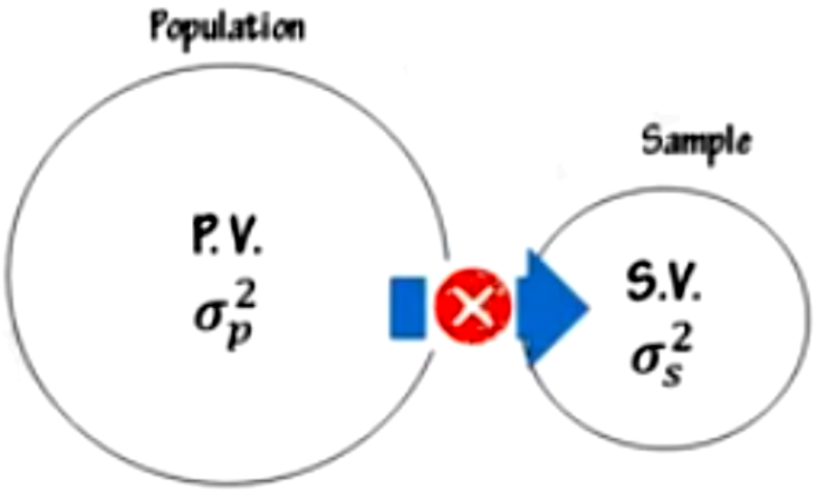
Population variance

Conclusion

If

$$\text{Calculated value of } \chi^2 \geq \text{Critical value of } \chi^2$$

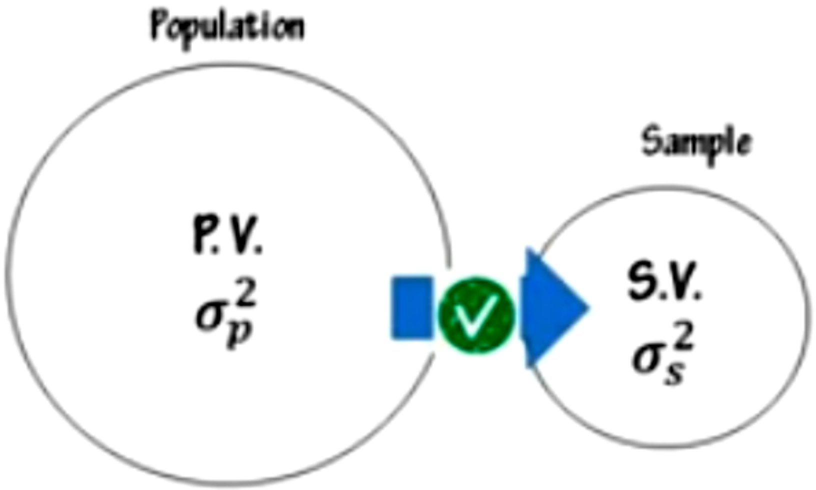
$$H_0: \sigma_s^2 = \sigma_p^2, \text{Rejected}$$



If

$$\text{Calculated value of } \chi^2 < \text{Critical value of } \chi^2$$

$$H_0: \sigma_s^2 = \sigma_p^2, \text{Continued}$$



Problem

Weight distribution of the sample of 10 Students

S. No.	1	2	3	4	5	6	7	8	9	10
Weight (Kg.)	38	40	45	53	47	43	55	48	52	49

- Variance of distribution of weight of all students of a school (σ_p^2) = 20 Kgs
- Can we conclude that the sample of 10 students is drawn from this school at 5% and 1% significance level?

Calculation of Sample Variance

38-47

S.No.	X_i (Weight in Kgs)	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
1	38	-9	81
2	40	-7	49
3	45	-2	4
4	53	6	36
5	47	0	0
6	43	-4	16
7	55	8	64
8	48	1	1
9	52	5	25
10	49	2	4
$n = 10$	$\Sigma X_i = 470$	$\Sigma(X_i - \bar{X})^2 = 280$	

$$\bar{X} = \frac{\sum X_i}{n} = \frac{470}{10} = 47$$

$$\sigma_s^2 = \frac{\Sigma(X_i - \bar{X})^2}{n-1} = \frac{280}{10-1} = 31.11$$

Calculating the Value of χ^2

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

$$\chi^2 = \frac{\sigma_s^2}{\sigma_p^2}(n-1)$$

$$\boxed{\sigma_s^2 = 31.11,\ \sigma_p^2 = 20,\ n = 10}$$

$$\chi^2 = \frac{31.11}{20} \times (10-1)$$

$$\Rightarrow \chi^2 = 13.99$$

Conclusion

Compare $\chi^2_{Cal.}$ to $\chi^2_{Crit.}$ for D.F.=(n-1)=9, at 5% and 1% Significance level

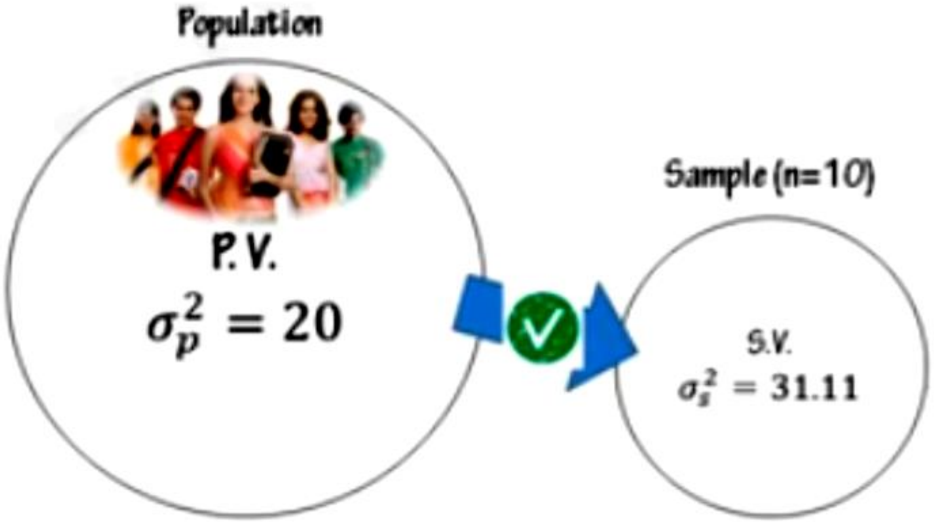
$\chi^2_{Crit.} = 16.92$ at 5%

$\chi^2_{Crit.} = 21.67$ at 1%

$\chi^2_{Cal.} = 13.99 < 16.92$

$\chi^2_{Cal.} = 13.99 < 21.67$

$H_0: \sigma_s^2 = \sigma_p^2$, Continued



Sample is taken from the population with variance 20Kgs

Critical Values of the Chi-Square Distribution

D.F	0.99	0.95	0.90	0.10	0.05	0.025	0.010	0.005
1	0.000157	0.00393	0.015791	2.706	3.841	5.024	6.635	7.879
2	0.0201	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.1148	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.2971	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.5543	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.8721	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	1.2390	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.6465	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	2.0879	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.5582	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	3.0535	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.5706	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	4.1069	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.6604	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	5.2293	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.8122	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	6.4078	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	7.0149	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	7.6327	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	8.2604	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.8972	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	9.5425	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	10.1957	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	10.8564	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	11.5240	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	12.1981	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	12.8785	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	13.5647	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	14.2565	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	14.9535	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	22.1643	26.509	29.051	51.805	55.758	59.342	63.691	66.766

T-Test

What is T-test and when to use it?

When to use T test?

- A *t* test can only be used when comparing the means of **two groups.**

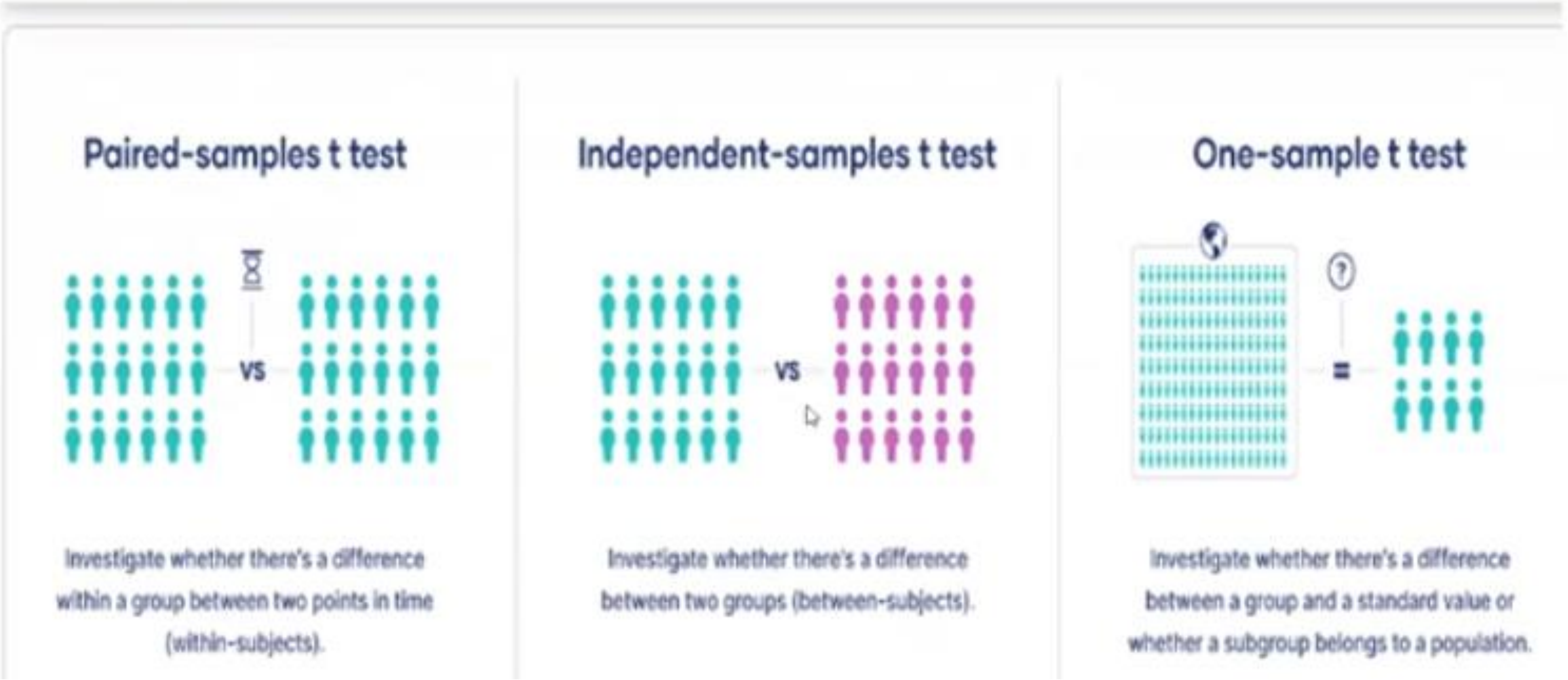
- If you want **to compare more than two groups**, or if you want to do multiple pairwise comparisons, **use an ANOVA** test or a post-hoc test

Introduction

- **T-test**, is a method of testing hypotheses about the mean of a **small sample drawn** from a normally distributed population when the Population **standard deviation is unknown**.

- The t-test tells you how significant the differences between groups are; In other words it lets you know if those differences (measured in means) could have happened by chance.

One Sample, Two Independent Samples & Paired Samples



One sample t test Examples

- For example, imagine a company wants to test the claim that their batteries last more than 40 hours. Using a simple random sample of 15 batteries yielded a mean of 44.9 hours, with a standard deviation of 8.9 hours. Test this claim using a significance level of 0.05.

Two Independent Samples t-test

- If the groups come from **two different populations** (e.g., two different species, or people from two separate cities), perform a **two sample t test (independent *t* test**).

Paired Sample t test example

- If the groups come from a single population (e.g., measuring **before and after** an experimental treatment like, the mean blood pressure is the same before and after using the drug), we perform a **paired *t* test**.

- Example, A doctor may want to know if some new drug leads to a significant reduction in blood pressure compared to the current standard drug used

Conditions for t-Test

- Population is normal
- Sample size is small (<30)
- Population standard deviation (σ_p) is unknown
- Sample standard deviation (σ_s) is known or can be calculated from the given sample.

Test Statistic for t-Test

$$t = \frac{\bar{X} - \mu_0}{\sigma_s / \sqrt{n}},$$

with degrees of freedom (d.f.) = (n-1)

- $\bar{X}$ sample mean
- μ_0 claimed population mean
- σ_s is sample standard deviation
- 'n' is the sample size
- Degrees of freedom, if n=20 then, d.f.= n-1=20-1 =19

Calculation of σ_S, if not given

$$\sigma_s = \sqrt{\frac{\sum(X_i - \bar{X})^2}{(n-1)}}$$

Sum of squared deviations of each sample value (X_i) from the mean of the sample ($\bar{X}$)

Z-Test and t-Test Conditions

Z-Test	t-Test
• Population is normal	• Population is normal
• Sample size is large or small (<30)	• Sample size is small (<30)
• σ_p is known	• σ_p is unknown

Concluding, Accept or Reject H_0

1. Process the test statistic and calculate the value of 't'

$$t = \frac{\bar{X} - \mu_0}{\sigma_s / \sqrt{n}}$$

2. $\boxed{\text{Find } t_{Cri.} \text{ from t-distribution at a given value of } \alpha \text{ against the d.f.}}$

3. Compare $t_{Cal.}$ and $t_{Cri.}$

$$t_{Cal.} < t_{Cri.} \mid \text{Accept } H_0$$

$$t_{Cal.} > t_{Cri.} \mid \text{Reject } H_0$$

Problem Description

Conclude that the Restaurant's sales have increased at 5% significance level in the given situation.

- A Restaurant near the railway station has been having average sales of 500 tea cups daily.
- Because of the development of the bus stand nearby, it expects to increase its sales.
- During first 12 days after the start of bus stand, the data of daily sales given as;
  550, 570, 490, 615, 505, 580, 570, 460, 600, 580, 530, and 526

Parameters Given

- Average sales before the start of the bus stand
  (μ_0) =500 tea cups/day
- First 12 days sales (tea cups/day) after the start of bus stand
  550,570,490,615,505,580,570,460,600,580,
  530,526
- Significance level
  $$\alpha = 5\%$$
- Conclusion to draw
  Average sales (tea cups/day) increased after the start of the bus stand or (μ) >500 tea cups/day

Null and Alternative Hypothesis

$H_0: \mu = 500$ tea cups/day

(it implies the average sales as 500 tea cups/day will continue after the start of the Bus stand)

$\boxed{H_a: \mu > 500 \text{ tea cups/day}}$

(it implies the average sales will be more than 500 tea cups per day after the start of the Bus stand)

- H_a is one sided and μ is greater than 500 tea cups/day
- Right tailed test condition

st Statistic

Given that

- Sample size $(n) = 12$ (<30)
- Claimed population mean $(\mu_0) = 500$

$$t = \frac{\bar{X} - \mu_0}{\sigma_s / \sqrt{n}}$$

- First 12 days sales (tea cups/day) after the start of bus stand
  550, 570, 490, 615, 505, 580, 570, 460, 600, 580, 530, 526
- Calculate from the given sample data
  - ✓ Sample mean $(\bar{X})$
  - ✓ Sample standard deviation (σ_s)

Sample Mean ($\bar{X}$) and Sample Std. Dev. (σ_s)

Day	X_i (Daily Sales)	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
1	550	550 – 548 → 2	2^2 → 4
2	570	22	484
3	490	-58	3364
4	615	67	4489
5	505	-43	1849
6	580	32	1024
7	570	22	484
8	460	-88	7744
9	600	52	2704
10	580	32	1024
11	530	-18	324
12	526	-22	484
ΣX_i	6576	$\Sigma(X_i - \bar{X})^2 = 23978$	
n	12		

$$\bar{X} = \frac{\Sigma X_i}{n} = \frac{6576}{12} = 548$$

$$\sigma_s = \sqrt{\frac{\Sigma(X_i - \bar{X})^2}{n-1}}$$

$$\sigma_s = \sqrt{\frac{23978}{12-1}} = 46.69$$

Calculation of t

- Sample mean $(\bar{X}) = 548$
- Claimed population mean $(\mu_0) = 500$
- Sample standard deviation (σ_s)=46.69
- Sample size $(n) = 12$

$$t = \frac{\bar{X} - \mu_0}{\sigma_s / \sqrt{n}} \Rightarrow t = \frac{548 - 500}{46.69 / \sqrt{12}} = \frac{48}{13.49} = 3.56$$

t-Test

- Sample size (n)=12
- d.f. = (n-1)=12-1=11
- H_a is right tailed, $\alpha = 5\%$
- $t_{0.05}$ against 11 d.f.

$$t_{Cal.} = 3.56 > t_{Cri.} = 1.796$$

H_0, Rejected

Critical Values of the t-Distribution (t_α)					
d.f.	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.988
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831

Conclusion

$$t_{Cal.} = 3.56 > t_{Cri.} = 1.796$$

$H_0: \mu = 500$ tea cups/day after the start of the bus stand, Rejected

$H_a: \mu > 500$ tea cups/day after the start of the bus stand,