**Insurance Claim Example**

Recently the director board of the Thakushika company has introduced a medical insurance policy for its employees. According to the experience in other companies the probability that any employee made an insurance claim in any given month is 0.4. If the company has 900 employees find the probability that more than 375 employees will made insurance claims in the next month.

**Employee Absenteeism Example**

Management in a company with 600 employees is very much concerned about the absenteeism of its employees and its impact on the production. They have thus estimated, using past records, the probability that any employee is absent for work in any given day is 0.02. Find the probability that there will be more than 5 employees absent in any given day.

**Hypothesis Testing**

- We estimate parameters by using
  - past records
  - similar experience

Are they correct??
DIPLOMA PROGRAM EXAMPLE

The coordinator of a Diploma program claims that 80 percent of the candidates who complete the program successfully got employed. This year a group of 15 students joined the program completed successfully. Find the probability that
i. All of them got employed
ii. None of them got employed
iii. At least 10 of them got employed

PLANT NURSERY EXAMPLE

In a plant nursery it is guaranteed that 90 percent of the seeds in a packet will germinate successfully. If a farmer bought a packet having 15 seeds, find the probability that
I. Exactly 10 seeds will germinate
II. More than 12 seeds will germinate
III. More than 5 seeds will not germinate

HYPOTHESIS TESTING

• We made certain claims
• We guaranteed certain things

Are they correct ??

NULL HYPOTHESIS

• Testing a Belief
• Testing a Claim

Claim / Belief Null Hypotheses - $H_0$
**EXAMPLES**

Testing whether the population mean is 100.

\[ H_0 : \mu = 100 \]

Testing whether the population proportion is 0.5.

\[ H_0 : \rho = 0.5 \]

Testing whether the population mean is less than or equal to 75.

\[ H_0 : \mu \leq 75 \]

**ALTERNATIVE HYPOTHESIS**

What will happen if null hypothesis is rejected

Alternative Hypotheses - \( H_1 \)

Testing whether the population mean is less than 100.

\[ H_0 : \mu \geq 100 \]

\[ H_1 : \mu < 100 \]

Testing whether the population proportion is greater than 0.5

\[ H_0 : \rho \leq 0.5 \]

\[ H_1 : \rho > 0.5 \]
HOW TO TEST

TESTING FOR POPULATION MEAN

- TEST STATISTIC
  \[ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \]

- HYPOTHESIS
  \[ H_0 : \mu = 100 \]
  \[ H_1 : \mu \neq 100 \]

- DECISION
  Reject \( H_0 \) if \( Z \) is very LARGE or very SMALL

\[ P(-1.96 < Z < 1.96) = 0.95 = 95\% \]
\[ P(Z < -1.96 \text{ or } Z > 1.96) = 0.05 = 5\% \]

If \( Z > 1.96 \) or \( Z < -1.96 \)
we reject \( H_0 \) with 95% confidence
(at 5% level of significance)
TESTING FOR POPULATION MEAN

- TEST STATISTIC
  \[ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \]

- HYPOTHESIS
  \[ H_0: \mu \leq 100 \]
  \[ H_1: \mu > 100 \]

- DECISION
  Reject \( H_0 \) if \( Z \) is very LARGE

DAILY MEAL EXAMPLE

According to past records, the average spending per person for the daily meals is Rs. 250 with a standard deviation of Rs. 40. In a study to find out whether the rates applied today will differ from the previous rates, a sample of 36 people were selected and their mean spending for daily meals was calculated as Rs. 265. Test whether the average spending has changed or not with 95% confidence.
**X = SPENDING FOR DAILY MEALS**

- **HYPOTHESIS**
  
  \[ H_0 : \mu = 250 \]
  
  \[ H_1 : \mu \neq 250 \]

  \[ n = 36 \] \[ \bar{X} = 265 \] \[ \sigma = 40 \]

- **TEST STATISTICS**

  \[ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \]

  Under the null hypothesis, \( Z = \frac{\bar{X} - 250}{\sigma/\sqrt{n}} \)

  By substituting the values for \( \sigma \), \( \bar{X} \) and \( n \)

  \[ Z = \frac{265 - 250}{40/\sqrt{36}} = 2.25 \]

  \[ Z > 1.96 \]

  we reject \( H_0 \) with 95% confidence

  (at 5% level of significance)

  \[ -2.57 < Z < 2.57 \]

  we cannot reject \( H_0 \) with 99% confidence

  (at 5% level of significance)

  **AVGAREGE SPENDING HAS CHANGED.**

  **CANNOT SAY AVERAGE SPENDING HAS CHANGED WITH 99% CONFIDENCE.**
In an attempt to test whether or not a special training program has improved the productivity of the employees in a company, a sample of 25 employees were selected and the time taken to do a particular job after the training was measured. The average time required to do the job by these trained employees was calculated as 18 minutes. The time required to do the same type of a job by an untrained employee is 20 minutes with a standard deviation 16 minutes. Test whether the training was able to reduce the time to do the job significantly.

\[ X = \text{TIME REQUIRED TO DO A JOB} \]

- **HYPOTHESIS**
  
  \[ H_0: \mu \geq 20 \]
  
  \[ H_1: \mu < 20 \]

- **TEST STATISTICS**

  \[ \begin{align*}
  n &= 25 \\
  \bar{X} &= 18 \\
  \sigma &= 16
  \end{align*} \]

\[ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \]

- Test Statistic = \[ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \]

- Under the null hypothesis, \[ Z = \frac{\bar{X} - 20}{\sigma/\sqrt{n}} \]

- By substituting the values for \( \sigma \), \( \bar{X} \) and \( n \)

\[ Z = \frac{18 - 20}{16/\sqrt{25}} = -0.625 \]

\[ Z > -0.625 \]

We cannot reject \( H_0 \) with 95% confidence (at 5% level of significance)

We cannot say that the training has reduced the time required to the job significantly.
ELECTRIC RESISTORS EXAMPLE

A firm producing small electric resistors claims that they will not allow more than 0.75 amperes to pass by under normal conditions. However, when a sample of 32 resistors was tested, it was found that the average amperes allowed through them is 0.9 amperes. Test the manufactures claim at 10% level assuming the population variance is 0.5 amperes.

\[ X = \text{CURRENT PASSES THROUGH A RESISTOR} \]

\[ \text{HYPOTHESIS} \]

\[ H_0 : \mu \leq 0.75 \]
\[ H_1 : \mu > 0.75 \]

\[ n = 32 \quad \bar{X} = 0.9 \quad \sigma^2 = 0.5 \]

\[ \text{TEST STATISTICS} \]

\[ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \]

\[ Z < 1.28 \]

we cannot reject \( H_0 \) with 90% confidence (at 10% level of significance)

\[ Z \]

\[ \begin{array}{c}
0.90 \\
0.10 \\
1.28 \\
1.2
\end{array} \]

WE CANNOT SAY THAT MORE THAN 0.75 AMPERES PASSES THROUGH A RESISTOR.
**UNKNOWN POPULATION VARIANCE**

- Test Statistic: \( Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \)
- How to Calculate?
- Estimate \( \sigma^2 \) by \( s^2 \)
- Then test Statistic becomes: \( \frac{\bar{X} - \mu}{s/\sqrt{n}} \)

**LIGHT BULBS EXAMPLE**

A manufacturer claims that the average life time of 60W light bulbs is 750 hours or more. In order to test this statement a sample of size 12 is selected and the following data were collected (life time in hours).

670  790  595  435  740  780  805  705
845  880  625  750

Test the above statement at 5% level of significance.

**X = LIFE TIME OF A LIGHT BULB**

- HYPOTHESIS
  \( H_0: \mu \geq 750 \)
  \( H_1: \mu < 750 \)

- TEST STATISTICS
  \[ t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \]

\( \bar{X} = 718.33 \quad s = 123.15 \)
\( X = \text{LIFE TIME OF A LIGHT BULB} \)

- Test Statistic = \( t = \frac{\overline{X} - \mu}{s/\sqrt{n}} \)
- Under the null hypothesis, \( t = \frac{\overline{X} - 750}{s/\sqrt{n}} \)
- By substituting the values for \( s \), \( \overline{X} \) and \( n \)
  \[
  t = \frac{718.333 - 750}{123.15/\sqrt{12}} = -0.89
  \]

\(-0.196 > -0.05\)

we cannot reject \( H_0 \) with 95% confidence
(at 5% level of significance)

WE CANNOT SAY THAT THE LIFE TIME OF LIGHT BULB IS SIGNIFICANTLY LESS THAN 750 HOURS
MICROSOFT EXCEL INSTRUCTIONS

1. Click on PHStat tab
2. Select One Sample Test, t-test for Mean, Sigma unknown
3. Enter Hypothesized Mean
4. Check "Sample Statistics unknown"
5. Check Test Statistics > Lower Tail Test

MINITAB INSTRUCTIONS

Instructions necessary for the above output are

1. Select Stat
2. Select Basic Statistics
3. Select 1-Sample t
4. Select Data Range
5. Click options
6. Enter Confidence Level
7. Click OK

MINITAB OUTPUT

One-Sample T: C1
Test of mu = 750 vs < 750

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SE Mean</th>
<th>95% Upper Bound</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>12</td>
<td>718.333</td>
<td>123.147</td>
<td>35.549</td>
<td>782.176</td>
<td>-0.89</td>
<td>0.196</td>
</tr>
</tbody>
</table>

POLLUTION COUNT EXAMPLE

Environmental Authorities has recently announced that any area will be suitable for living only if the average pollution count does not exceed 150. In order to test the suitability of a newly partitioned area for living the pollution counts were collected over a period of 20 consecutive days and the following results were obtained:

145 157 152 150 148 153 155 149
146 145 152 145 150 159 142 148
156 147 153 162

Test whether the area is suitable for living.
**X = Pollution Count for the City**

- **Hypothesis**
  
  \[ H_0: \mu \leq 150 \]
  
  \[ H_1: \mu > 150 \]

\[ N = 20 \quad \bar{X} = 150.7 \quad S = 5.25 \]

- **Test Statistics**
  
  \[ t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \]

**X = Life Time of a Light Bulb**

- **Test Statistic**
  
  \[ t = \frac{\bar{X} - 150}{S/\sqrt{n}} \]

- **Under the null hypothesis,**
  
  \[ t = \frac{150.7 - 150}{5.25/\sqrt{20}} = 0.59 \]

\[ Z < 1.729 \]

we cannot reject \( H_0 \) with 95% confidence 
(at 5% level of significance)

**Microsoft Excel Output**

<table>
<thead>
<tr>
<th>Data</th>
<th>( \mu \leq 150 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Significance</td>
<td>0.05</td>
</tr>
<tr>
<td>Sample Size</td>
<td>20</td>
</tr>
<tr>
<td>Sample Mean</td>
<td>150.7</td>
</tr>
<tr>
<td>Sample Standard Deviation</td>
<td>5.252568294</td>
</tr>
<tr>
<td>Intermediate Calculations</td>
<td></td>
</tr>
<tr>
<td>Standard Error of the Mean</td>
<td>1.174509976</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>19</td>
</tr>
<tr>
<td>t Test Statistic</td>
<td>0.595993235</td>
</tr>
<tr>
<td>Upper-Tail Test</td>
<td></td>
</tr>
<tr>
<td>Upper Critical Value</td>
<td>1.729132792</td>
</tr>
<tr>
<td>p-Value</td>
<td>0.279106899</td>
</tr>
</tbody>
</table>

Do not reject the null hypothesis

WE CANNOT SAY THAT THE AREA IS NOT
SUITABLE FOR LIVING
A small retailer was planning to buy a new vehicle for his business. He estimated the price of his dream vehicle as 2.4 million rupees, but unable to purchase it so far due to the financial constraints he faced. Now due to the changes in the import policies of the new government, different people argue that prices of the vehicles has also changed. Therefore, in order to test whether the price of his dream vehicle has changed or not, he checked the prices of 18 such vehicles (in million rupees) advertised in the weekend newspapers and the following data were collected.

2.3  2.5  1.9  2.2  2.4  2.3  1.8  2.0  2.2  
2.4  2.0  2.1  2.2  1.8  1.9  2.4  2.1  2.0

Test at 5% level whether the price of his dream vehicle has changed.

**HYPOTHESIS**

- \( H_0 : \mu = 2.4 \)
- \( H_1 : \mu \neq 2.4 \)

**TEST STATISTICS**

\[
t = \frac{\bar{X} - \mu}{s/\sqrt{n}}
\]

\[\bar{X} = 2.14 \quad s = 0.22 \]

\[n = 18\]

\[t = \frac{2.14 - 2.4}{0.22/\sqrt{18}} = -5.10\]

\[Z < -2.11\]

We reject \( H_0 \) with 95% confidence (at 5% level of significance)

**PRICE OF THE VEHICLE HAS CHANGED.**
**MICROSOFT EXCEL OUTPUT**

<table>
<thead>
<tr>
<th>Data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Null Hypothesis</td>
<td>( \mu = 2.4 )</td>
</tr>
<tr>
<td>Level of Significance</td>
<td>0.05</td>
</tr>
<tr>
<td>Sample Size</td>
<td>18</td>
</tr>
<tr>
<td>Sample Mean</td>
<td>2.138888889</td>
</tr>
<tr>
<td>Sample Standard Deviation</td>
<td>0.217306747</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intermediate Calculations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Error of the Mean</td>
<td>0.051219691</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>17</td>
</tr>
<tr>
<td>t Test Statistic</td>
<td>-5.097865759</td>
</tr>
</tbody>
</table>

Two-Tail Test

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Critical Value</td>
<td>-2.109815559</td>
</tr>
<tr>
<td>Upper Critical Value</td>
<td>2.109815559</td>
</tr>
<tr>
<td>p-Value</td>
<td>8.93343E-05</td>
</tr>
</tbody>
</table>

Reject the null hypothesis

---

**POPULATION VARIANCE**

- **TEST STATISTIC**
  \[ \chi^2 = \frac{(n - 1)S^2}{\sigma^2} \]

- **HYPOTHESIS**
  \[ H_0 : \sigma^2 = \sigma_1^2 \]
  \[ H_1 : \sigma^2 \neq \sigma_1^2 \]

- **DECISION**
  Reject \( H_0 \) if \( \chi^2 \) is very LARGE or very SMALL

---

α %

<table>
<thead>
<tr>
<th>D.F</th>
<th>0.975</th>
<th>0.95</th>
<th>0.50</th>
<th>0.10</th>
<th>0.05</th>
<th>0.025</th>
<th>0.01</th>
<th>0.005</th>
<th>0.001</th>
<th>0.0001</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.001</td>
<td>0.004</td>
<td>0.045</td>
<td>2.71</td>
<td>3.84</td>
<td>5.02</td>
<td>6.64</td>
<td>10.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.021</td>
<td>0.051</td>
<td>1.39</td>
<td>4.61</td>
<td>5.99</td>
<td>7.38</td>
<td>9.21</td>
<td>13.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.064</td>
<td>0.103</td>
<td>2.37</td>
<td>6.25</td>
<td>7.82</td>
<td>9.35</td>
<td>11.3</td>
<td>16.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.141</td>
<td>0.171</td>
<td>3.36</td>
<td>7.78</td>
<td>9.49</td>
<td>11.1</td>
<td>13.3</td>
<td>18.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.216</td>
<td>0.236</td>
<td>4.35</td>
<td>9.24</td>
<td>11.1</td>
<td>12.8</td>
<td>15.1</td>
<td>20.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.281</td>
<td>0.294</td>
<td>5.35</td>
<td>10.6</td>
<td>12.6</td>
<td>14.5</td>
<td>16.8</td>
<td>22.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.341</td>
<td>0.349</td>
<td>6.35</td>
<td>12.0</td>
<td>14.1</td>
<td>16.0</td>
<td>18.5</td>
<td>24.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.400</td>
<td>0.407</td>
<td>7.34</td>
<td>13.4</td>
<td>15.5</td>
<td>17.5</td>
<td>20.1</td>
<td>26.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.454</td>
<td>0.458</td>
<td>8.34</td>
<td>14.7</td>
<td>16.9</td>
<td>19.0</td>
<td>21.7</td>
<td>27.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.504</td>
<td>0.509</td>
<td>9.34</td>
<td>16.0</td>
<td>18.3</td>
<td>20.5</td>
<td>23.2</td>
<td>29.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.606</td>
<td>0.615</td>
<td>11.3</td>
<td>18.5</td>
<td>21.0</td>
<td>23.3</td>
<td>26.2</td>
<td>32.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.720</td>
<td>0.727</td>
<td>14.3</td>
<td>22.3</td>
<td>25.0</td>
<td>27.5</td>
<td>30.6</td>
<td>37.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.930</td>
<td>0.934</td>
<td>19.3</td>
<td>28.4</td>
<td>31.4</td>
<td>34.2</td>
<td>37.6</td>
<td>45.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>1.210</td>
<td>1.219</td>
<td>23.3</td>
<td>33.2</td>
<td>36.4</td>
<td>39.4</td>
<td>43.0</td>
<td>51.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1.601</td>
<td>1.610</td>
<td>29.3</td>
<td>40.3</td>
<td>43.8</td>
<td>47.0</td>
<td>50.9</td>
<td>59.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
EXAMPLE

In order to test the variation in the fat content of mutton, ten equal size pieces of mutton was collected and their percentage fat content was measured and given below.

30  26  29  19  24  35  28  36  
27  32

Test whether the variation of fat content is equal to 20% or not at 10% level of significance.

X = PERCENTAGE FAT CONTENT IN MUTTON

• HYPOTHESIS

$H_0 : \sigma^2 = 20$
$H_1 : \sigma^2 \neq 20$

$S^2 = 25.82$

• TEST STATISTICS

$$\chi^2 = \frac{(n-1)S^2}{\sigma_1^2}$$

• Test Statistic

$\chi^2 = \frac{(n-1)S^2}{\sigma_1^2}$

• Under the null hypothesis, $\chi^2 = \frac{(n-1)S^2}{20}$

• By substituting the values for s and n

$t = \frac{(10-1)25.82}{20} = 11.62$

3.33 < $\chi^2$ < 16.9
we cannot reject $H_0$ with 90% confidence
(at 10% level of significance)
**POPULATION PROPORTION**

• TEST STATISTIC
  
  \[ Z = \frac{\pi - p}{\sqrt{\frac{p(1 - p)}{n}}} \sim N(0, 1) \]

• HYPOTHESIS
  
  \[ H_0 : p = p_1 \]
  \[ H_1 : p \neq p_1 \]

• DECISION
  
  Reject \( H_0 \) if \( \chi^2 \) is very LARGE or very SMALL

---

**DEFECTIVE CUPS EXAMPLE**

Plastic cups produced by a manufacturer is packed as 50 cups per one packet. He says that there can be defective cups but it will not exceed a maximum of two cups per packet. In order to test his statement 10 such packets were checked and the number of defective cups was recorded as follows.

0 2 1 2 1 1 3 0 1 2

Test the manufacturer’s claim at 5% level of significance.

**X = PROPORTION OF DEFECTIVE CUPS PER PACKET**

• HYPOTHESIS
  
  \[ H_0 : \rho \leq 0.04 \]
  \[ H_1 : \rho > 0.04 \]

\( \rho = 0.026 \) \( n = 10 \)

• TEST STATISTIC
  
  \[ Z = \frac{\rho - p}{\sqrt{\frac{p(1 - p)}{n}}} \]

**P = PROPORTIONATE FAT CONTENT IN MUTTON**

• Test Statistic = \( Z = \frac{\pi - p}{\sqrt{\frac{p(1 - p)}{n}}} \)

• Under the null hypothesis, \( Z = \frac{0.026 - 0.04}{\sqrt{\frac{0.04(1 - 0.04)}{10}}} \)

• By substituting the values for \( p \) and \( n \) \( Z = \frac{0.026 - 0.04}{\sqrt{\frac{0.04(1 - 0.04)}{10}}} = -0.2259 \)
Test Statistic, $z = -0.2259$

If $Z < 1.64$
we cannot reject $H_0$ with 95% confidence
(at 5% level of significance)

---

**PACKAGE DESIGNER EXAMPLE**

A package designer wishes to determine customers’ preferences on two new package designs. He placed equal number of packages from each design on the shelf of a supermarket and observed that how many of each design will purchase by the customers during a period of one week. At the end of the week 150 and 175 packages were purchased from the two designs respectively. Test whether is there any difference in customers’ preferences at 5% level.

**X = PROPORTION OF CUSTOMERS PREFERRED FIRST DESIGN**

- **HYPOTHESIS**
  
  $H_0: \pi = 0.5$
  
  $H_1: \pi \neq 0.5$
  
  $\pi = 0.46 \quad n = 325$

- **TEST STATISTIC**

  $$Z = \frac{\pi - p}{\sqrt{p(1-p)}} \frac{1}{\sqrt{n}}$$

**P = PROPORTIONATE FAT CONTENT IN MUTTON**

- **Test Statistic**

  $$Z = \frac{\pi - 0.5}{\sqrt{\frac{0.5(1-0.5)}{n}}}$$

- Under the null hypothesis, $Z = \frac{0.46 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{325}}} = -1.44$
\[-1.96 < Z < 1.96\]

we cannot reject \( H_0 \) with 99\% confidence
(at 5\% level of significance)

\[0.025 \quad 0.95 \quad 0.025\]

\[-1.96 \quad 1.96\]

\[-1.44\]

CANNOT SAY AVERAGE SPENDING HAS CHANGED WITH 95\% CONFIDENCE.