Hypothesis Testing

Hypothesis testing allows us to use a sample to decide between two statements made about a Population characteristic. These two statements are called the Null Hypothesis and the Alternative Hypothesis.

Definitions

\( H_0: \text{The Null Hypothesis} \)
This is the hypothesis or claim that is initially assumed to be true.

\( H_A: \text{The Alternative Hypothesis} \)
This is the hypothesis or claim which we initially assume to be false but which we may decide to accept if there is sufficient evidence.
This procedure is familiar to us already from the legal system: “Innocent until proven guilty”.

The Null Hypothesis (Innocent) is only rejected in favour of the Alternative Hypothesis (Guilty) if there is sufficient evidence of this “beyond reasonable doubt”.

So to summarise $H_0$ is the status quo and $H_A$ is what we want to prove using the data we have collected.

$H_0$ and $H_A$ take the following form:

**Null Hypothesis**

$H_0 : \text{Population Characteristic} = \text{Hypothesised Value}$

**Alternative Hypothesis** ~ three possibilities

*Upper tailed test:*

$H_A : \text{Population Characteristic} > \text{Hypothesised Value}$

*Lower tailed test:*

$H_A : \text{Population Characteristic} < \text{Hypothesised Value}$

*Two tailed test:*

$H_A : \text{Population Characteristic} \neq \text{Hypothesised Value}$

NOTE: The same Hypothesised value must be used in the Alternative Hypothesis as in the Null Hypothesis
A first look at the test procedure

1. We initially assume the null hypothesis is true, this gives us some information about a population characteristic for example: \( H_0: \mu = 3 \).

2. We then take a sample from this population and measure the sample statistic which corresponds to the population characteristic under investigation. eg: \( H_0: \mu = 3 \) then we measure the sample mean \( \bar{x} \)

3. If the sample statistic is sufficiently far away from this hypothesized value then we conclude that the Null Hypothesis is probably false. IE: the sample we have chosen probably could not have come from a Population with the Hypothesized Characteristic.

Example:

**Null Hypothesis:** The average (mean) rent for houses in Dublin is €1400 per month

**Alternative Hypothesis:** The average (mean) rent for houses in Dublin is less than €1400

We take a random sample of houses and measure the sample mean rent. We compare this sample mean with €1400 and if it is sufficiently lower than €1400 we reject the null hypothesis ie we conclude that the alternative hypothesis is true.
Note: We can sometimes make mistakes, suppose the average rent for houses in Dublin really is €1400. Let’s also suppose that even though we choose a random sample it happens that this particular sample contains lots of houses with low rents and has a sample mean of €1250. Looking at the sample we might conclude that the real population mean rent is less than €1400 whereas in fact it is equal to €1400.

This is an example of one error (Type I) which can be made during Hypothesis Testing. There are two types of errors that can be made:

More Definitions

**Type I errors:** We reject the Null Hypothesis even though the Null Hypothesis is true.

**Type II errors:** We do not reject the Null Hypothesis when in fact the Null Hypothesis is false and the Alternative is true.

*LEGAL ANALOGY:*
Null Hypothesis : Innocent
Alternative Hypothesis: Guilty

Type I error: You are convicted and found guilty even though you are innocent.
Type II error: The jury finds you innocent even though you are guilty of the crime.
Test Statistics and Rejection Regions

We saw that we will reject the Null Hypothesis if the Sample Statistic is SUFFICIENTLY far away from the Hypothesized Population Characteristic.

How do we define Sufficiently far away?

Assuming the Null Hypothesis is true we can compute a TEST STATISTIC which belongs to a known distribution (eg the Normal distribution).

We then compute a REJECTION REGION so that there is a small predefined probability that the test statistic could by chance lie in this region.

If the test statistic lies in this region we then decide to reject the Null Hypothesis as the probability it lies there is below the threshold we initially set
HOW THIS WORKS - An Example

When we have a large sample the Central Limit Theorem tells us that $\bar{x}$ has approximately a NORMAL distribution.

The Mean of $\bar{x}$ is $\mu$ (the value in the Null Hypothesis)

The Standard Deviation of $\bar{x}$ is $\frac{\sigma}{\sqrt{n}}$

So we choose $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$ to be our TEST STATISTIC as it is Standard Normal.

Looking at the Standard Normal distribution we know that the 95th percentile is 1.645 ie: $\text{Prob}(Z > 1.645) = .05$
SO WHAT?

This tells us that if we

1. Choose a random sample from the Population of interest which has hypothesized mean \( \mu \)

2. Calculate the Mean \( \bar{x} \) for this Sample and use this to calculate the Test Statistic

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

3. The 95th Percentile of the Standard Normal Distribution is 1.645 (ie: there is only a 5% chance that the computed Z could be this large).

So if we find that \( Z > 1.645 \) we conclude that since there is only a 5% chance that \( \bar{x} \) could be that much bigger than \( \mu \) it is probable that the real population mean is actually bigger than the hypothesized value \( \mu \).

Errors Revisited

In the example above we rejected the null hypothesis because there was only a 5% chance that the Sample could have come from the Hypothesized Population. But 5% is not 0% so there was in fact a 5% chance that in rejecting \( H_0 \) we made a mistake - a Type I error.
Before we embark on a Hypothesis Test we must decide what probability of a Type I error we can live with. We cannot completely eliminate these Type I errors, one reason being that by choosing a rejection region to lower the chance of a Type I error we actually increase the chance of a Type II error.

**Definitions**

- The probability of a Type I error is called the **Level Of Significance** of the Hypothesis Test and is denoted by $\alpha$ - alpha.
- The probability of a Type II error is denoted by $\beta$.

We choose $\alpha$ to be small ( .01, .05 or .1) but we cannot completely eliminate the probability of a Type I error, as we mentioned already $\alpha$ and $\beta$ are related

*The only way to reduce $\beta$ without increasing $\alpha$ is to increase the sample size.*

Type I errors are generally considered the more serious so in our testing procedure we control the probability of these errors ( $\alpha$ ) and usually are unaware of the probability of Type II errors ( $\beta$ ).
Example

A car manufacturer offers a 60,000 mile warranty on its new cars. You plan on buying one of these cars and must decide whether to purchase an extended warranty. When the car first needs a major repair you will repair the car and then replace it, believing that it will probably need frequent repairs afterwards. A motoring magazine has reported the number of miles at which 30 of this manufacturer’s cars first needed repair. You will conduct a Hypothesis test to aid your decision.

What hypotheses would you test in each of the cases below:

1. The extended warranty is very expensive

2. The extended warranty is not very expensive

Explain.

1. $H_0 : \mu = 60,000 \text{ vs } H_A : \mu > 60,000$

As the extended warranty is very expensive you will not purchase it unless there is substantial evidence that the first time a major repair is needed is after the original warranty has lapsed.

2. $H_0 : \mu = 60,000 \text{ vs } H_A : \mu < 60,000$

As the extended warranty does not cost that much you might be inclined to buy it unless it can be shown that the first major repair tends to occur while the original warranty is in place.
Example
A new prescription drug is to be marketed as a tablet and is supposed to contain 5 mg of codeine in each tablet. The FDA must decide whether to allow the sale of this tablet. \( \mu \) is the true mean dosage of codeine per tablet.

The FDA decides to conduct a test of

\[ H_0 : \mu = 5 \quad \text{vs} \quad H_A : \mu \neq 5 \]

Why was the above alternative hypothesis chosen over the other two possibilities:

\[ H_0 : \mu = 5 \quad \text{vs} \quad H_A : \mu > 5 \quad \text{and} \quad H_0 : \mu = 5 \quad \text{vs} \quad H_A : \mu < 5 \]

For the null and alternative hypotheses given describe (in the context of the problem) Type I and Type II errors and discuss the consequences of making each type of error.

Answers: If the tablets contained too much codeine the person would be taking too high a dosage which could be dangerous. If the tablets contained too little codeine then the person would not be getting the desired relief. Therefore it is important for the FDA to investigate both possibilities and hence use the two tailed test. If the FDA makes a Type I error (reject \( H_0 \) when \( H_0 \) is true) they deprive the public of a potentially useful product and also force the manufacturer to redesign the medication incurring extra costs. If the FDA makes a Type II error (fail to reject \( H_0 \) when \( H_0 \) is false) then the drug is sold without the claimed dosage. In this case the consumer runs the risk of either ingesting too much codeine which could be dangerous or not getting the desired relief.
SECTION Large Sample Tests for the Population Mean

The elements

We will be conducting a test for the Population Mean so the Null Hypothesis will be

\[ H_0: \mu = a \]

and the Alternative Hypothesis will be one of the following

\[ H_A: \mu > a \]
\[ H_A: \mu < a \]
\[ H_A: \mu \neq a \]

While we do not know what the population distribution is we know that the sample is large. The Central Limit Theorem tells us that for large samples (\(n>30\)) \(\bar{x}\) is approximately Normally distributed and that

\[ \mu_{\bar{x}} = \mu \quad \text{and} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \]

so the test statistic

\[ z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \]

should be approximately Standard Normal.

P-Values
Definition

The P-Value is the smallest level of significance at which $H_0$ can be rejected.

How to compute a P-value

<table>
<thead>
<tr>
<th>Type of Test</th>
<th>$Z_{\text{CALC}}$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper-tailed</td>
<td></td>
<td>Area under Normal curve to the right of $Z_{\text{CALC}}$</td>
</tr>
<tr>
<td>Lower-tailed</td>
<td></td>
<td>Area under Normal curve to the left of $Z_{\text{CALC}}$</td>
</tr>
<tr>
<td>Two-tailed</td>
<td>$Z_{\text{CALC}} &gt; 0$</td>
<td>Twice the area to the right of $Z_{\text{CALC}}$</td>
</tr>
<tr>
<td>Two-tailed</td>
<td>$Z_{\text{CALC}} &lt; 0$</td>
<td>Twice the area to the left of $Z_{\text{CALC}}$</td>
</tr>
</tbody>
</table>

How to use P-values

Once the P-value has been determined, the conclusion at any particular level of $\alpha$ results from comparing the P-value to $\alpha$:

1. $P\text{-value} \leq \alpha \Rightarrow \text{Reject } H_0 \text{ at level } \alpha$
2. $P\text{-value} > \alpha \Rightarrow \text{Do Not Reject } H_0 \text{ at level } \alpha$
Easy Examples

1. For which of the given P-values would the Null Hypothesis be rejected when performing a level .05 test?
   a. .001  b. .021  c. .078  d. .047  e. .148

2. Pint glasses are supposed to hold one pint. Your friend is buying a round for 5 people and believes that the 5 glasses contain less than 5 pints. You a trainee statistician, say that you can convince him otherwise.
   Let \( \mu \) denote the true average amount of beer in 5 pint glasses filled in UCD bar. Find the P-value associated with each test statistic value below for testing \( H_0: \mu = 5 \) vs \( H_A: \mu > 5 \)
   a. 1.4  b. 0.9  c. 1.9  d. 2.4  e. -0.1

Answers:
   a) .808  b) .1841  c) .0287  d) .0082  e) .5398
A Better Example
A sample of 40 speedometers of a particular brand is checked for accuracy at 55 mph. The resulting sample average and sample standard deviation are 53.8 and 1.3 respectively. Let $\mu$ denote the true average speedometer reading when the actual speed is 55 mph. Compute a P-value and use it and a significance level of .01 to decide whether the sample evidence strongly suggests $\mu$ is not 55.

ANSWER
1. True average speedo reading when actual speed is 55mph
2. $H_0$: $\mu = 55$
3. $H_A$: $\mu \neq 55$
   \[ z = \frac{\bar{x} - 55}{\frac{s}{\sqrt{n}}} \]
4. \[ z = \frac{53.8 - 55}{1.3/\sqrt{40}} = -5.84 \]
5. P-value is less than .001 and so the Null Hypothesis would be rejected at significance level .01. The sample evidence suggests quite strongly that the true average reading is not 55 mph when the actual speed is 55 mph.
Another Example

The national mean cholesterol level is approximately 210. 100 people with high cholesterol levels (over 265) participated in a drug study and were treated with a new drug Cholestyramine. After treatment the sample mean was 228 and the sample standard deviation was 12. One question of interest is whether people taking this drug still have a mean cholesterol level that exceeds the national average. Compute the P-value for this data. If a .05 significance level is chosen what conclusion would you draw?

1. Population Characteristic: \( \mu = \) Average cholesterol level for all people taking this drug
2. \( H_0: \mu = 210 \)
3. \( H_A: \mu > 210 \)

\[
z = \frac{\bar{x} - 210}{\frac{s}{\sqrt{n}}} = \frac{228 - 210}{12/\sqrt{100}} = 15\]

5. P-value is less than .0005 which is less than .05 and so the Null Hypothesis would be rejected at significance level .05. The sample evidence suggests very strongly that the drug does not reduce cholesterol to the national mean level.
**EXAMPLE**
A neurologist wants to test the effects of a particular drug on the nervous system. She conducts a test to measure the drug’s affect on the response rates of rats. She knows that the mean response time for rats which are not injected with this drug is 1.2 seconds. She injects 100 rats with the drug and finds that the mean response rate for these rats is 1.05 seconds with a standard deviation of 0.5 seconds. The neurologist now wants to test whether the mean response time for the drug-injected rats differs from 1.2 seconds using a significance level $\alpha = 0.05$

**Example**

“Radio stations in Ireland have too much talk and not enough music, so people are forced to listen to pirate stations.”

- To test this statement we listened to three non-pirates 2FM, 98FM and FM104 and recorded the actual time music was played each hour. The experiment was performed on a random selection of 50 hours and yielded the following results:

The mean time per hour music was played was 43.03 minutes, with a standard deviation of 4.10 minutes.

Perform a test at significance level 0.05 to test the Hypothesis that true mean time spent by these 3 radio stations playing music per hour is less than 45 minutes.
Example
The Environmental Protection Agency (EPA) sets limits on the maximum allowable concentration of certain chemicals in water. for the substance PCB the level has been set at 5 ppm. A random sample of 36 water specimenns from a resevoir results in a sample mean concentration of 4.82 and a standard deviation of 0.6.

Is there sufficient evidence to substantiate a claim that the resevoir water is safe? Use a 0.01 level of significance.

Would you recommend using a significance level greater than 0.01 ? Why or why not?
Small Sample Hypothesis Tests for the Mean of a Normal Population

Introduction

In the previous sections we were able to perform Hypothesis tests for the mean of a population because we knew that the sample mean was approximately normally distributed.

This, we recall, was due to the Central Limit Theorem which applied in the case of a large sample (n ≥ 30) irrespective of the underlying population distribution.

If we do not have a large sample then the CLT does not apply and so we cannot use Z tables to find P-values or rejection regions. Another problem is that for small samples s - the sample standard deviation will not be a good approximation to σ - the population standard deviation.

All is not lost however if we make some initial assumptions about the population distribution. We will restrict ourselves to populations which are Normal.
**Definition**
For a random sample of size \( n \) from a Normal population
the test statistic:
\[
t = \frac{\bar{x} - \mu}{s/\sqrt{n}}
\]
belongs to a T distribution with \( n-1 \) df.

**How to perform a T - test**

Hypothesis tests using a T - test statistic are performed in almost exactly the same way as we saw for Z - tests. The only changes to the test procedure are that we use T - tables instead of Z - tables for calculating rejection regions or P - values.

Rejection regions are calculated using table 10 and P-values are calculated using table 9.

Each of these tables actually contains many different T-tables for different degrees of freedom \( \nu \).

We are always interested in \( \nu = n-1 \) where \( n \) is the sample size.
Example
Let \( \mu \) denote the true average surface area covered by 1 litre of a certain paint. \( H_0: \mu = 400 \) is to be tested against \( H_A: \mu > 400 \). Assuming the coverage is normally distributed, give the appropriate test statistic and rejection region for each given sample size and significance level.

a. \( n = 10, \alpha = .05 \) 

b. \( n = 18, \alpha = .05 \) 

c. \( n = 25, \alpha = .001 \) 

d. \( n = 50, \alpha = .1 \)
Answers:
\[
t = \frac{\bar{x} - 400}{s/\sqrt{10}} \quad 9 \text{ df} \quad \text{Reject } H_0 \text{ if } t > 1.83
\]
\[
t = \frac{\bar{x} - 400}{s/\sqrt{18}} \quad 17 \text{ df} \quad \text{Reject } H_0 \text{ if } t > 2.57
\]
\[
t = \frac{\bar{x} - 400}{s/\sqrt{25}} \quad 24 \text{ df} \quad \text{Reject } H_0 \text{ if } t > 3.47
\]
\[
t = \frac{\bar{x} - 400}{s/\sqrt{50}} \quad 49 \text{ df} \quad \text{Reject } H_0 \text{ if } t > 1.30
\]

Example
A researcher collected data in order to test \( H_0: \mu = 17 \) against \( H_A: \mu > 17 \). Place bounds on the P-value for each of the given test statistic values and associated degrees of freedom.

a. \( t = 1.84, \text{ df} = 14 \)  
   d. \( t = 1.3, \text{ df} = 8 \)

b. \( t = 3.74, \text{ df} = 24 \)  
   e. \( t = 2.67, \text{ df} = 40 \)

c. \( t = 2.42, \text{ df} = 13 \)

\[\begin{align*}
a. & \quad .0391 < \text{P-value} < .0467 \\
b. & \quad .0004 < \text{P-value} < .0006 \\
c. & \quad .0133 < \text{P-value} < .0160 \\
d. & \quad \text{P-value} = .1149 \\
e. & \quad .0051 < \text{P-value} < .0065
\end{align*}\]
Example
The IQ of adults is thought to be normally distributed with mean 100. Suppose 10 randomly selected prisoners convicted of felonies had IQ’s of 100, 135, 108, 94, 111, 96, 99, 104, 109, and 120. Using a level .05 test can you conclude that the mean IQ of those convicted of felony offences is significantly different from that of the general population?

Answer:
1. $\mu =$ Average IQ of all prisoners convicted of felonies
2. $\alpha = .05$
3. $H_0: \mu = 100$
4. $H_A: \mu \neq 100$
\[
t = \frac{\bar{x} - 100}{s/\sqrt{n}}
\]
with 9 degrees of freedom
5. Reject if $t > 2.26$ or if $t < -2.26$
\[
t = \frac{107.6 - 100}{12.394/\sqrt{10}} = 1.94
\]
6. $t_{calc}$ is not in the rejection region and so $H_0$ is not rejected i.e there is not enough evidence to conclude that the mean IQ for convicted felons is significantly different from the mean of the general population.
Example
The accompanying radiation readings were obtained from television display areas in a sample of 10 department stores.

.40 .48 .60 .15 .50 .80 .50 .36 .16 .89

The recommended limit for this type of radiation exposure is .5 mR/h. Assuming that the observations come from a normal distribution with mean \( \mu \), test \( H_0: \mu = .5 \) against \( H_A: \mu > .5 \) using a level .1 test.

Compute a P-value for Example above
Large Sample Hypothesis Tests for a Population Proportion

So far we have seen how to make inferences about a Population Mean using the Sample Mean. But we are not always interested in the Population Mean sometimes we are interested in the proportion of observations in the population with some property.

8.5.1 Definitions

\( p \) - Population Proportion: The proportion of individuals or objects in a specified population that possess a certain property.

\( \hat{p} \) - Sample Proportion: The proportion of individuals or objects in a random sample from the specified population that possess the property.

Both of these measure the same concept, \( p \) for the Population and \( \hat{p} \) for a Sample.

The Sample Statistic \( \hat{p} \) is the natural statistic for making inferences about the Population Characteristic \( p \).
The Sampling Distribution of \( \hat{p} \)

1. Mean of \( \hat{p} = p \) .... The mean (expected value) of the Sample Proportion is the Population Proportion.

2. \[ \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \]

3. When \( n \) is sufficiently large, the sampling distribution of \( \hat{p} \) is approximately normal. The sample is sufficiently large if \( p \pm 3\sigma_{\hat{p}} \) does not include 0 or 1, where \( p \) is the Hypothesised value.

**Definition**

For large samples the test statistic

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

is approximately standard normal.
How to perform a Hypothesis test for $p$
Large Sample Hypothesis tests for the population proportion $p$ are performed using the same procedure as before except that

$$z = \frac{x - \mu}{s/\sqrt{n}}$$

is replaced by

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

Again Tables 4 and 5 are used for P-Values and Rejection Regions.

Example
A store sells Video Recorders and DVDs. The trend seems to be towards DVDs and the manager is thinking about dropping Videos. $p$ is the proportion of Videos sold. It is decided to test $H_0 : p = .25$ vs $H_A : p < .25$ and drop videos if $H_0$ can be rejected at level .01. If $n=100$ recent purchases are sampled and 21 are videos what decision is appropriate.
ANSWER
1. \( p = \) proportion of recent sales which are videos
2. \( \alpha = .01 \)
3. \( H_0: p = .25 \)
4. \( H_A: p < .25 \)
5. \[ z = \frac{\hat{p}-.25}{\sqrt{.25*.75\over 100}} \]
6. Reject if \( z < -2.33 \)
7. \[ z = \frac{.21-.25}{\sqrt{.25*.75\over 100}} = -0.92 \]
8. \( z_{\text{calc}} \) is not in the rejection region and so \( H_0 \) is not rejected
   ie there is not enough evidence to conclude that the true proportion of video sales is less than .25. The manager should not drop the sale of videos.

Example
To test the ability of car mechanics to identify a simple engine problem, a car with a single such problem was taken in turn to 72 different garages. Only 42 of the 72 mechanics who examined the car identified the problem correctly. Does this indicate that the true proportion of mechanics who could identify the problem is less than .75? Compute a P-value and use it to draw your conclusion.
Confidence Intervals

Statistics is all about estimating population parameters based on sample data.

For a given Population we may want to estimate

- The Population Mean
- The Population Median
- A Population Proportion
- The Population Variance
- The 95th Percentile of Population
- Etc.

While providing a point estimate is part of the solution, it is not the complete solution.

How good is our estimate?
Is it reliable?

Confidence Intervals provide the solution.  
They give us the Margin of Error of our estimate.
LARGE SAMPLE CONFIDENCE INTERVAL FOR A POPULATION MEAN

GENERAL FORMULA

\[ \bar{x} \pm (z \text{ critical value}) \frac{\sigma}{\sqrt{n}} \]

The level of confidence determines the \( z \) critical value.

- 99% \quad 2.58
- 95% \quad 1.96
- 90% \quad 1.645

Since \( n \) is large the unknown \( \sigma \) can be replaced by the sample value \( s \).

\[ \bar{x} \pm (z \text{ critical value}) \frac{s}{\sqrt{n}} \]
Example
A random sample of 225 1st year chemistry practicals was selected from the past 5 years and the number of students absent from each one recorded. The results were $\bar{x} = 11.6$ and $s = 4.1$ absences.
Estimate the mean number of absences per practical over the past 5 years with 90% confidence.

90% CI for $\mu$ is

$$\bar{x} \pm 1.645 \left( \frac{s}{\sqrt{n}} \right)$$

$$11.6 \pm 1.645 \left( \frac{4.1}{\sqrt{225}} \right) = 11.6 \pm 0.45 = (11.15, 12.05)$$

INTERPRETATION:
It is incorrect to say that there is a probability of 0.90 that $\mu$ is between 11.15 and 12.05. In fact this probability is either 1 or 0 ($\mu$ either is or is not in the interval).
The 90% refers to the percentage of all possible intervals that contain $\mu$ i.e. to the estimation process rather than a particular interval.
It is also incorrect to say that 90% of all practicals had between 11.15 and 12.05 missing students.
SMALL SAMPLE CONFIDENCE INTERVAL FOR A POPULATION MEAN

• Consider only samples from populations which are (approx) normal.
  \[
  \frac{\bar{x} - \mu}{s/\sqrt{n}}
  \]

• Use the distribution of which is known as the \(t\) distribution with \((n-1)\) degrees of freedom (df).

  Note: As the df increases the \(t\) curve approaches the \(z\) curve.

Areas under the \(t\) curve are tabulated in tables 9 & 10 of NCST (note \(v=\text{df}\))

A small sample confidence interval for \(\mu\) is

\[
\bar{x} \pm (t \text{ critical value}) \frac{s}{\sqrt{n}} \quad \text{with df} = n - 1
\]

NOTE:
This confidence interval is appropriate for small samples \textbf{ONLY} when the population distribution is normal.
Example
Sample of 15 test-tubes tested for number of times they can be heated on Bunsen burner before they cracked gave $\bar{x} = 1,230$, $s = 270$. Construct 99% confidence interval for $\mu$.

df = n-1 = 15-1 = 14
for 99% confidence $t = 2.977$

$$\bar{x} \pm (t \text{ critical value}) \frac{s}{\sqrt{n}} = 1,230 \pm 2.977 \frac{270}{\sqrt{15}} = (1,020,1,440)$$
LARGE SAMPLE CONFIDENCE INTERVAL FOR A POPULATION PROPORTION

In many populations each item belongs to one of two categories (S & F). We are interested in estimating the proportion (or percentage) of the population who belong to each category.
e.g. proportion of adults who smoke cigarettes, proportion who vote for FF, proportion who drink Budweiser etc.

Example
Each year 1st year PhD students may (S) or may not (F) choose to study Statistics. To estimate the fraction who do study Statistics a sample of 1000 students was chosen from the past 10 years and 637 had chosen Statistics as a 1st year subject.

If the fraction of students who choose statistics is \( p \) then randomly selecting somebody is has a probability \( p \) of S and \( 1-p \) of F. The obvious solution is to use the sample proportion (\( \hat{p} \)) to estimate the population proportion (\( p \)) where
\[
\hat{p} = \frac{\text{number of S's in sample}}{\text{Sample size}} = \frac{637}{1000} = .637
\]
The sampling distribution of \( \hat{p} \) is approx normal with
\[
\mu_{\hat{p}} = p \text{ and } \sigma^2_{\hat{p}} = \frac{p(1-p)}{n}
\]
Consequently the confidence interval for $p$ is

$$\hat{p} \pm (z \text{ critical value}) \sqrt{\frac{p(1-p)}{n}}$$

Approximate $p$ with $\hat{p}$ to get

$$\hat{p} \pm (z \text{ critical value}) \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

for the approximate confidence interval.

For our example the 95\% confidence interval for $p$ is

$$\hat{p} \pm (z \text{ critical value}) \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (.607, .667)$$
Example
UCD takes a random sample of recent PhD graduates \((n=135)\) and finds that 12 of these are unemployed. Compute a 90% confidence interval for the proportion of all PhD graduates who fail to find a job.
\[
\hat{p} = \frac{12}{135} = .089
\]
Check for large sample
\[
\hat{p} \pm 3\sigma_{\hat{p}} \approx \hat{p} \pm 3\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (.015,.163)
\]
90% confidence interval for \(p\)
\[
\hat{p} \pm z\sqrt{\frac{p(1-p)}{n}} \approx \hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (.049,.129)
\]
Suppose we know that the unemployment rate in the country is 4%. Can we conclude that PhD Graduates are more likely to be unemployed than the population in general? Since .04 lies outside the confidence interval it is not consistent with our data and we would conclude that the unemployment rate among PhD graduates is above that of the population as a whole.