Chapter 6: Statistical Inference (Single Sample)

Shiwen Shen

University of South Carolina

2017 Summer
Introduction

- Statistical inference can be classified as estimation problem and testing problem.
- The goal of estimation is to make a proper guess of unknown parameter, e.g. population mean $\mu$, population proportion $p$, etc, using data.
- The goal of testing is to exam whether the estimated value for the unknown parameter is good, or whether some statistical argument is correct.
- In this chapter, we discuss the estimation and testing methods based on single sample, focusing on population proportion and population mean.
The population proportion $p$ emerges when the characteristic we measure on each individual is categorical, or simply binary (i.e., only 2 outcomes possible). Here are some examples:

- $p =$ proportion of airline has experienced exceedance
- $p =$ proportion of defective water filters in a factory
- $p =$ proportion of HIV positive in SC

We can connect these binary outcomes to the *Bernoulli trails* assumptions for each individual in the sample:

1. Each trial results in only two possible outcomes, labeled as “success” and “failure.”
2. The trials are independent.
3. The probability of a success in each trial, denoted as $p$, remains constant. It follows that the probability of a failure in each trial is $1 - p$. 
Suppose we define $Y =$ the number of successes out of $n$ sampled individuals so $Y \sim b(n, p)$. A natural point estimator for $p$, the population proportion, is

$$\hat{p} = \frac{Y}{n},$$

the sample proportion. $\hat{p}$ is read as p hat.
Property of $\hat{p}$

- $\hat{p}$ is an unbiased estimator of $p$. That is,
  
  $$E(\hat{p}) = p.$$  

- To quantify the precision of $\hat{p}$,
  
  $$\text{var}(\hat{p}) = \frac{p(1 - p)}{n}.$$ 

- Question: What is the (asymptotic) distribution of $\hat{p}$?
Sampling Distribution of $\hat{p}$ and CLT

- To derive the sampling distribution of $\hat{p}$, we need first introduce the central limit theorem.

- **Central Limit Theorem**: Suppose that $Y_1, Y_2, \ldots, Y_n$ is a random sample from a population distribution with mean $\mu$ and variance $\sigma^2$. When the sample size $n$ is large, we have

$$\overline{Y} \sim \mathcal{AN} \left( \mu, \frac{\sigma^2}{n} \right)$$

- $\mathcal{AN}$ is read as **Asymptotically Normal**.

- Recall HW9 problem 4.
Simulation Study of CLT Cont’d

single read

Frequency

0 1 2 3 4 5 6

y

sample size of 2

Frequency

0 1 2 3 4

results

sample size of 20

Frequency

0 50 100 150

0.5 1.0 1.5

results

sample size of 500

Frequency

0 50 100 150

0.85 0.95 1.05 1.15

results
Sampling distribution of $\hat{p}$

With the help of central limit theorem, we can derive an asymptotic distribution of $\hat{p}$. Recall that $Y =$ the number of successes out of $n$ sampled individuals and $Y \sim b(n, p)$. We can express $Y$ as the sum of $n$ independent Bernoulli trials with success probability $p$. That is

$$Y = \sum_{i=1}^{n} X_i,$$

where $X_i \sim \text{Bernoulli}(p)$. $E(X_i) = p$, and $\text{Var}(X_i) = p(1 - p)$. It follows that $\hat{p} = Y/n = \sum_{i=1}^{n} X_i/n = \bar{X}$. By CLT, we have

$$\hat{p} = \frac{Y}{n} = \bar{X} \sim \mathcal{AN} \left( p, \frac{p(1-p)}{n} \right).$$
Using a point estimator only **ignores important information**; namely, how variable the estimator is.

To avoid this problem (i.e., to account for the uncertainty in the sampling procedure), we therefore pursue the topic of interval estimation (also known as confidence intervals).

The main difference between a point estimate and an interval estimate is that

- a **point estimate** is a one-shot guess at the value of the parameter; this ignores the variability in the estimate.

- an **interval estimate** (i.e., **confidence interval**) is an interval of values. It is formed by taking the point estimate and then adjusting it downwards and upwards to account for the point estimate’s variability.
Recall that $\hat{p} \sim \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$.

Let us define $z_\alpha$ be the upper $\alpha$ percentage point of the standard normal distribution, i.e., $P(Z > z_\alpha) = \alpha$.

An approximate $100(1 - \alpha)\%$ confidence interval for $p$ is

$$\left(\hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}, \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}\right).$$
An approximate 100(1 − \( \alpha \))% confidence interval for \( p \) is

\[
\left( \hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}, \quad \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \right)
\]

The quantity \( z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \) is called the **margin of error**.

**Rule of thumb:** To use normal approximation, we need \( n\hat{p} \geq 15 \) and \( n(1 - \hat{p}) \geq 15 \).

**Note of the form of the interval:**

point estimate ± \( Z_{\alpha/2} \times \text{standard error} \)
Suppose that we are interested in parameter $p$ for certain population. We take a sample of size $n$ and calculate the sample proportion $\hat{p} = Y/n$. A 95% confidence interval is given by

$$\text{Point. Est } \pm 1.96 \text{ Standard Error.}$$

The 95% confidence comes from the fact that if we repeated this experiment over and over again, approximately, 95% of all samples would produce a confidence interval that contains the true proportion, and only 5% of the time would the interval be in error.

We call $100(1 - \alpha)\%$ the confidence level.
Here is a pictorial illustration of the confidence interval:
Definition: a *statistical hypothesis* is an assertion or conjecture concerning one or more population parameters.

Example:

1. The proportion of underweight milk is more than 3% in a local farm.
2. More than 7% of the landings for a certain airline exceed the runway.
3. The defective rate of the water filter is less than 5%.
4 Steps to a Hypothesis Test

1. State the null ($H_0$) and alternative ($H_a$ or $H_1$) hypotheses.

2. Collect the data and calculate test statistic assuming $H_0$ is true.

3. Assuming the $H_0$ is true, calculate the $p$-value.

4. Draw conclusion based on the $p$-value. We either reject $H_0$ or fail to reject $H_0$.

Let us look at an example to illustrate these steps...
Historically, the defective rate of water filters is 7% in a certain factory. A new quality control system is introduced to reduce the defective rate. Suppose that we randomly choose 300 water filters, and calculate \( \hat{p} = 0.041 \). We want to test whether the new system reduce the defective rate or not.

Let \( p \) = proportion of defective water filters in the factory after introducing the new system.
Null hypothesis is denoted by \( H_0 \), which represents what we assume to be true. Under null hypothesis, the exact value of the parameter is specified.

Alternative hypothesis is denoted by \( H_a \) or \( H_1 \), which represents the researcher’s interest.

In most situation, researchers want to reject null hypothesis in favor of the alternative hypothesis by performing some experiment.

In defective water filters example,

\[
H_0 : p = 0.07 \\
H_a : p < 0.07 \text{ (the new system reduces the defective rate)}
\]
Step 2: Calculate test statistic

- How should we make our decision based on the sample?
- We reject \( H_0 \), if \( \hat{p} \) is far less than 0.07, which is not likely to happen if assuming \( p = 0.07 \).
- We need the sampling distribution to quantify how far is far.
Recall that if $H_0 : p = p_0$ is true, then

$$
\hat{p} \sim \mathcal{N} \left( p_0, \sqrt{\frac{p_0(1-p_0)}{n}} \right).
$$

Therefore, the test statistic is (standardizing)

$$
Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim \mathcal{N}(0, 1)
$$

In defective water filter example, assuming $H_0$ is true, the test statistic is calculated as

$$
z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.041 - 0.07}{\sqrt{\frac{0.07(1-0.07)}{300}}} = -1.97.
$$

Question: is this number likely to appear if assume $p_0 = 0.07$ is true?
Step 3: Calculate p-value

- The *p-value* is the probability of getting the sample results you got or something more extreme assuming that the null hypothesis is true.

- If the *p*-value is small, we doubt the null hypothesis since it is not likely to observe such a "extreme" test statistic under $H_0$. There is evidence to against null hypothesis.

- On the other hand, if the *p*-value is large, we have a pretty good chance to observe the computed test statistic under $H_0$ in a single experiment, there is no reason to question the $H_0$.

- In other words, the p-value for a hypothesis test measures how much evidence we have against $H_0$, that is,

  \[ \text{the smaller the p-value} \implies \text{the more evidence against } H_0 \]
Step 3: Calculate p-value cont’d

- In defective water filter example, the p-value of the test is:

\[ P(Z < -1.97) = 0.024 \] (found in table)

- Why “<”? 

<table>
<thead>
<tr>
<th>Alternative hypothesis</th>
<th>Hypothesis type</th>
<th>p-value formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_a : p &lt; p_0 )</td>
<td>Left-tail hypothesis</td>
<td>( P(Z &lt; z_0) )</td>
</tr>
<tr>
<td>( H_a : p &gt; p_0 )</td>
<td>Right-tail hypothesis</td>
<td>( P(Z &gt; z_0) )</td>
</tr>
<tr>
<td>( H_a : p \neq p_0 )</td>
<td>Two-tail hypothesis</td>
<td>( 2P(Z &lt; -</td>
</tr>
</tbody>
</table>

- In our example, the \( p \)-value is 0.024, do you think it is large?
Step 4: Conclusion

- If the p-value is small, we reject the null hypothesis and conclude the alternative hypothesis.

- If the p-value is not small, we do not reject the null hypothesis and do not conclude the alternative hypothesis.

- There is one remaining question, how small should $p$-value be to be considered as "small"? We need level of significance to answer it.
We use $\alpha$ to denote the level of significance.

Level of significance is determined before you see the data.

In practice, we usually set $\alpha = 0.05$. Other common choices are $\alpha = 0.01$, or $\alpha = 0.1$.

We simply compare the $p$-value with the $\alpha$ level, we reject $H_0$ if the $p$-value is less than $\alpha$; and do not reject $H_0$ is the $p$-value is greater than or equal to $\alpha$.

In defective water filter example, $p$-value = 0.024 < 0.05 (pre-defined), therefore, we reject $H_0$, and conclude that we have sufficient evidence to conclude the new system reduces the defective rate (note: we conclude $H_a$ in the context of the question.)
A certain type of flu outbreaks in northern part of the USA. The historical records shows that there are 7% of the residences in Columbia carrying flu under usual condition. Researchers want to see whether there is an outbreak in Columbia, there are 30 out of 250 randomly chosen people in the sample carrying flu. Can we conclude that there is an outbreak in Columbia? Answer the following question using confidence interval approach (95%) and hypothesis testing approach (assuming level of significance is 0.05).
Recall that a $100(1 - \alpha)\%$ C.I. is

$$\left( \hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}, \quad \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \right).$$

▶ The point estimate $\hat{p} = 30/250 = 0.12$.

▶ $z_{\alpha/2} = z_{0.025} = 1.96$

▶ Standard error: $\sqrt{0.12(1 - 0.12)/250} = 0.021$

▶ 95% CI is: $(0.079, 0.161)$

▶ Conclusion?
Inference for $p$: Hypothesis Test

- **Step 1: State $H_0$ and $H_1$**
  
  $H_0 : p = 0.07$
  
  $H_a : p > 0.07$

- **Step 2: Calculate test statistic assuming $H_0$ is true**

  
  $$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.12 - 0.07}{\sqrt{\frac{0.07(1-0.07)}{250}}} = 3.10.$$ 

- **Step 3: Calculate p-value**

  
  $p$-value$=P(Z > 3.10) \approx 0.001.$

- **Step 4: Draw the conclusion**

  
  $\alpha = 0.05$, $p$-value is smaller than 0.05. We reject $H_0$, and conclude that there is an outbreak in Columbia.
There are two mistakes we can make in a hypothesis test.

- **Type I error**: $H_0$ is rejected but in reality $H_0$ is true
- **Type II error**: $H_0$ is not rejected but in reality $H_0$ is false

<table>
<thead>
<tr>
<th>$H_0$ is true</th>
<th>$H_0$ is false</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do not reject $H_0$</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>No Error</td>
<td>Type I error</td>
</tr>
<tr>
<td></td>
<td>Type II error</td>
</tr>
<tr>
<td></td>
<td>No Error</td>
</tr>
</tbody>
</table>
The probability of type I error is denoted by $\alpha$ (same as the level of significance), i.e.,

$$\alpha = \text{Prob}(\text{reject } H_0 | H_0 \text{ is true}).$$

The type II error is denoted by $\beta$, i.e.,

$$\beta = \text{Prob}(\text{fail to reject } H_0 | H_a \text{ is true at some value}).$$

The idea situation is both type I and type II error is 0, which means we can always make correct decision. However, only oracle knows the true. For us, every decision we make will have associated error probability.
In practice, if we try to decrease the type I error, the type II error will increase, and vice versa. Remember, there is no free lunch!

Researchers should consider the consequences of type I error and type II errors to help determine significance level.

**Example:** An environmentalist takes samples at a nearby river to study the average concentration level of a contaminant. He wants to find out, using a 0.1 level of significance, if the average concentration level exceeds the acceptable level for safely consuming fish from the river.
Describe a Type I error for this problem and the potential consequence.

Describe a Type II error for this problem and the potential consequence.
For binary random variable, we have discussed how to estimate the population proportion using point estimate and confidence interval.

Moreover, we built a 4-step procedure to test the hypothesis corresponding to the population proportion.

Now, it’s time to move on to the case when the random variable is numerical and learn some method to guess population mean $\mu$ in the proper way.
Sampling distribution of $\bar{Y}$

- Recall sample mean $\bar{Y}$ is a reasonable point estimator of the population mean $\mu$.

- **RESULT:** Suppose $Y_1, Y_2, \ldots, Y_n$ is a random sample from a $\mathcal{N}(\mu, \sigma^2)$ distribution. Then the sample mean $\bar{Y}$ has the following sampling distribution:

$$\bar{Y} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$$

- The above result reminds us
  - $\bar{Y}$ is an **unbiased** estimator of $\mu$.
  - $se(\bar{Y}) = \sigma / \sqrt{n}$
Example

- Let $Y =$ time (in seconds) to react to brake lights during in-traffic driving.

- We assume $Y \sim \mathcal{N}(\mu = 1.5, \sigma^2 = 0.16)$. We call this the population distribution.

- Suppose that we take a random sample of $n = 5$ drivers with times $Y_1, \ldots, Y_5$. What is the distribution of the sample mean $\bar{Y}$?
If we standardize $\overline{Y}$, we obtain

$$Z = \frac{\overline{Y} - \mu}{\sigma / \sqrt{n}} \sim \mathcal{N}(0, 1)$$

However, population standard deviation $\sigma$ is usually unknown. Replacing it with the **sample standard deviation** $S$, we get a new sampling distribution:

$$t = \frac{\overline{Y} - \mu}{S / \sqrt{n}} \sim t(n - 1),$$

a **$t$ distribution** with degrees of freedom $\nu = n - 1$. 
The $t$ distribution has the following characteristics:

- It is continuous and symmetric about 0.
- It is indexed by a value $\nu$ called the degrees of freedom.
- As $\nu \rightarrow \infty$, $t(\nu) \rightarrow \mathcal{N}(0, 1)$.
- When compared to the standard normal distribution, the $t$ distribution, in general, is less peaked and has more probability (area) in the tails.
Example

Hollow pipes are to be used in an electrical wiring project. The manufacturer claims that the outside diameter of this pipe is 1.29 inch. The data below were collected by a design engineer. The data are measurements of $Y$, the outside diameter of this type of pipe (measured in inches). These $n = 25$ pipes were randomly selected and measured—all in the same location.

<table>
<thead>
<tr>
<th>1.296</th>
<th>1.320</th>
<th>1.311</th>
<th>1.298</th>
<th>1.315</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.305</td>
<td>1.278</td>
<td>1.294</td>
<td>1.311</td>
<td>1.290</td>
</tr>
<tr>
<td>1.284</td>
<td>1.287</td>
<td>1.289</td>
<td>1.292</td>
<td>1.301</td>
</tr>
<tr>
<td>1.298</td>
<td>1.287</td>
<td>1.302</td>
<td>1.304</td>
<td>1.301</td>
</tr>
<tr>
<td>1.313</td>
<td>1.315</td>
<td>1.306</td>
<td>1.289</td>
<td>1.291</td>
</tr>
</tbody>
</table>

Under this assumption (which may or may not be true), calculate the value of

$$t = \frac{\overline{Y} - \mu}{s/\sqrt{n}}$$
We use R to find the sample mean $\bar{y}$ and the sample standard deviation $s$:

```r
> mean(pipes)  ## sample mean
[1] 1.29908
> sd(pipes)    ## sample standard deviation
[1] 0.01108272
```

With $n = 25$, we have

$$t =$$
If the manufacturer’s claim is true (that is, if $\mu = 1.29$ inches), then

$$t = \frac{\bar{y} - \mu}{s/\sqrt{n}}$$

comes from a $t(24)$ distribution. The $t(24)$ pdf is displayed below:
**Key question:** Does $t = 4.096$ seem like a value you would expect to see from this distribution? If not, what might this suggest? Recall that $t$ was computed under the assumption that $\mu = 1.29$ inches (the manufacturer’s claim).

**QUESTION:** The value $t = 4.096$ is what percentile of the $t(24)$ distribution?

\[
> \text{pt}(4.096, 24)
\]

[1] 0.9997934
Recall if $Y_1, \ldots, Y_n$ is a random sample from a $\mathcal{N}(\mu, \sigma^2)$ distribution, then

$$t = \frac{\bar{Y} - \mu}{s/\sqrt{n}} \sim t(n-1)$$

An obvious question arises: “What if $Y_1, \ldots, Y_n$ are non-normal?”

Answer: The $t$ distribution result still approximately holds. That is, the $t$ distribution is robust to the normality assumption.

How to assess the normal distribution assumption? Normal quantile-quantile (qq) plot
The plot is constructed as follows:

- On the vertical axis, we plot the observed data, ordered from low to high.
- On the horizontal axis, we plot the (ordered) theoretical quantiles from the distribution assumed for the observed data (here, normal).

For the pipe diameter data, below is the qq plot.

```R
> qqnorm(pipes, pch=16, main="")
> qqline(pipes)
```
The ordered data do not match up perfectly with the normal quantiles, but the plot doesn’t set off any serious alarms.

Fat pencil test: Imagine a “fat pencil” lying along the line. If all the points are covered by this imaginary pencil, a normal distribution adequately describes the data.

If there is no gross departure from the straight line on the plot, we should accept that the model describes the data well.
CI for $\mu$. Assume normality and known $\sigma^2$

- Recall if $Y_1, \ldots, Y_n$ is a random sample from a $\mathcal{N}(\mu, \sigma^2)$ distribution and $\sigma^2$ is known, then

$$Z = \frac{Y - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

- Similar to CI derivation in population proportion, a $100(1 - \alpha)\%$ CI for $\mu$ is given by

$$\left( \bar{y} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad \bar{y} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right)$$
Go back to pipe diameter data. Based on past experience, the engineers assume a normal population distribution (for the pipe diameters) with known population standard deviation 0.02. We want to find a 95% CI for $\mu$, the mean pipe diameter.

**Solution:** We have the sample mean $\bar{Y} = 1.30$. So a 95% C.I. interval is

$$\left( \bar{y} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \; \bar{y} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right)$$

$$= \left( 1.30 - 1.96 \times \frac{0.02}{\sqrt{25}}, \; 1.30 + 1.96 \times \frac{0.02}{\sqrt{25}} \right)$$

$$= \left( 1.292, \; 1.308 \right)$$

**Practical Interpretation:** Based on the sample data, with 95% confidence, the outside diameters for the hollow pipe is between 1.29 and 1.31 inches.
Recall if $Y_1, \ldots, Y_n$ is a random sample from a $\mathcal{N}(\mu, \sigma^2)$ distribution and $\sigma^2$ is unknown, then

$$
t = \frac{\bar{Y} - \mu}{S/\sqrt{n}} \sim t(n-1)
$$

A $100(1 - \alpha)$% CI for $\mu$ is given by

$$
\left( \bar{Y} - t_{n-1, \alpha/2} \frac{S}{\sqrt{n}}, \quad \bar{Y} + t_{n-1, \alpha/2} \frac{S}{\sqrt{n}} \right)
$$

where $S$ is the sample standard deviation.
Go back to pipe diameter data. Assume the pipe diameters are normally distributed with unknown population variance. We want to find a 95% CI for $\mu$, the mean pipe diameter.

Solution: We have the sample mean $\overline{Y} = 1.30$ and sample standard deviation $S = 0.011$. So a 95% C.I. interval is

$$
\left( \overline{Y} - t_{n-1,\alpha/2} \frac{S}{\sqrt{n}}, \quad \overline{Y} + t_{n-1,\alpha/2} \frac{S}{\sqrt{n}} \right)
$$

$$
= \left( 1.30 - 2.064 \times \frac{0.011}{\sqrt{25}}, \quad 1.30 + 2.064 \times \frac{0.011}{\sqrt{25}} \right)
$$

$$
= (1.295, \ 1.305)
$$

Practical Interpretation: Based on the sample data, with 95% confidence, the outside diameters for the hollow pipe is between 1.295 and 1.305 inches.

Note: 2.064 can be found in t table, or it can be calculated in R by $qt(0.975, \ 24)$. 
Sample size determination

▶ In a planning stages of an experiment or investigation, it is often of interest to determine **how many individuals** are needed to write a confidence interval with a given level of precision.

▶ For example, we might want to construct a 95 percent confidence interval for a population mean, so that the **interval length** is no more than 5 units (e.g., days, inches, dollars, etc.).

▶ **Sample size determination** is also associated with the practical issues like cost, time spent in data collection, personnel training, etc.
Suppose that $Y_1, Y_2, \ldots, Y_n$ is a random sample from a $\mathcal{N}(\mu, \sigma^2)$ population with $\sigma^2$ known. Recall that a $100(1 - \alpha)\%$ CI for $\mu$ is given by

$$
\bar{Y} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = B, \text{ say}
$$

$B$ is called the **margin of error**.
It is possible to determine the sample size $n$ once we specify these three pieces of information:

- the value of $\sigma^2$ (or an educated guess at its value; e.g., from past information, etc.)
- the confidence level, $100(1 - \alpha)$
- the margin of error, $B$.

This is true because

$$B = z_{\alpha/2} \left( \frac{\sigma}{\sqrt{n}} \right) \iff n = \left( \frac{\sigma z_{\alpha/2}}{B} \right)^2$$
Sample size determination example

In a biomedical experiment, we would like to estimate the population mean remaining life of healthy rats that are given a certain dose of a toxic substance. Suppose that we would like to write a 95 percent confidence interval with a margin of error equal to $B = 2$ days. From past studies, remaining rat lifetimes have been approximated by a normal distribution with standard deviation $s = 8$ days. How many rats should we use for the experiment?

**Solution:** With $z_{\alpha/2} = z_{0.025} \approx 1.96$, $B = 2$ and $\sigma = 8$, the desired sample size to estimate $\mu$ is

$$n = \left( \frac{\sigma z_{\alpha/2}}{B} \right)^2 = \left( \frac{8 \times 1.96}{2} \right)^2 \approx 61.46$$

We would sample $n = 62$ rats to achieve these goals.
Hypothesis Test on population mean

1. State the null ($H_0$) and alternative ($H_a$) hypotheses.
   - Null hypothesis $H_0 : \mu = \mu_0$
   - Alternative hypothesis
     - Right-tail $H_a : \mu > \mu_0$
     - Left-tail $H_a : \mu < \mu_0$
     - Two-tail $H_a : \mu \neq \mu_0$

2. Collect the data and calculate test statistic assuming $H_0$ is true.
   \[
   \sigma \text{ known: } z_0 = \frac{\overline{Y} - \mu_0}{\sigma / \sqrt{n}} \quad \text{OR} \quad \sigma \text{ unknown: } t_0 = \frac{\overline{Y} - \mu_0}{s / \sqrt{n}}
   \]

3. Assuming the null hypothesis is true, calculate the $p$-value.

<table>
<thead>
<tr>
<th>Alternative</th>
<th>Type</th>
<th>p-value (based on the knowledge of $\sigma$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_a : \mu &gt; \mu_0$</td>
<td>Right-tail</td>
<td>$P(Z &gt; z_0)$ or $P(t &gt; t_0)$</td>
</tr>
<tr>
<td>$H_a : \mu &lt; \mu_0$</td>
<td>Left-tail</td>
<td>$P(Z &lt; z_0)$ or $P(t &lt; t_0)$</td>
</tr>
<tr>
<td>$H_a : \mu \neq \mu_0$</td>
<td>Two-tail</td>
<td>$2P(Z &lt; -</td>
</tr>
</tbody>
</table>

4. Draw conclusion based on the $p$-value. We either reject $H_0$ or fail to reject $H_0$. 
Go back to pipe diameter data. Based on past experience, the engineers assume a normal population distribution (for the pipe diameters) with known population standard deviation 0.02. Researchers what to find out whether the pipe diameter is 1.31. Assume a significance level 0.05.
Hypothesis test example: known $\sigma^2$

Solution:

- **Step 1: State hypothesis**
  
  $H_0 : \mu = 1.31$
  
  $H_a : \mu \neq 1.31$

- **Step 2: Test statistic**

  \[
  z_0 = \frac{1.30 - 1.31}{0.02/\sqrt{25}} = -2.5
  \]

- **Step 3: p-value**

  \[
  p - value = 2P(Z < -| -2.5|) = 2P(Z < -2.5) = 0.012
  \]

- **Step 3: Conclusion**

  $p$-value $= 0.012 < \alpha$, we reject $H_0$. We have enough evidence to conclude the mean pipe diameter is not 1.31 inches.
Go back to pipe diameter data. Assume the pipe diameters are normality distributed with unknown population variance. Researchers want to find out whether the pipe diameter is less than 1.308. Assume a significance level 0.05.
Hypothesis test example: unknown $\sigma^2$

Solution:

▷ Step 1: State hypothesis

▷ Step 2: Test statistic

▷ Step 3: $p$-value

▷ Step 3: Conclusion
```r
> t.test(pipes, alternative="less", mu=1.308)

One Sample t-test

data: pipes
t = -4.0243, df = 24, p-value = 0.0002478
alternative hypothesis: true mean is less than 1.308
95 percent confidence interval:
   -Inf 1.302872
sample estimates:
mean of x
   1.29908
```