Inference about a Population Proportion

Focal Questions:
When sampling binary data population, how do we form a point estimate, a confidence interval and conduct a test of significance about a population proportion?

Binary Data – when each individual observation is one of two possible outcomes.

Population parameter:

\[ p = \text{proportion of successes in the population} \]

We learn about the population from a simple random sample.
We summarize the data via:

\[ x = \text{(\# of Successes in Sample)} \]

and

\[ \hat{p} = \frac{x}{n} = \text{fraction of successes in the sample} \]

The \textbf{point estimate of the population parameter} \( p \) is:

\[ \hat{p} \]

It is an \textbf{unbiased estimator} of the population proportion \( p \).

The \textbf{standard error of} \( \hat{p} \) is:

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

When the \textbf{sample size} \( n \) is large, the \textbf{sampling distribution of} \( \hat{p} \) is approximately a \textbf{normal distribution} with mean \( p \) and standard deviation \( \sqrt{\frac{p(1-p)}{n}} \).
So it would be natural to form a large sample confidence interval for the population proportion $p$ that would take the form:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
Example: A University of Michigan study showed that many adults have experienced lingering fright from a movie or TV show they saw as a teenager. In a survey of 150 college students, 39 said they still experience this type of “residual anxiety” from a movie or TV show.

In this setting the population parameter is:

\[
p = \text{proportion of college student who experience this residual anxiety.}
\]
The point estimator of this parameter is:

\[ \hat{p} = \frac{39}{150} = .26 \]

A 98% confidence interval for the population parameter is:

\[ Z^* = 2.326 \]

C.I. \[ .26 \pm (2.326) \sqrt{\left( \frac{.26 \cdot .74}{150} \right) } \]

\[ .26 \pm .083 \]

\[ (.177, .343) \]

*Interpretation of the interval.*

We are 98% confident that the true prop. with residual anxiety is between 18% and 34%.
Example: Acid rain is a growing problem throughout the U.S. and indeed all over the world. Forty samples of rain water are collected at a particular location throughout the year. Each sample is analyzed for its pH (a measure of acidity vs alkalinity). The rain samples showed an average sample pH level of:

\[ \bar{x} = 4.7 \]

and a sample standard deviation of

\[ s = 0.5 \]

Find a 99% confidence interval for \( \mu \), the average pH level of rain at this particular location.

\[ \bar{x} \pm (2.58) \left( \frac{s}{\sqrt{n}} \right) \]

\[ 4.7 \pm (2.58) \left( \frac{0.5}{\sqrt{40}} \right) \]

\[ 4.7 \pm 0.2 \]

\[ (4.5, 4.9) \]
Interpretation of the interval:

For a 99% interval based on a random sample, the probability the interval will contain $\mu$ is .99.
Tests of Hypotheses

The research hypothesis is called the

\[ \text{alternative hypothesis} \]

It is often believed (or hoped) to be true. It is the one that the study **must prove to be true**. Thus it requires strong evidence to establish its validity.

The other hypothesis is called the

\[ \text{null hypothesis} \]

It often represents the current viewpoint or the "status quo".

Notation: We use

\[ H_0: \text{null hypothesis} \]
\[ H_A: \text{alternative hypothesis} \]

**Fundamental Characteristic of a Test of Hypothesis:**

The null hypothesis must be assumed to be true unless the data presents compelling evidence indicating the alternative hypothesis (not the null) is true.
Test Statistic -

measures the compatibility of the data with the null hypothesis.

How do we measure this compatibility?

p-value

The probability computed assuming Ho is true, that the test statistic takes a value as extreme or more extreme (in the direction of Ha) than the actual observed value.

The p-value shows how extreme the observed data was compared to the null hypothesis.

The smaller the p-value, the stronger the evidence is that the alternative hypothesis is true and that the null hypothesis is not true.
Researchers often select a small number $\alpha$. If $p\text{-value} \leq \alpha$ then they will believe that the alternative hypothesis has been confirmed. The number $\alpha$ is called the **significance level** of the test. $\alpha = .05$ is commonly used.

**Stating Proper Conclusions -**

The conclusion should include two parts.

1. The **Decision Reached** (one or the other):

   - We reject $H_0$ in favor of $H_A$ at $\alpha = \bigcirc$.
   - We fail to reject $H_0$ at $\alpha = \bigcirc$.

2. **A Conclusion about the Parameter of Interest:**

   - We are convinced that (describe the alternative hypothesis)
   - We are NOT convinced that (describe the alternative hypothesis)
Example: Blood coagulation measurements were taking before and after a sample of patients with blood coagulation disorders was placed on heparin therapy. One particular measurement was on the antithrombin III activity. For each patient, what was recorded was the change in antithrombin III activity (after treatment level – before treatment level).

If the treatment was effective, the antithrombin III activity should be greater following treatment and hence these changes should be positive. If the treatment had no effect, then the changes should vary around zero (basically increasing or decreasing due to natural variation. Conduct a test of hypothesis.

The parameter of interest in this study is:

\[ \mu = \text{Ave. change in ATIII activity (after - before)} \]
\[ \text{treatment with heparin.} \]

The null and alternative hypotheses are:

\[ H_0: \mu = 0 \quad \quad H_A: \mu > 0 \]

The test statistic in this setting is:

\[ Z = \frac{\bar{X} - 0}{(S/\sqrt{n})} \quad \text{Under } H_0, \text{ this } Z\text{-score has an approximate } N(0,1) \text{ dist. when } n \text{ is large} \]
Extreme values of the test statistic in the positive direction will cause us to believe $H_A$ instead of $H_0$.

Data:

\[ n = 38 \quad \bar{x} = 47.6 \quad s = 40.2 \]

\[ Z = \frac{47.6 - 0}{\left(\frac{40.2}{\sqrt{38}}\right)} = 7.3 \]

\[ \text{p-value} = P\left[Z > 7.3 \mid Z \sim N(0,1)\right] \]

\[ \approx .000 \]

This is clearly very strong evidence against $H_0$ and supporting $H_A$. 
Steps in a Test of Hypothesis:
1. Define the parameter of interest in the study.

2. State the null and alternative hypotheses in terms of this parameter.

3. Calculate the value of a test statistic, measuring compatibility of the data with the null hypothesis.

4. Find the p-value of the data, measuring how extreme the data is in the direction of the alternative hypothesis.

5. Compare the p-value to $\alpha$ and state the conclusion of the test of hypothesis.
Chart of conclusions and consequences:

<table>
<thead>
<tr>
<th>Reality</th>
<th>$H_0$ is true</th>
<th>$H_A$ is true</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conclusions:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reject $H_0$ in</td>
<td>type I error</td>
<td>correct decision</td>
</tr>
<tr>
<td>favor of $H_A$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fail to reject $H_0$</td>
<td>correct decision</td>
<td>type II error</td>
</tr>
</tbody>
</table>

In a test of hypothesis, we control the probability of making a type I error.

$$P[\text{type I error}] = \alpha$$

The level of significance, $\alpha$, determines how compelling the data must be to convince us to reject $H_0$. 
Example: Can people tell the difference between diet coke and diet pepsi? A sample of 24 students was asked to taste test three cola soft drinks. They were told that two of the three drinks were the same and the other was different. Some students had two cokes and one pepsi while others had two pepsi's and one coke. The cups were coded, but otherwise identical.

Population parameter:

\[ p = \text{proportion of people who could correctly identify different brand.} \]

Hypotheses:

\[ H_0: p = \frac{1}{3} \quad H_a: p > \frac{1}{3} \]

The test statistic is therefore

\[ z = \frac{\hat{p} - \frac{1}{3}}{\sqrt{\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) \frac{1}{24}}} \]
We reject $H_0$ in favor of $H_a$ if

$z$ is extreme in $+$ direction

Data: $n = 24 \quad x = 10 \quad \hat{p} = \frac{10}{24} = .4167$

$$z = \frac{\frac{10}{24} - \frac{1}{3}}{\sqrt{\frac{\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)}{24}}} = .87$$

$p$-value $= 1 - .8078$

$=.1922$

Conclusion:

We fail to reject $H_0$ at $\alpha = .05$. We are not convinced that the proportion who can identify the difference exceeds $1/3$. 