I. Confidence intervals for \( \mu \) when \( \sigma \) is unknown (Section 7.3)

Recall:

- If the population distribution is normal, then
  \[
  \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}
  \]
  has a standard normal distribution.

- If \( \sigma \) is known, this justifies the use of
  \[
  \bar{X} \pm c \frac{\sigma}{\sqrt{n}}
  \]
  as a confidence interval for \( \mu \), where \( c \) comes from the standard normal table, B.2.

- For example, \( c = 1.96 \) for a 95% confidence interval.
What if $\sigma$ isn’t known?

- Would like to replace $\sigma$ by an estimate $S$ computed from the data.

- **Question:** Can we use the interval
  \[ \bar{X} \pm c \frac{S}{\sqrt{n}} , \]
  where $c$ is still computed from the standard normal table?

- **Answer:** No. The intervals won’t have the correct confidence levels. (You’ll see this in computer lab this week.)

- **Idea:** Replacing $\sigma$ (a number) by $S$ (a random variable) introduces more uncertainty, so we need to use a bigger multiplier $c$. 
How do we find $c$?

- **Fact:** If the population distribution is normal, then
  \[
  \frac{X - \mu}{S/\sqrt{n}}
  \]
  has “a $t$ distribution with $n - 1$ degrees of freedom.”

- Some algebra will justify using the interval
  \[
  X \pm c \frac{S}{\sqrt{n}}
  \]
  where $c$ is computed not from the standard normal table, but from the $t$ table.

- Still we want $c$ to satisfy “the area between $-c$ and $c$ is equal to the confidence level,” but now it’s area under a $t$ density.
Using the $t$ table

- The $t$ table is a bit different from the standard normal table.
- Each row of the table corresponds to a different degree of freedom. If the sample size is $n$, use the $n - 1$ degrees of freedom row.
- Each column in the table corresponds to a different area $\alpha$ to the right.
- Each number in the table has area $\alpha$ to its right.
- A few examples will make this much more clear.
<table>
<thead>
<tr>
<th>Degrees of Freedom</th>
<th>Area $\alpha$ in right tail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>1</td>
<td>1.376</td>
</tr>
<tr>
<td>2</td>
<td>1.061</td>
</tr>
<tr>
<td>3</td>
<td>0.978</td>
</tr>
<tr>
<td>4</td>
<td>0.941</td>
</tr>
<tr>
<td>5</td>
<td>0.920</td>
</tr>
<tr>
<td>6</td>
<td>0.906</td>
</tr>
<tr>
<td>7</td>
<td>0.896</td>
</tr>
<tr>
<td>8</td>
<td>0.889</td>
</tr>
<tr>
<td>9</td>
<td>0.883</td>
</tr>
<tr>
<td>10</td>
<td>0.879</td>
</tr>
<tr>
<td>80</td>
<td>0.846</td>
</tr>
</tbody>
</table>

Table 1: A portion of the $t$ table
Examples: We’ll solve the kind of problem needed for a confidence interval.

- **Problem:** Find the number $c$ such that the area between $-c$ and $c$ under the $t$ density with 7 degrees of freedom is equal to 0.90.
- **Solution:** The basic picture is the same as before. See Figure 1.
- Since we want the area between $-c$ and $c$ to be 0.90, the area to the right of $c$ is 0.05.
- So we look in the 0.05 column of the $t$ table, and get the answer 1.895.
- (Note that the corresponding answer from the standard normal table is 1.64.)
Figure 1: $c$ has area 0.05 to its right
• **Problem:** Find the number $c$ such that the area between $-c$ and $c$ under the $t$ density with 5 degrees of freedom is equal to 0.95.

• **Solution:** The basic picture is the same as before. See Figure 2.

• Since we want the area between $-c$ and $c$ to be 0.95, the area to the right of $c$ is 0.025.

• So we look in the 0.025 column of the $t$ table, and get the answer 2.571.

• (Note that the corresponding answer from the standard normal table is 1.96.)
Figure 2: $c$ has area 0.025 to its right.
Putting it all together

To find a $t$ confidence interval:

- Find the degrees of freedom, which is $n - 1$.
- Find $c$ from the $t$ table, using the $n - 1$ degrees of freedom row.
- Plug $c$, $S$, $\bar{X}$, and $n$ into the formula

$$\bar{X} \pm c \frac{S}{\sqrt{n}}.$$ 

- Remember that this procedure is valid if
  the population distribution is normal.
Confidence interval example

- In 1882, Michelson performed an experiment to measure the speed of light.
- He obtained 23 measurements of the speed of light in km/sec.
- The mean of his measurements is 299715.1 km/sec.
- The standard deviation of his measurements is 158.2 km/sec.
- **Problem:** Compute a 98% confidence interval for the true speed of light.
• **Solution:** First, since there are \( n = 23 \) measurements, we will use the 22 degrees of freedom row of the \( t \) table.

• Since we want a 98% confidence interval, we want the area between \(-c\) and \(c\) to be 0.98.

• This makes the area to the right of \(c\) equal to 0.01.

• From the \( t \) table we find \( c = 2.508 \).

• Plug everything into the formula:

\[
299715.1 \pm 2.508 \frac{158.2}{\sqrt{23}}.
\]

• This is

\[
299715.1 \pm 82.7
\]

• Note that in contrast to what we did previously, we didn’t need to assume a value for the population standard deviation \(\sigma\). We let the data give us an estimate \(S\) for this, and modified our procedure accordingly.
II. Testing hypotheses about proportions (Sections 8.1 and 8.2)

- Confidence intervals are used for estimation.
- Hypothesis testing is used for making decisions.
- (But we’ll see that confidence intervals and hypothesis tests are closely related.)
- We’ll learn about hypothesis tests via an example related to ESP.
III. **Example: Rhine’s ESP experiments**

In the 1930s J.B. Rhine and others conducted experiments to test whether a person had ESP.

**A. The procedure**

- A deck of cards with 5 designs (square, circle, star, plus sign, wavy lines) was used.

![Designs](image)

- The cards were shuffled thoroughly by the experimenter.

- A card was turned over each minute, and the subject had to write down the design without seeing the card.
B. The hypotheses, informally

- We are to decide between the two competing claims:
  - The subject does not have ESP
  - The subject has ESP
- We want to rewrite these in the context of a probability model.
- Let \( \pi \) stand for the probability that the subject correctly identifies a card.
- If the subject does not have ESP, we’d expect \( \pi = 0.2 \).
- If the subject does have ESP, we’d expect \( \pi > 0.2 \).
C. The hypotheses, formally

- We are to decide between \( \pi = 0.2 \) and \( \pi > 0.2 \) based on the data.
- We’ll call the claim \( \pi = 0.2 \) the *null hypothesis*, denoted \( H_0 \).
- We’ll call the claim \( \pi > 0.2 \) the *alternative hypothesis*, denoted \( H_a \).
- More concisely, the hypotheses are
  
  \[
  H_0: \pi = 0.2 \\
  H_a: \pi > 0.2
  \]
• Note that the burden is on the experimenter to disprove $H_0$.

• We’ll always try to set up hypotheses this way, where we’ll stick with $H_0$ unless there’s strong evidence against it.

• Useful to think of these in legal terms:
  – $H_0$: the defendant is innocent
  – $H_a$: the defendant is guilty
D. Possible errors

- In making a decision, we risk 2 possible errors.

- Deciding against $H_0$ when it is in fact true.
  - This is called a “Type I error.”
  - This is usually a more serious error, because of the way we choose our hypotheses.
  - In the legal analogy, a Type I error means convicting an innocent person.

- Deciding to stick with $H_0$ when $H_a$ is in fact true.
  - This is called a “Type II error.”
  - This is usually less serious.
  - In the legal analogy, a Type II error means acquitting a guilty person.
E. THE TEST STATISTIC

- We need to use the data to decide.
- Let $p$ stand for the sample proportion of correct answers.
- Intuitively if $p$ is “significantly” greater than 0.2, we’ll decide in favor of $H_a$.
- Define

$$Z = \frac{p - 0.2}{\sqrt{0.2(0.8)/n}}.$$ 

- Saying that $p$ is “significantly” greater than 0.2 is the same as saying that $Z$ is “significantly” greater than 0. (As long as we define “significantly” properly in both cases.)
- It’s more convenient to work with $Z$, because we know that if $H_0$ is true, then $Z$ (approximately, for large $n$) has a standard normal distribution.
- We call $Z$ the test statistic.
F. The p value

- How do we decide whether $Z$ is “significantly” larger than 0?
- We’ll assume $H_0$ is true, and see how likely it is that we’d get such a large value of $Z$ as the one we got in the experiment.
- The answer is the “p-value.”
G. Computing the p-value

- A specific experiment of the type described was performed in 1938.
- A large number of students were used as subjects.
- There were a total of 60000 cards used.
- The subjects got 12489 of the 60000 correct, which is a proportion of 0.20815 correct.
• The observed proportion 0.20815 is bigger than 0.2, but is it large enough to choose $H_a$?

• To answer this, we compute the observed value of the $Z$ statistic:

$$z_{obs} = \frac{0.20815 - 0.2}{\sqrt{.2(.8)/60000}} \approx 4.99.$$

• The p-value is the probability that a standard normal random variable is greater than 4.99, which is a very small number (about 0.0000003 from Table B.2).
H. Drawing conclusions

- Such a small p-value (0.0000003) provides strong evidence against $H_0$, so we would probably choose the alternative hypothesis $H_a$, that $\pi > 0.2$.

- Based on the data, I would be comfortable concluding that the true probability of correctly identifying the card is greater than 0.2.

- We have to be careful, though: This higher probability might be due to ESP, but it might also be due to other factors, such as cheating. I would not be comfortable deciding in favor of ESP unless I could be convinced that these other factors weren’t present.
• Typically we decide on a cutoff p-value $\alpha$ before the data are collected.

• If the calculated p-value is less than $\alpha$, we reject $H_0$.

• If the calculated p-value is not less than $\alpha$, we don’t reject $H_0$.

• Smaller cutoffs $\alpha$ provide more protection against Type I errors, but less protection against Type II errors.