Chapter 9: TESTS OF HYPOTHESES FOR A SINGLE SAMPLE

Part 2:
More on hypothesis testing for $\mu$
Type I & II errors in hypothesis testing
Using critical values instead of a $p$-value
Sections 9-1 to 9-3

• Statistical Hypothesis
  A statistical hypothesis is a statement about the parameters of one or more populations.

I should NEVER see a hypothesis stated in terms of a statistic:

$$H_0 : \bar{x} = 8$$

After we collect the data, I KNOW WHAT $\bar{x}$ IS! I don’t have to make a hypothesis about it.
Plus, I’m not interested in inferences about the sample, I’m interested in inferences or statements about the population (specifically, the parameters). I should see something like:

\[ H_0 : \mu = 8 \]
Types of errors in hypothesis testing
Part of section 9.1

- We could reject $H_0$ when it was actually true
  - **type I error** (the rate for this error is $\alpha$)
  - false positive
  - thought you found something interesting, but there really wasn’t

- We could fail to reject $H_0$ when the null was actually false
  - **type II error**
  - false negative
  - there was something interesting, but YOU MISSED IT!

Table showing possible outcomes

<table>
<thead>
<tr>
<th></th>
<th>$H_0$ is false</th>
<th>$H_0$ is true</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td>Correct</td>
<td>Type I error</td>
</tr>
<tr>
<td>Do not reject $H_0$</td>
<td>Type II error</td>
<td>Correct</td>
</tr>
</tbody>
</table>
Using the critical value method/rejection region method for hypothesis tests
See 9-1.2 and Figure 9-2.

Suppose you are testing the hypothesis...

\[ H_0 : \mu = 50 \]
\[ H_1 : \mu \neq 50 \]

If the null is true, \( \mu = 50 \) and \( n \) large, then

\[ \bar{X} \sim N(50, \frac{\sigma^2}{n}) \quad \text{[see graphic p.6]} \]

Under \( H_0 \) being true, \( \bar{X} \) has a normal distribution centered at 50.

\textit{But} if the alternative is true, \( \bar{X} \) is still normally distributed, but centered somewhere else (at the true \( \mu \), whatever that may be).

So, getting a sample with an \( \bar{X} \) far from 50 implies we should reject \( H_0 \) because \( \mu \) is probably something other than 50.
How far from $\mu_0 = 50$ must $\bar{x}$ be to reject $H_0$?

The answer depends on things like $n$, and $\sigma^2$, and also our choice of $\alpha$.

Let’s set the probability of a type I error $\alpha = 0.05$.

Now, suppose $n = 10$ and $\sigma = 2.4$ is known.

Under $H_0$ true, $P(48.5 \leq \bar{X} \leq 51.5) = 0.95$.

Or we could re-state it as, under $H_0$ true...

$P(\bar{X} < 48.5) = 0.025$ and $P(\bar{X} > 51.5) = 0.025$

“If the null is true (i.e. $\mu = 50$), rarely will we see an $\bar{x}$ that’s less than 48.5 or greater than 51.5 (just too far from the true mean).”

We can use these values to set a rejection region (shown in red in the graphic) for our hypothesis test that will maintain our type I error rate...
When $H_0$ is true, there’s only a 5% chance that $\bar{X}$ will fall outside of the range 48.5 to 51.5.

So, if $\bar{X}$ falls outside this range, we reject $H_0$.

We call the range where we reject $H_0$ the rejection region (shown in red for $\alpha = 0.05$ for this example).
There’s only a 5% chance of mistakenly rejecting \( H_0 \) when it was true (this means, if we use this rejection region, we’ve set \( \alpha = 0.05 \)).

If \( \bar{X} \) falls in the rejection region and \( H_0 \) true, we will mistakenly reject \( H_0 \). But this has only a 5% chance of happening (if \( H_0 \) actually true).

Some terminology...

- The rejection region is also called the critical region.

- The values that give the thresholds for rejection (like 48.5 and 51.5) are called the critical values.
Back to hypothesis testing for $\mu$
Sections 9.2 and 9.3

• Recall in a hypothesis test for $\mu$

  – When $\sigma^2$ is unknown, our test statistic will be

    \[ T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \]

    and we rely on the population being nearly normal for our probabilities to be correct.

  – When $\sigma^2$ is known, our test statistic will be

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

• There are two different procedures for making a decision in a hypothesis test:

  1. The $p$-value method
  2. The critical value method
• The two procedures:

1. Compute the \textit{p-value} for the test and reject if the \textit{p}-value < \( \alpha \).

2. Determine \textbf{critical value(s)} for a given \( \alpha \) (i.e. set the rejection region by determining the test statistic thresholds), and reject if the test statistic falls in the rejection region.

• These two methods lead to the same conclusion.

- I prefer the \( p \)-value method, but the \textit{t}-table in our book is very \textit{sparse}, so we can’t get very much accuracy on our \( p \)-value from the book when using \( T \) as the test statistic. So, you may prefer the critical value method.

More examples of hypothesis testing on handout...