We now have an 8-step procedure for hypothesis testing with p-values. The computation of the p-value itself (step 7) depends on $H_0 / H_1$:

If $H_0: \mu = \mu_0$ (or $\mu = \mu_0$) Then

1. Rightailed: $p$-value $= P(X > X_{\text{obs}} | \mu = \mu_0)$

Contrary to $H_0$

Graphically, $p$-value:

2. Leftailed: $p$-value $= P(X < X_{\text{obs}} | \mu = \mu_0)$

Contrary to $H_0$

Graphically, $p$-value:

3. Twoailed: $p$-value $= \sum_{|t| > t_{\text{obs}}} F(t)$

Then if $p$-value < $\alpha \Rightarrow$ Reject $H_0$ in favor of $H_1$.
else $\Rightarrow$ Cannot reject $H_0$ in favor of $H_1$.

= Accept $H_0$.!
In prev. example, we had $n = 64$, $\mu_{\text{obs}} = 34.4$, $s = 1.1$, and asked “Does data provide evidence to support $\mu > 34$?” Thus

$H_0: \mu \leq 34$

I always write these so that $H_0$ and $H_1$ have opposite directions, because it’s logical. The book does not.

$H_1: \mu > 34$

But, remember, it’s sufficient to test $H_0: \mu = 34$

“Blue note”

$d f = 64 - 1$

$p$-value $= P(\bar{x} > \bar{x}_{\text{obs}}) = P(t > t_{\text{obs}}) = P(t > 2.91) \approx 0.0025$,

$L = \frac{\bar{x}_{\text{obs}} - \mu_0}{s/\sqrt{n}} = \frac{34.4 - 34}{1.1/\sqrt{64}} = 2.91$

Since $p$-value $< \alpha$, Thus “There is evidence to support $\mu > 34$.”

It is tempting to say the above “conclusion” (at $\alpha = 0.05$), that $\mu > 34$, is obvious and trivial. After all the sample gave $\bar{x}_{\text{obs}} = 34.4$, which is greater than 34 already.

It’s NOT obvious! Suppose the sample/data gave $\bar{x}_{\text{obs}} = 34.1$, i.e. still larger than 34 - Then

$\mu_0$

$\frac{34.1 - 34}{1.1/\sqrt{64}} = 0.73 \Rightarrow p$-value $= P(t > 0.73) = 0.24$

This $p$-value is larger than any reasonable $\alpha$. So, we cannot reject $H_0$ in favor of $H_1$, even though the obs. sample mean is bigger than 34. 34.1 is larger than 34, but just not enough (in units of standard error, $\frac{s}{\sqrt{n}}$) to justify rejecting $H_0 (\mu \leq 34)$ in favor of $H_1 (\mu > 34)$.

[Once again, this $t_{\text{obs}}$ is NOT $t^*$ in CI’s.]
There are many ways to rephrase the statement/question in a problem. Here are some of them:

\[ n = 64, \ \bar{x} = 34.4, \ s = 1.1 \]

\[ t_{o.b.s} = \frac{34.4 - 34}{1.1 / \sqrt{64}} = 2.91 \]

**Does data support** \( \mu > 34 \)?

\( H_0: \mu < 34 \)

\( H_1: \mu > 34 \)

\[ p\text{-value} = \text{prob}(\bar{x} > \bar{x}_{o.b.s} | \mu = 34) = \text{prob}(t > t_{o.b.s}) \]

\[ = \text{prob}(t > 2.91) = 0.0025 < \alpha \]

\( \therefore \) **Reject** \( H_0 (\mu < 34) \) **in favor of** \( H_1 (\mu > 34) \).

\( \therefore \) **Data does support** \( \mu > 34 \).

**Does data support** \( \mu < 34 \)?

\( H_0: \mu > 34 \)

\( H_1: \mu < 34 \)

\[ p\text{-value} = \text{prob}(\bar{x} < \bar{x}_{o.b.s} | \mu = 34) = \text{prob}(t < t_{o.b.s}) \]

\[ = \text{prob}(t < 2.91) = 1 - \text{prob}(t > 2.91) = 0.9975 > \alpha \]

\( \therefore \) **Cannot Reject** \( H_0 (\mu > 34) \) **in favor of** \( H_1 (\mu < 34) \).

\( \therefore \) **Data does not support** \( \mu < 34 \).

**Does data contradict** \( \mu > 34 \)?

\( H_0: \mu > 34 \)

\( H_1: \mu < 34 \)

\[ p\text{-value} = \text{prob}(\bar{x} < \bar{x}_{o.b.s} | \mu = 34) = \text{prob}(t < t_{o.b.s}) \]

\[ = \text{prob}(t < 2.91) = 1 - \text{prob}(t > 2.91) = 0.9975 > \alpha \]

\( \therefore \) **Cannot Reject** \( H_0 (\mu > 34) \) **in favor of** \( H_1 (\mu < 34) \).

\( \therefore \) **Data does not contradict** \( \mu > 34 \).

**Does data contradict** \( \mu < 34 \)?

\( H_0: \mu < 34 \)

\( H_1: \mu > 34 \)

\[ p\text{-value} = \text{prob}(\bar{x} > \bar{x}_{o.b.s} | \mu = 34) = \text{prob}(t > t_{o.b.s}) \]

\[ = \text{prob}(t > 2.91) = 0.0025 < \alpha \]

\( \therefore \) **Reject** \( H_0 (\mu < 34) \) **in favor of** \( H_1 (\mu > 34) \).

\( \therefore \) **Data does contradict** \( \mu < 34 \).
We've done hyp. testing with p-values for testing a single μ:

\[ H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0 \]

But, recall from the CI days, we also had 2 μ's. The hypotheses are then:

\[ H_0: \mu_2 = \mu_1 \quad H_1: \mu_2 \neq \mu_1 \quad \text{(i.e.: } \mu_2 - \mu_1 \neq 0) \]

It turns out we can solve a more general problem:

\[ H_0: \mu_2 - \mu_1 = 0 \quad H_1: \mu_2 - \mu_1 \neq 0 \]

E.g., you may want to see if \( \mu_2 - \mu_1 \) exceeds 13, then \( \Delta = 13 \).

If you want to see if \( \mu_2 \) and \( \mu_1 \) are different, then \( \Delta = 0 \).

The point: \( \Delta \) is determined by the question, not by data.

It's the 2-sample analog of \( \mu_0 \).

⇒ Then, if 2-samples are independent, then assuming \( H_0 = T \),

\[
Z = \frac{(\bar{x}_2 - \bar{x}_1) - \Delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)
\]

Recall:

\[
t = \frac{\bar{x}_2 - \bar{x}_1}{S_{\bar{x}_1}} \sim t, \quad df=n-1
\]

\[ t = \frac{(\bar{x}_2 - \bar{x}_1) - \Delta}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim t - \text{dist. with } df = \text{Welch.} \]

Then, p-values are computed just as before.

See the "blue note":

\[
p-value = \begin{cases} \text{prob}(t > \text{tobs } | \mu_2 - \mu_1 = \Delta) & \text{if } H_1: \mu_2 - \mu_1 > \Delta \\ \text{prob}(t < \text{tobs } | \mu_2 - \mu_1 = \Delta) & \text{if } H_1: \mu_2 - \mu_1 < \Delta \\ \text{twice "tail"} & \text{if } H_1: \mu_2 - \mu_1 \neq \Delta \end{cases}
\]

⇒ If the two samples are paired: Make a new column:

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( d = x_1 - x_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>( \overline{d} )</td>
<td>( S_d )</td>
<td></td>
</tr>
</tbody>
</table>

\[
t = \frac{\overline{d} - \Delta}{S_d / \sqrt{n}} \sim t - \text{dist. } df = n-1
\]

p-value computed as before.
Reconsider this past example. Let’s do it with CI (now, with \( t \)) and with p-value.

**Example:** 82 students have picked up their test, but 30 have not, even 1 week after the test was returned. Call these 2 groups “Attendees” and “Non-Attendees.”

<table>
<thead>
<tr>
<th></th>
<th>( n )</th>
<th>( \bar{x} )</th>
<th>( s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-attend</td>
<td>30</td>
<td>11.8</td>
<td>3.32</td>
</tr>
<tr>
<td>Attend</td>
<td>82</td>
<td>13.25</td>
<td>3.04</td>
</tr>
</tbody>
</table>

\( \mu_1 \) = mean of test for non-attend students who have ever taken a test.

\( \mu_2 \) = mean of test for attend students.

Is there evidence from data that \( \mu_1 \) and \( \mu_2 \) are different?

We built the 2-sample (2-sided) 95% CI for \( \mu_2 - \mu_1 \):

\[
(13.25 - 11.8) \pm 1.96 \sqrt{\frac{(3.32)^2}{30} + \frac{(3.04)^2}{82}} \quad \text{from Welch}
\]

\[
1.45 \pm 1.26 (0.693) = 1.45 \pm 1.36 (0.06, 2.86) = (0.09, 2.81)
\]

Zero is not included in the CI \( \Rightarrow \) evidence that \( \mu_2 - \mu_1 \) are different.

We also commented that because the whole CI is to the right of 0, there is evidence that \( \mu_2 > \mu_1 \).

Now, let’s use the p-value method to answer the question “Is there evidence that \( \mu_2 > \mu_1 \)?”

\[
H_0: \mu_2 - \mu_1 \leq 0 \\
H_1: \mu_2 - \mu_1 > 0
\]

\[
t_{\text{obs}} = \frac{1.45 - 0}{0.693} = 2.1
\]

\( t \text{ is not 2.01} \) same conclusion

\[
p\text{-value} = \text{prob}(t > 2.1 | \mu_2 - \mu_1 = 0) \approx 0.0205
\]

At \( \alpha = 0.05 \), p-value < \( \alpha \), reject \( H_0 \) in favor of \( H_1 \).

\( \mu_2 < \mu_1 \), \( \mu_2 > \mu_1 \)

Not important to remember this (Welch’s) formula.

"In English": There is evidence that \( \mu_2 > \mu_1 \).
Do NOT Read This page until you are quite comfortable with CIs and p-values.

I have said repeatedly that $t^* \neq t_{obs}$. And I stick by that!
But, it turns out there are some relationships between them.
Here is an illustration:

We have seen 2 ways of answering the question: Does data support $m_2 > m_1$?

1) Find 95% obs LCB for $m_2 - m_1$
   
   $1.45 - 1.68(0.693) = 0.286$
   
   So 95% confident that $m_2$ exceeds $m_1$ by at least 0.286

2) Test $H_0: m_2 - m_1 < 0$ $\quad H_1: m_2 - m_1 > 0$
   
   p-value = $p(t > t_{obs}) = p(t > 2.1) = 0.0205$
   
   $\frac{1.45 - 0}{0.693} = 2.1$
   
   So at $\alpha = .05$, there is evidence for $m_2 - m_1 > 0$

But consider the following related questions:

3) At what conf. level is $LCB_{obs}$ for $m_2 - m_1$ equal to zero?
   
   $1.45 - t^*(0.693) = 0 \Rightarrow t^* = \frac{1.45}{0.693} = 2.1$
   
   So conf. level = $p(t < t^*) = 1 - 0.0205 = 0.9795$

4) Does data support $m_2 - m_1 > 0.29$?
   
   $H_0: m_2 - m_1 < 0.29 \quad H_1: m_2 - m_1 > 0.29$
   
   p-value = $p(t > t_{obs}) = p(t > 1.68) = 0.05$
   
   $\frac{1.45 - 0.286}{0.693} = 1.68$

So 97.95% confident that $m_2$ exceeds $m_1$.

Note the relationships between $t_{obs}$ & $t^*$.
But, again, until you are very comfortable with the 2 methods, keep $t^*$ and $t_{obs}$ separate.
We are done with 1-sample and 2-sample, z and t-tests, for paired and unpaired data, but all of that has dealt with pop. means.

Let's not forget pop. props. The procedure is the same

\[
Z = \frac{\hat{p} - \mu}{\sqrt{\frac{\mu(1-\mu)}{n}}} \sim N(0,1) \quad \text{but} \quad \frac{\hat{p} - \mu}{\sqrt{\frac{P_e(1-P_e)}{n}}} \quad \text{not} \ t
\]

\[\frac{\bar{x} \pm z^* \frac{S_x}{\sqrt{n}}}{\frac{\bar{x} \pm t^* \frac{S_x}{\sqrt{n}}}{\sqrt{n}}} \quad \text{df = n-1}
\]

Test for \( \mu \):

\[
H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0
\]

\[
z_{obs} = \frac{\bar{x}_{obs} - \mu_0}{\sigma_x/\sqrt{n}} \quad t_{obs} = \frac{\bar{x}_{obs} - \mu_0}{S_x/\sqrt{n}}
\]

\[
p-\text{value} = p(\bar{x} \boxtimes \bar{x}_{obs}) \quad df = n-1
\]

\[\text{Test for } \pi_x:
\]

\[
H_0: \pi = \pi_0 \quad H_1: \pi \neq \pi_0
\]

\[
z_{obs} = \frac{P_{obs} - \pi_0}{\sqrt{\pi_0(1-\pi_0)}} \quad \text{no } t! \text{ Because we assumed } H_0 = \pi, \text{ it: } \pi_x = \pi -
\]

\[
p-\text{value} = p(\pi \boxtimes \pi_{obs})
\]

C.I. for \( \mu_2 - \mu_1 \):

\[
\bar{x}_2 - \bar{x}_1 \pm z^* \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \quad \text{df = Welch}
\]

Test for \( \mu_2 - \mu_1 \):

\[
H_0: \mu_2 = \mu_1 \quad H_1: \mu_2 \neq \mu_1
\]

\[
z_{obs} = \frac{(\bar{x}_2 - \bar{x}_1)_{obs} - \Delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad t_{obs} = \frac{(\bar{x}_2 - \bar{x}_1)_{obs} - \Delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}
\]

\[
p-\text{value} = p(\bar{x}_2 - \bar{x}_1 \boxtimes ((\bar{x}_2 - \bar{x}_1)_{obs})) \quad df = \text{Welch}
\]

Again, no t!

C.I. for \( \pi_x - \pi_{y} \):

\[
\pi_x - \pi_{y} \pm z^* \sqrt{\frac{\pi_x(1-\pi_x)}{n_x} + \frac{\pi_{y}(1-\pi_y)}{n_y}}
\]

Test for \( \pi_x - \pi_{y} \):

\[
H_0: \pi_x = \pi_{y} \quad H_1: \pi_x \neq \pi_{y}
\]

\[
z_{obs} = \frac{(\pi_x - \pi_{y}) - \Delta}{\sqrt{\frac{\pi_x(1-\pi_x)}{n_x} + \frac{\pi_{y}(1-\pi_y)}{n_y}}}
\]

\[
p-\text{value} = p(\pi_x - \pi_{y} \boxtimes (\pi_x - \pi_{y})_{obs})
\]

(paired vs. unpaired)

\[
p(\pi \boxtimes z_{obs})
\]
In hw_lect16_1 we used a CI to answer the question Does it appear that the true proportion of detective screws is not 2.5%. Here, answer the same question with the p-value approach. Specifically,

a) Set-up the appropriate hypotheses.

b) Compute the p_value (using the data in that hw)

c) At alpha =0.01, state the conclusion? Is it consistent with the conclusion from the CI approach?

Suppose you are asked if there is evidence that \( \mu_x > \bar{x}_{obs} - 1.645 \frac{S}{\sqrt{n}} \)?

FVR: The right-hand side is the 95\% LCB (which we are skipping).

a) Set-up the appropriate \( H_0 / H_1 \)

b) Compute the p-value. Assume \( n = \infty \) in Table VI.
This document was created with Win2PDF available at http://www.win2pdf.com.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.
This page will not be added after purchasing Win2PDF.