One-sided tests

I. So what is a one sided test?

1) in our original set-up, we formulated a hypothesis such as:

\[ H_0: \mu_1 = \mu_2 \]

our alternative was:

\[ H_1: \mu_1 \neq \mu_2 \]

we also discussed, briefly, that one could have:

\[ H_1: \mu_1 < \mu_2 \quad \text{or} \quad H_1: \mu_1 > \mu_2 \] (not both).

2) The last two alternatives give rise to what is known as a one-sided test.

a) not all tests that you will see have one-sided alternatives (an example is ANOVA with more than two samples).

b) So what happens in a one sided test?

i) we REJECT if we conclude \( \mu_1 < \mu_2 \) or \( \mu_1 > \mu_2 \) depending on which \( H_1 \) we use.

ii) we do not test both alternatives at once.

3) How do you decide if you need to do a one sided test?

Either based on:

1) logic: Do you know something about the situation that would lead you to believe that a one sided test would do better?

   - for example - you test a new medication to lower blood pressure.

   - what outcome are you interested in?

   - that the medicine LOWERS blood pressure. If \( \mu_1 = \text{“average blood pressure before medication”} \), then you're interested in \( H_1: \mu_1 > \mu_2 \).

     - you want to show that the medication lowers blood pressure; you're not interested if it somehow raises blood pressure.

   or

2) subject matter knowledge: Do you know something about the biology of the situation that let's you use a one sided test?

   - for example, you want to compare heights in men vs. women.
- you know that men are taller on average, so if $\mu_1 =$ “average height of women”, then you're interested in $H_1: \mu_1 < \mu_2$.

- you know that women are shorter on average, so you would expect the outcome of this experiment to show that.

- what can be confusing here is that sometimes one person knows something another doesn't.

- For example, if I wanted to compare the length of male vs. female kestrels (a small hawk), I know that females are larger, so I get to do a one sided test.

- Unless you knew this ahead of time, you would have to do a two sided test (since you (probably) don't know that much about these birds.

In any case, you make your decision based on what you know about the problem (logic or subject matter knowledge).

**NEVER, NEVER** decide to do a one sided test by comparing $\bar{y}_1$ with $\bar{y}_2$. That's cheating.

4) Here's an outline for Welch's test - other tests are very similar (all the tests we've done so far can be one or two sided):

a) Set up our hypotheses:
   i) $H_0: \mu_1 = \mu_2$
   ii) $H_1: \mu_1 < \mu_2$ (or $H_1: \mu_1 > \mu_2$)

b) decide on $\alpha$.

c) **verify that your results agree with your alternative hypothesis**

   - if your alternative hypothesis is $H_1: \mu_1 < \mu_2$, make sure $\bar{y}_1 \leq \bar{y}_2$ (or vice-versa)
   
   - if this is not true, STOP - don't do anything else. Don't calculate anything else. Your results contradict your alternative hypothesis!

d) calculate $t^*$ in the usual way

   - make your comparison in the usual way:
     
     if $|t^*| \geq t_{\text{table}}$ then reject (otherwise “fail to reject”).

   - make sure that you now use the top row of your $t$-tables to select $\alpha$ (up until now we've only used the bottom row).

   or (as usual):

     if $p$-value $\leq \alpha$ then reject (otherwise “fail to reject”)
e) note that this only works if you do part (c) first; otherwise this can lead you to make a mistake
(the more traditional way of doing things is given in your text).

6) An example:

a) exercise 7.42 on p. 269 [7.52, p. 265] \( \frac{7.5.7, p. 259}{7.5.7, p. 265} \).

i) lettuce seedlings were grown on either standard soil, or soil with extra nitrogen.

ii) what are you interested in?
   - standard soil is better?? Not likely.
   - nitrogen soil is better?? This makes sense.

b) So form your hypotheses:

   \[ H_0: \mu_1 = \mu_2 \]
   \[ H_1: \mu_1 < \mu_2 \]

   - why? because \( \mu_1 \) represents plants grown in standard soil.

   - also note: we could pick \( H_0: \mu_1 \geq \mu_2 \) (some statisticians prefer this). We still do
     not, however, “accept” \( H_0 \).

c) pick \( \alpha = .10 \) (book picks \( \alpha \) for us).

d) verify that our results agree with the alternative hypothesis;
   - yes, \( \bar{y}_1 < \bar{y}_2 \) is true.

d) calculate \( t^* \). Here's the math for Welch's test (notice it doesn't change):

\[
t^* = \frac{3.62 - 4.17}{\sqrt{\frac{0.54^2}{5} + \frac{0.67^2}{5}}} = -1.43
\]

e) so now figure out our d.f. (same as previously):

\[
v = \frac{(SE_1^2 + SE_2^2)^2}{SE_1^4/n_1 - 1 + SE_2^4/n_2 - 1}
\]

\[
= \frac{(0.2415^2 + 0.2996^2)^2}{0.2415^4 + 0.2996^4}
\]

\[
= 7.67
\]
f) now we look up $t$ with 7 d.f. and $\alpha = .10$ and get 1.415.

g) Our comparison:

$$|t^*| = 1.43$$

is less than or equal to 1.415,

so we reject $H_0$.

(if we had done a two sided test we would have had to FTR - one sided tests have more power!!)

h) We conclude that we have evidence to show that adding nitrogen to soil affects the weight of lettuce seedlings.

7) Some comments:

a) note that your $p$-values are now half of what they were before. Why? because you don’t have to put half of your probability in the other tail!! This is good, because you have a much better chance of rejecting $H_0$ (you get more power).

b) remember: if you look at your data after performing the test and find that $\bar{y}_1 \gg \bar{y}_2$, you just can’t go back and test for $\mu_1 > \mu_2$. You’d be cheating!!

c) if you decide to do a one-tailed test, you must decide to do so before looking (snooping) through your data. You should have a reason for choosing a one-tailed test before you even perform the experiment.

II. The one-sided Mann-Whitney U-test.

1) No long discussion. Just a brief outline - you should be able to figure this out yourself.

2) Outline:

a) Develop your hypotheses (just as for the one-tailed t-test, except remember you’re testing distributions, not means)

b) decide on $\alpha$.

c) Calculate $U^*$. But note (like in the $t$-test) that if your alternative is that distribution 1 > distribution 2 (or $\mu_1 > \mu_2$), and $K_1 < K_2$, then you can stop since you’re $p$-value will be over 0.5.

d) Look up your $U^*$ in table 6. Your comparison is the same as for the two-tailed test (if $U^* \geq U_{table}$, reject), except that you need to use the appropriate column for a one-tailed test (remember to modify your U-table if you haven’t!)

e) make your decision to reject $H_0$ or fail to reject $H_0$.

3) You’ll have one of these for homework.

III. As mentioned, the other tests we've learned (so far!) can also all be one sided (one sample t-test, sign test, signed rank test, etc.)