Math 316

Section 7.8 Student's t: Conditions and Summary

So what is this "Student" thing anyway? Student is the pen name of the guy (William Sealy Gosset) who first came up with and used the t-tests in hypothesis testing, so it is often called Student's t-test. (No, it's not in honor of you students.) See **Conditions** and **Verification of Conditions** on **Page 281** and **Consequences of Inappropriate Use of Student's t** on **Page 282**. Regarding Consequence 1, it is basically saying that if we assume normality (either because we think the populations are normally distributed, or because we think the samples sizes are large enough to result in normally distributed sample means) when we should not, we will end up with a smaller P-value than we should. (It's like we're sort of cheating.) It has a negative effect similar to the effect there would be if we were to assume that we have a larger sample size than we actually do. Regarding Consequence 2, these other tests are what we spend much of our time over the next several weeks learning about. See the **Summary of the t Test** on **Page 284**.

Section 7.9 More on Principles of Testing Hypotheses

First, let me remind you that area under a probability distribution function (for example, a normal curve) represents probability. Look at **the four ways of describing P** on **Page 287**. Here is a fifth description (really, the most accurate), very similar to the other four.

First, remember how we compute P. We take a sample from each population and find the mean of each sample. Then for the two samples we compute

$$t_s = \frac{\bar{y}_1 - \bar{y}_2}{SE_{\bar{Y}_1 - \bar{Y}_2}}$$

which measures the relative difference of the two sample means. Next, suppose that the two populations have the same mean (that is, the null hypothesis $H_0: \mu_1 = \mu_2$ is true), and suppose that we took many other samples from these two populations, and compared them, one pair of samples (one from each population) at a time, as illustrated in **Figure 7.3.4**. Then the curve seen in Table 4 is the distribution of the t_s values, and:

> P is the fraction of all pairs of samples whose test statistic t_s would be as large or larger than the t_s computed for the samples we actually have. Visually, this is the area in the tail(s).

Section 7.10 The Wilcox-Mann Whitney Test

The Wilcox-Mann-Whitney Test is used to determine whether two population means are different or not, just like the t test. When the t test can be used (as detailed on page 273), it is generally a better way to compare samples. The main weakness of the t Test is that it cannot always be used, since the population(s) may not be normally distributed and/or the sample size(s) may not be very large. The Wilcox-Mann-Whitney Test is not as powerful (that is, we might not end up concluding that the population means are different when they actually are; that is, we are more likely to make a Type II error), but it can always be used, regardless of normality and sample size. In general, the less restrictive the conditions (normally distributed data? large sample sizes?) are for using that test, the less powerful the test is. See the book discussion of this on pages 297 – 298. Note: we didn't have any HW on the randomization test, so you probably didn't pay much attention to it. For those of you who a bit more serious about research (both performing and interpreting it), you might spend a few minutes reading Section 7.1.

Here is **HW 7.10.8.** Here are the data, now ordered:

# smaller	Joggers
1	13
2	19
4	28
5	32
5	32
6.5	37
7	39
7	40
8	41
12	52
14	60

$K_1 = 71.5$

Fitness	# smaller
9	0
18	1
23	2
27	2
31	3
33	5
37	5.5
41	8.5
41	8.5
42	9
47	9
49	9
54	10
59	10
70	11

$K_2 = 93.5$

Compute the test statistic (easy!):
$U_s = \max(71.5, 93.5) = 93.5$

Table 6 with $n = 15, n' = 11$
$\Rightarrow P > .20$

Note that the P values in Table 6 are non-directional (two-tailed) hypothesis testing where H_A is $\mu_1 \neq \mu_2$ (rather than the directional $\mu_1 < \mu_2$ or $\mu_1 > \mu_2$). Divide the given P value in half if your test is directional. Note that $K_1 + K_2 = n \cdot n'$. For the above Wilcox-Mann Whitney Test, while we compute U_s differently than t_s, U_s has several qualities similar to t_s. A larger test statistic U_s:

- Means the samples are more different.
- Gives us a smaller P value.
- Means it is more likely that we will conclude that the samples come from populations with different means.

From Table 6, it's not totally obvious that larger samples are beneficial, but (as always) they are. See **Figure 7.10.3** regarding what the smallest and largest U_s can be. If time: **Class Observation 1** regarding how larger sample sizes affect U_s (and its resulting P value) and **Class Observation 2** regarding what would happen if we had used the t Test for the above problem. If time: **HW 7.10.9**.