COMPARING SAMPLE MEANS (PART 1 UNPAIRED DATA)

CHEM 251 SDSU

COMPARING MEANS

- There are many cases where the results of two sets of measurements need to be compared.

- In these instances, it is not necessary to know the "true" value of the what is being measured, we simply need to have information about the quality of the measurements of the samples.

- The question that is typically asked (the <u>null hypothesis</u>) is wether or not the means of each set of measurements are the same (or not).

PAIRED OR UNPAIRED

- In comparing means we need to first identify the type of data set that we are dealing with.

- The measurements could be **paired**, or **unpaired**.

- With **paired** data, the same samples are measured with each analysis.

- With **unpaired** data, the sample is split and part is measured by each analysis.

COMPARING MEANS OF UNPAIRED DATA

- We can modify the t test once again to test the null hypothesis (H_0: $\mu_A = \mu_B$).

- We can then compare t_{exp} to the desired tabulated confidence level.

$$t_{exp} = \frac{|\overline{X}_A - \overline{X}_B|}{\sqrt{\dfrac{s_A^2}{n_A} + \dfrac{s_B^2}{n_B}}}$$

Trials	Student A	Student B
1	14.602	14.408
2	14.782	14.517
3	14.668	14.322
4	14.534	14.477
5	14.721	14.398
6	14.596	
Average	14.6505	14.4244
Std. Dev.	0.091	0.075

HOW MANY DEGREES OF FREEDOM?

- The degrees of freedom for a comparison of means depends on the difference in the standard deviations of each set of measurements.

- If s_A^2 and s_B^2 provide estimates of the same population variance, we can pool the standard deviations.

- This is determined with a simple F-test.

Trials	Student A	Student B
1	14.602	14.408
2		
3		
5		
6	14.596	
Average	14.6505	14.4244
Std. Dev.	0.091	0.075

$$F_{exp} = \frac{s_A^2}{s_B^2}$$

PERFORMING AN *F*-TEST

Trials	Student A	Student B
1	14.602	14.408
2	14.782	14.517
3	14.668	14.322
4	14.534	14.477
5	14.721	14.398
6	14.596	
Average	14.6505	14.4244
Std. Dev.	0.091	0.075

$F(0.05, \nu_{num}, \nu_{denom})$ for a Two-Tailed F-Test

$\frac{\nu_{num}}{\Downarrow \nu_{denom}}$	1	2	3	4	5	6
1	647.8	799.5	864.2	899.6	921.8	937.1
2	38.51	39.00	39.17	39.25	39.30	39.33
3	17.44	16.04	15.44	15.10	14.88	14.73
4	12.22	10.65	9.979	9.605	9.364	9.197
5	10.01	8.434	7.764	7.388	7.146	6.978
6	8.813	7.260	6.599	6.227	5.988	5.820

$H_0: s_A^2 = s_B^2$

$F_{exp} \leq F_{(0.05,5,4)}$ accept H_0

$F_{exp} > F_{(0.05,5,4)}$ reject H_0

TEST WITH POOLED VARIANCE

- If the null hypothesis (H_0: $s_A^2 = s_B^2$) is accepted, the variances from both measurements may be pooled.

- The degrees of freedom are then calculated as:

$n_A + n_B - 2$

$$t_{exp} = \frac{|\overline{X}_A - \overline{X}_B|}{s_{pool}} \times \sqrt{\frac{n_A n_B}{n_A + n_B}}$$

$$s_{pool} = \sqrt{\frac{(n_A - 1) s_A^2 + (n_B - 1) s_B^2}{n_A + n_B - 2}}$$

TEST WITHOUT POOLED VARIANCE

- If the variances are sufficiently different (the null hypothesis is rejected) a different calculation is used to determine the total degrees of freedom (v).

- Round v to the nearest integer value.

$$t_{exp} = \frac{|\overline{X}_A - \overline{X}_B|}{\sqrt{\dfrac{s_A^2}{n_A} + \dfrac{s_B^2}{n_B}}}$$

$$v = \frac{\left[\dfrac{s_A^2}{n_A} + \dfrac{s_B^2}{n_B}\right]^2}{\dfrac{\left(\dfrac{s_A^2}{n_A}\right)^2}{n_A+1} + \dfrac{\left(\dfrac{s_B^2}{n_B}\right)^2}{n_B+1}} - 2$$

SAMPLE CALCULATION

If Student A and B measured samples from the same population, are their reported means different at the 95% confidence level?

Trials	Student A	Student B
1	14.602	14.408
2	14.782	14.517
3	14.668	14.322
4	14.534	14.477
5	14.721	14.398
6	14.596	
Average	14.6505	14.4244
Std. Dev.	0.091	0.075