Summary Statistics for Quiz 6A and 6B

Quiz A	Quiz B
$n_1 = 11$	$n_2 = 11$
$\bar{x}_1 = 7.5$	$\bar{X}_2 = 6.8$
$S_1 = 2.4$	$S_2 = 1.8$

Lets say I make the claim that Quiz 6B was harder than Quiz 6A. In other words on the average Quiz A > Quiz B grades. I will test at 95% confidence

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_a : \mu_1 - \mu_2 > 0$$

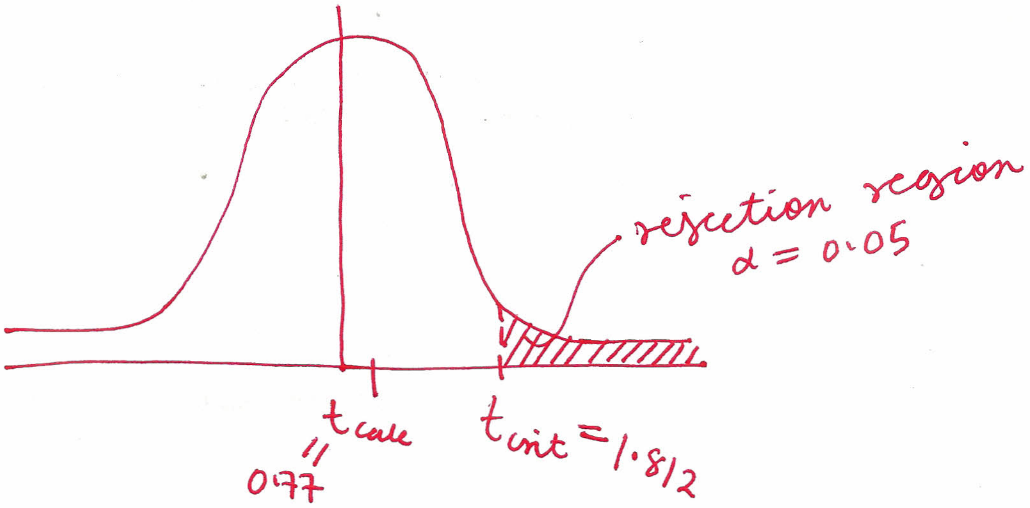
$$\alpha = 0.05$$

Since we have small samples, we shall use t-value and t-tables. So our test statistic is,

$$t_{calc} = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_0}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{n}}}$$

$$t_{calc} = \frac{(7.5 - 6.8) - 0}{\sqrt{\frac{2.4^2}{11} + \frac{1.8^2}{11}}} = 0.77$$

So we plot this value in relation to our rejection region. We may use critical values. (or if you like p-values).



How did I get t_{crit} ?

$$t_{crit} = t_{\alpha, df=n-1} = t_{0.05, 10} = t_{0.05, 10}$$

So go to the t-table, look for column $\alpha = 0.05$, go to row 11 degrees of freedom and read the value.

Also for a 2-sample test, $df = \min(n_1 - 1, n_2 - 1)$

So our t_{calc} does not fall in rejection region, so fail to reject H_0! So my claim that Quiz A scores are (statistically) higher than Quiz B, inferring that Quiz B was more difficult that Quiz A, is not valid. Both quizzes were of equal difficulty. And I am 95% confident with my conclusion.

Students: Please rework using p-values.

Now let's say I make the claim

"quiz A and quiz B grades/difficulty are the same"

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_a: \mu_1 - \mu_2 \neq 0$$

$$\alpha = 0.05$$

So now we have a 2-tail test.

$$t_{calc} = \frac{(7.5 - 6.8) - 0}{\sqrt{\frac{2.4^2}{11} + \frac{1.8^2}{11}}} = 0.77$$

This time let's use p-value. So go to the t-table at 10 degrees of freedom and read the α-value for a t_{calc} value of 0.77.

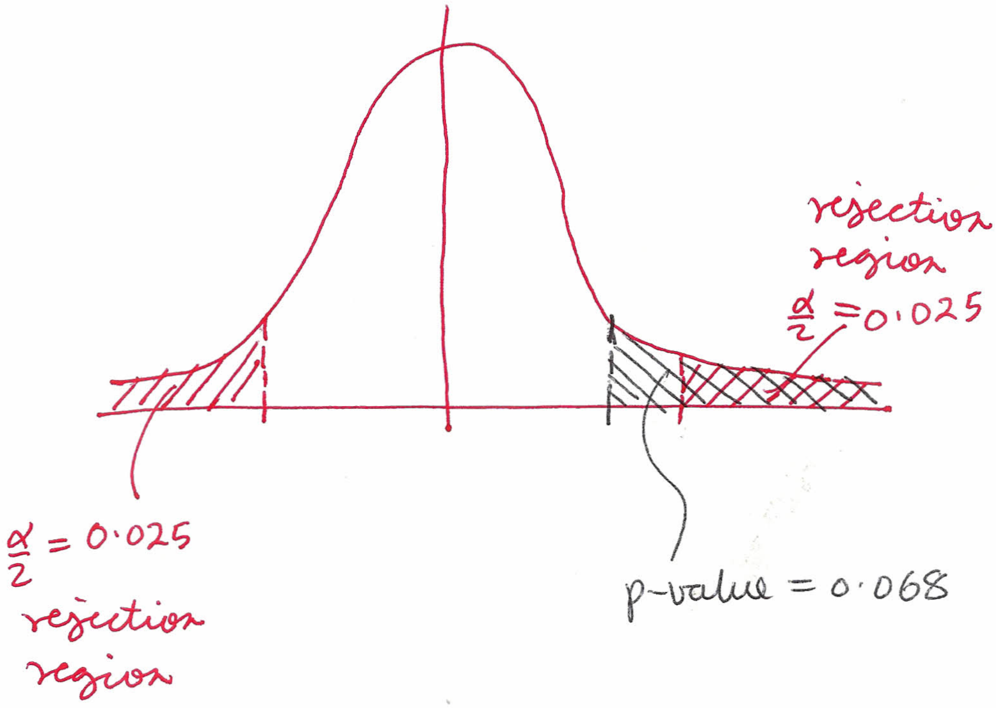
There is not one so we have to interpolate on the 10 df row.

t	1.812	1.372	0.77
	$\downarrow$	$\downarrow$	$\downarrow$
α	0.05	0.10	$? = x$

$$\frac{x - 0.1}{0.77 - 1.372} = \frac{0.1 - 0.05}{1.372 - 1.812}$$

$$\Rightarrow \quad x = \frac{0.1 - 0.05}{1.372 - 1.812} \cdot (0.77 - 1.372) = 0.068$$

This is our observed α-value, commonly called the p-value. So let's draw the picture.



Why did I shade the p-value from the positive end? Because I had positive t_{calc}. If I had a negetive t_{calc} I would have shaded from the negetive end toward the center.

So p-value > α

We are not in the rejection region.

Fail to reject H_0! Quiz A and Quiz B are the same. My claim is correct.

For a 2-tail test, we may alternately conduct the test using a confidence interval

If $\mu_0 \in CI$, then we fail to reject H_0, otherwise we reject H_0.

95% CI for 2-sample means

$$= (\bar{X}_1 - \bar{X}_2) \pm t_{\substack{\alpha=\frac{0.05}{2} \\ @10df}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

note that if we had large samples we would use Z-value.

$$t_{\alpha=\frac{0.05}{2},10df} = t_{\alpha=0.025,10df} \xrightarrow{\text{table}} 2.228$$

$$95\%\ CI = (7.5 - 6.8) \pm 2.228 \sqrt{\frac{2.4^2}{11} + \frac{1.8^2}{11}}$$

$$= 0.7 \pm 2.015$$

$$= [-1.315, 2.715]$$

So our $\mu_0 = 0$, and it is in the CI.

Fail to reject H_0!

So, again, the claim is valid.

Just a thought.

If we had $CI = [-3.5, -0.1]$, all negetive values, we would reject H_0. Also because we have all negetive values, we would have

$$\mu_1 - \mu_2 \rightarrow \text{negetive values}$$

This would mean that Quiz 2 grades are consistently higher than Quiz 1 on average

If we had all positive values, we would reject H_0 and also

$$\mu_1 - \mu_2 \rightarrow \text{positive values}$$

This would mean Quiz 1 grades are consistently higher than Quiz 2 on average.