MA207 Two Sample Population Mean Test Name:______________________________

Ex 1 A buyer is comparing cars and wants to know if the Chevy Malibu has significantly better gas mileage than the Ford Focus. To find the gas mileage of a car, a dealership test-drives the car on 20 different days and records the mileage each day; the data for each car are normally distributed. The Malibu gets 25 mpg in the city with standard deviation 5 mpg. The Focus gets 22 mpg in the city with standard deviation 4 mpg. Test the buyer's hypothesis at $\alpha = 0.05$.

	n	$\bar{x}$	s
1) Malibu	20	25	5
2) Focus	20	22	4

a. State the hypotheses.

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 > \mu_2$

b. Is it a left-tailed, right-tailed, or two-tailed test?

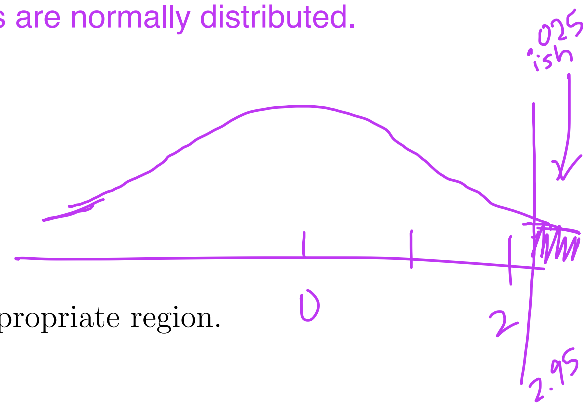
c. Explicitly check any assumptions needed to use the test.

Random- yes, practically. Drove on 20 different days. As long as they drove on days pre-determined, not just when they woke up and it was sunny or something.)

Independent- yes, the cars have no overlap or relationship
Normal population- yes, this is implied because the samples are normally distributed.

d. Compute the test statistic.

$$t = \frac{25 - 22}{\sqrt{\frac{5^2}{20} + \frac{4^2}{20}}} \approx 2.095$$



e. Determine the P-value. Draw a labeled curve and shade the appropriate region.

$df = \min(20-1, 20-1) = 19$

2.095 is very close to 2.093, so the right tail area is very close to .025. $\boxed{P \text{ is almost } .025}$

f. Interpret the P-value in the context of this problem. (You are not using α yet.)

The probability of obtaining a difference in sample means that is even more than 3 (or a test statistic that is even more than 2.95) is only about .025, assuming that the overall mileage for each car is the same.

g. (a) Are the results statistically significant or not statistically significant?

$\underset{.025}{P} < \underset{.050}{\alpha}$

(b) Do we therefore reject the null hypothesis or not reject the null hypothesis?

(c) Explain in complete sentences what this means in the context of this problem.

There is sufficient evidence to conclude that the mean gas mileage for the Malibu over all days is greater than the mean gas mileage for the Focus over all days.

Ex 2 A buyer is comparing cars and wants to know if the Chevy Malibu has significantly better gas mileage than the Ford Focus. To find the gas mileage of a car, a dealership test-drives the car on 20 different days and records the mileage each day; the data for each car are normally distributed. The Malibu gets 25 mpg in the city with standard deviation 5 mpg. The Focus gets 22 mpg in the city with standard deviation 4 mpg. Compute the 90% confidence interval for the difference of the two population means.

a. Find a ~~98~~ 90% confidence interval for the difference between the two population proportions.

df	90%
19	→ 1.729

$$25-22 \pm 1.729\sqrt{\frac{5^2}{20}+\frac{4^2}{20}}$$

$$3 \pm 2.5$$

$$(0.5, 5.5) \text{ mpg}$$

b. Identify the point estimate, critical value, the standard error, and the margin of error.

3 1.729 1.43 2.5

c. Explain what the confidence interval means in practical terms.

I am 90% confident that the difference in overall gas mileages between Malibu cars and Focus cars is between 0.5 and 5.5 mpg.

d. Does the null hypothesis value for the difference between the two population proportions lie in the confidence interval? Does this match what you expected based on your significance test? Explain.

No, 0 is not in the interval. The difference is not 0, meaning there is a difference in average mileage for the two cars. This matches the results from the test which showed sufficient evidence that the Malibu has higher average gas mileage over all days.

Ex 3 A member of the math faculty is trying to determine if students who took a Calculus course in high school perform differently in the math courses they take at the university. A sample of 45 students who did not take Calculus in high school results in a math GPA of 2.03 with standard deviation 0.68. A sample of 35 students who did take Calculus in high school results in a math GPA of 2.21 with standard deviation 0.59. Using $\alpha = 0.05$, does the evidence support the faculty member's hypothesis?

	n	$\bar{x}$	s
1) didn't take Cal	45	2.03	.68
2) did take Cal	35	2.21	.59

a. State the hypotheses.

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \neq \mu_2$

b. Is it a left-tailed, right-tailed, or two-tailed test?

c. Explicitly check any assumptions needed to use the test.

Randomly obtained data-hopefully, though it didn't specify (it could be a flawed study)
Independent samples- yes, a student can't belong to both populations, and there aren't matched pairs
The sample sizes are both greater than 30.

d. Compute the test statistic.

$$t = \frac{2.03 - 2.21}{\sqrt{\frac{.68^2}{45} + \frac{.59^2}{35}}} \approx -1.27$$

e. Determine the P-value. Draw a labeled curve and shade the appropriate region.

$df = \min(45-1, 35-1) = 34$

df	$t_{\square}$	$t_{.100}$	$t_{.050}$
30	$\uparrow$	1.31	1.697

34 isn't on the table. You always round down in such cases..

The right tailed area is even bigger than .100
The P-value is even bigger than .200 (since this is a two-tailed test).

f. Interpret the P-value in the context of this problem. (You are not using α yet.)

The probability of obtaining a difference in sample GPAs of -.18 (a test statistic of -1.27) or a difference even more extreme is greater than .200, assuming the two population mean GPAs are the same.

g. Explain in complete sentences what this means in the context of this problem.

There is not enough evidence to reject the claim that the average GPA in math courses taken at the university is the same for students who took calculus in high school as it is for those who didn't take calculus in high school.

Ex 4 A member of the math faculty is trying to determine if students who took a Calculus course in high school perform differently in the math courses they take at the university. A sample of 45 students who did not take Calculus in high school results in a math GPA of 2.03 with standard deviation 0.68. A sample of 35 students who did take Calculus in high school results in a math GPA of 2.21 with standard deviation 0.59. Compute the 95% confidence interval for the difference of the two population means.

a. Find a 9~~0~~5% confidence interval for the difference between the two population proportions.

df = 34 use 30 on chart & 95% confidence

df	95%
30	→ 2.042

$$(2.03 - 2.21) \pm 2.042 \sqrt{\frac{.68^2}{45} + \frac{.59^2}{35}}$$

$$-.18 \pm .290$$

$$(-.47, .11)$$

b. Explain what the confidence interval means in practical terms.

I am 95% confident that the difference between the average gpa of all students who took calculus in high school and the average gpa of all students who didn't take calculus in high school is between -0.47 and .11.

c. Does the null hypothesis value for the difference between the two population means lie in the confidence interval? Does this match what you expected based on your significance test? Explain.

The null hypothesis is that there is that the two population means are the same. Put another way, the null hypothesis is that there is no difference in the two population means. So the null hypothesis value for comparing two means is always 0.

Yes, 0 is in this interval so it is possible that the difference is 0. This means that there could be no difference overall for the average GPA of all students who took calculus in high school with the average GPA of all students who didn't take calculus in high school. This matches the results of the significance test which did not reject the null hypothesis.