1. A trainer, Sarah, is researching summer camps designed to help enhance endurance. One such camp submitted video footage of 9 different people on the first day of the camp versus the last day of the camp. In each video, the athlete is tasked with running as many laps as possible on a track, which is then recorded by Sarah. Using that data, Sarah tests whether a person is able to run more laps after the camp than before. n

$L_1 \rightarrow$ Before	4	6	8	5	4	7	6	3	5
$L_2 \rightarrow$ After	9	11	12	9	7	9	9	7	8

a. What type of question is this? Is it dealing with mean or proportion?

$L_3 : L_2 - L_1$ 2 sample dependent w/ mean

b. State the Hypothesis.

$H_0 : \mu_d = 0$ $H_1 : \mu_d > 0$

c. What is the direction?

Right-sided

d. Calculate the Test Statistic. 1-Var Stat on $L_3 : \bar{d} = 3.67 ; S_d = 1$

$$TS = \frac{\bar{d} - d_0}{S_d / \sqrt{n}} = \frac{3.67 - 0}{1 / \sqrt{9}} = 11.01$$

e. Calculate the p-value. What is the decision? $\alpha = 0.05$

$$p = tcdf(11.01, E99, 9-1) = 2.06 \times 10^{-6} \approx 0.0000 < 0.05$$

Reject H_0

f. Choose the correct interpretation for the decision.

- [x] At 5% level of significance, we have sufficient evidence to say that the difference of mean laps on the first day versus the last day differ.
- [ ] At 5% level of significance, we have insufficient evidence to say that the difference of mean laps on the first day versus the last day differ.

2. Using the same information from the previous question, find the confidence interval and interpret.

$PE = \bar{d} = 3.67$ $MOE = t_{\alpha/2} \frac{S_d}{\sqrt{n}} = \left| \underbrace{invT\left(\frac{0.05}{2}, 9-1\right)}_{2.31} \right| \frac{1}{\sqrt{9}} = 0.77$

$(3.67-0.77, 3.67+0.77)$
$= \boxed{(2.9, 4.44)}$

We are 90% sure that the true mean difference is between 2.9 & 4.44 kps.

3. With the knowledge that the confidence interval is (142.95, 325.43), find the point estimate and the margin of error.

L-CI U-CI

$$PE = \frac{L\text{-}CI + U\text{-}CI}{2} = \frac{142.95 + 325.43}{2} = 234.19$$

$$MOE = \frac{(U\text{-}CI) - (L\text{-}CI)}{2} = \frac{325.43 - 142.95}{2} = 91.24$$

4. Shelly saw that of the 425 people involved in a survey about a restaurant's quality, 305 people said they were satisfied with the quality. Solve the following questions with this information.

   a. What type of question is this? Is it dealing with mean or proportion?

   Confidence Interval w/ proportions

   b. Is the sample large enough?

   $n\hat{p} \geq 10$
   $\cancel{(425)}\left(\frac{305}{\cancel{425}}\right) \geq 10$
   $305 \geq 10$ ✓

   $n\hat{q} \geq 10$
   $\cancel{(425)}\left(\frac{425-305}{\cancel{425}}\right) \geq 10$
   $120 \geq 10$ ✓

   Yes

   c. What is the point estimate?

   $$PE = \hat{p} = \frac{305}{425} = 0.72$$

   d. What is the margin of error? $Z_{\alpha/2} = \left|invNorm\left(\frac{\alpha}{2}, 0, 1\right)\right|$

   $$MOE = Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = \underbrace{\cancel{\left|invNorm\left(\frac{0.01}{2}, 0, 1\right)\right|}}_{2.58} \times \sqrt{\frac{(0.72)(1-0.72)}{425}} = 0.06$$

   e. Construct a 99% confidence interval.

   $$(0.72-0.06, 0.72+0.06) = (0.66, 0.78)$$

   f. Fill in the blanks of the interpretation for the decision.

We are 99% confident that the true unknown population proportion lies in the interval (0.66, 0.72).

5. Veronica, a federal grant official, needs to decide between two programs for which to give additional funding. Program A has 124 students enrolled and based off of their retention rates, an average of 88 people is expected to return for their next year. Program B has 173 students enrolled with an average of 103 people expected to return for their next year, based off of their retention rates. Assume that the standard deviation is 23 for program A and 48 for program B. Determine whether the scenario would be labeled as pooled or non-pooled during hypothesis testing. Once you determine that, figure out whether the retention rates of program A is better than program B.

Q#1

Q#2

A: $n=124$, $\bar{x}=88$, $s=23$

B: $n=173$, $\bar{x}=103$, $s=48$

a. What type of question is this? Is it dealing with mean or proportion?

Q#1 = Bartletts w/mean Q#2 = 2 sample ind. w/mean

b. State the Hypothesis for the first part of the question. (Q#1)

$H_0: \sigma_B^2 = \sigma_A^2$ $H_1: \sigma_B^2 > \sigma_A^2$

B then A b/c $S_B > S_A$

c. Calculate the Test Statistic.

$$TS = \frac{S_B^2}{S_A^2} = \frac{48^2}{23^2} = 4.36$$

d. Calculate the p-value. What is the decision when the confidence level is 90%? → $\alpha = 0.1$

$$p = \text{fcdf}(4.36, E99, 173-1, 124-1) = 1.84 \times 10^{-16} \approx \underbrace{0.0000 < 0.1}_{\text{Reject } H_0}$$

e. Choose the correct interpretation.

☒ At 10% level of significance, we have sufficient evidence to say that the variances of programs A and B differ.

❑ At 10% level of significance, we have insufficient evidence to say that the variances of programs A and B differ.

f. Based off that decision, is the scenario pooled or non-pooled?

= ≠

Rejecting H_0 means rejecting $\sigma_A^2 = \sigma_B^2$ so its non-pooled

g. State the hypotheses for the second part of the question. (Q#2)

$H_0: \mu_A = \mu_B$ $H_1: \mu_A > \mu_B$

h. What is the direction?

Right-sided

2-SampTTest

i. Determine the test statistic and p-value. What is the decision? $\alpha = 0.1$ still

$TS = t = -3.58$ $p = \underbrace{0.9998 > 0.1}$

Fail to Reject H_0

j. Choose the correct interpretation.

- ❑ At 10% level of significance, we have sufficient evidence to say that the mean retention rate of program A is better than program B.
- ☒ At 10% level of significance, we have insufficient evidence to say that the mean retention rate of program A is better than program B.

6. Suppose that out of 93 people ordering drinks from a concession stand, only 5 of them ordered Canada Dry. Estimate the proportion of people that order Canada Dry each game.

a. What type of question is this? Is it dealing with mean or proportion?

Confidence Interval w/ proportion

b. Is the sample large enough?

$n\hat{p} \geq 10$

$\cancel{(93)}\left(\frac{5}{\cancel{93}}\right) \geq 10$

$5 \geq 10$ ✗

No → Wilson's

c. What is the point estimate?

$$PE = \tilde{p} = \frac{X+2}{n+4} = \frac{5+2}{93+4} = \frac{7}{97} = 0.07$$

d. What is the margin of error? $Z_{\alpha/2} = \left|invNorm\left(\frac{\alpha}{2}, 0, 1\right)\right|$

$$MOE = Z_{\alpha/2}\sqrt{\frac{\tilde{p}\tilde{q}}{n+4}} = \underbrace{\cancel{\left|invNorm\left(\frac{0.1}{2}, 0, 1\right)\right|}}_{1.64} \times \sqrt{\frac{(0.07)(1-0.07)}{93+4}} = 0.04$$

e. Construct a 90% confidence interval. $\alpha = 0.1$

$(0.07-0.04, 0.07+0.04) = (0.03, 0.11)$

f. Fill in the blanks of the interpretation for the decision.

We are 90% sure that the unknown population proportion of people that order Canada Dry is in the interval (0.03, 0.11).

7. Seniors in high school have heard repeatedly that applying to multiple colleges is recommended because the average rate of acceptance to colleges are 68% in the United States. Assuming that collected data from a survey shows 252 acceptances when 534 seniors apply to the same college, test whether the true rate of acceptance is less than 68%.

a. What type of question is this? Is it dealing with mean or proportion?

1 sample w/proportion

b. State the Hypothesis.

$H_0: p = 0.68 \qquad H_1: p < 0.68$

c. What is the direction?

Left-sided

d. Calculate the test statistic. $\hat{p} = \frac{252}{534} = 0.47$

$$TS = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.47 - 0.68}{\sqrt{\frac{(0.68)(1-0.68)}{534}}} = -10.40$$

e. Calculate the p-value. What is the decision? $\alpha = 0.05$

$p = \text{normcdf}(-E99, -10.40, 0, 1) = \underbrace{0 < 0.05}_{\text{Reject } H_0}$

f. Choose the correct interpretation for the decision.

☒ At 5% level of significance, we have sufficient evidence to say that the proportion of seniors getting accepted into college in the US is less than 68%

- [ ] At 5% level of significance, we have insufficient evidence to say that the proportion of seniors getting accepted into college in the US is less than 68%

8. A publishing company wants to determine whether the reviews of an action fantasy novel with a male lead or a novel with that same plot but with a female lead is different. To test this, 30 people are assigned to each book of which they will rate on a scale from 1 to 5 in satisfaction. Using the results below, determine whether the male lead novel preforms differently from the female lead novel. Assume that the variance is equal with a 1% significance level.

L_1

Male Lead Novel: 3, 2, 4, 5, 1, 3, 2, 4, 4, 5, 2, 1, 4, 5, 3, 2, 2, 1, 3, 4, 5, 3, 4, 2, 4, 5, 2, 4, 2, 3

Female Lead Novel: 4, 5, 3, 3, 5, 3, 2, 4, 3, 5, 2, 1, 4, 1, 3, 4, 5, 3, 2, 4, 5, 3, 4, 2, 5, 1, 3, 4, 2, 3

L_2

a. What type of question is this? Is it dealing with mean or proportion?

2 sample independent w/mean

b. State the hypothesis.

$H_0: \mu_{ML} = \mu_{FL}$ $\quad H_1: \mu_{ML} \neq \mu_{FL}$

c. What is the direction?

Two-sided

d. Is this scenario pooled or non-pooled?

Told variances are equal → pooled

e. Determine the test statistic and p-value. 2-SampTTest w/ data

$TS = t = -0.41$ $\quad p = 0.6854$

f. What is the decision? $\alpha = 0.01$

$0.6854 > 0.01$ → Fail to Reject H_0

g. Choose the correct interpretation.

- [ ] At 1% level of significance, we have sufficient evidence to say that the mean score for novels with a male lead versus with a female lead are different.
- [x] At 1% level of significance, we have insufficient evidence to say that the mean score for novels with a male lead versus with a female lead are different.

9. A group of students are conducting an experiment on how quickly a dog will perform a task after a release command is given. Using the data given below, answer the questions below.

$\bar{x} = 17.69$ & $s = 9.87$ ← 1-Var Stat

Rep.	1	2	3	4	5	6	7	8
Time	36.39	21.13	23.18	19.43	14.74	13.74	8.64	4.23

← L_1

a. What type of question is this? Is it dealing with mean or proportion?

Confidence Interval w/ mean

b. Is the sample large enough?

$n > 30$ No → t-distribution

c. What is the point estimate?

$PE = \bar{x} = 17.69$

d. What is the margin of error? $t_{\alpha/2} = \left| invT\left(\frac{\alpha}{2}, df\right) \right|$

$MOE = t_{\alpha/2} \frac{s}{\sqrt{n}} = \underbrace{\left| invT\left(\frac{0.01}{2}, 8-1\right) \right|}_{3.50} \frac{9.87}{\sqrt{8}} = 12.21$

e. Construct a 99% confidence interval. $\alpha = 0.01$

$(17.69 - 12.21, 17.69 + 12.21) = (5.48, 29.9)$

f. Fill in the blanks of the interpretation for the decision.

We are 99% confident that the true unknown population mean lies between 5.48 and 29.9.

10. A teacher asks her class of 23 to record the exact amount of time it took for them to complete an essay for a bet she has with a few other teachers. Each teacher is betting that their class takes an average of 3.5 hours to complete an essay from

start to finish. Using the data below, determine whether the class of 23 takes longer than the agreed average time to complete an essay.

1	2	3	4	5	6	7	8	9
3.15	4.05	1.50	6.32	2.27	5	1.57	4.01	4
10	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
1.59	4.07	2	3.11	5.25	3	6.08	4.27	3
19	**20**	**21**	**22**	**23**				
6	3.31	2.14	5.21	4.23				

← 4 (all of white)

1-Var Stat

$\bar{x} = 3.70$

$s = 1.48$

a. What type of question is this? Is it dealing with mean or proportion?

1 sample w/ mean

b. State the Hypothesis.

$H_0: \mu = 3.5$ $\quad H_1: \mu > 3.5$

c. What is the direction?

Right-sided

d. Calculate the test statistic.

$$TS = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{3.70 - 3.5}{1.48/\sqrt{23}} = 0.65$$

e. Calculate the p-value. What is the decision with a 90% confidence level?

$\hookrightarrow \alpha = 0.1$

$p = \text{tcdf}(0.65, E99, 23-1) = 0.2612 > 0.1$

fails to Reject H_0

f. Choose the correct interpretation for the decision.

- ☐ At 10% level of significance, we have sufficient evidence to say that the average time spent completing an essay is more than 3.5 hours.
- ☒ At 10% level of significance, we have insufficient evidence to say that the average time spent completing an essay is more than 3.5 hours.

11. A professor is trying to find the optimal range a student should study before a test. To do so, the professor sends out surveys to every student that takes the course and maintains at least a B in the class, which amounts to 59 students. The survey

reveals that an average of 43 hours with a standard deviation of 2.5 hours are spent studying leading up to the exam.

a. What type of question is this? Is it dealing with mean or proportion?

Confidence Interval w/mean

b. Is the sample large enough?

$n = 59 \rightarrow n \geq 30$ Yes $\rightarrow$ normal dist.

c. What is the point estimate?

$PE = \bar{x} = 43$

d. What is the margin of error? $Z_{\alpha/2} = |\text{invNorm}(\frac{\alpha}{2}, 0, 1)|$

$$MOE = Z_{\alpha/2}\frac{s}{\sqrt{n}} = \underbrace{\cancel{\left|\text{invNorm}\left(\frac{0.01}{2}, 0, 1\right)\right|}}_{2.58}\frac{2.5}{\sqrt{59}} = 0.84$$

e. Construct a 99% confidence interval. $\alpha = 0.01$

$(43 - 0.84, 43 + 0.84) = (42.16, 43.84)$

f. Fill in the blanks of the interpretation for the decision.

We are 99% sure that the population mean lies in the interval (42.16, 43.84).