CREDIT: The questions on this document were written by Erik Packard, PhD, Associate Professor of Mathematics at Colorado Mesa University.

1. (18 pts) On the website www.listsofjohn.com it is reported that there are 59,817 ranked peaks in Alaska. I took a random sample of 8 peaks and found a sample mean elevation of 4029.1 feet and a sample standard deviation of 2329.8 feet. Give a 95% confidence interval for the mean elevation of all 59,817 peaks.

$$95\%\text{ CI} = 4029.1\text{ ft} \pm (1.960)\frac{2329.8\text{ ft}}{\sqrt{8}} \quad\longrightarrow\quad 2414.63\text{ ft} < \mu < 5643.57\text{ ft}$$

2. (19 pts) Data given are 4 pairs of games CMU men's basketball played home and away against the same teams. At the 1% significance level can we conclude any difference points scored on average at home and away for all possible games CMU men's basketball team could play?

Team Played	**NM Highlands**	**Adams St**	**Western**	**Ft. Lewis**
Points scored by CMU at home	96	83	86	89
Points scored by CMU as visitor	84	90	78	80

$H_0 : \mu_{\text{Home}} = \mu_{\text{Visitor}} \quad\longrightarrow\quad \mu_{\text{Home}} - \mu_{\text{Visitor}} = 0$

$H_a : \mu_{\text{Home}} \neq \mu_{\text{Visitor}} \quad\longrightarrow\quad \mu_{\text{Home}} - \mu_{\text{Visitor}} \neq 0$

Critical Value(s): Given that $\alpha = 0.01$ and $\text{df} = 3$, $t_{\text{Table}} = \pm 5.841$

Test Statistic:

$$t_{\text{Data}} = \frac{(\overline{x}_{\text{Home}} - \overline{x}_{\text{Visitor}}) - 0}{\frac{(s_{\text{Home}} - s_{\text{Visitor}})}{\sqrt{n}}} = \frac{5.5 - 0}{\frac{8.50}{\sqrt{4}}} = 1.29$$

Yes/No Answer: Since $t_{\text{Data}} < t_{\text{Table}}$, No.

3. (12 pts) Suppose somebody else with different data for the previous problem and gets a test statistic of $t = 1.250$. Give their p-value and say what it means in everyday terms.

Given that $t_{\text{Data}} = 1.250$ and $\text{df} = 3$, $p\text{-value} = 2(0.15) = 0.30$

If there's no difference between the mean number of points scored at home and the mean number of points scored away for all possible games CMU men's basketball team could play, then the chance of finding as strong or stronger evidence suggesting that there is a difference between the mean number of points scored at home and the mean number of points scored away for all possible games CMU men's basketball team could play is 0.30.

4. (19 pts) Data was collected from samples from the website www.listsofjohn.com for how many people had climbed peaks in Mesa County and Garfield County. At the 5% significance level can we conclude that Mesa County peaks are climbed more on average than Garfield County peaks on this website?

	Sample Mean	**Sample Standard Deviation**	**Sample Size**
Garfield County	5.4	3.050	5
Mesa County	8.0	8.188	8

$H_0 : \mu_{\text{Mesa County}} \leq \mu_{\text{Garfield County}} \quad \longrightarrow \quad \mu_{\text{Mesa County}} - \mu_{\text{Garfield County}} \leq 0$

$H_a : \mu_{\text{Mesa County}} > \mu_{\text{Garfield County}} \quad \longrightarrow \quad \mu_{\text{Mesa County}} - \mu_{\text{Garfield County}} > 0$

Critical Value(s): Given that $\alpha = 0.05$ and $\text{df} = 4$, $t_{\text{Table}} = 2.132$

Test Statistic:

$$t_{\text{Data}} = \frac{\left(\overline{x}_{\text{Mesa County}} - \overline{x}_{\text{Garfield County}}\right) - 0}{\sqrt{\dfrac{\left(s^2\right)_{\text{Mesa County}}}{n_{\text{Mesa County}}} + \dfrac{\left(s^2\right)_{\text{Garfield County}}}{n_{\text{Garfield County}}}}} = \frac{(8.0 - 5.4) - 0}{\sqrt{\dfrac{(8.188)^2}{8} + \dfrac{(3.050)^2}{5}}} = 0.81$$

Yes/No Answer: Since $t_{\text{Data}} < t_{\text{Table}}$, No.

5. (8 pts) Suppose somebody else collects data for #4, what is the chance they won't conclude Mesa Peaks are more climbed when in fact they are?

Unknown.

6. (8 pts) Suppose somebody else collects data for #4, what is the chance they will conclude Mesa Peaks are climbed more when in fact they are not?

5% or less.

7. (21 pts) So far this year CMU baseball player Keenan Eaton has got 17 hits in 42 attempts. Treat this as a random sample and give a 95% confidence interval for the percentage of all attempts that will result in hits (i.e., what his batting average will be at the end of the season).

$$95\%\ \text{CI} = \frac{17}{42} \pm (1.960)\sqrt{\frac{(17/42)(25/42)}{42}} \quad \longrightarrow \quad -0.57 < \hat{p} < 1.37$$

8. (17 pts) So far this year CMU baseball players Keenan Eaton has got 17 hits in 42 attempts and Reagan Todd has got 11 hits in 38 attempts. At the 10% significance level, can we conclude any difference in their hitting percentage for all possible attempts?

$H_0 : p_{\text{Keenan Eaton}} = p_{\text{Reagan Todd}} \quad \longrightarrow \quad p_{\text{Keenan Eaton}} - p_{\text{Reagan Todd}} = 0$

$H_a : p_{\text{Keenan Eaton}} \neq p_{\text{Reagan Todd}} \quad \longrightarrow \quad p_{\text{Keenan Eaton}} - p_{\text{Reagan Todd}} \neq 0$

Critical Value(s): Given that $\alpha = 0.10$, $z_{\text{Table}} = \pm 1.645$

Test Statistic:

$$z_{\text{Data}} = \frac{\left(\hat{p}_{\text{Keenan Eaton}} - \hat{p}_{\text{Reagan Todd}}\right) - 0}{\sqrt{\dfrac{\hat{p}_{\text{Pool}}\,\hat{q}_{\text{Pool}}}{n_{\text{Keenan Eaton}}} + \dfrac{\hat{p}_{\text{Pool}}\,\hat{q}_{\text{Pool}}}{n_{\text{Reagan Todd}}}}} \quad \text{where} \quad \hat{p}_{\text{Pool}} = \frac{x_{\text{Keenan Eaton}} + x_{\text{Reagan Todd}}}{n_{\text{Keenan Eaton}} + n_{\text{Reagan Todd}}}$$

$$\hat{p}_{\text{Pool}} = \frac{(17+11)}{(42+38)} = \frac{7}{20} \quad \therefore \quad \hat{q}_{\text{Pool}} = 1 - \hat{p}_{\text{Pool}} = \frac{13}{20} \quad \therefore$$

$$z_{\text{Data}} = \frac{\left(\dfrac{17}{42} - \dfrac{11}{38}\right) - 0}{\sqrt{\dfrac{(7/20)(13/20)}{42} + \dfrac{(7/20)(13/20)}{38}}} = 1.08$$

Yes/No Answer: Since $z_{\text{Data}} < z_{\text{Table}}$, No.

The rest are worth 1 point each.

9. If H_0 is true what is the probability that you will reject it by mistake?

α

10. If H_0 is not true what is the probability that you mistakenly not reject it?

Unknown.

11. If you mistakenly don't reject H_0, what type of error is it?

Type II Error.

12. What is the chance of making a type I error?

α

13. What is the notation for the total area of the rejection region?

α

14. The edge(s) of the rejection region(s) are called what?

    Critical Values.

15. Are the critical value(s) found by a table or calculation in this class?

    Table.

16. Is the test statistic found by a table or calculation?

    Calculation.

17. What is the area under a *t* curve?

    One.

18. To figure the sample size, *n*, needed for a CI for a proportion, if you have a reasonable value for *p*' and use it then your CI may have a margin of error a little too big, but your sample size will be smaller making collecting the data easier.