**Name:** *Caleb McWhorter — Solutions*
**MAT 222**
**Fall 2019**
**Homework 5**

*"Brian! There's a message in my cereal. It says 'Ooooooooo'. [Brian] Peter, those are Cheerios."*

*–Peter & Brian Griffin, Family Guy*

**Problem 1:** A random sample of $n$ observations is selected from a normal population to test the null hypothesis that $H_0 : \mu = 98.9$. For each of the following alternative hypotheses below, specify for which $t$-values you would reject the null hypothesis.

(a) $H_a : \mu \neq 98.9, \quad \alpha = 0.05, \quad n = 12,$ $|t| > 2.201$

(b) $H_a : \mu > 98.9, \quad \alpha = 0.10, \quad n = 14,$ $t > 1.350$

(c) $H_a : \mu < 98.9, \quad \alpha = 0.01, \quad n = 30,$ $t < -2.462$

(d) $H_a : \mu > 98.9, \quad \alpha = 0.05, \quad n = 91,$ $t > 1.664$

(e) $H_a : \mu \neq 98.9, \quad \alpha = 0.01, \quad n = 23,$ $|t| > 2.819$

**Problem 2:** What does it mean for a statistical inference to be robust? Are $t$-procedures robust?

***Solution.*** *A statistic is robust if violations of the assumptions required for probability calculations are not 'greatly' affected by these violations. Yes, $t$-procedures are robust against non-normal populations.*

**Problem 3:** Assessment of highway conditions is important for their maintenance. One way of measuring the safety of highways is to measure crack intensity—a measurement of the average number of cracks per some fixed distance of road. A local government sends out a crew to measure the crack intensity in a local strip of highway. They took 15 samples of 50-meter stretches of the highway and found a mean crack intensity of 0.129 with standard deviation 0.23.

(a) Construct a 95% confidence interval for the crack intensity of this highway. Interpret your result.

*As long as the sample is not skewed, a $t$-procedure is appropriate. We have $n = 15$, $\overline{x} = 0.129$, and $s = 0.23$. We have degrees of freedom $15 - 1 = 14$. Then with a 95% confidence interval, we have $t^* = 2.145$. Then we have*

$$\overline{x} \pm t^* \frac{s}{\sqrt{n}} = 0.129 \pm 2.145 \cdot \frac{0.23}{\sqrt{15}} = 0.129 \pm 0.127$$

*so that we have a 95% confidence interval of $(0.002, 0.256)$. Therefore, we are 95% confident that the mean number of cracks in this highway is between 0.002 and 0.256 cracks per 50 m.*

(b) The American Association of State Highway and Transportation Officials (AASHTO) recommends a maximum crack intensity of 0.100 before repairs. Using an appropriate null and alternative hypothesis and a significance level of 5%, determine if this highway exceeds this standard.

*We have*

$$\begin{cases} H_0 : \mu = 0.100 \\ H_a : \mu > 0.100 \end{cases}$$

*The test statistic is*

$$t = \frac{0.129 - 0.100}{0.23/\sqrt{15}} = \frac{0.029}{0.05939} = 0.488 \overset{\text{dof } 14}{\rightsquigarrow} p > 0.25$$

*The* $p$*-value is greater than* $0.25 > \alpha = 0.05$*. Therefore, we fail to reject the null hypothesis. There is not enough evidence to suggest that this highway exceeds the AASHTO recommended crack intensity standards, i.e. the data is consistent with the null hypothesis.*

**Problem 4:** A study of business students was performed to determine if students felt having a printed version of lecture notes was (or would be) helpful in understanding the material. Eighty-six students from a class with lecture notes and thirty-five students from a class without such notes were surveyed. Responses were measured on a scale of 1–9 with 1 being "strongly disagree" and 9 being "strongly agree." The data is summarized below.[1]

| | $n$ | $\overline{x}$ | $s$ |
|---|---|---|---|
| Notes | 86 | 8.48 | 0.94 |
| No Notes | 35 | 7.80 | 2.99 |

(a) Is a pooled two-sample $t$-test appropriate? Why or why not?

*We have* $\dfrac{s_{NL}}{s_L} = \dfrac{2.99}{0.94} = 3.18$*, which is greater than 2 so that a pooled* $t$*-procedure would not be appropriate.*

(b) Construct a 99% confidence interval for the difference in opinions between these two groups. Interpret your results.

*We have* $n_{NL} = 35$*,* $\overline{x}_{NL} = 7.80$*,* $s_{NL} = 2.99$ *and* $n_L = 86$*,* $\overline{x}_L = 8.48$*, and* $s_L = 0.94$*. We have degrees of freedom* $\min\{35-1, 86-1\} = 34 \rightsquigarrow 30$*. For a 99% confidence interval with degrees of freedom 34, we have* $t^* = 2.750$*. Then*

$$(8.48 - 7.80) \pm 2.750\sqrt{\frac{0.94^2}{86} + \frac{2.99^2}{35}} = 0.68 \pm 2.750(0.515) = 0.68 \pm 1.42$$

*so that we have confidence interval* $(-0.74, 2.1)$*. Therefore, we are 99% certain that students receiving printed lecture notes rated the usefulness of the notes on average 0.74 less useful to 2.1 more useful than those that did not receive the notes.*

[1] Gray, J.I., and Abernathy, A.M. "Pros and cons of lecture notes and handout packages: Faculty and student opinions." *Marketing Education Review*, Vol 4, No. 3, p.25, 1984.

(c) Use an appropriate hypothesis test with significance level 1% to determine if there is any significant difference between the mean response rates. Interpret your results.

*We test the hypothesis*

$$\begin{cases} H_0 : \mu_{NL} = \mu_L \\ H_a : \mu_{NL} \neq \mu_L \end{cases}$$

*From the previous part, we know that the degrees of freedom is* $34 \rightsquigarrow 30$. *We have* $\alpha = 0.01$ *and test statistic*

$$t = \frac{8.48 - 7.80}{\sqrt{\dfrac{0.94^2}{86} + \dfrac{2.99^2}{35}}} = \frac{0.68}{0.515} = 1.32 \overset{\textit{dof } 30}{\rightsquigarrow} 0.05 < p < 0.10$$

*Using this two-sided test, this gives* $p$*-value* $0.10 < p < 0.20$. *Therefore because* $p > \alpha = 0.01$, *we fail to reject the null hypothesis. There does not seem to be a perceived usefulness difference in the notes between students that were given printed lecture notes and those that did not.*

(d) Explain how you could have used (b) to answer (c).

*Because we are performing a two-sided test, a confidence interval can be used to perform the hypothesis test. If there was no difference, 0 would be in the confidence interval corresponding to our significance level. Observe that 0 is in the interval* $(-0.74, 2.1)$ *so that we would fail to reject the null hypothesis.*

(e) What if the researchers had a reason to believe the difference in the average response was 1? Recompute the test statistic in (c) using this assumption.

*Technically in (c), we had* $t = \dfrac{(8.48 - 7.80) - 0}{\sqrt{\dfrac{0.94^2}{86} + \dfrac{2.99^2}{35}}}$. *If we instead assume that the difference is 1 instead of 0, we have*

$$t = \frac{(8.48 - 7.80) - 1}{\sqrt{\dfrac{0.94^2}{86} + \dfrac{2.99^2}{35}}} = \frac{-0.32}{0.515} = -0.621$$

**Problem 5:** Which is more important for $t$-procedures: that the sample be normal or have no skewness or that the samples be SRS?

*It is more important that the samples be simple random samples (SRS).*

**Problem 6:** Must degrees of freedom (especially in a computer system) for two-sample $t$-procedures always be integers?

*No. Our methods are approximations. There are many better methods that give more appropriate degrees of freedom which do not give integer values, e.g. the Satterthwaite approximation.*

**Problem 7:** Most people believe that there is a difference between male and female heights, on average. The traditional 'wisdom' is that men are, on average, taller than women. To test this, you take a simple random sample of 23 men and 27 women. You find the men have an average height of 178.4 cm and the women have an average height of 164 cm with standard deviations 7.59 cm and 7 cm, respectively. Use an appropriate $t$-procedure to test the hypothesis that, on average, men are taller than women using a significance level of 0.1%.

*Let $M$ denote male and $F$ denote female. We are using $\alpha = 0.001$. We choose null and alternative hypotheses*

$$\begin{cases} H_0 : \mu_M - \mu_F = 0 \\ H_a : \mu_M - \mu_F > 0 \end{cases}$$

*Observe a pooled $t$-procedure is appropriate because $7.59/7 = 1.08 < 2$. We have*

$$s_p^2 = \frac{(23-1)7.59^2 + (27-1)7.0^2}{23+27-2} = \frac{2541.38}{48} = 52.9454$$

*Therefore, $s_p = \sqrt{52.9454} = 7.27$. We have degrees of freedom $23 + 27 - 2 = 48$. The test statistic is*

$$t = \frac{178.4 - 164}{7.27\sqrt{\dfrac{1}{23} + \dfrac{1}{27}}} = \frac{14.4}{2.06} = 6.99 \overset{\textit{dof } 48}{\rightsquigarrow} 0$$

*Then the $p$-value is $p \approx 0 < \alpha = 0.001$. Therefore, there is sufficient evidence to reject the null hypothesis in favor of the alternative, $\mu_M - \mu_F > 0$, i.e. $\mu_M > \mu_F$, i.e. there is sufficient to suggest that, on average, males are taller than females.*