STAT 3202: Homework 08

Spring 2019, OSU

Due: Monday, April 8 (During Lab)

Please see the **detailed homework policy document** for information about homework formatting, submission, and grading.

Exercise 1

Four chemical plants, producing the same products and owned by the same company, discharge effluents into streams in the vicinity of their locations. To monitor the extent of pollution created by the effluents and to determine whether this differs from plant to plant, the company collected random samples of liquid waste, five specimens from each plant. The data are given in the accompanying table.

Plant	Polluting Effluents (lb/gal of waste)				
A	1.65	1.72	1.50	1.37	1.60
B	1.70	1.85	1.46	2.05	1.80
C	1.40	1.75	1.38	1.65	1.55
D	2.10	1.95	1.65	1.88	2.00

Do the data provide sufficient evidence to indicate a difference in the mean weight of effluents per gallon in the effluents discharged from the four plants? Test using $\alpha = 0.05$. You may use `R` to do this.

Your answer should include:

- The null and alternative hypotheses you are testing.
- The full ANOVA table filled in – i.e., a table formatted like the one below, with values in all starred entries:

Source	df	SS	MS	F-test statistic	p-value
Treatment	**	**	**	**	**
Error	**	**	**		
Total	**	**			

- A one sentence conclusion in words.

Exercise 2

Do SAT scores for high school students differ depending on the students' intended field of study? A study compared 15 students who intended to major in engineering with 15 students who intended to major in language and literature. Given in the accompanying table are the means and standard deviations of the scores on the verbal and mathematics portion of the SAT for the two groups of students:

	Verbal	Math
Engineering	$\bar{y} = 446$, $s = 42$	$\bar{y} = 548$, $s = 57$
Language/literature	$\bar{y} = 534$, $s = 45$	$\bar{y} = 517$, $s = 52$

Perform the ANOVA overall F-test to determine whether there is sufficient evidence to claim a difference in the mean verbal SAT scores between high school students who intend to major in engineering and and those who intend to major in language/literature. Report the:

- The value of the **test statistic** for the test.
- The p-**value** for the test. (You should find the p-value using R.)

Exercise 3

Return to the SAT data in Exercise 2. Perform a two-sample t-test to determine whether there is sufficient evidence to claim a difference in the mean verbal SAT scores for high school students who intend to major in engineering and language/literature. Report:

- The value of the **test statistic** for the test.
- The p-**value** for the test. (You should find the p-value using R.)
- A comparison of the p-values in Exercise 2 and 3.

Note that it is a mathematical/statistical fact that:

$$\text{If } X \sim t_{\text{df}} \text{ then } Y = X^2 \sim F_{1,\text{df}}$$

So in particular, if you square the test statistic you find in this exercise, it should equal the test statistic you found in Exercise 2.

Exercise 4

A researcher evaluated maneuver times for vehicles of various sizes that were involved in making a left turn at an intersection with a separate left-turn lane but without a separate left-turn phase on the traffic light governing the intersection (an "unprotected" left-turn maneuver). The maneuver time was measured from the instant that a vehicle entered the opposing lanes of traffic until it completely cleared the intersection. Four-cylinder automobiles were classified as "small cars" and six- or eight-cylinder automobiles as "large cars." Trucks and buses were combined to form a third category identified as "truck or bus." Other motorized vehicles (motorcycles, etc.) were ignored in the study. A summary of the data, giving maneuver times (in seconds) for vehicles that attempted the left-turn maneuver from a standing stop, appears in the accompanying table.

Vehicle Type	Sample Size	Mean	Standard Deviation
Small Car	45	4.59	0.70
Large Car	102	4.88	0.64
Truck or bus	18	6.24	0.90

Is there sufficient evidence to claim that the mean maneuver times differ for the three vehicle types? Your answer should include:

- A completed ANOVA table, which includes the p-value of the test.

Exercise 5

Return to the vehicle data in Exercise 4. Suppose the researcher wants to test whether the difference in maneuver time between small and large cars was statistically significant. Enumerate the groups in the order

of the table. (That is μ_1 is the mean for "small cars" and μ_2 is the mean for "large cars.") The researcher is testing:

$$H_0 : \mu_1 - \mu_2 = 0 \text{ vs. } H_A : \mu_1 - \mu_2 \neq 0$$

Using the fact that, when σ^2 is estimated using the MSE,

$$\frac{\bar{y}_1 - \bar{y}_2}{\text{SE}\left[\bar{y}_1 - \bar{y}_2\right]} \sim t_{n-K},$$

perform the researcher's hypothesis test at $\alpha = 0.01$. Report:

- The value of the **test statistic** for the test.
- The p**-value** for the test. (You should find the p-value using R.)
- A **decision**. (i.e., accept or reject the null of no difference between small and large cars, at level $\alpha = 0.01$.)

Along the way, you will need to:

- Derive an expression for the variance of the estimator $\bar{y}_1 - \bar{y}_2$.
- Use the MSE from ANOVA to estimate the population variance σ^2 of the Y_{ki}, then use this to estimate the standard deviation of $\bar{y}_1 - \bar{y}_2$, also known as the standard error.
