ST517 Sample Midterm 2—Solutions

1. True
2. False—"was made" vs. "could have been made"
3. False
4. D
5. C
6. C
7. A
8. B
9. D
10. A
11. C
12. C
13. B
14. A
15. B

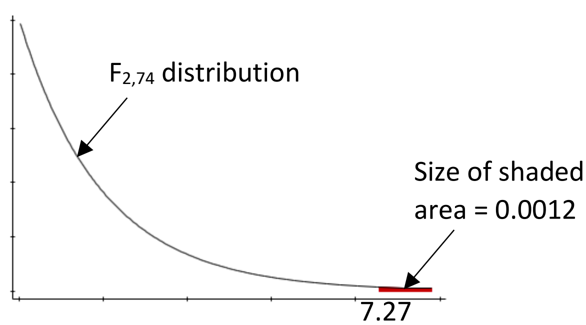
16. A Type I Error would be if the defendant was found guilty when in fact they were innocent.

17.
   a. 3—shelves = groups, so number of shelves = df_{groups} +1 = 2 + 1
   b. 77—n = df_{total} +1 = 76 + 1
   c.

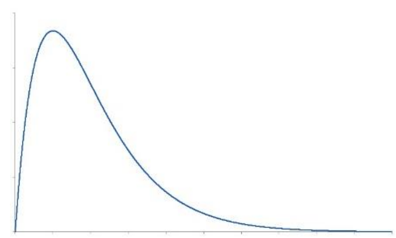
	DF	SS	MS	F	p-value
Shelf	2	**244.2**	122.1	**7.27**	0.0012
Error	**74**	**1242.5**	**16.8**		
Total	76	1486.7			

df_{error} = 76 – 2 OR df_{error} = 77 – 3
$SS_{shelf} = SS_{groups}$ = 122.1 * 2
SS_{error} = SSE = 1486.7 – 244.2
MS_{error} = MSE = 1242.5 / 74
F = 122.1 / 16.8

d.



Note: it is ok if you draw a right-skewed distribution that has more of a "hill" shape; just make sure you do not draw the F-distribution (or the Chi-square distribution) as bell-shaped!



e. Yes, it would be appropriate to conduct a multiple comparisons procedure since the ANOVA F-test is statistically significant, providing evidence that at least one of the population means is different from the rest.
There is evidence that the true average sugar content is higher for shelf 2 than either shelves 1 or 3. Based on the confidence intervals for $\mu_2 - \mu_3$ and $\mu_2 - \mu_1$, all of the reasonable estimates for these parameters are positive (which indicates that $\mu_2 > \mu_3$ and $\mu_2 > \mu_1$). True average sugar content for shelves 1 and 3 do not appear to be significantly different as the zero is a reasonable estimate for $\mu_3 - \mu_1$.

18. $H_0: p = 0.5$ vs. $H_A: p > 0.5$

Since we are interested in learning about the proportion for a single population, we would like to use the large-sample z-test.

Fortunately, the conditions for using this test are met: we have a large sample since $np_0 = 50$ and $n(1 - p_0) = 50$ are both greater than 10 and we are told the sample was randomly selected.

The test statistic is:

$$z = \frac{0.57 - 0.5}{\sqrt{\frac{(0.5)(1 - 0.5)}{100}}} = 1.40$$

This tells us that the sample proportion was 1.4 standard errors above the null value of 0.5.

The p-value (0.081) represents the proportion of the standard normal distribution that is greater than or equal to 1.4. This probability is relatively large (greater than 5%), so we cannot reject the null hypothesis.

Thus there is not enough evidence that the majority of students at this university are in favor of the new garage.

19. $H_0: \mu_1 - \mu_2 = 0$ vs. $H_A: \mu_1 - \mu_2 \neq 0$

Since we are interested in learning about the difference in means, we would like to use the 2-sample t-test.

We have two large, random, & independent samples, and it is not reasonable to assume the population variances are equal, so the test statistic is:

$$t = \frac{(6.5 - 4.8) - 0}{\sqrt{\frac{33.64}{235} + \frac{18.49}{340}}} = 3.825$$

The difference in sample means was 3.825 standard errors above what was expected under the null.

The null distribution is a t-distribution with 234 degrees of freedom.

There is a 0.02% chance we would have observed a difference in sample means that was 1.7 or more extreme if in fact there were no difference in the population means.

We reject H_0 (at the 5% level) and conclude that there is evidence of a statistically significant difference in the amount of exercise times for males and females.

$$(6.5 - 4.8) \pm 1.96\sqrt{\frac{33.64}{235} + \frac{18.49}{340}} \Rightarrow 1.7 \pm 0.871 \Rightarrow (0.829, 2.571)$$

We are 95% confident that the true mean amount of time exercising for males is between 0.829 hours and 2.571 hours* higher than the true mean amount of time exercising for females. *Note: this corresponds to between 49.74 and 154.26 minutes.

Based on both the statistically significance results of the test and the estimates in the confidence interval, there would appear to be a meaningful difference between the amount of time men spend exercising compared to women.

Note:

- The above statement considered additional time from (almost) 1 hour to 2.5 hours per week to be meaningful in the context of exercise. If you disagree with this, you could have said something like: "Even though the statistically significant results of the hypothesis test indicate there is a difference, the estimates from the confidence interval do not appear to be meaningful in context, as differences that could be only 1 hour a week more are not that much."
- Or maybe you are not familiar enough with the context to judge how meaningful the values are, so you could have said something like: "The statistically significant results of the hypothesis test indicate there is a difference; we should consult a context expert to determine how meaningful a difference that could range anywhere from (almost) 1 hour to 2.5 hours per week of exercise is."
- The key for this final statement is to consider both the results of the hypothesis test and the values in the confidence interval <u>in context</u>.