Stat 204 Exam 2 ANSWERS Due 4/16/20

1. The book's **Atlanta Commute (distance)** dataset provides the commute distances in miles for 500 randomly chosen people who work in the Atlanta metropolitan area.

a. Use *StatKey* to obtain the mean and standard deviation of the commute distances in this sample.

The sample mean is $\bar{x} = 18.156$ **miles, and the sample's standard deviation is** $s = 13.798$ **miles.**

b. Use *StatKey* to create 5000 bootstrap samples and a bootstrap distribution of bootstrap sample means. What are the mean and standard error of your bootstrap distribution?

This will vary a bit from simulation to simulation; the bootstrap distribution I created had an average of 18.166 **miles and a standard error of** 0.601 **miles.**

c. What does the formula in Section 6.2 say the standard error should be? How does this theoretical result compare to the standard error of your bootstrap distribution?

Using the formula,

$$\mathbf{SE} = \frac{s}{\sqrt{n}} = \frac{13.798}{\sqrt{500}} = 0.617\,,$$

which is very close to the bootstrap distribution's SE.

d. Use *StatKey* and your bootstrap distribution to find 90%, 95%, and 99% confidence intervals for the average commute distance of workers in the Atlanta metropolitan area. Determine the margin of error for each of these confidence intervals. What happens to the margin of error as the confidence level increases?

Again, these will vary a bit from simulation to simulation; here are the intervals for my bootstrap distribution:

$$90\% \;:\; 17.181 \le \mu \le 19.149\;,\; \mathbf{MOE} \approx \frac{19.149 - 17.181}{2} = 0.984\,,$$

$$95\% \;:\; 17.020 \le \mu \le 19.321\;,\; \mathbf{MOE} \approx \frac{19.321 - 17.020}{2} = 1.1505\,,$$

$$99\% \;:\; 16.67 \le \mu \le 19.791\;,\; \mathbf{MOE} \approx \frac{19.791 - 16.67}{2} = 1.5605\,.$$

As the confidence level increases, the margin of error increases.

e. Someone claims that the average commute distance for workers in the Atlanta metropolitan area is 20 miles. State the corresponding null and alternative hypotheses, compute the relevant t-statistic using this sample's average and standard deviation, and use *StatKey* to compute the p-value. What do you conclude about this person's claim?

The hypotheses are

$$H_0 \,:\, \mu = 20 \quad \textbf{versus} \quad H_a \,:\, \mu < 20\,,$$

so the t-statistic and p-value are

$$t \;=\; \frac{18.156 - 20}{0.617} \;=\; -\,2.988 \approx -3\,, \quad p \;=\; .0015\,.$$

We therefore have evidence against the person's claim; we have evidence that the average commute distance is less than 20 miles. Note: a 2-sided alternative hypothesis is fine, too, and would lead to a p-value of $2 \times .0015 = .0030$, same conclusion.

f. In the previous question, you computed a t-score to test a hypothesis. How many degrees of freedom are there for this sample? If you interpret the t-score you just computed as a z-score, what is the p-value? Explain.

There are 499 degrees of freedom, so the t-score is indistinguishable from a z-score; the one-sided p-value for a z-score of -2.988 is $.0014$, so we reach the same conclusion.

---

2. To analyze how well lie detectors perform when subjects are stressed, 48 randomly chosen subjects were connected to a lie detector and asked to read true statements out loud while receiving an electric shock. The lie detector incorrectly reported that 27 of the 48 participants were lying.

Note: For this problem, you will have to enter the data into *StatKey*. In particular, do not use the book's ***Lie Detector (Missed a lie)*** dataset to answer the following questions!

a. Use *StatKey* to determine the best estimate, based on this data, of the proportion of times the lie detector yields false positives, i.e., inaccurately reports deception.

The sample proportion is $\hat{p} = \frac{27}{48} = 0.5625 = 56.25\%$.

b. Use *StatKey* to create 5000 bootstrap samples and a bootstrap distribution of bootstrap sample proportions. What are the center and standard error of your bootstrap distribution?

This will vary a bit from simulation to simulation; the bootstrap distribution I created had a center of 0.562 and a standard error of 0.071 .

c. What does the formula in Section 6.1 say the standard error should be? How does this theoretical result compare to the standard error of your bootstrap distribution?

Using the formula,

$$\textbf{SE} \;=\; \sqrt{\frac{0.5625 \times 0.4375}{48}} \;=\; 0.0716\,,$$

which is very close to the bootstrap distribution's SE.

d. Use *StatKey* and your bootstrap distribution to find a 95% confidence interval for the overall percentage of false positives reported by the lie detector.

Again, this will vary a bit from simulation to simulation; for my bootstrap distribution, the 95% confidence interval for the population proportion p was

$$0.417 \leq p \leq 0.708\,.$$

e. Does this sample provide evidence that lie detectors give innacurate results more than half the time when subjects are stressed? State the relevant null and alternative hypotheses, use *StatKey* to create a randomization distribution based on this sample and the null hypothesis, obtain the p-value, and state your conclusion clearly.

The hypotheses are

$$H_0\,:\,p = 0.5 = 50\% \quad \textbf{versus} \quad H_a\,:\,p > 0.5\,.$$

For my randomization distribution, 23.8% of the simulated proportions were greater than or equal to $\hat{p} = 0.5625$, making the p-value 0.238 . We therefore cannot reject the null hypothesis, i.e., we do not have evidence that lie detectors give innacurate results more than half the time when subjects are stressed.

f. You just used a randomization distribution to test whether we have evidence that lie detectors give innacurate results more than half the time when subjects are stressed. Conduct the same hypothesis test by applying the theoretical method from Section 6.1 : compute the relevant z-score, obtain the p-value from the normal distribution, and state your conclusion.

Using the results from Section 6.1, the relevant z-score is

$$z \;=\; \frac{0.5625 - 0.5}{\sqrt{\frac{.5\times.5}{48}}} = 0.866\,,$$

yielding a p-value of 0.193 ; we therefore have the same conclusion as in part (e). Note that the SE in the denominator for this z-score uses the value from the null hypothesis, as this calculation is based on the assumption that the null hypothesis is true.

---

3. Is sleep or caffeine better for memory? To answer this question, 24 randomly chosen adults were divided equally into two groups and given a list of 24 words to memorize. During a break, members of one group took a 90-minute nap while members of the other group each took a caffeine pill. They were then all tested to see how many of the words they could remember; the results are in the book's ***Sleep Caffeine Words*** dataset.

a. State the relevant null and alternative hypotheses.

The hypotheses are

$$H_0\,:\,\mu_1 = \mu_2 \quad \textbf{versus} \quad H_a\,:\,\mu_1 > \mu_2\,,$$

where μ_1 is the average number of words recalled by those who took a nap and μ_2 is the average number of words recalled by those who took a caffeine pill. These are the group labels used by *StatKey*.

b. Use *StatKey* to create a randomization distribution based on this data and the null hypothesis, obtain the p-value, and state your conclusion clearly.

For my randomization distribution, only 2.3% of the simulated differences were greater than or equal to 3, the observed difference in the sample. This yields a p-value of .023, so we reject the null hypothesis and have evidence that sleep is better than caffeine for memory.

---

4. To investigate the influence of smoking on pregnancy, a study collected data on 678 women who had gone off birth control with the intent of becoming pregnant. The data are given in Table A.8 (after problem A.51 in the Review Exercises for Unit A) and Table B.8 (after problem B.32 in the Review Exercises for Unit B) and are available in the book's ***Get Pregnant (by Smoker Status)*** dataset.

a. Was this an experiment or an observational study?

This was necessarily an observational study, not an experiment, for obvious ethical reasons.

b. Determine the following:

- the overall proportion of women who became pregnant $\frac{244}{678} = 0.36 = 36\%$
- the proportion of smokers who became pregnant $\frac{38}{135} = 0.281 = 28.1\%$
- the proportion of non-smokers who became pregnant $\frac{206}{543} = 0.379 = 37.9\%$
- the difference in proportions between smokers and non-smokers $0.281 - 0.379 = -0.0979 = -9.79\%$

c. Use *StatKey* to obtain a 95% confidence interval for the relevant difference in proportions.

By creating a bootstrap distribution based on this data, we find that 95% of the simulated differences in proportions are between -18.1% and -.91%, i.e., the 95% confidence interval is

$$-.181 \leq p_1 - p_2 \leq -.0091\,.$$

As usual, these endpoints will vary a bit from simulation to simulation.

d. Based on this data, is there evidence that the proportion of successful pregnancies is lower among smokers than non-smokers? State the relevant null and alternative hypotheses, use *StatKey* to create a randomization distribution based on this data and the null hypothesis, obtain the p-value, and state your conclusion clearly.

The hypotheses are

$$H_0 \,:\, p_1 = p_2 \quad \textbf{versus} \quad H_a \,:\, p_1 < p_2\,,$$

where p_1 is the proportion of smokers who get pregnant and p_2 is the proportion of non-smokers who get pregnant. For my randomization distribution, only .92% of the simulated differences were less than or equal to the observed difference of -9.8%, making the p-value .0092. We therefore have strong evidence that the proportion of successful pregnancies is lower among smokers than non-smokers.

---

5. Do problem 6.124 from the Exercises for Section 6.2-HT. Be sure to state the relevant hypotheses, show the steps involved in computing your t-score, and explain your answers with complete sentences.

The hypotheses are

$$H_0 \,:\, \mu = 12.5 \quad \textbf{versus} \quad H_a \,:\, \mu < 12.5\,,$$

the sample mean is $\bar{x} = 11.1$**, the sample standard deviation is** $s = 0.4$**, and the sample size is** $n = 11$**. Consequently, the** t**-score is**

$$t \;=\; \frac{11.1 - 12.5}{\frac{0.4}{\sqrt{11}}} \;=\; -\,11.6$$

and the p**-value is practically 0. We therefore have strong evidence against the null hypothesis, i.e., strong evidence that the average air pressure in balls used by the Patriots was significantly less than the allowable limit.**