

## 209 Exam 2 – Practice

### **Question 1 (Monday Breakups Redux)**

Facebook is a social networking website. One piece of data that members of Facebook often report is their relationship status: single, in a relationship, married, it's complicated, etc.

With the help of Lee Byron of Facebook, David McCandless examined changes in peoples' relationship status, in particular, breakups. A plot of the results showed that there were repeated peaks on Mondays. Based on this initial examination of data, McCandless speculated that breakups are reported at higher frequency on Mondays.

To test this research hypothesis, McCandless collected a random sample of 75 breakups reported on Facebook within the last year. Of these sampled breakups, 20 occurred on a Monday.

**Research Question:** Are people more likely to break up on Mondays than other days of the week?

**Part A:** Describe the parameter in context and provide the appropriate symbol.

**Part B:** Is this an observational study or an experiment? How does this affect our conclusions? Explain.

**Part C:** State the null and alternate hypothesis in words **and** using the notation we've seen in class.

**Part D:** Describe what type-1 and type-2 errors are in the context of this study.

**Part E:** Are the conditions met to conduct a hypothesis test for this scenario? State the conditions and whether they are met or not met (show work or no credit given)

**Part F:** Find the value of the test-statistic for this hypothesis test.

**Part G:** The p-value for this test is 0.0010. Write up a conclusion to answer the research question using the strength of evidence approach.

### Question 2 (Sleep Deprivation Redux)

A CDC report on sleep deprivation rates shows that the proportion of California residents who reported insufficient rest or sleep during each of the preceding 30 days is 8.0%, while this proportion is 8.8% for Oregon residents. These data are based on simple random samples of 11,545 California and 4,691 Oregon residents.

**Goal:** Conduct a hypothesis test to determine if these data provide strong evidence that the rate of sleep deprivation is different for the two states.

**Part A:** Describe the parameter in context and provide the appropriate symbols.

**Part B:** State the null and alternate hypothesis in words **and** using the notation we've seen in class.

**Part C:** Compute  $\hat{p}_{pool}$ .

**Part D:** Compute the test-statistic Z.

**Part E:** Suppose the p-value for this test is  $\approx 0.00001$ . Write a conclusion to this question using ‘strength of evidence’.

### **Question 3 (Lead Exposure)**

Researchers interested in lead exposure due to car exhaust sampled the blood of 52 police officers subjected to constant inhalation of automobile exhaust fumes while working traffic enforcement in a primarily urban environment. The blood samples of these officers had an average lead concentration of 124.32 micro-grams per liter (micro-g/L) and a SD of 37.74 micro-g/L; a previous study of individuals from a nearby suburb, with no history of exposure, found an average blood level concentration of 35 micro-g/L. (Mortada et al. 2000)

**Part A:** Write down the hypotheses that would be appropriate for testing if the police officers appear to have been exposed to a different concentration of lead.

**Part B:** Describe type-1 and type-2 errors in the context of this study. Is one worse to make than the other? Explain.

**Part C:** Explicitly state and check all conditions necessary for inference on these data (show work or no credit given)

**Part D:** Compute the test-statistic for the hypothesis in Part A (show work for credit).

**Part E:** What is the degrees of freedom corresponding to this test-stat?

**Part D:** The p-value for this test is  $4.96 * 10^{-23}$ . Write a conclusion to the researcher's question in context at a significance level of  $\alpha = .05$ .

#### Question 4 (Fuel Economy)

Each year the US Environmental Protection Agency (EPA) releases fuel economy data on cars manufactured in that year. Below are summary statistics on fuel efficiency (in miles/gallon) from random samples of cars with manual and automatic transmissions manufactured in 2021.

**Research Question:** Do these data provide strong evidence of a difference between the average fuel efficiency of cars with manual and automatic transmissions in terms of their average highway mileage?

MPG	Mean	Std. Dev.	n
Automatic	28.7	3.90	25
Manual	30.1	5.13	25

**Part A:** What type of hypothesis test do we want to conduct for this scenario? Circle one of the following: (proportion / diff in proportions / mean / difference in means)

**Part B:** Check the conditions to conduct a hypothesis test for this research question. (show work or no credit given)

**Part C:** Calculate the test-statistic and state the corresponding df.

**Part D:** The p-value corresponding to this test is 0.0504. Explain how your conclusion to this research question would differ if you used the ‘strength of evidence’ approach or by using the ‘decision making’ approach with a significance level of  $\alpha = .05$ .