Cieran Wong

Prof. McDonald

MATH 263H

October 26, 2022

Assignment #6

#6.1.32

1. I don’t think there is an association, however, within the groups themselves, the non-coffee drinkers had a wider range of heights and were more distributed whereas the coffee drinkers were more collected around a certain point and had a smaller standard deviation.
2. Since it’s a dot plot, the distribution indicates that there were more of those occurrences.
3. Mean height for coffee: 66.688. Mean height for non-coffee: 67.657. Perhaps there is an association, however, I would say that 1 inch is not that big of a difference to comfortably say that there is an association between drinking coffee and height.
4. Median height for coffee: 67. Median height for non-coffee: 67. No change in my thoughts, still think that there is no association between height and drinking coffee.
5. Perhaps, however, we have not considered other confounding variables such as genetics, diet, place of birth, country of origin, etc. All these variables could contribute to a stunted growth, drinking coffee is just the added cherry on top.

#6.1.38

1. Experiment, it’s stated.
2. Explanatory: taking notes on paper vs on a computer. Categorical.
3. Response: memorization ability. Quantitative.
4. For paper: The mean was 6.925, the median was 7, and the standard deviation was 1.067. Looking at the boxplot, the center line was almost in the center, moving slightly to the right. The volume or area of the box was also a lot smaller than the computer’s boxplot, and this may indicate that the range of scores is a lot smaller. For computer: The mean was 5.5, the median was 5, and the standard deviation was 2.039. Center line for the boxplot was pretty far left, and the area of the boxplot was quite large compared to the paper’s boxplot. The data is more spread out compared to paper.
5. Paper tended to have a higher score on the quiz, and this is determined by looking at the median, the mean, and the standard deviation. Paper had a higher mean, median, and a smaller standard deviation compared to computer, indicating that there is an association between taking notes on paper and memorization ability.

#6.2.16

1. Observational units: Population of Milwaukee guessed.
2. Explanatory: Being told the population of Chicago or Green Bay. Quantitative. Response: Guessing the population of Milwaukee. Quantitative.
3. Experiment. Subjects were randomly assigned a “treatment” and were then asked to perform a task, to see if there was any influence by the treatment.
4. No, all the students were kids at a four year university in California who were all in an introductory statistics course.
5. Yes? It’s stated that some students were randomly chosen to be told the population of Chicago while the rest were told the population of Green Bay.
6. Table summary.

|  |  |  |  |
| --- | --- | --- | --- |
| City | Sample Size | Sample Mean | Sample SD |
| Chicago | 35 | 1357.34 | 802.21 |
| Green Bay | 34 | 271.38 | 370.96 |

1. (Chicago – Green Bay) = Observed Difference = 1085.961.
2. First possible explanation would be the anchor phenomenon discussed in the description. Perhaps psychologically, when told a fact or information before asking to derive something, it influences a person’s thoughts and they try to link the two, thinking there must be some sort of correlation if they’re being told this piece of information and then being asked a question that appears to be somewhat linked. The second possible explanation would be the size of the population they were told. Big difference between 3 million and 100,000. Those who were told about the population of Chicago, reported higher numbers than those who were told about Green Bay.

#6.2.20

1. Observational unit: age the person died and whether they had children. Explanatory: if a person had children or not. Categorical. Response: how long they lived. Quantitative.
2. Observational study. No one is being assigned any sort of treatment or being randomly assigned to groups for treatment. They are simply observing data to see if there is any correlation.
3. Table summary.

|  |  |  |  |
| --- | --- | --- | --- |
| Children? | Sample Size | Sample Mean | Sample SD |
| Had | 70 | 78.43 | 25.81 |
| No | 20 | 63.90 | 14.36 |

1. (Had – No) = Observed Difference = 14.529.
2. Diet and stress. Diet could significantly affect the life span of an individual. Stress could also determine the life span as a lot of stress is unhealthy and may cause someone to age exponentially.

#6.3.18

1. Null: No difference in spatial score when for men and women. Alternate: There is a difference between the spatial score in men and women.
2. Using a theory-based method makes sense as the sample sizes are 30 a piece and the data is not strongly skewed, as stated.
3. P-value: 0.0007. Probability of obtaining t = 3.57, -3.57, smaller or larger is 0.0007 so long the null is true.
4. (1.0551, 3.7449) 95% confidence interval. Average spatial score for men is between 1.055 and 3.745 higher than women.
5. We have reason to believe the average spatial score for men is higher than women, and 95% confident that it is between 1.055 and 3.745. Can’t determine it’s a cause-and-effect nor can it be generalized as the sample was not taken randomly from a larger population.
6. Null: Average verbal score for men and women is the same. Alternate: Average verbal score is different for men and women. Similar to spatial, sample size is 30 and data is not strongly skewed, as stated. P-value = 0.0000 < 0.0001. Probability of obtaining a t = 5.07, -5.07, smaller or larger is less than 0.0001 if null is true. 95% CI: (2.5423, 5.8577) Women have an average verbal score that is between 2.542 and 5.878 higher than men. Strong evidence and reason to believe that the average verbal score for women is higher compared to men, 95% confident that it is between 2.542 and 5.878 higher. Again, can’t be determined as a cause and effect nor can it be generalized as the sample was not taken randomly from a larger population.

#6.3.20

1. If p-value is 0.0162 which is greater than 0.01, we have not enough evidence.
2. P-value = 0.0162 < 0.05. Enough evidence since the p-value is smaller.
3. True: B, C. False: A, D.