Cieran Wong

Professor McDonald

MATH 263H

September 17, 2022

Assignment #2

#1.5.20

1. -Null hypothesis: 77.5% of monkeys will choose a box when directed by human gestures. π0 = 77.5%

-Alternate hypothesis: More than 77.5% of monkeys will choose a box when directed by human gestures. πa > 77.5%.

b. P-value = 2.52E-4 = 0.0462

c. Since the p-value is so small, we reject the null hypothesis and can conclude that more than 77.5% of the monkeys will choose a box when directed by human gestures.

d. Z-statistic = 3.479

e. The new p-value would be twice the original p-value, new p-value = 2(2.52E-4) = 0.092

#2.1.22

1. This will yield a biased estimate as there is no information regarding the various species of birds that either visit the yard or are native to the location of the yard. Therefore, it will be a little more difficult to estimate the proportion of finches that visit the bird feeder due to lack of information.
2. Similar to part a), this yields a biased estimate as there is a lack of sufficient data or information that backs up the hypothesis of finches preferring to eat with other finches. What happens if there are only two finches in the entire sample of birds that come by to eat in the yard? Perhaps they would definitely prefer eating with one another since they most likely know and associate with each other.
3. The proportion of birds that may be male could differ due to the type of food or the setting of the feeder. Perhaps for certain species if a feeder is a certain color, one gender would prefer it over the other. Using the finches, say for example, a female finch likes the color blue, however a male finch, dislikes blue and refuses to eat from something that is blue in color. This will cause only female finches to visit the feeder while the males will avoid it at all costs.
4. Calling on part c), if a feeder is a certain color and either the male or female prefer that color, that would affect the data collection when determining the male population of birds in the area. The data would not only be inaccurate, but it would also be unreliable and biased as the scenario favors one over the other.

#2.2.20

1. Skewed to the right.
2. The mean would most likely be greater than the median as the data is skewed to the right. This is caused by all the data on the far right of the graph would throw off the mean, while the median stays relatively lower as we have more data collected around where the center would be if the data and the graph was symmetrical.
3. Median = 35, Mean = 45.68.
4. I think the mean would be higher as there is data that is now going to be on the far-right side of the graph. Just a few pieces of data, as seen earlier, already threw off the graph and caused it to skew to the right. Adding on this bigger number would make it go way farther, causing the mean to increase. When plugged in, the new mean is now 48.68, an increase of 3, and the median stays the same at 35.

#2.2.26

1. Mean = 100, Standard Deviation = 15
2. 20 times the number of samples in part a).
3. 20 times the number of samples in part b).

#2.3.10

1. Parameter describes the whole population, while statistic refers to the sample. For this, it would be a statistic as it states that a random sample of students were asked the question. σ = 1.97 hours.
2. Standardized statistic = (3.01-2.84)/1.97 = 0.086
3. We fail to reject the null hypothesis as the standardized statistic is less than 2.

#2.3.12

I think the number would still be 2 as regardless of where the distribution is, the standardized statistic should remain the same to be consistent throughout. This will be useful when doing comparisons between data, for example, comparing normal distributions and t-distributions. If the standardized statistic were to differ, it would be a little more difficult to come to a conclusion when comparing the data as we are then unsure if a greater or less than 2 means the same thing for both distributions. Consistency and reducing the need for computing power would benefit from keeping it the same as a normal distribution.

#2.3.16

1. The variable of interest is whether or not the diameter of the needle is equal to 1.65mm. This variable would be quantitative as the variable represents an amount, in this case, the diameter of the needles.
2. π0 = 1.65mm, πa ≠ 1.65mm.
3. n = 35, x-bar = 1.64 mm, σ = 0.07mm
4. Bell graph with expected mean = 1.66mm and median = 0.04mm.

#2.3.30

1. The parameter would be the diameter of the needles produced and if they are equal to 1.65mm. π = 1.65mm
2. The null hypothesis: π0 = 1.65mm. The alternate hypothesis: πa ≠ 1.65mm.
3. A Type I error (false positive) would be if the needles were measured as not equaling 1.65mm and the entire batch was tossed, when in reality, the needles did not have a significant deviation from 1.65mm, enough to cause bodily harm. The consequence in this scenario is 1) wasted batches, and 2) halted production due to an unneeded check of the production line. This results in wasted time, resources, and money.
4. A Type II error (false negative) would be if the needles from the sample were measured, and it was concluded that there were no needles that varied significantly from the 1.65mm benchmark, when in fact there were a few that significantly deviated from the benchmark. A result of this error would be that someone could get seriously hurt from the needle being too big or too small.