An Observational Study on Parking at Columbia College

By Bailey Souza, Aaron Hankins, and Michael King

November 23, 2016

Introduction:

The purpose of this project is to estimate the difference in long-term parking availability for non-handicapped students on weekdays between the morning and midday at Columbia College.

To accomplish this, we took a stratified random sample of ten parking spaces each in the five major parking areas at the college (the upper and lower sections of Temporary Parking, and Lots A, B, and C of the Dormitory Parking) from 7:30 a.m. to 8:00 a.m. and from 11:30 a.m. to 12:00 p.m.. These data were then recorded in the form of a proportion (i.e. 7/10), with positive values signifying an available parking space.

Definitions and Assumptions:

Our population is defined as the total number of long-term, non-handicapped student parking spaces at Columbia College. Our sample proportions and sample standard deviations will be computed as a proportion from the simple random samples gathered for the given time. We assume that:

1. The availability of parking between weekdays remains relatively constant.
2. The people using these parking spaces - to avoid outliers - are non-handicapped students parking for longer than an hour. This is because there are a number of handicapped, faculty, and/or short-term parking spaces located throughout the college that may skew the distributions.
3. Temporary parking and Dormitory Parking Lots A-C are the most widely used student parking areas, and therefore the most relevant.
4. The side of Dormitory Parking Lot B closest to the dormitories would be frequently filled by students living in said dormitories, so it was excluded.
5. The other “halves” of each lot of Dormitory Parking are virtually empty throughout the course of the day, so we did not include them in order to prevent skewing of the data.

Our initial hypothesis from our personal experiences parking at Columbia College is that there will be more available parking spaces at midday than there will be in the morning.

Sampling Design and Methodology:

In order to gather this data, we labeled each parking space within each area with a numerical value ranging from 1 to *x, x* being the number of parking spaces in the given area. Next, using random.org, we randomly chose ten parking spaces in each parking area to check for their availability and marked them inconspicuously for ease of tracking before checking their availability. We decided that the best way to gather this data would be a simple success/fail tallying of the samples in the form of a proportion, with a "success" meaning an available parking space. As the sample sizes were too small to be reliable for a large two-sample proportion test, we used the "plus four" method for more accuracy in our confidence interval, and a two-sample proportion test to determine the probability that parking availability is greater in the afternoon than in the morning.

Problems We Faced:

Our original issue in gathering our data began with how we should sample. At first, we had individual samples of total parking availability in each of the major lots at both times, but we later realized that our data was in no way random since we consecutively gathered the data for our samples, and there was no test we could use that didn't require simple random samples. Also, the accuracy of our testing was difficult to secure because taking samples of data at a given time across five parking lots on the Columbia College campus takes enough time that spaces may be filled or vacated during the actual sampling process. We dealt with this by having our samples accord to a time frame as opposed to a single time such as 7:45 a.m., but the data may have been skewed due to this.

Results:

Before running our 95% "plus four" confidence interval, we computed our different proportions according the plus four method: "proportion 1" is the afternoon time, with 20/52 or 0.3846, and "proportion 2" is the morning time, with 17/52 or 0.3269. The 95% confidence interval that the mean difference between parking availability at midday and in the morning is (-0.126, 0.2414). This points towards a very small difference in times.

For the hypothesis test, we used a significance level of 0.05. We performed this test assuming the null hypothesis to be that there is no difference in the proportions of available parking spaces between the two time frames, and the alternative hypothesis to be that there is more available parking during midday than in the morning. Our

respective values for proportion 1, proportion 2, and the pooled proportion are 19/50 (0.38), 16/50 (0.32), and 35/100 (0.35). Note that these values are slightly different than our plus four confidence interval proportions because there is no added data for this test. We computed a z statistic of 0.629, which corresponds to a p-value of 0.2643. This means that, given there is no difference between the time frames, there is a 0.2643 probability of finding these data in a similar sample drawing. Since 0.2643 is substantially larger than our significance level of 0.05, we have no reason to reject the null hypothesis.

Discussion:

Our tests gave us corresponding results for no difference between the two times. However, we believe that our samples were not large or numerous enough to be entirely accurate. In order to get a more accurate statistic, we would have needed to draw more samples of greater size and average them for comparison.

Conclusion:

With our gathered data, our results point toward there being no substantial difference in parking availability between midday and morning at Columbia College. However, we believe that, due to insufficient data, these conclusions may not be accurate enough to rely on. From personal experience, our group as well as our friends and associates have all believed midday to have a greater availability for parking. We suggest that this topic be studied further, with larger sample and more numerous sample sizes.