Gabriel Taveras

Dr. Ira Gerhardt

MATH 491

December 2, 2018

Statistics Project

Pre-Health Advisors have long wondered whether knowledge of an undergraduate’s entering SAT score helps to predict the students MCAT score. The hypothesis created is that yes, knowledge on an undergraduate’s entering SAT scores helps predict the students MCAT score. The data given had seven variables all pertaining to a sample of 556 students; Total MCAT score, MCAT BS score, MCAT PS score, MCAT VR score, Total SAT score (Verbal and Math), Verbal SAT score, and Math SAT score. There are some caveats to the data; the data is not recent as the MCAT scores are measured using the old (pre-2005) system.

Before calculating the given MCAT scores, research was done to see what the true average MCAT scores were with scaled scores from the MCAT scores from the years 2006-2011 in comparison to the given data’s MCAT scores; The Scores were given by [AAMC.org](https://students-residents.aamc.org/advisors/article/percentages-and-scaled-score-tables/). Comparing the legacy (2006-2011) MCAT score average with the given MCAT score average, The given MCAT data is approximately 3 points higher than the legacy MCAT score average; the legacy mean is 25.05 with a standard deviation of 6.43 and the given MCAT data’s mean is 28.18 with a standard deviation of 4.97 which could mean the students have either gotten smarter or it was a lucky batch that year.

One calculations was done through MATLAB by taking finding the average SAT score and splitting the sample to where one group (group L) was comprised of students with SAT scores being lower than the average and the second group (group H) was comprised of students with SAT scores being higher than the average. On figure 1, the side-by-side box-plot shows that Group H (number 2) has a narrower spread compared to Group L (number 1); the whiskers begin at 20 and end at a near 40 value while Group L has a wide spread that begins at a low to mid 10 value and ends at a high 30. What this Box-plot shows is that those with higher than average SAT scores will have a greater chance of earning a high MCAT score compared to those with a lower than average score that will have a lower chance.

Through a website called [magoosh.com](https://magoosh.com/mcat/new-mcat-score-conversion/), a MCAT test prep/assistance blog, determined the conversion from the new system to the old system what MCAT score would be what percentile. Most websites that provide help on MCAT scoring say that 50 percentile is “good”, which in the old system means a MCAT score of 25, is the cutoff point. By calculating the 99% confidence interval, for Group L their MCAT scores will be between 27.72 and 28.67 while with Group H their MCAT scores will be between 30.75 and 31.87. The 90% confidence interval with the difference of the 2 means is 12.8 and 13.3 which is extremely wide and must be about the overall spread of the MCAT scores.

Using 3 different two-sample Hypothesis Tests to determine if there is a significant difference between the mean MCAT score obtained by students in Group L versus the students in Group H at left tailed, right tailed, and both tailed, the null hypothesis of no difference will no be rejected if there is no difference. There is a significant difference as the p-value for the both tailed test is an even 0. What this means is that through the analysis of MCAT scores and SAT scores being related is highly possible as the MCAT data was separated by the Average SATs scores.

Before using a Linear Regression model to compare the total SAT scores with the total MCAT scores, a scatter plot was created with a best-fit line applied in order to see if higher SAT scores mean higher MCAT scores; figure 2 shows that the cloud of data moves mostly diagonally up and to the right which helps prove that SAT scores are tied to MCAT scores in a way. After running the linear regression model with the Total MCAT variable being the response variable, the model shows that the SAT scores do affect MCAT scores to where for every approximately 0.02 SAT points the MCAT score should go up by one point.

Another dataset was created where the SAT scores were only unique scores which meant no repeated scores, and the MCAT scores were the means of all scores where the MCAT scores had the same SAT scores. Creating a scatter plot for the dataset showed that the thin data cloud follows the previous dataset’s scatter plot where it moved diagonally up and to the right. The linear regression model created showed similar results with the previous dataset but it has a weaker correlation compared to the previous dataset as the p-value of previous dataset is about 28 powers of 10 more accurate than the unique dataset.

With all the analysis and calculations in place it is safe to say that with a 99% confidence there is a correlation where SAT scores do affect MCAT scores.