The Teacher's Aid

Matt Cheatham
Advisor: Dr. Richard Cooper
April 25, 2002

1 Introduction

Every year in the Fall a new class of first-years enters Trinity University for their quest through college. Many of these first-years may have never taken a math class more difficult than algebra or geometry. This means that professors in the mathematics department have a tough job ahead of them, not to mention the students. To prepare both sides for that dreaded calculus class, Trinity has required that most incoming first-years take an entrance level math exam. Those that have passed the A.P. exam or have received credit for calculus by other means are exempt from the test. This exam helps administrators and faculty advise each student individually on whether he/she should take pre-calculus (MATH-1301) or Calculus I (MATH-1311). A passing grade on the exam is a thirteen out of twenty-five. The exam basically gives professors a fine dividing line to evaluate student's abilities in calculus. We will later see if this dividing line is a good indicator of some of the student's grades in calculus. When a student passes the entrance exam their advisor should, safely be able to encourage the student to attempt Calculus I. Moreover, with the information contained in The Analysis section of this project, the advisor will be able to defend his/her position with the advisee by showing statistical information to the student. By providing this information, the professor should be able to show and explain about how well the student will do among peers. The purpose of this project is to show that the exam is fair and unbiased towards the students, and also provide statistical analysis to guide the professor into the correct position to advise the student.

I believe I have found a few problems with the exam by looking through the sample data, and I found that there are some procedures that I think could be changed to ensure that the decision made by the student and their advisor will be the best and most accurate, speaking in terms of course placement. My first task was to build a database that can be queried by a professor or administrator to allow them to gain valuable knowledge about the student's abilities in this analytically minded field.

2 The Database

Databases are a quick and easy way to find information on a particular subject. They are used in everyday life, and every person is in a database. These databases are used in places such as the CIA, IRS, Trinity University, ESPN or BarnesandNoble.com. However, each database has its own characteristics and different information contained in it. Databases are a powerful tool in the success of many companies today. Their abilities reach as far as searching a list of websites with that one "AOL keyword", to searching a small list of people to find certain information about an individual person. You can even search for groups if the database allows it. For example, looking through BarnesandNoble.com to find books written by an author that you like or searching eBay.com for that ugly lamp that you just cannot find anywhere else.

The database that I have produced, not nearly as sophisticated as the examples above, allows a user, in this case a professor or administrator, to search a list of incoming first-year students and the details of their exam. First the professor must have the student's identification number and enter it into the database's main query. In this case only, the query will return a list of information back to the user. This information consists of the student's I.D. number once again, their total score on the exam, their letter grade earned in calculus, the GPA of that letter grade, the course, section, and term they took the course. The course is given because some of the students in the database took MATH-1301, however for clarity of the information these students are not included in the statistical analysis that will be shown later. On a side note, if this database is used to advise next year's new class, the course, grade and term would, obviously, not be applicable. Included with these categories is a list of the questions the student missed along with their subject(s). For example if the user enters student I.D. number 297269 the database's query it will return Figure 1a in the appendix. In another instance, when the user enters student I.D. number 322644 the query will return Figure 1b in the appendix. The reason for constructing such a database is to allow professors to view how well some of their students did on the entrance level math exam. Since not all students follow the mean and not all students miss the same type of questions, the database allows the professor to see what particular subjects the student had trouble with.

3 The SQL

There are five separate tables in the database. Tables are uniquely identified by their names and are comprised of columns and rows. Columns contain the column name, data type, and any other attributes for the column. Rows contain the records or data for the columns. The student's I.D. number, which is one column, links four of these tables. The relationships of these tables can be seen in Figure 2 in the appendix. One table contains the student's total score on the exam, labeled TotalScoreTable, another has the grade and GPA the student made in their course, labeled GPA, a third table has the course, section and term the student took calculus or pre-calculus, labeled Course/Grade, the forth table, labeled AnswerTable, has a list of the questions and whether the students individually got each question correct or incorrect. The last table, labeled Question/Subject, has the subjects paired with each question. This table is linked to the four others using a different process. In this case I used an SQL, or Structured Query Language, statement format to link the question numbers with the table labeled AnswerTable. SQL statements are used to perform tasks such as update data on a database, or retrieve data from a database. In this case, we will only be retrieving data, however the features of the SQL can always be updated. Here is the format of a simple select statement:

select "column1" [, "column2", etc] from "tablename" [where "condition"];
and the things inside the [ ] are optional.

The column names that follow the select keyword determine which columns will be returned in the results. You can select as many column names as you would like, or you can use a " $*$ " to select all columns. Basically the database searches to see if the number in the AnswerTable is a zero, if it is then the search goes to the next question and then checks again. If the question was answered incorrect then the database outputs the number and subject of the question. The SQL statements for questions number one and two look like:

1. ((([Question/Subject].Question)=[AnswerTable].[1]),
2. ((([Question/Subject].Question)=[AnswerTable].[2]).

This process is the same for each question.

4 The Analysis

More important than just being able to look at the student's abilities individually is to see how the students compare with the group as a whole. The best and most efficient way to do this, with a large group, is to use a program like Microsoft Excel.

First what I did was have excel plot each student on a graph using their GPA in calculus ($y - axis$) and test score on the exam ($x - axis$). Then excel fit a line to the points, as you can see in Figure 3 in the appendix. What this fit-line does is give us a prediction of the student's grade in the class while knowing what the student made on the exam. Of course, you can see on the graph there will always be outliers. The dividing line for a passing grade on the exam is a thirteen, thus if you make less than a thirteen a student is advised to try pre-calculus first and then move directly on to calculus the next semester. The line fitted to the data predicts that a student choosing calculus instead of pre-calculus while making under a thirteen on the entrance exam will make between a C and C- in calculus, not too great. Basically this line-fit model gives us a graph of our success and something that we can base improvement on by moving the line higher and higher each year.

Another plot that provides us with information about the data and backs up the predictions in the line-fit model is a residual plot. To plot the residuals we take the observed grade point average made in calculus minus the predicted grade point average given by the line-fit model. Basically the residuals are an exact error for the predictions. When looking at the plot, consistencies or patterns mean that the data may be related in some way. By related we mean there is no randomness. As we can see by looking at the residual plot, labeled Figure 4 in the appendix, this does not seem to be the case. We graph both the line-fit model and the residual plot to look for normality.

Statistics are all around us. In fact it would be difficult to go through a full week without using statistics. Without statistics we could not plan our budgets, pay our taxes, enjoy games to their fullest, or evaluate classroom performance. Now we need to look at each set of data, the test scores and grade point averages, as a whole. We do this by having Microsoft Excel find descriptive statistics, on the data in two different groups. This branch of statistics lays the foundation for all statistical knowledge. One group consists of just the test scores and the other has the student's grade point averages. However, both groups have the same number of students in them, 122. We had to remove some of the students because they withdrew from calculus or did not have some of the information needed to fully include them in the study. From Figure 5 in the appendix we see that the average test score was fairly high at 16.3, passing, but the average calculus grade for a student that took the exam is fairly low at 2.25, lower than a C+. If we look back at the line-fit model for the test scores we see that the predicted GPA for a test score of about 16.3, is about 2.25. This further backs-up the line-fit model 's accuracy with its estimations of the student's grades. Also the standard deviation of both the GPA and the test scores were quite large, meaning that many of the students did not finish close to the mean of the group as a whole.

Our original idea is to decide on a hypothesis and test whether or not that hypothesis holds. We set H_0 to be that the average of the student grades on the test are equal. Then we set H_A to be the opposite of H_0:

$H_0 : \mu_1 = \mu_2 = \mu_3 = ... = \mu_k$
$H_A :$ The are not equal.

For the sake of this exam we would like to be able to accept H_0 but this does not look promising because of the scatteredness of the data on the line-fit plot. With no statistical information I would be lead to believe that we will reject H_0 but not necessarily accept H_A. H_0 will never be totally accepted because we would need the entire set of students and their test scores to accept. However, if we did have all of the information needed to accept, all of the data points would also have to fall correctly into place. We first use an excel program from the regression analysis, called ANOVA or analysis of variance, that gives us immediate verification on whether or not we can reject H_0 or not reject H_0. If we look at the analysis of variance chart in Figure 6 in the appendix we see that we can actually reject H_0 because of two separate reasons. One, we have an $F_{observed}$ value equal to 13.45 and an $F_{\alpha=0.05}$ value equal to 3.67×10^{-4}, this value comes from using an F distribution chart, which you can find it in the back of just about any statistics book, using degrees of freedom of $m = 1$ and $n = 121$. ¿From all statistics books we know that:

if $F_{observed} > F_{\alpha=0.05}$, then we can always reject H_0.

Two, we have a p-value, from the X variable, of 3.67×10^{-4} which is much less than $\alpha = 0.05$. We use this α value because it allows us to be fairly confident about our decision to reject H_0, however we see that we could reject H_0 even if $\alpha = 0.01$. Another interesting point about the ANOVA is the t-test. The t-test provides us with a test of the slope of the line in the line-fit model from Figure 3. The equation for the line is:

$y = a + \beta X$. a is the intercept and β is the slope.

The hypothesis for the slope is:

$H_0 : \beta = 0$
$H_A : \beta \neq 0$

We can reject the hypothesis for the slope because the $t_{\alpha/2} = 1.658 < t_{observed} = 3.67$. However, the value for R^2 is only 10.1% which means that the we are not confident about the rejection at all. Thus this may not be the best way to test the data for inconsistencies. The t-test also shows that there is no relation between the variables X and y.

When I initially looked at the results of the exam I found some problems that I thought might need to be addressed. First, it looked as though students were not being given enough time to finish the exam. The last five questions on the exam were not being attempted by many of the students. However, this will always be the case since the exam is timed and this problem is an expected shortcoming of the exam. Second, we see from the list in Figure 7 in the appendix that out of the top five most difficult questions, four of them fall in the category of trigonometry, numbers 21, 23, 24, and 25. Now with the trigonometry questions being grouped together and placed at the end of the exam we are not given a clear view of student's abilities in this field of mathematics. This means that a student must budget his/her on the first twenty questions and hope that when the trigonometry comes up they have the time and the know how to do them. This can potentially play a major role in showing that the incorrect answers to the questions may be linked or related by their subject or by their position in the exam. Actually we can prove this is not the case. Now we will use a more in-depth approach to the questions and the subjects of the exam individually using a X^2 distribution by testing them for homogeneity. First we form an original hypothesis that says every question, proportionally, is missed in a pattern forming matter. Next we set our alternative hypothesis to be that the questions are missed randomly. This looks like:

H_0 : The proportions correct of Q1=Q2=...=Q25,
H_A : These proportions are not equal.

Next we must next produce two charts. One chart will contain questions one through twenty-five in a row and then two columns for each question, one for the total correct answers and one for the total incorrect answers. This chart can be seen in the appendix as Figure 8. The other chart, Figure 9 in the appendix, will have the eleven subjects in one row and then, similarly to the first chart, the total correct and incorrect answers in two columns for each subject. Under the two columns will be another row that totals the number of attempts at either the certain subject or individual question. After this step we will find the estimate value, this is given by multiplying the total number of correct answers with the total number of correct and incorrect responses for each question, this is found in the third row and is the number 121 that is given for every question since the same number of people attempted the same number of questions, and then we divide that value by the total correct and incorrect answers for the entire test. The equation looks like this:

$$E_{i,j} = \frac{X_{i,\bullet} \cdot X_{\bullet,j}}{X_{\bullet,\bullet}} \qquad 1 \leq i \leq 2, \quad 1 \leq j \leq 25.$$

This equation produces values in two rows of twenty-five. The single bullet is the sum over the variable i or j. The double bullet is the sum over both variables. Then to get the value of the X^2 observed we take the sum of each individual correct and incorrect value minus each estimated value and then square it. Then we divide by the estimated and we are given the X^2 observed value. This equation looks like:

$$X^2 = \sum_{i=1}^{2} \sum_{j=1}^{25} \frac{(X_{i,j} - E_{i,j})^2}{E_{i,j}} = 255.27.$$

Now we compare the value of the X^2 observed and the X^2 critical. If the X^2 observed is larger than the X^2 critical then we can reject H_0 safely. In the example above, the questions themselves are being tested. We have a X^2 value equal to 255.27 and a X^2 critical value of 36.42. The X^2 critical value is given by looking at the X^2 distribution chart and using degrees of freedom equal to twenty-four, the number of questions minus one, and an α of 0.05. The original hypothesis states that the questions, proportionally, are missed equally, however we know that the questions are missed randomly since we can reject the original hypothesis. This conclusion does not completely rule out my first problem with the test, which was that students were finding the end of the test more difficult, but it surely does not help it either. Thus we can say that a question at the end of the test is not missed for just that reason. My next problem with the test was that the trigonometry questions were being missed more frequently, not only because of their placement but because of their difficulty as well. Well, we know that it is not the placement from the first X^2 test. This time we test the subjects of the questions only. The same process from above works when testing this part of the exam. The hypothesis and alternative hypothesis are respectively:

H_0 : The proportions correct of S1=S2=...=S11,
H_A : These proportions are not equal.

However, as with the questions, I found that I could positively reject H_0 because the X^2 observed value equals 226.1, this comes from the equations:

$$E_{i,j} = \frac{X_{i,\bullet} \cdot X_{\bullet,j}}{X_{\bullet,\bullet}} \qquad 1 \leq i \leq 2, \quad 1 \leq j \leq 11,$$

$$X^2 = \sum_{i=1}^{2} \sum_{j=1}^{11} \frac{(X_{i,j} - E_{i,j})^2}{E_{i,j}} = 255.27,$$

and the X^2 critical equals 18.31, with degrees of freedom equal to 10, the number of subjects minus one, and α equal to 0.05. Thus the problems I saw with the exam, both structural and the time limitation, were not a factor as shown by the X^2 distribution.

5 Conclusion

With this analysis and database professors should have a better idea of a student's abilities in mathematics. Through the statistics and hypothesis, professors can now compare a student with the entire entering class and explain to them, with the information given, why the path that the professor advises them to travel is a wise one. This will ease some of the students into precalculus that need to be and maybe even give some of them the chance to let the information sink into their brains, instead of rushing through calculus. When I was first motivated to work on this project I wanted to make sure that everything contained in it would be easily accessible to professors, the information would be useful, easy to read, and the database would be easy to use. I chose a project like this because it pulled together my three favorite subjects, those being, in no particular order, mathematics, computer science, and statistics. I feel that I have learned a great deal from my experience with the database and statistics. I found it was slightly difficult to learn the SQL statements but using them in the database was much easier than any other way. The database that I designed is quick, easy to use and holds interesting information. I feel that I have met many of the goals that I set out for myself and this project. My hope is that by using the database and the analysis, professors can help the students make the right choice in the future. All of my thanks to Dr. Cooper for taking time out of his schedule to help me with the project.