What's My Grade?

Michael Lloyd, Ph.D.
Professor of Mathematics and Computer Science

Abstract

The distribution of students' course grades for a variety of classes will be examined and a probabilistic prediction of their course grades based on their current performance will be derived. Besides being statistically interesting, the data and techniques could be used pedagogically.

Introduction

"What's my grade?" is a common question instructors are asked by their students before they take the final. In this paper, I will investigate some predictors for a student's course grade.

The following table lists the courses that are studied in this paper and the corresponding level.

Course	Level
College Algebra	1000
Plane Trigonometry	1000
Precalculus	1000
Statistical Methods	2000
Probability/Statistics	3000

These courses were selected because I teach them regularly and thus have sufficiently large sample sizes. Also, only data from the summer 2001 semester through the fall 2003 semester were used because it was in 2001 that I started giving short homework assignments to all of the aforementioned classes instead of quizzes. Withdrawals were ignored because I do not keep partial grade information for such students; also, there were very few incompletes, so these were also omitted.

Dependence on Semester

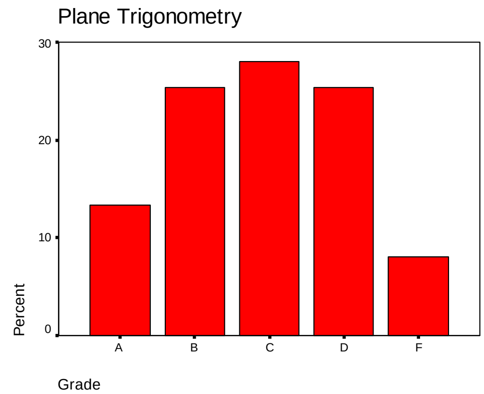
The following table show the results of three 1-way ANOVAs on the course grade where 4 was assigned to an A, 3 for a B, etc. Statistical Methods and Probability/Statistics are only taught one semester per year so they could not be included in these ANOVAs.

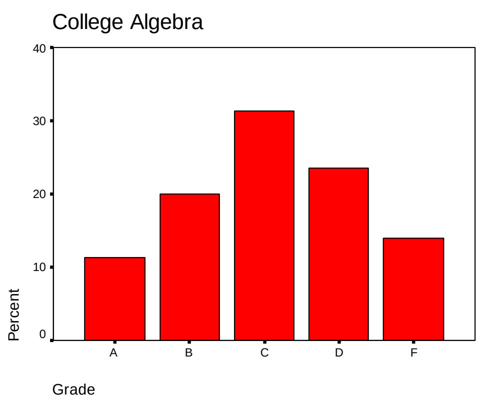
There may only be a significant difference in the average course grade between semesters for Precalculus. Grades are probably higher in the spring because most students in that course are fresh out of high school in the fall. The students who fail Precalculus in the fall are more serious about studying when they retake that course in the spring. These ANOVAs support the decision that there is no semester dependence and thus lumping all the data together for each class is justifiable.

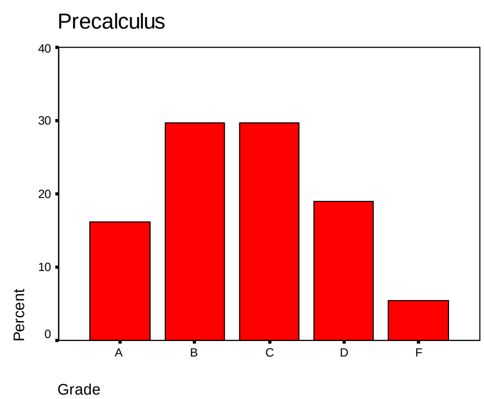
	Semester	Significance
College Algebra	Fall, Spring, Summer	0.47
Plane Trigonometry	Fall, Spring, Summer	0.46
Precalculus	Fall < Spring	0.06

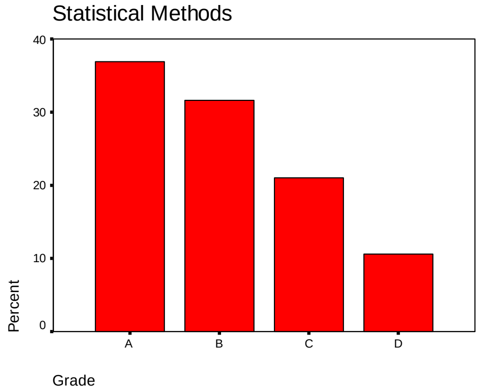
Distribution of Grades

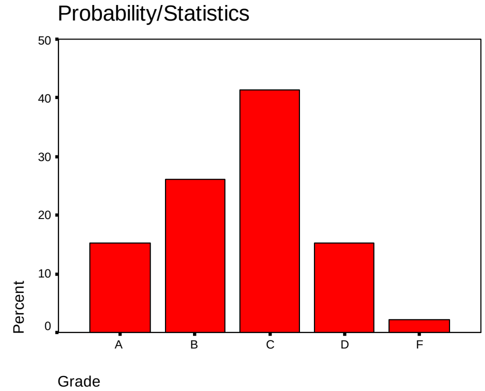
Before investigating how to predict students' grades, I thought it was a good idea to examine the distribution of course grades. The distribution for the lower level classes Plane Trigonometry, College Algebra, and Precalculus are similar. Statistical Methods stands out as having the property that lower course grades are earned with lessening probability. There are also many high course grades in Probability/Statistics.









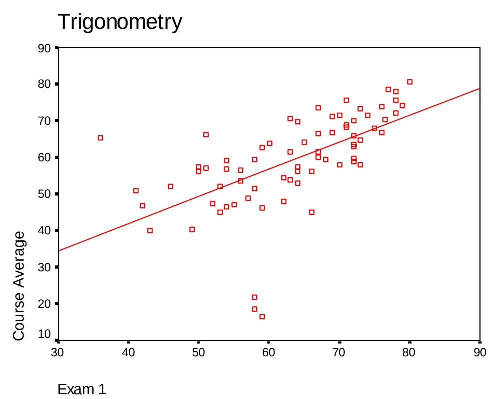


The following table gives 95% confidence intervals for passing each course, where passing is defined to be an A, B or C. Note that the pass rate for College Algebra is significantly less than both Statistical Methods and Probability/Statistics. Also, the Trigonometry pass rate is significantly less than Statistical Methods.

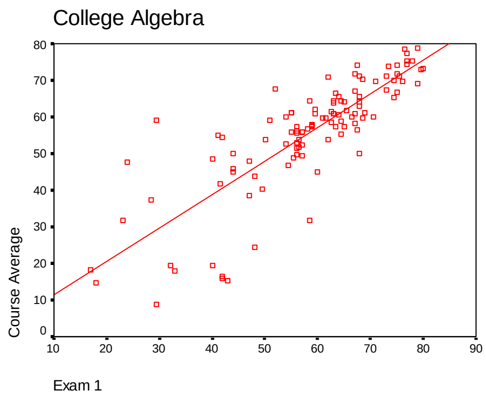
	Number students	Probability
College Algebra	115	63 ± 9% = (54, 71)%
Trigonometry	75	67 ± 11% = (56, 77)%
Precalculus	111	76 ± 8% = (68, 84)%
Statistical Methods	38	89 ± 10% = (80, 99)%
Probability/Statistics	46	83 ± 11% = (72, 94)%

Course Average Prediction Based on the First Exam

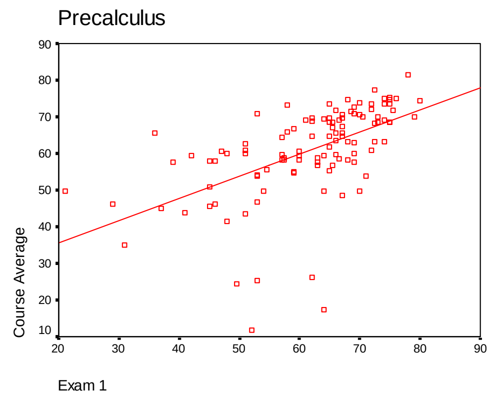
The following are scatter plots and regression equations of (exam 1, course average). (Students with a missing exam 1 score were omitted.) All significance levels were 0.000 except Probability/Statistics with a level of 0.062. Note that the coefficient of determination R^2 almost decreases as course level increases. Thus, the first exam is more indicative of a student's ultimate course grade in the lower-level courses. The variables are C = course average, E1 = exam 1 average, and n = sample size. Note that the exams are all out of 80 points.



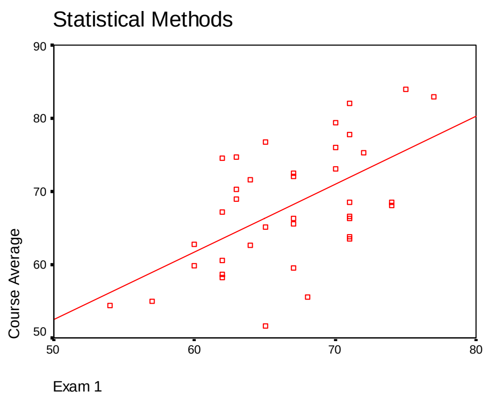
$C = 12.1 + 0.743E1$, $R^2=0.35$, $n=72$



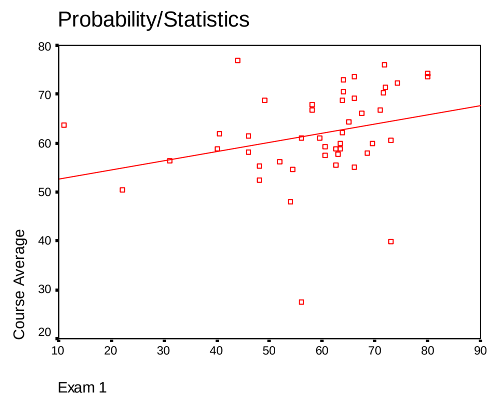
$C = 2.31 + 3.64E1$, $R^2=0.68$, $n=112$



$C = 23.3 + 0.608E1$, $R^2=0.31$, $n=111$



$C = 6.1 + 0.926E1$, $R^2=0.34$, $n=38$



$C = 50.9 + 0.185E1$, $R^2=0.05$, $n=46$

The following tables give the probability of passing or failing the course conditioned on passing or failing the first exam. A perfect correlation would correspond to the probability matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Plane Trigonometry	Pass Course	Fail Course
Pass first Exam	0.78	0.22
Fail first Exam	0.41	0.59

Precalculus	Pass Course	Fail Course
Pass first Exam	0.89	0.11
Fail first Exam	0.38	0.62

College Algebra	Pass Course	Fail Course
Pass first Exam	0.84	0.16
Fail first Exam	0.20	0.80

Statistical Methods	Pass Course	Fail Course
Pass first Exam	0.92	0.08
Fail first Exam	0.00	1.00

Probability/ Statistics	Pass Course	Fail Course
Pass first	0.91	0.09

Note that only 1 individual failed his or her first Statistical Methods exam.

Exam		
Fail first Exam	0.64	0.36

The following tables give the probability of passing or failing the course conditioned on the earning an A or B on first exam. For College Algebra and Plane Trigonometry, the first row of the probability matrix is closer to [1,0] than when conditioning on the first exam.

Plane Trigonometry	Pass Course	Fail Course
A or B on first Exam	0.95	0.05
C or worse on first Exam	0.39	0.61

College Algebra	Pass Course	Fail Course
A or B on first Exam	0.95	0.05
C or worse on first Exam	0.44	0.56

Precalculus	Pass Course	Fail Course
A or B on first Exam	0.90	0.10
C or worse on first Exam	0.58	0.42

Statistical Methods	Pass Course	Fail Course
A or B on first Exam	0.92	0.08
C or worse on first Exam	0.83	0.17

Probability/ Statistics	Pass Course	Fail Course
A or B on first Exam	0.89	0.11
C or worse on first Exam	0.79	0.21

The following tables give the probability of passing or failing the course conditioned on earning an A on first exam. In every course except Probability/Statistics, an A on the first exam implies passing the course. In fact for Probability/Statistics, passing the first exam is almost independent of passing the course.

Plane Trigonometry	Pass Course	Fail Course
Exam 1 = A	1.00	0.00
Exam 1 < A	0.58	0.42

College Algebra	Pass Course	Fail Course
Exam 1 = A	1.00	0.00
Exam 1 < A	0.57	0.43

Precalculus	Pass Course	Fail Course
Exam 1 = A	1.00	0.00
Exam 1 < A	0.70	0.30

Statistical Methods	Pass Course	Fail Course
Exam 1 = A	1.00	0.00
Exam 1 < A	0.88	0.12

Probability/ Statistics	Pass Course	Fail Course
Exam 1 = A	0.83	0.17
Exam 1 < A	0.82	0.18

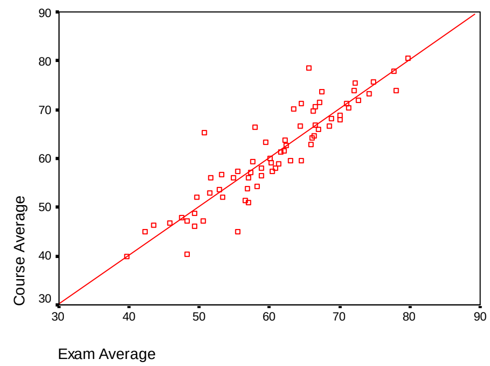
Course Average Predictions Based on the First Three Exams

Near the end of the course, students become increasingly concerned with what their course grade will be. The current homework average was not included because I was seeking a convenient method for predicting the final grade. Also, it was inconvenient to determine the current average for this study.

The following scatter plots and regression equations are for (3-exam, course average). The significance levels for the courses were all 0.000. Note that R^2 decreases as the course level increases. I think this is because upper-level students are more flexible in improving their study habits if they do poorly on the first exam.

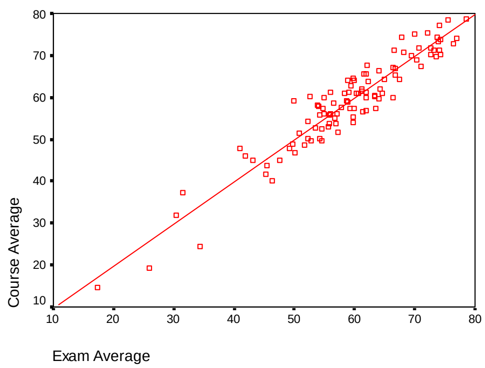
The variable exam average (EA) is obtained by adding the first three exams and dividing by three. Any student with a missing exam was omitted.

Trigonometry



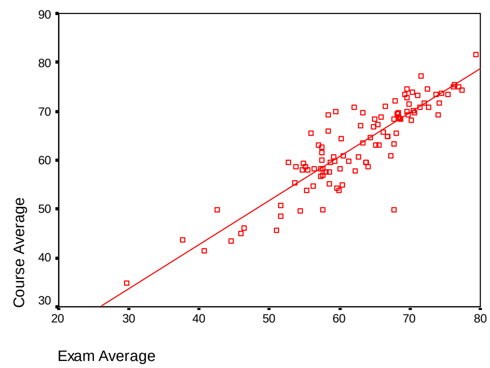
C = 0.202 + 1.001EA, R^2 = 0.83, n=69

College Algebra



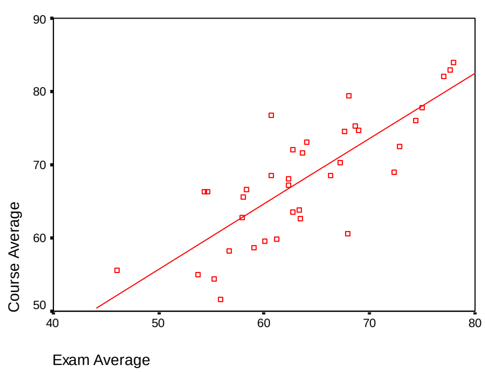
C = -0.398 + 1.002EA, R^2 = 0.90, n=103

Precalculus



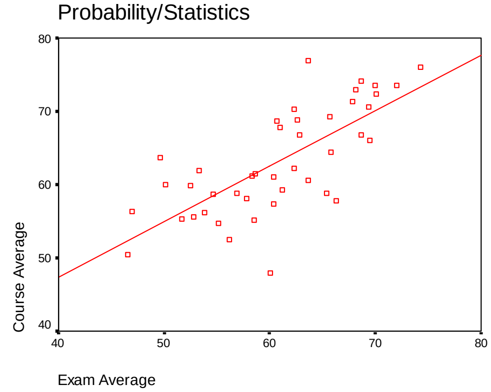
C = 6.87 + 0.898EA, R^2 = 0.79, n=104

Statistical Methods



C = 11.0 + 0.895EA, R^2 = 0.64, n=37

Probability/Statistics



C = 17.0 + 0.759EA, R^2 = 0.53, n=43

The following tables give the probability of passing or failing the course conditioned on the 3-exam average. Except for Statistical Methods, about 8% of everyone who has a passing 3-hour exam average will fail the course.

Plane Trigonometry	Pass Course	Fail Course
Passing Avg.	0.92	0.08
Failing Avg.	0.25	0.75

College Algebra	Pass Course	Fail Course
Passing Avg.	0.93	0.07
Failing Avg.	0.26	0.74

Precalculus	Pass Course	Fail Course
Passing Avg.	0.92	0.08
Failing Avg.	0.35	0.65

Statistical Methods	Pass Course	Fail Course
Passing Avg.	1.00	0.00
Failing Avg.	0.33	0.67

Probability/ Statistics	Pass Course	Fail Course
Passing Avg.	0.91	0.09
Failing Avg.	0.73	0.27

The following table gives the probability that the 3-exam average underestimates, predicts exactly, or overestimates the course grade.

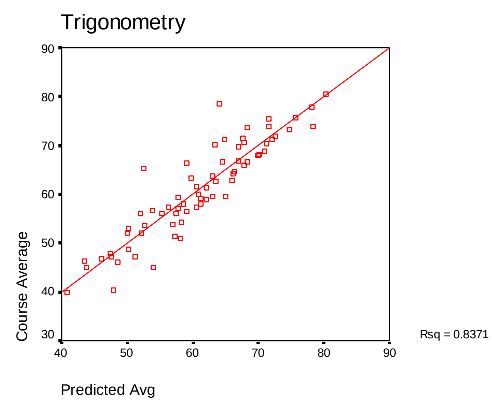
	Underestimates	Same	Overestimates
College Algebra	0.17	0.66	0.17
Plane Trigonometry	0.12	0.78	0.10
Precalculus	0.19	0.66	0.15
Statistical Methods	0.43	0.49	0.08
Probability/Statistics	0.42	0.44	0.14

The adjacent table indicates the rare event when the 3-exam average was off by two or more letter grades.

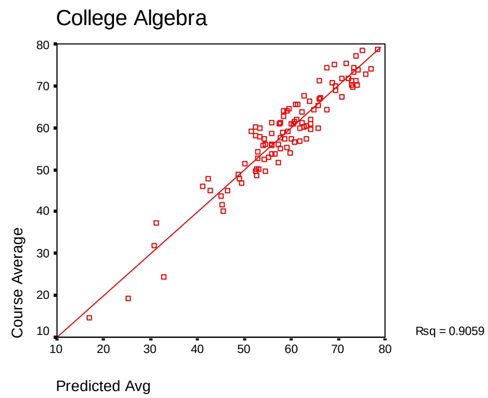
		probability
College Algebra	never	0.00
Plane Trigonometry	underestimated once	0.01
Precalculus	overestimated once	0.01
Statistical Methods	underestimated 4 times	0.11
Probability/Statistics	underestimated once	0.02

A more sophisticated method for using the first three exams would be to use multiple linear regression. The following are scatter plots of (multilinear prediction, course) average.

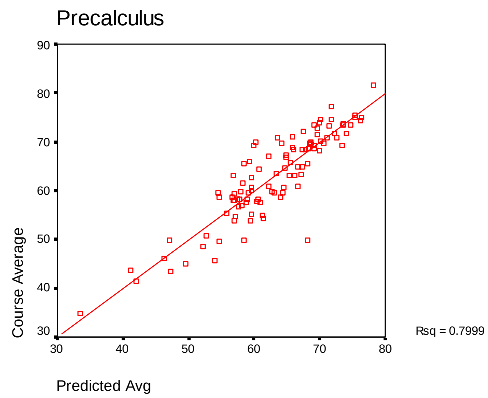
Note that for the semesters used in the study, College Algebra and Precalculus used essentially the same book. Also, the middle exam (E2) covered logarithms and exponential functions had the most influence in the regression model.



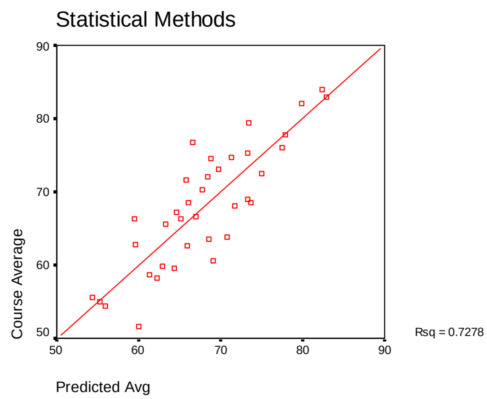
$C = -0.6+0.311E1+0.300E2+0.403E3, R^2=0.84$



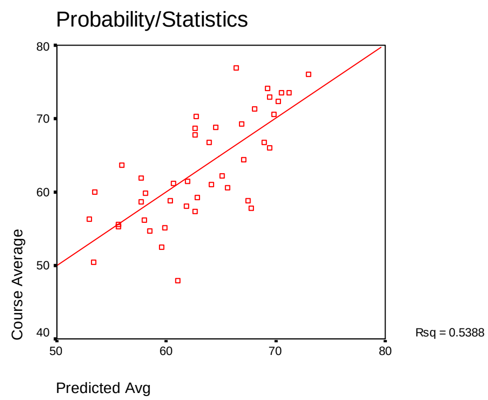
$C = -0.4+0.280E1+0.373E2+0.350E3, R^2=0.91$



$C = 7.5+0.251E1+0.373E2+0.266E3, R^2=0.80$



$C = -8.9+0.661E1+0.334e2+0.189E3, R^2=0.73$



$C = 17.5+0.225E1+0.313E2+0.212E3, R^2=0.54$

For trigonometry, the last exam (E3) included the law of sines, law of cosines and vectors was most influential variable.

For Statistical Methods, the first exam (E1) was the most important. This exam covers descriptive statistics and interpretation of basic statistical graphs and measures.

For Probability/Statistics, the second exam (E2) was the most important. That exam primarily covers probability word problems.

There was only a slight improvement in R^2 over the using the 3-exam average, so the use of this more sophisticated approach is not justifiable.

The following tables give the probability of passing or failing the course conditioned on the grade predicted by the multilinear regression model's prediction. As expected, this does not improve much on the 3-exam average approach.

Plane Trigonometry	Pass Course	Fail Course
Predict Passing	0.92	0.08
Predict Failing	0.21	0.79

College Algebra	Pass Course	Fail Course
Predict Passing	0.92	0.08
Predict Failing	0.32	0.68

Precalculus	Pass Course	Fail Course
Predict Passing	0.91	0.09
Predict Failing	0.14	0.86

Statistical Methods	Pass Course	Fail Course
Predict Passing	0.97	0.03
Predict Failing	0.00	1.00

Probability/ Statistics	Pass Course	Fail Course
Predict Passing	0.89	0.11
Predict Failing	0.67	0.33

The adjacent table gives the probability that the grade predicted by multiple linear regression underestimates, predicts exactly, or overestimates the course grade. Note that the table is more balanced and improved for upper level courses.

	Underestimates	Same	Overestimates
College Algebra	0.17	0.64	0.19
Plane Trigonometry	0.13	0.77	0.12
Precalculus	0.13	0.70	0.17
Statistical Methods	0.24	0.52	0.24
Probability/Statistics	0.22	0.54	0.24

The multiple regression model was off more than 2 letter grades only once each, and those were in Plane Trigonometry and Precalculus.

Conclusion

The following were determined in this study:

- Precalculus students have on average higher grades in the spring than the fall.
- In every course except Probability/Statistics, an A on the first exam implies passing the course.
- For Probability/Statistics, passing the first exam is independent of passing the course.
- The 3-exam average is more correlated to the ultimate course grade for lower-level courses.
- A student with a passing 3-exam average is at least 91% likely to pass the course.

Some of the patterns discovered in this paper may be included on my syllabi or on the class website. This would be done with the intention of helping the students make decisions like whether it would be in their best interests to drop a course. However, revealing this information may adversely affect their attitudes.

Ideas for Further Study

The summer classes could be removed and the current homework average used for the prediction. (I do not assign homework in the summer.) Also, it might be interesting to compare the students' actual course grades with what they think they will receive. Finally, it might be determined if gender has any effect on a student's course grade.

Biography

Michael Lloyd received his B.S in Chemical Engineering in 1984 and accepted a position at Henderson State University in 1993 shortly after earning his Ph.D. in Mathematics from Kansas State University. He has presented papers at meetings of the Academy of Economics and Finance, the American Mathematical Society, the Arkansas Conference on Teaching, the Mathematical Association of America, and the Southwest Arkansas Council of Teachers of Mathematics. He has also been an AP statistics consultant since 2002.