Statistical analysis is behind a push to discontinue use of the exam for admission to the University of California.

Is the SAT I Exam Irrelevant?

MARK SCHILLING
California State University, Northridge

The SAT examination is the key instrument, along with high school performance, that for several decades has determined admission to the majority of four-year institutions of higher learning in the United States. But the exam is increasingly under fire on two counts.

First, the College Board and the Educational Testing Service (which creates and administers the exam) have consistently held that the exam is an objective measure of "innate" aptitude; consequently applicants cannot improve their scores by taking a test preparation course. Yet the success of Stanley Kaplan's courses and the Princeton Review argue otherwise, and increasing numbers of students are enrolling in such courses in order to increase their chances of admission.

The second and perhaps more critical charge is that the SAT, particularly the SAT I exam, is a poor predictor of college success. Significantly, the President of the University of California and a key faculty committee have each recently suggested discontinuing the use of SAT I scores in admission decisions. Their rationale is based on statistical evidence—specifically, information from a *multiple regression analysis* of data obtained from tens of thousands of applicants to UC. In this article I will give a short primer on regression and then present some of the results and implications of the UC study.

Regression

The primary statistical meaning of the word "regression" is *prediction.* Given a set of data on two or more variables, the goal is to be able to predict as well as possible the value that one variable (the *response variable*) may have, given the values of the remaining variables (the *explanatory variables*). Prediction is normally done by the *method of least squares.* Certain equations are solved to find the unique function f (typically linear) of the explanatory variables, known as the *regression function*, for which the sum of the squared differences between each of the values of the response variable in the data set and their values as predicted by f is as small as possible. Here is an illustration:

Student	1	2	3	4	5	6	7	8
X	70	80	62	90	45	30	65	50
Y	35	30	41	28	60	80	55	50

Table 1. Math skills scores (X) and math anxiety scores (Y) for eight students.

Eight students were given two tests, one measuring their math skills, the other measuring their level of math anxiety. Their scores are shown in Table 1. Possible scores for each test are 0 to 100. Choosing anxiety (Y) as the response variable and skills (X) as the explanatory variable, a regression line was computed to predict Y from X. The results are shown in the scatterplot below.

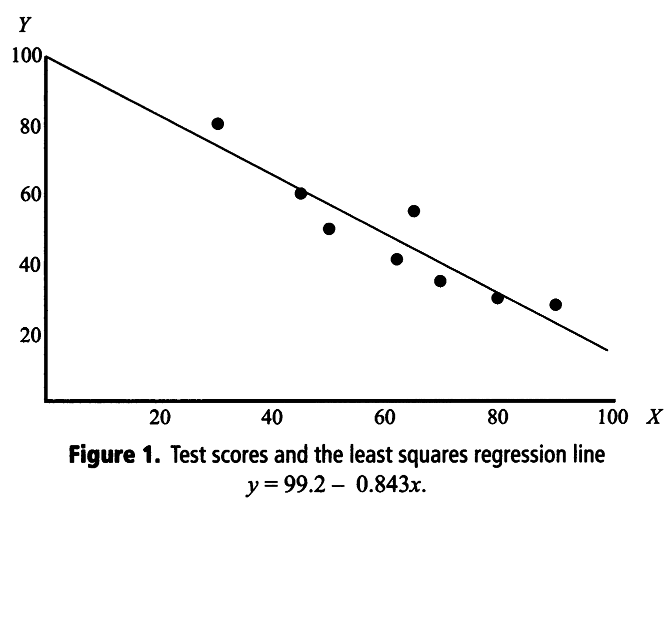


Figure 1. Test scores and the least squares regression line $y = 99.2 - 0.843x$.

The regression line makes some very sensible predictions: A student who scores 0 on the skills test is predicted to have nearly maximal anxiety ($Y = 99.2$), while a student who scores 100 on the skills test should have very little anxiety ($Y = 14.9$).

For the eight students in the study, the predictions are good but not perfect (since the points are not exactly on the line). The fundamental measure of how well the predictions reflect the actual response variable data is known as the *correlation coefficient* (or *coefficient of determination*) R^2. R^2 represents *the proportion of variation in the response variable values that is eliminated by the regression*. Understanding this tricky concept is essential to understanding the SAT study results presented below, and regression in general.

To illustrate, let's look first at how much variation there is in the original Y data above. Figure 2a shows the Y values projected onto the Y-axis. The variance of these eight values is

$$\frac{1}{7}\sum_{i=1}^{8}(y_i - \bar{y})^2 = 44.08.$$

If we think of the regression line as partially "explaining" (predicting) the Y values, then what is left unexplained is the set of vertical discrepancies between each of the eight points and the line. For example student #1, who scored 70 on the skills test, has a discrepancy of $35 - (99.2 - 0.843 \times 70) = -5.2$, meaning that the regression line overpredicted the student's anxiety score by 5.2 points. These discrepancies, known as *residuals*, are plotted in Figure 2b. Note that there is much less variation in the residuals than in the original y values. The variance of these residuals turns out to be 5.87, thus $R^2 = 1 - (5.87/44.08) = .867$. That is, 86.7% of the variation in math anxiety is "explained" by the differing levels of math skill that the subjects had. Clearly math skill is a critical predictor of a person's level of math anxiety.

Along with R^2, the other key end product of a regression analysis is the *regression coefficients*, which are simply the partial derivatives (slopes) of the regression function with respect to each of the explanatory variables. In the example above there is only one regression coefficient, -0.843. Its interpretation is this: suppose, for example, that one student scores 10 points higher on the skills test than another student. Then the regression analysis predicts that the first student will score 8.43 points lower on the anxiety test than the second student.

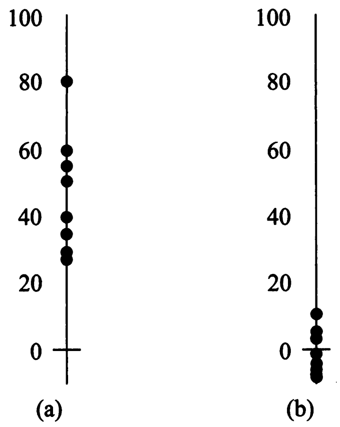


Figure 2. (a) Math anxiety scores. (b) Residuals after regressing on math skills scores.

The UC SAT Regression Study

The University of California requires all prospective applicants to take both the standard SAT I math + verbal exam and the SAT II. The SAT II consists of three achievement tests—one in writing, one in mathematics, and one in a third subject of the student's choosing. UC researchers employed *multiple* regression analysis, meaning there were several explanatory variables involved, to learn how well the SAT tests and high school grades predict the response variable, freshman GPA. The conclusions of their study are based largely upon looking at the effect on R^2 of including/excluding each of the explanatory variables in the model. The data came from the almost 78,000 first-time freshmen who entered any of UC's eight undergraduate campuses between Fall 1996 and Fall 1999. The regression results are shown in Table 2. Note that all of the R^2 values are small. This indicates that it is difficult to predict a college student's freshman GPA (at least from high school grades and SAT scores, which are considered to be among the best predictors available). The low R^2 values also turn out to be partly a consequence of the fact that only students with high grades and SAT scores tend to apply to the rather selective UC system.

Now look only at the last two rows of Table 2. Currently UC utilizes both SAT I and II as well as high school grades to determine admission. However, Table 2 shows that excluding the SAT I exam decreases R^2 only from 22.3% to 22.2%, a trivial reduction. That is to say, once high school grades and the SAT II tests are taken into account, the SAT I score that a student achieves does not appear to help predict the student's level of success as a college freshman.

Explanatory variables	R^2
HSGPA	15.4%
SAT I	13.3%
SAT II	16.0%
SAT I + SAT II	16.2%
HSGPA + SAT I	20.8%
HSGPA + SAT II	22.2%
HSGPA + SAT I + SAT II	22.3%

Table 2. Values of R^2 for predicting freshman GPA from all combinations of explanatory variables.

Regression analyses are a bit like good works of art or music—they must be viewed from many different angles before their full meaning can be appreciated. In regression, variables can form a complex interplay with each other, and those *not* included in the model (*lurking variables*) can hide fundamental relationships or give a false impression of the actual relationship between variables that *are* in the model.

With these facts in mind, the UC researchers looked deeper into the data in several ways. They looked at the results for each individual campus. They compared the model for the different racial and ethnic groups and different majors that students were intending. They looked for differences according to the high school of origin. The results in each case were quite consistent: Once high school grades and the SAT II are taken into account, the low importance of the SAT I exam applies across the board.

The researchers also analyzed an expanded model that includes two new explanatory variables, the logarithm of the student's family income and a measure of the parent's level of education. The purpose of this analysis was to assess the role of socioeconomic factors. Adding new explanatory variables necessarily increases the value of R^2. For this new model, however, R^2 only rose to 22.8%, indicating that socioeconomic variables are unimportant once the other variables have been accounted for.

Table 3 gives standardized regression coefficients for each variable in this new model. Standardized regression coefficients are a modification of the ordinary regression coefficients that show the number of standard deviations the predicted response changes for each one standard deviation change in an explanatory variable when all other variables are held constant.

The clear conclusion from Table 3 is that under this model high school grades and the SAT II are by far the best predictors of freshman college grades. Once these two factors have been taken into account, the importance of SAT I and the socioeconomic factors is negligible. Based on these results UC researchers were able to conclude that when all other factors are held constant, a 100 point increase on the SAT II exam adds about .21 grade points to predicted freshman GPA, whereas a 100-point increase on SAT I has almost no effect. They also showed through additional analyses that the SAT II is a fairer test in that it is less sensitive than SAT I to differences in family income and parents' education.

HSGPA	.28
SAT I	.02
SAT II	.24
Log of Family Income	.03
Parents' Education	.06

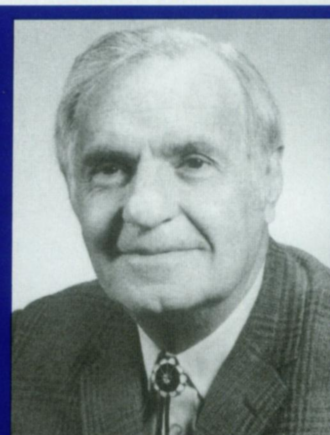
Table 3. Standardized Regression Coefficients for the Expanded Model

Even before this study, the University of California had decided to give less weight to the SAT I exam and more weight to SAT II in the eligibility index used for admission to UC. Further research will investigate response variables other than freshmen grades, such as student persistence and graduation rates and cumulative GPA at graduation. But now the evidence is strong enough that both the UC President and a systemwide faculty panel have called for the complete elimination of SAT I scores from the admissions process. And as the 800-pound gorilla among U.S. colleges, UC has gotten the attention of the College Board, which now plans to make significant changes in the SAT I beginning in 2005. These include adding a writing examination which may be similar to the writing test that is currently part of the SAT II. ■

References

Saul Geiser with Roger Studley, *UC and the SAT: Predictive Validity and Differential Impact of the SAT I and SAT II at the University of California*, University of California, Office of the President, October 29, 2001.

Murray S. Klamkin Retires as Problems Editor



Problems, it has been said, are the lifeblood of mathematics. Indeed, mathematical progress is driven by people consumed by the desire to solve problems. A mathematics magazine without a problem section is inconceivable. This particular magazine has been blessed by the presence of Murray S. Klamkin as Problems Editor since the September 1995 issue. In that time Murray has assembled over 200 problems (a great many of which he created himself), read thousands of submitted solutions, and produced a column we have been proud to publish. Murray's goal, in his own words, has been to produce problems elegant in statement, elegant in result, and elegant in solution. He has accomplished this even though he's been constrained by *Math Horizons*'s goal of being accessible to nearly everyone. Murray is retiring as *Math Horizons*'s Problem Editor with this issue. We can't thank him enough for his years of devoted, and, yes, elegant service to *Math Horizons*'s readership.

Steve Kennedy and Deanna Haunsperger, Editors