Michigan Law Review

Volume 63 | Issue 8

1965

Predicting Court Cases Quantitatively

Stuart Nagel
University of Illinois

Recommended Citation

Stuart Nagel, *Predicting Court Cases Quantitatively*, 63 MICH. L. REV. 1411 (1965).
Available at: https://repository.law.umich.edu/mlr/vol63/iss8/6

PREDICTING COURT CASES QUANTITATIVELY

*Stuart Nagel**

THIS article illustrates and systematically compares three methods for quantitatively predicting case outcomes. The three methods are correlation,[1] regression,[2] and discriminant analysis,[3] all of which involve standard social science research techniques. Two prior articles[4] have generated requests for a study dealing with the problems involved in handling a larger number of cases and predictive variables. The present article is also designed to provide such a study. It does not presuppose that the reader has read the earlier articles, although such a reading might help to clarify further some of the points made here. The cases used to illustrate the methods consist of 149 civil liberties cases decided by the United States Supreme Court from 1956 through 1960. The list of cases was obtained from a series of articles written by Sidney Ulmer and Glendon Schubert.[5] Technical aspects have been eliminated from the body of this article, leaving a simple explanation that should be sufficient to enable the non-technical reader to employ the methods in his legal research.

---

* Assistant Professor of Political Science, University of Illinois; member of the Illinois Bar; currently a Senior Specialist at the East-West Center of the University of Hawaii.—Ed.

The author is very grateful to Dorothy Anna Shipley for allowing him to use some of the data she compiled for her bachelor's thesis. Predictive Generalizations in Civil Liberties Cases, August 1963 (in the University of Illinois Library). Thanks are also owed to John Gilbert, statistical consultant at the Behavioral Sciences Center, for his helpful comments.

1. For further details than this article provides on correlation analysis, see Nagel, *Applying Correlation Analysis to Case Prediction*, 42 TEXAS L. REV. 1006 (1964); and Nagel, *Using Simple Calculations To Predict Judicial Decisions*, Practical Lawyer, March 1961, p. 68.

2. For further details on regression analysis, see BLALOCK, SOCIAL STATISTICS 273-358 (1960); GUILFORD, FUNDAMENTAL STATISTICS IN PSYCHOLOGY AND EDUCATION 365-72, 390-434 (1956).

3. For further details on discriminant analysis, see COOLEY & LOHNES, MULTIVARIATE PROCEDURES FOR THE BEHAVIORAL SCIENCES 116-33 (1962); TINTNER, ECONOMETRICS 95-102 (1952).

4. See the articles cited in note 1 *supra*.

5. Schubert, *The 1960 Term of the Supreme Court—A Psychological Analysis*, 56 AM. POL. SCI. REV. 90, 98 (1962); Ulmer, *A Note on Attitudinal Consistency in the United States Supreme Court*, 22 INDIAN J. POL. SCI. 195, 201 (1961); Ulmer, *Scaling Judicial Cases*, 4 AM. BEHAVIORAL SCIENTIST 31, 32 (April 1961); Ulmer, *The Analysis of Behavior Patterns on the United States Supreme Court*, 22 J. POLITICS 629, 649 (1960); and Ulmer, *Supreme Court Behavior and Civil Rights*, 13 WESTERN POL. Q. 288, 297-99 (1960).

I. The Methods

A. *The Basic Procedure*

All three methods rely on the relationship between case outcomes and various predictor variables. In this article, the outcome to be predicted is whether a given civil liberties case will be decided in the direction of narrowing civil liberties or in the direction of broadening civil liberties. The civil liberties cases deal with freedom of speech and religion, civil liberty aspects of criminal procedure, and equal protection under law. The predictor variables relate to the law, evidence, prior events, and other matters relevant to the cases and are set out in Table 1.

TABLE 1
The Weights for 8 Variables in 149 Civil Liberties Cases

Variables of the Cases	Weights of the Variables		
	Correlation	Regression	Discriminant
1. Having been decided negatively in the lower court rather than positively	.07	.23	.61
2. Having originated in the South, West, or District of Columbia, rather than the East or Midwest	.35	.32	.85
3. Involving state, local, private, or military action rather than federal civilian action	.11	.20	.55
4. Mainly involving legislative, police, or administrative action rather than chief executive, regulatory agency, or judicial action	.20	.12	.31
5. Mainly involving 1st, 4th, 6th, or 8th amendments rather than article 1, or 5th or 14th amendments	.19	.24	.64
6. Mainly involving free speech or equal protection rather than freedom of religion or criminal procedure matters	.13	.04	.11
7. Mainly involving murder, theft, fraud, assault, or no crime rather than contempt or other crimes	.21	.15	.39
8. Involving amicus curiae briefs by a combination of pressure groups or none rather than just one pressure group	.19	.31	.83
Minimum *S* Score (equals 0 if minimum *X* always equals 0)	0	0	0
Maximum *S* Score (equals sum of weights if maximum *X* always equals 1)	1.45	1.61	4.30
a-coefficient		—.49	
Cut-Off Score (equals average *S* Score of the cases for correlation and discriminant analysis; equals .5 for regression analysis if outcome is a 0 vs. 1 dichotomy)	.859	.500	2.734

Each of the three methods applies a variation of the following formula to the individual case in order to predict its outcome:

$$S = (W_1 \cdot X_1) + (W_2 \cdot X_2) + \cdots + (W_n \cdot X_n)$$

In the formula, X_1 represents the number of units of variable one that are present in the case being predicted; X_2 represents the number of units of variable two present; and so on to X_n, the last variable. If a variable only provides for two degrees of units (*e.g.*, absent *versus* present), its X can have only a value of zero or one. The W_1 represents the numerical prediction weight or importance of variable one; W_2 represents the weight of variable two; and so on to W_n, the last weight. The S represents the score obtained by summing the WX products in the formula.[6] If there are two possible case outcomes involved (*e.g.*, narrowing civil liberties *versus* broadening), a summation score less than a certain amount (the cut-off level) indicates a narrowing decision is more likely to be reached; a score greater than a certain amount indicates a broadening decision is more likely to be reached.

B. *The Raw Data*

The three methods differ mainly in their manners of calculating the weights of the variables employed in the above formula. All three methods, however, require the same kinds of data in order to calculate the weights. The data consist of how each one of a set of past cases is positioned with regard to the outcome and with regard to each of the variables thought to be relevant for prediction.[7] The categories on the outcome variable and on each of the predictor variables should either be dichotomies or be capable of being arranged in a meaningful ascending or descending order. Region of the litigation-provoking incident (*e.g.*, East, Midwest, West, South) is a variable that generally must be dichotomized (*e.g.*, East and Midwest *versus* West and South).[8] Age of the defendant is a variable that can be arranged in equal one-year or ten-year categories.

In correlation analysis, the values of the predictor and outcome variables should range only from zero (absent) to one (present), unless a variation called partial correlation analysis is used. In regression analysis, both the predictor and outcome variables can have

6. S is a general symbol. The S for correlation analysis is symbolized T, and the S for discriminant analysis is symbolized Z. The S for regression analysis plus the a-coefficient described later is symbolized Y.

7. There is nothing practical to be gained by reducing the variables to their underlying factors, since one can more efficiently predict with the variables themselves. FRUCHTER, INTRODUCTION TO FACTOR ANALYSIS (1954).

8. *But see* notes 14-15 *infra* and the accompanying text.

any range including millions, decimals, and negative numbers. In discriminant analysis, the predictor variables again can have any range, but the outcome variable is usually dichotomized (zero and one), although discriminant analysis can, if necessary, handle more than two groups.

C. *Assigning Weights to the Variables*

1. *Correlation Analysis*

For the weight of a given predictor variable, correlation analysis uses the correlation coefficient (symbolized r) between the variable and the outcome. When both the predictor variable and the outcome variable provide for only two categories or degrees apiece, the correlation coefficient is approximately equal to (1) the percentage of cases *positive* on the predictor variable that are also positive on the outcome variable minus (2) the percentage of cases *negative* on the predictor variable that are positive on the outcome variable.[9] Correlation analysis is simpler, but somewhat less accurate than the other two methods. It is simple enough to be readily usable without a computer, although a desk calculator would be helpful. But, if one has access to a computer, it would generally be wasteful to do a less accurate correlation analysis instead of a regression or discriminant analysis. The correlation analysis method was designed for the practicing lawyer or legal scholar who does not have easy access to a computer.

2. *Regression Analysis*

For a given weight, regression analysis uses what is called the unstandardized partial regression weight (symbolized b) between the predictor variable and the outcome. The user of this method need not know how to calculate a b-weight because there are standard computer programs available at most computing centers that can quickly and accurately calculate the b-weights for a set of variables in a set of cases. It is enough to know that, given the data, regression

---

9. See Nagel, *Testing Empirical Generalizations in Legal Research*, 15 J. LEGAL ED. 365, 372 (1963). One is led to erroneous results if one merely uses the percentage of cases *positive* on the predictor variable that are also positive on the outcome variable (without considering the percentage of cases *negative* on the predictor variable that are also positive on the outcome variable) as was done in JURY VERDICT RESEARCH CORP., HOW TO PREDICT PERSONAL INJURY VERDICTS 26 (1963). Their prediction formula in effect is $S = (P_1/P)(P_2/P) \ldots (P_n/P)(P)$, where P equals the percentage of cases won by the affirmative position in the total sample and where P_1, P_2, . . . and P_n equal the percent of cases won by the affirmative position when variable 1, 2, . . . or n is present. S thus supposedly equals a percentage or empirical probability of victory. This approach is inconsistent with the Bayes probability theorem. See MOSTELLER, ROURKE & THOMAS, PROBABILITY WITH STATISTICAL APPLICATIONS 143-50 (1961).

analysis will yield predictions that statistically minimize error in what statisticians call a "least squares sense." This roughly means that, if each case were plotted as a point on a multi-dimensional graph with the outcome on one axis and the predictor variables on the other axes, then the regression weights would determine a line through these points such that the sum of the squares of the deviations from the points to the line would be minimized.

As a highly simplified example of how regression analysis assigns a weight to a variable, assume one is trying to use regression analysis to predict the height (the outcome variable) of a specific person (the case at bar) from his width (the predictor variable). Suppose further that the only basis for prediction is the height and the width of three prior persons (the precedents). What regression analysis in effect does by numerical formulas is draw a two dimensional graph with various units of height (starting with zero) marked on the left side of the graph and various units of width marked on the bottom. The height and width of each prior person is then shown as a dot drawn over from the appropriate marking on the left side and up from the appropriate marking at the bottom. Then, one slanted line (called a regression line) is drawn as close to all three dots as is possible, with overall proximity measured by the sum of the distances squared from each point to the line. If a right triangle is drawn anywhere along this regression line so that all or part of the line forms the diagonal side, the *b*-weight is the ratio between the length of the vertical side of the triangle and the horizontal side of the triangle. The closer this ratio is to 1.00, the closer is the relation between the two variables. The *a*-coefficient represents the number of units up from the O point on the graph (*i.e.*, the intersection of the left side and the bottom) to the point where the regression line crosses the left side of the graph.

3. *Discriminant Analysis*

For a given weight, discriminant analysis uses a value (symbolized *k*) which maximizes the difference between (1) the *S* scores of the cases decided in a positive direction and (2) the *S* scores of the cases decided in a negative direction. Computer programs for discriminant analysis are now increasingly available at computing centers. The discriminant approach generally provides the most accurate predictions in the sense of quantity right divided by predictions made. However, when all the predictor variables and the outcome variable have been dichotomized, it produces practically the same predictions as does regression analysis or a variation of correlation analysis called partial correlation analysis.[10]

Again a highly simplified example will serve to illustrate the method of assigning a weight to a variable. Suppose one is trying to use discriminant analysis to predict the sex of a person from his or her width. As before, assume that the only bases of prediction are the sex and width of three prior persons. What discriminant analysis in effect does by numerical formulas is to find a number (called a discriminate weight) by which to multiply the width of each prior male. The average of these products is then determined. The width of each female is also multiplied by the number, and the average of those products is determined. The number or weight is ideal for discriminant analysis if there is no other number that will produce a wider divergence between the average of the male products and the average of the female products.

To use a computer program for regression or discriminant analysis, one punched card per case is needed unless the number of variables necessitates the use of more. Certain columns on each card should be set aside for each variable. Thus, if hole 1 is punched on column 12 of the card corresponding to case 23, this punch might indicate that a certain variable was present. If hole 3 is punched on column 19 of the same card, this punch might indicate that case 23 fell into interval 3 on the amount of medical expenses claimed by the plaintiff. Once the cards are punched by a typewriter-like keypunch, the personnel at the particular computing center being used can quickly inform the user how to combine his deck of data cards with a deck of standard program cards in order to obtain the kind of printed output shown in the columns of Table 1.

Both regression and discriminant analysis provide weights for each variable, while the other variables are statistically held constant or partialed out. Thus, in regression and discriminant analysis (but not ordinary correlation analysis), the weight of a given variable depends upon what other variables are included. If two similar variables are included, at least one will have a lower weight than it would have had if only that one had been included.

D. *Determining the Cut-Off Level*

The three methods also differ with regard to how they determine what *S* score is likely to lead to a negative decision and what *S* score is likely to lead to a positive decision. Both correlation analysis and

10. BLALOCK, *op. cit. supra* note 2, at 343-46; Rulon, *Distinctions Between Discriminant and Regression Analysis*, 21 HARVARD EDUCATION REV. 80, 89 (1951).

discriminant analysis use the average S of the cases as a cut-off score. Any future case having an S greater than this average is likely to be decided in a positive direction; similarly, if the S value is less than the determined average, the case is likely to be decided in a negative direction.[11]

In regression analysis, if the S plus an a-coefficient is greater than or equal to 0.5 (*i.e.*, closer to 1 than to 0), then a positive decision is more likely; the converse is true if the S plus a is less than 0.5 (*i.e.*, closer to 0 than to 1). Since any computer program that does regression analysis will also supply the appropriate a-coefficient, the object of regression analysis is to predict the outcome (Y) value, which can be 0 or 1 if the outcome has been dichotomized. If the outcome value were height, regression analysis would give an answer in inches or feet depending on the unit of measurement used, whereas correlation and discriminant analysis would give an answer that would be a purely abstract number above or below the cut-off between being tall or short.

II. The Illustration

A. *Dichotomizing the Variables*

Table 1 shows the predictor variables that were used, the correlation, regression, and discriminant weights for each variable, and the cut-off scores for each method.[12] For the sake of simplicity, each variable was dichotomized by collapsing together into one category the categories in which a relatively high percentage of the cases were decided in a positive direction, and by collapsing together into a second category the categories in which a relatively high percentage of the cases were decided in a negative direction. Most computing centers have cross-tabulation programs that will quickly indicate the percentage of cases decided in a positive direction for each category on each variable, although these percentages can be calculated by hand. One should also attempt to create categories that are internally homogeneous and externally different from each other regardless of

11. In Nagel, *Applying Correlation Analysis to Case Prediction*, 42 Texas L. Rev. 1006, 1015 (1946), the cut-off score was calculated by adding the average S of the cases that were decided in a positive direction to the average S of the cases that were decided in a negative direction and by dividing this sum by two. Merely determining the average S of all the cases is simpler, and the difference in accuracy is very slight. Likewise, with a large sample of cases it is unnecessary to have an upper cut-off score (equal to the S score of the case immediately above the average S) and a lower cut-off score (equal to the S score of the case immediately below the average S).

12. The discriminant weights and the discriminant cut-off score were originally in thousandths, but they were multiplied by one hundred to eliminate the zeros to the right of the decimal point. The cut-off scores are carried to one decimal place more than the weights so as to prevent some summated S scores from exactly equaling the cut-off scores.

their relation with outcome. The dichotomous categories are also more useful if they contain approximately equal quantities of cases.

If there are *N* categories on a predictor variable, the quantity of ways in which the variable can be dichotomized can be readily calculated by applying the formula for determining the number of combinations that can be made out of *N* things taken *J* at a time.[13] For example, if there are six categories on a variable, the six can be taken five at a time in six ways, four at a time in fifteen ways, and three at a time in twenty ways (half of which produce duplicate dichotomies). Thus a variable having six categories can be dichotomized into six plus fifteen plus ten, or thirty-one different ways. Only one dichotomy per variable can be fed into the correlation analysis. However, all the dichotomies can be fed into a regression or discriminant analysis if the computer program is capable of initially taking that many input variables and if the program is capable of throwing out the dichotomies it determines to be weak in relative predictive power. It is more efficient, however, not to feed the regression or discriminant analysis any dichotomies that are obviously relatively weak.

If desired, the dichotomizing process can be avoided by using only variables that are natural dichotomies (*e.g.*, sex) or that have ascending or descending categories (*e.g.*, social class) or by using the Glueck prediction method, which is, in effect, a method for weighting categories on predictor variables rather than a method for weighting the variables themselves.[14] One can also avoid the dichotomizing process with the correlation method by ranking the categories from the one having the highest percentage of victories for the affirmative outcome down to the one having the lowest percentage. The *X* value of a case with regard to such a variable then equals $(C\text{-}R) \div (C\text{-}1)$, where *C* equals the number of categories (*e.g.*, region was previously mentioned with four categories) and *R* equals the rank of the category in which the case is located. For example, if the region variable with four categories has a correlation weight of 0.11 with outcome and if the categories are associated with outcome in the order of East, West, Midwest, and South, then a case from the East would receive an *X* in the basic summation formula equal to $(4\text{-}1) \div (4\text{-}1)$ or 1. A case

13. The formula is N!/J! (N-J)! where N! (*i.e.*, N factorial) means 1 times 2 times 3 on up to N, and likewise with J!. ADAMS, INTERMEDIATE ALGEBRA 295, 336-38 (1960).

14. The Glueck prediction method is discussed in GLUECK & GLUECK, PREDICTING DELINQUENCY AND CRIME 18-32 (1959) and in Nagel & Goodman, *Judicial Prediction With Multinominal Variables* (forthcoming), which compares and discusses in detail the non-dichotomizing prediction methods briefly mentioned in this paragraph and elsewhere.

from the West would receive an X of (4-2) ÷ (4-1) or 0.67; a case from the Midwest would get a 0.33; and a case from the South would get a 0. The basic summation formula involves summing the product of the W and X for each variable in order to determine the total points of each case.[15] It should also be mentioned that the dichotomizing process can even be avoided while the regression or discriminant method is used by substituting for the value of X the percentage of cases falling into the affirmative outcome position for each category on each variable.

B. *Inapplicable and Missing Information*

The weights and cut-off scores in Table 1 were calculated on the basis of 149 of the 175 civil liberties cases on which data was gathered.[16] The original group of 175 cases had to be reduced to 166 because 9 of the cases could not be objectively classified as having broadened or narrowed civil liberties. The 166 cases were then reduced to 149 because for 17 cases complete information was not available or the categories were inapplicable for some of the variables. As an alternative to eliminating all 17 cases, one or more variables could have been eliminated, or missing information for a case could have been replaced with the average category on the variable involved, or the category "unknown" could have been collapsed into a positive or negative category depending on the percentage of positive decisions among the cases falling into the unknown category on the variable. The technique chosen for each variable should be the technique that will maximize the sample size, will retain powerful variables and non-duplicative variables, and will make no unreasonable assumptions.[17]

15. An appropriate correlation coefficient in this context is a contingency coefficient. SIEGEL, NON-PARAMETRIC STATISTICS FOR THE BEHAVIORAL SCIENCES 196-202 (1956). To give this coefficient a zero to 1.00 range, it should be divided by 0.71, because it can be shown algebraically that, when the outcome variable has been dichotomized, 0.71 is the maximum contingency coefficient that can be obtained. Another appropriate coefficient in this context is Cramer's multinominal coefficient. CRAMER, MATHEMATICAL METHODS OF STATISTICS 282 (1946).

16. The 175 cases do not contain all the civil liberties cases decided by the Supreme Court from 1956 through 1960, because the articles from which the list was obtained excluded unanimous cases (see note 5 *supra*), although a small, roughly random sample of unanimous civil liberties cases were found in the court reports and added to the original list bringing it to 175 cases. This incompleteness, however, does not affect the nature of the methods described in this article, and it probably does not substantially affect the weights shown in Table 1. Nevertheless, if additional unanimous cases had been used, the rate of correct predictions would probably have been substantially higher since the outcome of unanimous cases is generally easier to predict than the outcome of sharply divided cases.

17. By "powerful variable" is meant a predictor variable that has a high correlation with the outcome variable. By "non-duplicative variable" is meant a variable that has a low correlation with the other predictor variables. Good variables, in addition

C. *Why the Variables Seem To Be Relevant*

Variable 1 seems to be relevant because of a tendency by the Supreme Court to use its certiorari power more to reverse than to affirm lower court decisions. Variable 2 may be relevant because civil liberties violations may be greater in the South and West and because the Supreme Court may apply higher standards to the District of Columbia. Likewise, civil liberties violations may be greater at the state, local, private, and military level than at the federal civilian level (explaining variable 3) and greater in the legislative, police, and other administrative spheres than in the executive, regulatory agency, or judicial spheres (explaining variable 4); or, it may be that the Supreme Court has a more negative bias toward the former levels and spheres in comparison with the latter ones.

Variables 5 and 6 are closely related, although fourteenth amendment cases might cover all the categories in variable 6. Apparently, the Supreme Court was more solicitous of violations of free speech and equal protection than of violations of freedom of religion and criminal procedure during the base period. When a crime was involved, it seems that the pro-civil liberties position was more likely to win when the crime was more serious and the defendant had more at stake. The last variable may be relevant because the most important civil liberties cases in terms of the interests at stake have the most amicus curiae briefs filed and nearly all are filed by pro-civil liberties groups. When only one pressure group was involved, it was generally the American Civil Liberties Union attempting to influence a criminal procedure case.

D. *Applying the Results of the Analysis*

To illustrate how the data in Table 1 is applied, take the 1954 school desegregation case of *Brown v. Board of Education*[18] as an example. Applying the formula previously described for correlation analysis, *Brown* had all the variables present except variable 5, giving it a summation score of (.07) (1) + (.35) (1) + (.11) (1) + (.20) (1) + (.19) (0) + (.13) (1) + (.21) (1) + (.19) (1), which equals 1.26 out of a maximum of 1.45 and a minimum of zero. Since this summation score exceeds the cut-off score of 0.859, one would predict that *Brown* would be decided in a broadening direction, as it unanimously was. Similarly, *Brown* gets a score of 1.37 using the

to being powerful and non-duplicative, should be ones on which data can generally be found previous to the making of the decision being predicted.

18. 347 U.S. 483 (1954). Since *Brown* was decided prior to 1956, it was not one of the cases used to calculate the weights.

regression weights, and this S plus the a-coefficient of $-.49$ exceeds the 0.500 cut-off score. Likewise, *Brown* gets a discriminant S of 3.66, which is greater than the 2.734 cut-off score.

III. THE UTILITY OF QUANTITATIVE CASE PREDICTION

Table 2 indicates that ninety-five per cent of the cases like *Brown* having a discriminant S that falls into the interval 3.65 to 4.20 were decided in favor of the broadening position. Knowing the rough probability of victory in cases before the Supreme Court might be helpful in rationing scarce resources or revising the briefs of a law firm, a pressure group, or the solicitor general's office, although even if one has a case that falls into the extreme intervals in Table 2, it may still be worth participating in an appeal if the gain to be achieved in case of victory is enough to offset the low probability of victory, or if there is some special characteristic present that indicates the probability of victory is much higher than calculated (*e.g.*, the other side had no standing to sue).

The seven intervals in Table 2 all contain as equal a quantity of cases as is possible, in spite of many tied scores and the fact that 149 is not evenly divisible by seven. A more detailed table could be created using more than seven intervals, although the denominator of the percentages may become too small to give meaningful percentages. Theoretically, as many intervals or classes as there are combinations of characteristics among the cases could be provided.[19] A similar table and similar intervals could likewise be created for the regression scores or for the correlation scores. If the variables, the prediction method, and the clerical work were perfect, then all the intervals up to a score of 2.734 (the cut-off score) would have zero per cent of their cases decided in a broadening direction, and all the intervals beyond 2.734 would have one hundred per cent of their cases so decided.

---

19. For each combination, one could also indicate a predicted percentage of victory (as well as an actual percentage) by applying the regression formula with the a-coefficient. See Tanenhaus, Muraskin, Rosen, & Schick, *The Supreme Court Certiorari Jurisdiction—Cue Theory*, in JUDICIAL DECISION-MAKING 111, 129 (Schubert ed. 1963). As a simpler alternative, one could create a single regression equation from the data contained in a table such as Table 2. GUILFORD, *op. cit. supra* note 2, at 365-72. The data shown there, for instance, yields the equation $P = -.36 + .33\ S$, where P is the probability of victory for the broadening position and S is the summation score of the case being predicted. Such an equation, however, is a summarizing device; thus it may lose some of the detail provided by the percentages in Table 2, especially if the percentages ascend unevenly rather than in a smooth incline.

TABLE 2

PERCENTAGE OF VICTORY AT VARIOUS LEVELS OF DISCRIMINANT SCORES

Interval of Discriminant Scores	Number of Cases in Interval	Percentage of Cases Decided in Broadening Direction
.61 to 1.84	19	11%
1.86 to 2.20	22	23
2.38 to 2.69	21	48
2.70 to 2.83	23	48
2.85 to 3.24	22	68
3.26 to 3.56	23	78
3.65 to 4.20	19	95

"Predicting" cases from which the weights were calculated is, of course, really postdicting. It is, however, interesting to note that, given the variables and cases used, the correlation method correctly postdicted 104 cases or 70 per cent of the 149 cases; the regression analysis correctly postdicted 109 or 73 per cent; and the discriminant analysis was correct on 111 cases or 74 per cent of the 149. If a method cannot postdict reasonably well, it is unlikely to predict reasonably well.

By chance alone, one could only predict approximately 75 cases (or 50 per cent) accurately. By always predicting the most frequent outcome (which was a broadening decision), one could only predict 79 cases accurately or 53 per cent. Obtaining 111 correct postdictions and 38 incorrect ones, as the discriminant analysis did, could occur purely by chance only about once in a billion times.[20] A similar accuracy probability can be determined for the correlation and regression analysis postdictions.[21]

The postdiction power, and thus the prediction power, might have been increased even further if the author had used additional or different variables, a narrower set of cases, a greater proportion of unanimous cases, or more precise categorization. Nevertheless, this demonstration allows the conclusion that the use of quantitative prediction of court cases *plus* traditional prediction techniques is probably better than the use of the latter alone.

20. Such a probability can be determined by calculating $(111\text{-}38)^2/149$ and then reading the probability corresponding to this quotient in the first row of a chi-square table. GUILFORD, *op. cit. supra* note 2, at 238, 540. If the above quotient goes beyond the maximum of the chi-square table, then determine the square root of this quotient and read the probability corresponding to this square root from a large normal-curve table. *Id.* at 534.

21. If the discriminant analysis were used on the 149 cases and flipping a coin were used on the 26 eliminated cases, then approximately 124 cases (*i.e.*, 111 plus 13) or 72 % of the 175 would be postdicted correctly. One can, of course, improve on the coin-flipping technique by predicting a broadening outcome on the 26 eliminated cases or by replacing the missing information for 17 of the 26 cases with the average category or other reasonable estimate on the predictor variables involved.