Indiana Law Journal

Volume 41 | Issue 3

Article 3

Spring 1966

# Judicial Prediction and Analysis From Empirical Probability Tables

Stuart S. Nagel
*University of Illinois*

# JUDICIAL PREDICTION AND ANALYSIS FROM EMPIRICAL PROBABILITY TABLES

STUART S. NAGEL†

A previous article[1] described how some statistical techniques could be applied to synthesizing judicial precedents and to predicting the outcomes of future cases. For simplicity of presentation, the characteristics or variables in the precedent cases were made into either-or dichotomies. For example, the region of the country where the litigation-provoking incident occurred was dichotomized to read "originated in (1) the South, West, or District of Columbia, rather than (2) the East or Midwest." It is the purpose of this article to describe some related techniques that are particularly applicable to working with variables that consist of more than two non-numerical categories. The present techniques not only represent an improvement in broadened applicability but also in predictive power, simplicity, and in utility for building a behavioral science of law. The article presupposes no prior reading, and technical matters have been relegated to the footnotes and to an appendix.

The set of cases used to illustrate the methodology consists of 137 disputes involving international relations in which the United States was a party. These 137 cases represent all cases of the above type which are included in four leading international law casebooks.[2] The field of international law was used because of its potential social importance and because it is an important field of law that has had almost no quantitative analysis. The outcome to be predicted in these illustrative cases is whether the United States will lose or win. In the sample of 137 cases, the United States won 65 percent of the time.

---

† Associate Professor of Political Science at the University of Illinois; currently on leave to serve as Assistant Counsel to the U.S. Senate Judiciary Committee, but this article is not part of his Senate work. The author is very grateful to Martin Goodman for allowing him to use some of the data he compiled for his bachelors thesis, The United States as a Party in International Law Disputes, 1964 (in the University of Illinois Library). The author is also grateful to John Gilbert, the statistical consultant at the Center for Advanced Study in the Behavioral Sciences, for his many helpful comments made while the author was a fellow at the Center. Thanks for helpful comments are also owed to the members of the 1965 seminar on judicial prediction at the U.C.L.A. Law-Science Research Center.

1. Nagel, *Predicting Court Cases Quantitatively*, 63 MICH. L. REV. 1411 (1965).

2. BISHOP, INTERNATIONAL LAW (1952); HUDSON, INTERNATIONAL LAW (1951); KATZ & BREWSTER, INTERNATIONAL TRANSACTIONS & RELATIONS (1960); ORFIELD & RE, INTERNATIONAL LAW (1955).

## I. How to Use Empirical Probability Tables

Tables 1 and 2 are empirical probability tables in that they show for given sets of cases what percentage of the cases actually resulted in victory for the United States. For example in the set of 53 cases mentioned at the top of Table 1 in which United States domestic law was the main body of law being applied, the United States won 85 per cent of the time. If winning or losing a case were like flipping a coin, the *a priori* or the by-chance probability of victory as contrasted to the empirical probability would always be 50 percent. *A priori* probabilities, however, have little value in judicial prediction or analysis unless they are used to determine how much an empirical probability or other statistic deviates from chance.

In Table 1 each of the seven variables or characteristics on which the cases are positioned is designated by an arabic numeral. Each of the categories or positions on the variables is designated by a small letter. Only categories into which some of the 137 cases fell are included in Table 1. Thus on variable 6 relating to the industrial power of the opponent of the United States, one could provide a category for cases in which the opponent had greater industrial power than the United States had, but there were no such cases in the sample of 137 although a larger and more historically-oriented sample might have had such cases. The categories are supposed to be mutually exclusive so that no case can fall into more than one category on the same variable. The variables in Table 1 are arranged in the order of the correlation weight (symbolized $r$) that was assigned to each variable by a method that will be described shortly. The categories on each variable are arranged in the order of their respective victory percentages.

### A. *For Prediction Purposes*

Table 2 shows the percentage of cases decided in favor of the United States when the cases are grouped into seven approximately equal sets depending on what summation score each case has received. To calculate an unweighted summation score for a given case, one merely determines into what category the case falls on each of the seven variables, and then sums the seven corresponding empirical probabilities or percentages.[3]

3. To make Table 1 more usable it should be supplemented by more detailed descriptions of the categories so as to make the positioning of future cases easier and more objective. Because of the methodological nature of this article, the full substantive details are not provided here although some elaboration is made at various points in this article. One can find greater substantive detail in the bachelors thesis mentioned above which is available on loan from the University of Illinois Library.

TABLE 1. THE EMPIRICAL PROBABILITY OF VICTORY FOR THE UNITED STATES FOR EACH CATEGORY ON EACH VARIABLE

| Variables, Weights, and Categories (Arranged in rank order) | Quantity of Cases in Category | Percent of Cases Won by U.S. in Category |
|---|---|---|
| 1. Main Source of Law (r = .35, K = .24) | | |
| a. U. S. domestic law | 53 | 85% |
| b. International law and custom | 63 | 56 |
| c. Treaty | 17 | 47 |
| d. Domestic law of other country | 4 | 25 |
| 2. Main Subject Matter & U. S. Position (r = .30, k = .27) | | |
| a. National territory and U. S. seeking change | 3 | 100 |
| b. National territory and U. S. seeking status quo | 1 | 100 |
| c. Jurisdiction and U. S. seeking it | 26 | 85 |
| d. Hostilities bet. countries and U. S. repelling interference | 6 | 67 |
| e. Diplomatic or consular intercourse and U. S. seeking redress | 3 | 67 |
| f. Individual in int. law and U. S. seeking curtailment of his rights (e.g., deportation and alien rights) | 26 | 65 |
| g. Responsibility of other country asserted by U. S. generally for damages | 17 | 65 |
| h. Individual in international law and U. S. seeking to expand or maintain his rights (e.g., U. S. representing citizen in suit) | 20 | 60 |
| i. Treaty and U. S. seeking broad definition | 10 | 60 |
| j. Diplomatic and consular intercourse and redress sought from U. S. | 2 | 50 |
| k. Responsibility of U. S. asserted generally for damages | 15 | 47 |
| l. Treaty and U. S. seeking narrow definition | 7 | 43 |
| m. Hostilities between countries and U. S. seeking to interfere or obtain reparations | 1 | 0 |
| 3. Decision Making Tribunal (r = .27, k = .17) | | |
| a. U. S. Supreme Court | 41 | 76 |
| b. Lower federal court | 32 | 69 |
| c. Ad hoc tribunals | 56 | 61 |
| d. British court | 2 | 50 |

TABLE 1. (Continued)

| Variables, Weights, and Categories | Quantity of Cases in Category | Percent of Cases Won by U.S. in Category |
|---|---|---|
| e. Non-British foreign court | 3 | 33 |
| f. Hague | 3 | 0 |
| 4. Economic Interests and U. S. Position (r = .24, k = .25) | | |
| a. U. S. liberal in anti-trust cases | 9 | 100 |
| b. U. S. liberal in debtor-creditor cases | 3 | 100 |
| c. U. S. conservative in business v. labor cases | 1 | 100 |
| d. No economic interests involved or U. S. position neutral | 119 | 61 |
| e. U. S. conservative in debtor-creditor cases | 5 | 60 |
| 5. Civil Liberty Interests & U. S. Position (r = .23, k = .38) | | |
| a. U. S. liberal in non-criminal procedure matters | 4 | 100 |
| b. U. S. liberal in criminal procedure | 6 | 83 |
| c. U. S. conservative in non-criminal procedure matters | 10 | 80 |
| d. No civil liberty interest involved or U. S. position neutral | 99 | 65 |
| e. U. S. conservative in criminal procedure | 16 | 44 |
| 6. Industrial Power of U. S. Opponent (r = .21, k = .20) | | |
| a. Same | 6 | 100 |
| b. Not countries | 70 | 70 |
| c. Less | 66 | 56 |
| 7. Nature of Plaintiff (r = .19, k = .34) | | |
| a. State in U. S. | 1 | 100 |
| b. Alien in U. S. | 10 | 80 |
| c. Resident of foreign country | 11 | 73 |
| d. U. S. itself | 75 | 68 |
| e. U. S. citizen | 14 | 57 |
| f. Business firm | 18 | 50 |
| | 8 | 50 |

TABLE 2. THE EMPIRICAL PROBABILITY OF VICTORY FOR THE UNITED STATES FOR EACH INTERVAL OF SUMMATION SCORES

| Interval of Summation Scores | Number of Cases in Interval | Percentage of Cases Decided in Favor of the U. S. |
|---|---|---|
| I. Unweighted Scores | | |
| 3.27 to 4.17 | 18 | 22% |
| 4.18 to 4.31 | 19 | 37 |
| 4.32 to 4.41 | 21 | 62 |
| 4.42 to 4.61 | 19 | 68 |
| 4.62 to 4.76 | 20 | 75 |
| 4.79 to 5.03 | 20 | 85 |
| 5.05 to 5.84 | 20 | 100 |
| II. Weighted Scores | | |
| .903 to 1.112 | 19 | 16% |
| 1.127 to 1.156 | 19 | 42 |
| 1.158 to 1.177 | 20 | 60 |
| 1.179 to 1.222 | 21 | 62 |
| 1.223 to 1.263 | 20 | 85 |
| 1.266 to 1.347 | 20 | 90 |
| 1.348 to 1.580 | 18 | 100 |

For example, in 1949 a creditor of the I. G. Farben Company brought suit in Switzerland to stop the confiscation of I. G. Farben's Swiss assets by the United States and the Allied Powers unless payment were made to I. G. Farben's legitimate creditors.[4] In terms of the seven variables in Table 1, the case involves the domestic law of another country (P = .25), the U. S. seeking reparation (P = 0),[4a] a non-British foreign court (P = .33), the U. S. taking a position that neither favors the creditor nor the debtor (P = .61), no civil liberty interest (P = .65), a non-country opponent of the U. S. (P = .70), and a resident of a foreign country as the plaintiff (P = .73). The sum of these percentages is 3.27. If tables 1 and 2 were available prior to the I. G. Farben decision, one would say that the case had less than a .22 probability of being decided in favor of the United States as is indicated in the top row of Table 2. The case incidentally was decided against the position of the United States.[5]

For a more accurate prediction method, one can sum the products of

4. German Assets in Switzerland (I. G. Farben) Case, 45 SCHWEIZERISCHE JURISTEN-ZEITUNG 341 (1949).

4a. Where a category has a victory percentage of zero or 100, one should check the relevant judicial opinions and legislation to determine whether the judges or legislators have indicated that cases in that category are always supposed to be decided in a certain direction regardless how the cases are positioned on the other variables.

5. The unweighted summation method is essentially the same as the prediction system used by Sheldon and Eleanor Glueck for predicting delinquency and other forms of criminal behavior. GLUECK & GLUECK, PREDICTING DELINQUENCY AND CRIME 18-32 (1959).

each percentage times the discriminant weight (symbolized *k*) of each variable instead of merely summing the percentages. Using that alternative, the I. G. Farben case would receive a summation score of (.24) · (.25) + (.27) · (0) + (.17) · (.33) and so on summing to .903. According to Table 2, if an international dispute case like those in the sample of 137 has a weighted summation score of .903, the United States has less than a 16 percent chance of winning. If the score were 1.130, the United States would have between a .16 probability and a .42 probability of winning.

If all 137 cases are given weighted summation scores and are listed in rank order, then predicting a victory for the United States in all cases with a score greater than 1.157 (and a defeat in all cases with a lower score) would have resulted in 105 accurate predictions and 32 inaccurate ones. A cut-off score of 1.157 represents the best cut-off from the point of view of minimizing inaccurate predictions of victory or defeat. If the unweighted scores and a cut-off based on them had been used, 32 inaccurate predictions would have also resulted although weighting the variables normally improves predictability especially if there are only a few variables or only a few categories per variable.[6]

B. *For Analytic Purposes*

In addition to being potentially useful for prediction purposes, an empirical probability table like Table 1 can perhaps be even more useful for analyzing or understanding why a set of cases was decided the way they were. A good predictive method is generally a good analytic method although not necessarily. An example of a conflict between the two perspectives is a prediction method that is based on predicting the most frequently occurring outcome. If, for instance, the United States had won 90 perecnt of the cases and the future were like the past, one could predict accurately nine times out of ten by always predicting victory for the United States, but such a method provides no understanding of why some cases are won by the United States and why some cases are lost. It says nothing about what the predictor variables are or what their relative importance is.

How to weight the variables also depends partly on whether one is mainly interested in prediction or analysis. The k-weights or discriminant weights in Table 1, for instance, indicate the importance of each

6. GUILFORD, FUNDAMENTAL STATISTICS IN PSYCHOLOGY & EDUCATION 422-25 (1956). Weighting the seven variables in this example, was also less important in view of the fact that the seven variables were chosen from a larger set of variables by a method called stepwise regression analysis that substantially decreases the presence of overlapping variables. See the text material following note 14 *infra*.

variable primarily for predictive rather than analytic purposes. By using such weights as multipliers in calculating summation scores, one can minimize prediction errors. The discriminant weights, however, have no meaning out of the context of the seven predictor variables used. This is so because each discriminant weight is lowered by the presence of other partially overlapping predictor variables. On the other hand, the r-weights or correlation coefficients do indicate the extent to which each variable correlates with victory for the United States irrespective of what other variables are included.

Correlation coefficients are especially helpful for indicating the direction (positive or inverse) as well as the intensity (high or low) of the relation between a predictor variable and case outcomes where the categories on the variable have some inherently meaningful ascending or descending order. Intervals of length or time as used by physicists have such a character, but the kinds of variables most relevant to analyzing court cases are not so measureable. Since the correlation coefficients in Table 1 do not indicate what categories are associated with victory and what categories are associated with defeat, one has to determine this by looking at the category percentages under each variable.

An examination of the percentages tends to reveal that there are two important underlying elements that determine whether the United States will win or lose in this set of 137 international cases. One element has to do with playing on one's home grounds. Thus the United States is much more likely to win where American domestic law is the main body of law being applied rather than international custom, treaty, or foreign law. Likewise the United States is more likely to win before the United States Supreme Court and lower federal courts than before international tribunals or foreign courts. In this sense granting broader jurisdiction to international tribunals may be partly detrimental to the United States, but the net result may still be highly beneficial if some conflicts are thereby resolved that might otherwise lead to international hostilities.[7]

The other important underlying element has to do with the liberalism of the position taken by the United States as is shown in variables 4 through 7. The United States thus fares better when it is taking the liberal position in cases involving economic interests (*i.e.*, a pro-antitrust, pro-debtor, or pro-labor position), or the liberal position in cases involving civil liberties (*i.e.*, pro-safeguards for the innocent, pro-equal

---

7. For a discussion of the desirability of granting broader jurisdiction to international and foreign tribunals, see A.B.A. SECTION OF INTERNATIONAL & COMPARATIVE LAW, REPORT ON THE SELF-JUDGING ASPECT OF THE UNITED STATES DOMESTIC JURISDICTION RESERVATION TO ITS ADHERENCE TO THE STATUTE OF THE INTERNATIONAL COURT OF JUSTICE (1959).

protection, or pro-freedom of speech or religion). Likewise the United States does better when it has in past history faced an opponent of approximately equal industrial power rather than lesser industrial power. Along related lines the United States does slightly better when it is the plaintiff than when it is being sued.

Some of the effects of taking a more liberal position also show up on variable 2 although many of the percentages on that variable are unmeaningful because of the small number of cases on which they are based, which illustrates a disadvantage of providing too many categories on a variable. Nevertheless if one concentrates on just those categories having 10 or more cases, it appears that the United States is more likely to win (1) if it is seeking to establish rather than reject the jurisdiction of some tribunal, (2) if it is seeking to collect damages rather than seeking to avoid paying them; and (3) if it is seeking a broad interpretation rather than a narrow interpretation of a treaty. The United States fares about equally well when it is (1) curtailing an individual's rights in a deportation or alien proceeding before an American tribunal or (2) expanding an individual's rights by suing on his behalf before a claims commission or other tribunal. The equality in percentages between these two categories is possibly due to the fact that one involves the home-grounds element and the other involves the liberalism element.

Being on its home grounds favors the United States possibly because of relatively pro-American attitudes in American domestic law and among American judges. Taking a liberal position possibly helps the United States because of the liberalism of the United States Supreme Court since the 1930's and the possible liberalism of international tribunals. These explanations for the probability patterns observed may not be the best of all possible explanations. The important thing, however, is that the patterns of behavior revealed by an empirical probability table like Table 1 might not have been observed by merely briefing the cases in the traditional law school way.[8]

## II. How to Produce Empirical Probability Tables

In order to produce tables like tables 1 and 2 either by hand or with

8. Instead of ordering the categories of a multiple-category variable in terms of the extent to which each category is associated with victory for the United States, Hubert Blalock advocates ordering the categories in terms of the extent to which each category is associated with the underlying explanatory elements. Blalock's type of ordering thus emphasizes analysis rather than prediction. Ordering the categories in terms of their association with the outcome variable as is done in Table 1, however, does help to reveal what the underlying explanatory elements are. See Blalock, Nominal Scales, Generalizations to Populations and the Formulation of Scientific Laws (1965) (a paper presented at the annual meeting of the American Sociological Association).

the aid of computers, one must first determine (1) what set of cases are to be used as precedents, (2) what variables and categories on the variables might be useful in predicting or explaining the alternative outcomes that one is interested in, and (3) what category each case belongs in on each variable. For these three initial steps one mainly needs a knowledge of traditional legal research techniques, a little creative imagination, and some careful clerical work. The clerical work will of course benefit from having a few people checking on the work of each other in order to obtain a consensus on the categorizing done.

On the matter of picking the cases and the variables, a predictive goal might conflict with an analytic goal. For example, if one wanted to predict the outcome of a case involving an international boundary dispute one would probably only use other boundary dispute cases as precedent cases not a sample of 137 diverse international law cases. If, however, one wanted to analyze a variety of international law cases for the purpose of inferring some broad generalizations, then one would probably not confine himself to boundary disputes. Likewise for predictive purposes exclusively legalistic variables might be the most appropriate ones; but for theoretical-analytic purposes, one might want to see the relation to outcome of some sociological, psychological, or political variables.

A. *By Hand*

After each precedent case has been positioned on each variable, the processing of the data differs depending on whether or not one has access to a computer. If one does not have access to a computer, it is recommended that only the unweighted summation method be used unless there are only a small number of variables and cases. One index card can be set aside for each case. On the card can be indicated how each case was positioned on each variable using whatever symbols are convenient. For each variable, the cards can be sorted into as many piles as there are categories on the variable. Then one can count the total number of cards in each pile and the subtotal of cards which were decided in an affirmative direction on the outcome variable. By dividing this subtotal by the total for each pile, one can determine the victory percentage for each category and thus the contents of a table like Table 1.

After a table like Table 1 is created, one can use it along with the classification symbols on each index card to determine the summation score for each case which can also be written on the index card. The cards can then be rearranged in rank order from the case with the lowest summation score to the one with the highest. Next the cards can be divided into approximately seven piles as equal in number as is possible in spite of ties and the fact that the number of cases may not be evenly divisible by seven.[9] For each such pile of cards, one can count the total number of cards and the subtotal of cards decided in an affirmative direction. Dividing this subtotal by the total for each pile yields the victory percentage for each summation interval and thus the contents of a table like Table 2.

B. *With a Computer*

If a computer is available, then one can use one IBM-type card per case instead of one index card. On each IBM card set aside one column for each variable and punch hole 1 on the column if the case falls into category 1 and hole 2 if it falls into category 2 and so on. If a variable has more than 10 categories (which is more than the maximum number of usable positions on an IBM-card column), then use two columns for that variable and for a case in category 11 punch hole 1 on the first column and hole 1 in the second column and likewise for category 12 and so on. The variables and categories can be in any order. Somewhere on each card a case identification number can also be punched. This process will result in as many data cards as there are cases unless there are so many variables that one needs more than one data card per case.

To obtain the victory percentages for each category for making a table like Table 1, one merely puts the data cards along with some appropriate control cards in the input box at whatever computer center he has access to. In order to know how to prepare the control cards, one needs to consult whatever statistical users manual is available at the computing center. For example, at the Stanford University Computing Center where the widely-used *BMD* manual is available,[10] one would place five control cards in front of the data cards and two control cards behind them. The first two cards are different forms of user identification cards. They mainly indicate the users name and the fact that he wants the computer to apply the BMDO2S cross-tabulation program to his data cards. The next control card punched in accordance with the instructions in the manual mainly indicates to the computer how many data cards and how many variables it is to process. Then comes a control card that indicates how many categories are present on each variable followed by a

9. Equal piles are used to avoid having some of the denominators become too small to yield statistically meaningful percentages and to avoid having others become too large for differentiating among the many cases within the group. With 137 cases, seven piles seem appropriate although one could use fewer piles with fewer cases or more piles with more cases.

10. BMD—Biomedical Computer Programs (1964). This manual and its computer programs are available at many university computing centers.

control card that indicates what columns have been set aside for each variable. Next come the data cards followed by a card which indicates that each predictor variable is to be cross-tabulated or correlated with the outcome variable not with each other. The final card simply has the word "finish" punched on it.

After one has obtained the percentages output and made a table like Table 1, then the same data cards can be resubmitted with a set of BMD09S transgeneration control cards designed to create a new data deck. The control cards for this run mainly indicate what victory percentages correspond to each category on each variable.[11] In the new deck which the computer will create, each card still corresponds to one case. Every two columns correspond to a different variable. The big difference between the new data deck and the old one is that if case 12 falls into category 3 on variable 7, then a 3 is not punched on the new deck in the columns set aside for variable 7. Instead the victory percentage for category 3 on variable 7 is punched.[12] This repunching, like the calculation of the percentages, can be handled by the IBM 7090 computer at Stanford in less than two minutes.

The final computer step involves submitting the new data deck along with a few simple control cards for a discriminant analysis run calling for a program like BMD04M. The program will not only give discriminant weights like those shown in Table 1, it will also calculate a weighted summation score for each case and then print out the cases arranged in the rank order of their summation scores along with how each case was decided.[13] From this rank order listing one can block off approximately seven sets of cases equal in number and then calculate the percentages needed for a table like Table 2. One can also look at this listing to determine what summation score represents the best cut-off level such that there are a minimum number of negatively-decided cases above the level and affirmatively-decided cases below the level.

As an alternative or supplement to the discriminant analysis step, one can run the new data deck through a regression program like BMD02R or a correlation program like BMD03D to obtain a different

11. For greater predictive accuracy the percentages can be indicated to more than two decimal places.

12. Where the outcome variable is a numerical score rather than a win-lose dichotomy, one can substitute the average outcome score of each category (rather than the percent of wins) for the category number.

13. If the new data deck is run through a Pearsonian product moment correlation program the resulting correlation coefficients are identical to those obtained by applying Cramer's V for multiple-category variables to the old data deck. For a description of Cramer's V, see BLALOCK, SOCIAL STATISTICS 230 (1960).

set of weights.[14] Prior to using the discriminant analysis program, one can have the data cards applied to a stepwise regression program like BMD02R which is designed to aid in reducing the number of variables. The stepwise program will rank the variables in terms of their relative potency in accounting for the variation in outcome among the cases and will indicate how much additional variation each additional variable accounts for. Additional variables after approximately the first half-dozen generally tend to be redundant. Personnel at the computing center can be helpful in setting up the control cards for any of these runs.

## III. Some Broader Implications

Empirical probability tables in combination with the traditional non-quantitative techniques developed by lawyers can probably be helpful in predicting the outcome of court cases especially if the tables are prepared and periodically updated in consultation with the lawyers who will be using them. Such tables are currently prepared on a subscription basis for personal injury lawyers by the Jury Verdict Research Service of Cleveland. Unfortunately, however, their tables do not provide any data on intervals of summation scores and the corresponding empirical probabilities. Their category percentages thus represent a mass of numbers that cannot be meaningfully integrated unless summation data is provided or unless the user is informed of the applicability of the Bayes prediction method.[15] Nevertheless it is heartening to see this kind of commercial venture into the realm of case prediction. Perhaps in time this firm and other firms will expand into other fields of law using a more sophisticated and more useful methodology.

The possibly more important implications of quantitative case analysis have to do with the development of a behavioral science of law. In developing such a science, it seems necessary at first to emphasize the development of scientific tools for the analysis of legal phenomena. The next stage logically seems to be the application of such tools to the testing of empirical generalizations. When a lot of such testing has been done a body of empirical propositions will be developed. In the third stage,

14. For an alternative set of weights, the percentages on the new data deck can be transferred into logarithims or squate roots by exercising certain transgeneration options in the BMD04M discriminant analysis program. This sometimes improves predictability particularly if the distirbution of percentage-scores on a predictor variable is unbalanced away from one side of the average percentage. See Walker & Lev, Statistical Inference 423-25 (1953).

For a more detailed description of discriminant, regression, and correlation weighting, see Nagel, *supra* note 1, and the references cited therein.

15. See the appendix to this article for a description of the Bayes method of prediction.

one can inductively build broader theories of law from these empirical propositions.

Although one can talk conceptually of three stages, in reality all three types of work can be going on simultaneously by different persons or by the same person in different studies or even in the same study. Methodology helps to test propositions and propositions help to generate theory, but theory can also stimulate hypotheses which when tested become propositions. Likewise the desire to test certain hypotheses may generate new methodologies. Nevertheless in the early development or redevelopment of a behavioral or social science, methodology may be emphasized more than it later is.

Over the last ten years a number of articles and books have appeared calling for a more scientific analysis of the legal process, and additional ones will probably appear in the next few years.[16] This article has advocated the use of empirical probability tables, correlation coefficients, and various summation methods for analyzing court cases. It should, however, be pointed out that the same techniques can be used to analyze such behavior as that of legislators, administrators, attorneys, and the general public. The techniques can also be applied to entities other than individuals such as countries and time periods. Nevertheless, it is hoped that the near future will see less emphasis on methodological discussions and more emphasis on the application of such techniques to the testing of empirical generalizations and to the building of inductive theories of the legal process.

## APPENDIX

### ALTERNATIVE PREDICTIVE AND ANALYTIC TECHNIQUES APPLICABLE TO MULTIPLE-CATEGORY VARIABLES

The purpose of this appendix is to discuss some alternative techniques for predicting with multiple-category variables other than the summation methods presented in the body of the article. This discussion is placed in an appendix because it is more technical than the rest of the article and because the techniques discussed are generally inferior to the summation methods with regard to both predictive power and ease of application.

The first alternative technique is the Bayes method.[17] It involves applying the following formula to the empirical probabilities of Table 1

16. SCHUBERT, QUANTITATIVE ANALYSIS OF JUDICIAL BEHAVIOR (1959); Cowan, *Decision Theory in Law, Science, and Technology*, 17 RUTGERS L. REV. 499 (1963); Nagel, *Testing Empirical Generalizations in Legal Research*, 15 J. LEGAL ED. 365 (1963).

17. MOSTELLER, PROBABILITY WITH STATISTICAL APPLICATIONS 143-50 (1961).

in order to arrive at a composite probability of victory for the United States or

$$P(W \text{ given } S) = \frac{P(W) \cdot P(S \text{ given } W)}{[P(L) \cdot P(S \text{ given } L)] + [P(W) \cdot P(S \text{ given } W)]}$$

where P = probability, W = win for U. S., L = loss for U. S., S = any set of categories. For example, suppose all one knows about a case is that it involves a treaty and the United States is seeking the jurisdiction of some tribunal. With this information the numerical value of the formula becomes:

$$.74 = \frac{(\frac{89}{137})(\frac{(17)(.47)}{89} \cdot \frac{(26)(.85)}{89})}{(\frac{48}{137})(\frac{(17)(1-.47)}{48} \cdot \frac{(26)(1-.85)}{48}) + (\frac{89}{137})(\frac{(17)(.47)}{89} \cdot \frac{(26)(.85)}{89})}$$

The formula takes into consideration the fact that there were 137 cases, 89 of which were decided in favor of the United States and 48 of which were decided against the United States. The quantities 17 and 26 and the decimals .47 and .85 correspond to categories 1c and 2c. By analogy, one could apply the formula to a case where one knows how it is positioned on all seven variables rather than just two variables.

A big defect in this method is that it assumes all variables are independent or non-overlapping relative to each other. The weighted summation method, on the other hand, allocates a discriminant weight to each variable which statistically eliminates or partials out the overlap between the variables used. Another defect is the laboriousness of the calculations especially if many variables are used, although the calculations can be made easier and more accurate by using a desk calculator.[18]

Another alternative might be called the every-combination method. This method requires the preparation of an empirical probability table that shows the percentages of cases won by the United States in every possible combination of categories. For example if one were just working with variables 1 and 2, an empirical probability would have to be shown for cases that were in category *a* on 1 and *a* on 2; *a* on 1 and *b* on 2; and so on. Since there are four categories on variable 1 and thirteen categories on variable 2, there are 52 (*i.e.*, 4 times 13) combinations to

18. Given the assumption of independence of the variables, the Bayes method can be deductively proved to be valid. *Ibid.* The Jury Verdict Research Corporation has proposed a method for handling multiple category variables which at first glance looks like a version of the Bayes method, but the JVR method can be shown algebraically to be inconsistent with the Bayes method. See JURY VERDICT RESEARCH CORP., HOW TO PREDICT PERSONAL INJURY VERDICTS 26 (1963); and Nagel, *supra* note 1, at 1414.

be provided for with just the first two variables. Given the number of categories on each of the seven variables, 163,800 combinations have to be provided for if all seven variables are used. In effect the method converts seven variables having from three to 13 categories apiece into one variable having 163,800 categories.

An obvious defect in this approach is that with only 137 cases there are many combinations that have not occurred; and for those combinations that have occurred, the denominator of the percentages may be too small to give meaningful percentages. The percentages are also difficult to interpret for analytical purposes. In spite of these serious defects, a variation of this every-combination method has been proposed by James Coleman.[19]

John Sonquist and James Morgan of the University of Michigan have proposed a method related to the every-combination method.[20] Their approach is in effect a selective-combination method since their computer program generates a maximum of 63 combinations. The particular combinations that the program presents in its output are the combinations that are most different from each other with regard to their positioning on the outcome variable, with each combination being required to meet a minimum level of predictive importance. The combinations are worded in such a way that every case falls into one of them and only one of them. The computer program splits multiple-category variables (*e.g.,* variable 1) into one category and its opposite (*e.g.,* treaty cases versus non-treaty cases).

The Sonquist-Morgan approach is probably the best alternative discussed in this appendix. It is, however, substantially more difficult to apply than the summation method discussed in the body of the article. Indeed it is almost impossible to apply if one does not have access to their computer program plus an IBM 7090 computer that operates like the one at the University of Michigan. The output of the program is also more difficult to interpret, particularly for analytic purposes. Moreover, the program's ratio of accurate predictions to predictions made may be less than the weighted summation method possibly because of information lost in dichotomizing the variables. Their approach though is quite new and needs more use to be fully evaluated.

There are a number of alternative methods that might be useful for

19. COLEMAN, INTRODUCTION TO MATHEMATICAL SOCIOLOGY (1964) 219-24. His variation of the every-combination method also has the defect of laborious complexity, and it is able to handle only one multiple-category variable among the predictor variables.

20. Sonquist & Morgan, The Detection of Interaction Effects—A Report on a Computer Program for the Selection of Optional Combinations of Explanatory Variables (1964).

obtaining accurate predictions but which have little value for understanding why some cases are decided one way and why some are decided another way. Predicting the most frequently occurring outcome was previously mentioned as being such a method. Trend analysis is closely related since it in effect involves predicting the most frequently occurring recent outcome. Thus if in the series of 137 cases, the United States won the first 89 but has lost the last 48, then one might logically predict the United States will lose the next case. Predicting the most frequent outcome is an especially poor predictive technique if the most frequent outcome of two possible outcomes occurs only a little more than 50 percent of the time. Likewise projecting trends is an especially poor predictive technique if the trend line is flat such that the United States or other affirmative position has been consistently winning or losing approximately the same proportion of cases over time.[21]

Another predictive method that has little analytical or causal value is the judgmental method. It involves making a small or large survey of lawyers or other experts in the field of law involved in order to determine how they think a case or set of cases will be decided. Such a prediction method may be quite powerful depending on the expertise of those asked and the sophistication of the handling of the data,[22] but unless systematic probing questions are asked with regard to why the cases will be decided as predicted, this method is not likely to reveal anything with regard to the variables correlated with or responsible for the decisions reached.

Finally it should also be mentioned that there are a number of approaches to case prediction that are only applicable if the multiple-category variables are numerical (like age) or if they are converted into either-or

21. Bruce Jacobs of the Massachusetts Institute of Technology has devised a form of trend analysis that involves plotting over time the number of judges on the Supreme Court who agreed with a given argument in a given type of case. He then combines the trend lines for a number of arguments in order to predict the voting split on the Supreme Court for a case at a future point in time. His method in effect involves plotting decisions on issues rather than decisions on cases, but it reveals nothing with regard to what variables correlate with the alternative decisions on the issues plotted. Bruce Jacobs, A Quantitative Analysis of Supreme Court Decision-Making in Congressional Reapportionment Cases (unpublished manuscript).

22. One way to handle the data might involve presenting the experts with a description of the set of cases to be predicted and asking them to classify each case into five outcome categories. If the cases are international disputes in which the United States is a party, the outcome categories might be (1) U, S. is very likely to win, (2) U. S. is more likely to win than lose, (3) U. S. has an even chance, (4) U. S. is more likely to lose than win, and (5) U. S. is very likely to lose. Each case can be given an average category score (which is equal to the sum of the category numbers in which it was placed divided by the number of experts involved) or a more sophisticated scale score can be given to each case in accordance with the principles of psychometric measurement. GUILFORD, PSYCHOMETRIC METHODS 223-44 (1954).

dichotomies. This is true of such techniques as factor analysis,[23] simultaneous equations,[24] or Boolean algebra.[25] The information lost in dichotomizing may be of value to improved predictive power and especially to a theoretical interpretation of the elements responsible for the decisions. These three methods are also unduly complex. Boolean algebra, in addition, seems to be particularly incapable of yielding empirical probabilities. Thus it can only say that a future case will be won or lost rather than give a probability of victory like those associated with the summation intervals in Table 2, or with the Bayes method, or the selective-combinations method of Sonquist and Morgan.[26]

23. Kort, *Content Analysis of Judicial Opinions and Rules of Law*, Judicial Decision-Making 133-97 (1963).
24. Kort, *Simultaneous Equations and Boolean Algebra in the Analysis of Judicial Decisions*, 28 Law & Contemp. Prob. 143-63 (1963).
25. Lawlor, *Foundations of Logical Legal Decision Making*, 63 Modern Uses of Logic in L. 98 (1963).
26. Reed Lawlor is currently developing a technique that combines the summation approach with Boolean algebra so as to enable his Boolean algebra approach to yield empirical probabilities of victory.

## DOCUMENTS

*The Indiana Law Journal instituted a new department in the Spring Issue of 1951 (see 26 Ind. L.J. 348) entitled Documents. The stated purpose of the Documents section was to provide space to documentary material of legal and historical value which might not otherwise become available to scholars and general readers. In recent years the Documents section has been dormant but not forgotten. We here reassert our belief that an outlet for document items of interest related to legal, historical or modern problems is needed. We are therefore pleased to keep that idea alive by presenting the following Document.*

The Editors