Glossary of Terms and Techniques for Political Analysis

Approved For Release 1999/09/25 : CIA-RDP85T00875R001100020001-3

Preface

For some years a move has been underfoot to turn POLITICAL science into political SCIENCE. The traditional historical-descriptive approach has been labeled inadequate and sophisticated behavioral and management science tools have been brought to bear on political problems. In the rush toward science, jargon has proliferated. Mathematical, statistical, and computer terms, the holy water of the new era, are sprinkled generously throughout articles in the professional journals. At best the new political science vocabulary increases precision, at worst it hinders communication. Even when it is to be deplored, however, it cannot be dismissed because many of the leading figures in the discipline use this jargon and many of the important new works are couched in such language. Thus this glossary.

The glossary is divided into two alphabetic sections. The first section defines approximately 500 terms—a vocabulary list of words and phrases which are used by political scientists, particularly when describing many of the new techniques of political analysis. The second section explains approximately 100 techniques—methods or approaches for analyzing political data. The terms appear in lower case letters, the techniques are capitalized.

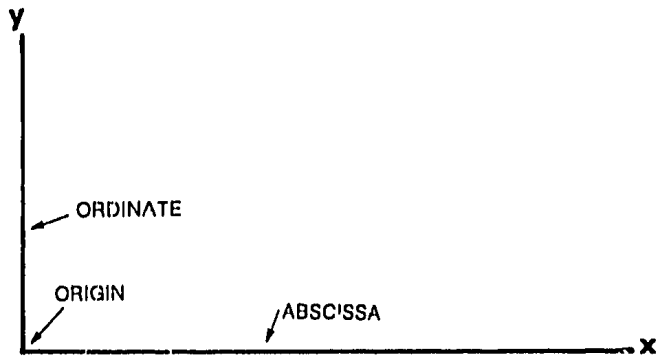
An effort has been made to be comprehensive and it is hoped that this glossary will ease the way through most political science books or journal articles. Comments and additions, however, should be sent to [redacted] or 25X1A
25X1A [redacted] of the Office of Current Intelligence, room 7G07, 143-5204.

TERMS

a posteriori--determined after the fact. An a posteriori probability is one determined by empirical observation. For example, experience tells us that a flipped coin lands heads one half of the time.

a priori—determined in advance. An a priori probability is set strictly by definition of the mathematical system. For example, a football coach could specify before the game that each of his two quarterbacks will play 50 percent of the game.

abscissa—the horizontal axis on a graph, designated X. (see: *ordinate*)



absolute value--the value of a number with the positive or negative sign omitted. Absolute value is identified by two parallel lines around the number. Thus /8/ (read "the absolute value of eight") and / –8/ (read, "the absolute value of negative eight") are both equal to 8. Absolute value measures the size of a number or movement, without measuring its direction. For example, $|6|-3+|-4|-|-5|=2$

acceptance region—one of two mutually exclusive regions in a sample space used in the testing of statistical hypotheses. If the sample point falls into one (the region of acceptance) the hypothesis is accepted as true; if it falls into the other, the hypothesis is rejected. (see: *sample space, hypothesis testing*)

accuracy—in statistics the closeness of computations or estimates to the exact or true values.

actor—any individual or social group that affects the decision process within a political system. As in the theater, the political actor "plays a role" and this role and attendant behavior can be described, analyzed, explained, and predicted for political systems at all levels—local, state, national, and international. The rational actor model focuses on the whole state as a single actor, interacting with other states in the international system. (see: *rational actor model*)

admissible strategy—a strategy with at least as good an overall outcome as any other possible strategy but which may be surpassed by a better strategy for some particular situation. (see: *GAME THEORY*)

ADP (automatic data processing)—the name used to denote the type of computer equipment which processes data automatically as opposed to manually. (see: *computer*)

aggregate data—data which summarizes the characteristics of a number of cases, such as trade or census figures, but provides no information on individual cases, such as separate trade transaction figures or specific counts of people per household. (see: *SURVEY RESEARCH*)

aggregation of interests—the process by which two or more political parties, interest groups, governmental bodies, or individuals combine their demands to build a consensus in support of a particular policy or political objective.

aggregative index—an index number which is constructed by combining a number of items as distinct from picking out a representative item. An aggregative price index, for example, gives the total of prices on all commodities as opposed to picking out certain representative commodities.

aggregative model—a model of a political or economic system in which the variables are constructed from groups of individual variables, as when an average value of military weapons is substituted for the actual prices of the weapons. (see: *model*)

algorithm—a set of mathematical rules for a given procedure; a sequence of statements defining a computer program.

alpha error—in hypothesis testing, the probability of rejecting a hypothesis when it is actually true. (see: *beta error*)

alphanumeric—a name or label which contains both letters and numerals used in computer programs.

alternative hypothesis—in the testing of hypothesis, any admissible hypothesis alternative to the one under test.

amplification—(see: *distortion*)

analog computer—a device which simulates some mathematical process or relationship so that the results of the process can be observed as physical quantities, such as voltage or current. (see: *digital computer*)

analysis—the methodical investigation of a problem, and the separation of the problem into smaller related units for more detailed study.

analytical model—a model in which interactions are expressed as mathematical equations. (see: *model*)

anomic group—largely spontaneous, temporary groups formed to express discontent through riots or demonstrations. (see: *GROUP THEORY*)

antilogarithm or antilog—the number from which a logarithm (or log) is derived. For example 100 is the antilog of log 2 since $10^2=100$. (10 is called the base.) (see: *logarithm*)

approach—a strategy of analysis which provides a unique viewpoint and a set of tools or methods for the examination of political phenomena. The psychological approach, for example, focuses on the influence of personality traits on political behavior and uses psychoanalytic and other techniques to study personality. (see: *method*)

argument—an independent variable (see: *independent variable*)

array—a table for the presentation of data in a row and column format. A correlation matrix, for example, is a two dimensional description of the association between variables. The intersection of row B and column 4 shows a correlation of .6 between these two variables. (see: *contingency table, matrix, variable*)

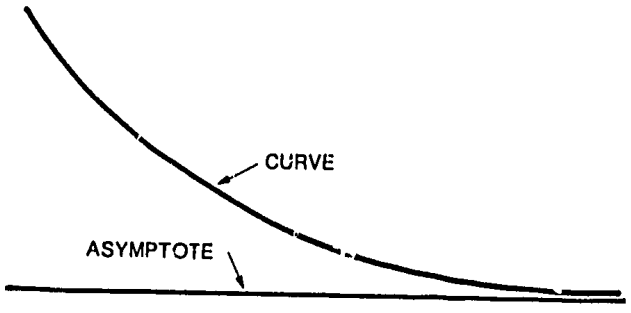
	1	2	3	4
A	1.00	.32	-.80	.22
B	.32	1.00	-.08	.60
C	-.80	-.08	1.00	-.10
D	.22	.60	-.10	1.00

articulation of interests—the making of a demand on a political system; the means by which attitudes, opinions, and beliefs are transformed into expressed demands for governmental action. In structural-functional analysis, the articulation of interests is recognized as a basic input activity in a typical political system. (see: *input, SYSTEMS ANALYSIS, STRUCTURAL FUNCTIONAL ANALYSIS*)

artificial intelligence—the capability of a machine to perform functions that are normally associated with human intelligence, such as reasoning, learning, and self-improvement.

association—the degree of dependence or independence which exists between two or more variables whether they are measured quantitatively or qualitatively. If two variables are dependent, one upon the other, a change in the value of one will effect a change in the value of the other. (see: *CORRELATION*)

asymptote—a straight line always approaching but never meeting a curve; tangent at infinity.



attitude scaling—measures and compares political and social attitudes by arranging the answers to written or oral questionnaires according to the intensity of feeling revealed on a scale of "strongly agree" to "strongly disagree." When attitude scales are applied to legislators and judges, for example, the results are used to identify voting blocs, the impact of party on voting decisions, and the reasons for inconsistent voting patterns. (see: *GUTTMAN SCALING*)

attribute—in political science, a particular kind of qualitative political variable that refers to the structural characteristics of nations, for example, the power of the executive. In statistics, a qualitative characteristic of an individual person or case that doesn't change—such as sex—as distinguished from a variable which can take different values at different times—such as age. (see: *variable*)

Aumann-Maschler theory—a type of game theory which does not attempt to predict which coalition will form but only to determine what the payoffs will be once a coalition is formed. The theory takes only the strengths of the players into account; all considerations of fair play and equity are put aside. (see: *GAME THEORY*)

authority—political influence derived from a willing acceptance by others of one's legitimate right to make rules or issue commands and to expect compliance. The authority relationship is voluntary, as contrasted with power which is based on the use of physical coercion or material resources. (see: *power*)

autocorrelation—a measure of the degree to which a series of numbers correlates with itself. It shows, for example, the amount of change over time measured by the consistency in a series of numbers, and indicates, therefore, whether the present values are significantly different from past ones. If a variable changes little over time, autocorrelation will be high. For example, the number of people who vote in a certain precinct will vary according to the weather, the issues, etc., but it may remain basically the same from year to year. In this case, autocorrelation will be high. (see: *CORRELATION*)

autonomous equations—in econometrics, an equation which describes the behavior of one particular group or sector of the economy. (see: *ECONOMETRICS*)

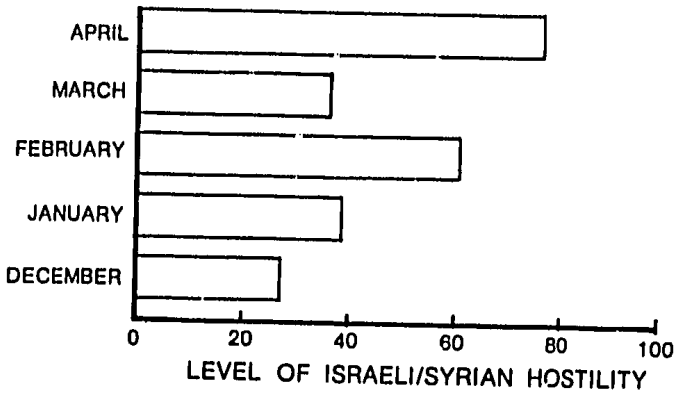
autoregression—the generation of a series of observations whereby the value of each observation is partly dependent upon the values of those which have immediately preceeded it.

average—(see: *mean*)

Bales' typology—a system designed for the general study of small groups which classifies observed behavior into positive reactions, attempted answers, questions, and negative reactions. (see: *GROUP THEORY*)

band—a group of circular recording tracks on a computer's data storage device, such as a drum or disc. (see: *computer*)

bar chart—a graphical representation of frequencies or magnitudes by rectangles drawn with lengths proportional to the frequencies or magnitudes concerned.



base—a number or magnitude used as a standard of reference.

base line—(see: *abcissa*)

base period—the period of time for which data used as the base of an index number, or other ratio, have been collected. It can be as short as one day, or as long as the average of a group of years.

Bayes Theorem—(see: *BAYESIAN ANALYSIS*)

behavior pattern—any regular or recurring form of human activity. Political behavior patterns may range from those based on largely internal, psychological responses (thought, perception, judgment, attitude) to overt, observable physical responses (voting, protesting, lobbying, campaigning). Human behavior patterns provide the basis for the scientific study of politics.

behavioral science—(see: *social science, BEHAVIORALISM*)

behaviorism—which should be distinguished from "behavioralism," is a formal school of psychology which holds that only overt, observable data in the form of stimuli and responses are meaningful for analysis, whereas mental factors such as attitudes, opinions, and beliefs are meaningless. (see: *BEHAVIORALISM*)

benchmark problem—a problem used to evaluate the performance of computers relative to each other. (see: *computer*)

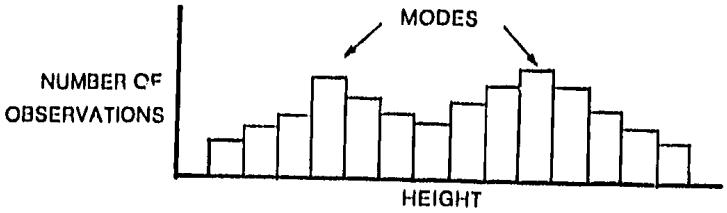
Bernouli Theorem—sometimes called the law of averages or law of large numbers, it states that as the number of trials of an event rises indefinitely, the probability of each observation will come to approach the established probability. For example, if you flip a coin long enough, heads will eventually occur 50% of the time, even if you begin with long irregular runs of tails.

beta coefficient—in statistics, the coefficient in regression equations, often denoted by the letter "B." The value of the beta coefficient is a measure of the influence that the independent variable has upon the dependent variable. (see: *coefficient, REGRESSION*)

beta error—when testing a hypotheses, an error incurred by accepting a hypothesis when it is actually incorrect. (see: *alpha error*)

bias—in statistics, any effect which distorts a statistical result totally as distinct from a random error which may distort on any one occasion but balances out on the average.

bimodal distribution—a frequency distribution in which two intervals have more observations than any other interval. For example, a distribution of people's heights may have two modes, one primarily of men and the other primarily of women. (see: *distribution, J-curve, mode*)



binary notation—a system of writing numbers using only 0 and 1. The numbering system we are most familiar with uses base 10. Thus the number 111 is read: one one, one ten, and one hundred. The number 1000 is read: no ones, no tens, no hundreds, and one thousand. Binary notation uses the base 2 and each column to the left of the decimal point is 2 raised to a successively higher power. The one's column in the normal number system is equivalent to 2^0, the ten's column is 2^1, the hundreds column is 2^2, the thousands column 2^3, and so on. Thus in binary notation, 111 is read: one 2^0, one 2^1, and one 2^2. Adding these, 111 is equal to the normal number 7 ($2^0=1+2^1=2+2^2=4$). Similarly the number 8 is written 1000 in binary notation: no 2^0, no 2^1, no 2^2, and one 2^3 ($2\times2\times2=8$).

binary sequence—any sequence, each number of which can take one of two possible values.

binomial distribution—the probabilities of getting a certain number of successes from a given number of trials. For example, the probabilities of getting 0, 1, 2, or 3 heads when a balanced coin is flipped 3 times are ⅛, ⅜, ⅜, and ⅛ respectively. This is found by listing all the possible combinations (eight) of the three flips and then counting how often 0, 1, 2, or 3 heads occur:

H	H	H	T	H	H
H	H	T	T	T	H
H	T	H	T	H	T
H	T	T	T	T	T

bipolar factor—in factor analysis, a factor which is positively correlated with some variables but negatively correlated with others. When such a factor is identified, it is regarded as expressing a property which may have a negative as well as positive intensity. For example, cowardice as opposed to bravery, cowardice being regarded as a quality in itself and not as the mere absence of bravery. (see: *FACTOR ANALYSIS*)

biserial correlation—a coefficient designed to measure the correlation between two qualities, one of which is measurable, the other a simple dichotomy according to the presence or absence of an attribute. A correlation measuring the strength of a government, for example, according to whether it is a military or non-military one, is biserial. (see: *attribute, CORRELATION*)

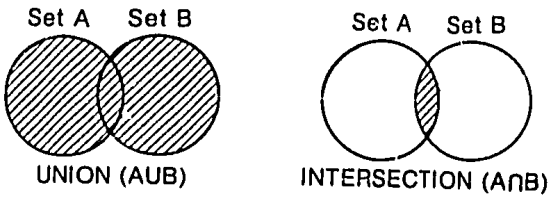
bit—an abbreviation for binary digit, which is a digit of a number written in the scale of two, usually using only 0 and 1. In computer language, the term also refers to a single piece of information stored and conveyed by an electrical impulse. (see: *binary notation*)

bivariate analysis—a generic term for any analysis which takes into account two variables. (see: *variable*)

block—any group of items under treatment or observation. A block may comprise a group of contiguous plots of land, all the votes of a given area, all the psychological characteristics of a given family. The general purpose of dividing all the material under study into blocks is to isolate the sources of differences.

block diagram—a diagram of a system, instrument, computer, or program in which selected portions are represented by annotated boxes and interconnecting lines. (see: *computer*)

Boolean algebra, Boolean logic—a branch of algebra in which the elements are sets instead of numbers, and a branch of logic in which the elements are sets instead of sentences. Boolean operations include union and intersection instead of addition or subtraction. (see: *Venn diagram*)



bootstrap—a technique or device designed to bring itself into a desired state by means of its own action. A machine routine, for example, whose first few instructions are sufficient to bring the remainder into the computer from an input device. (see: *computer*)

capability analysis—an assessment of a state's ability to achieve an objective vis-a-vis other states through the application of military, political, economic, psychological, and other forms of power and influence. Capability analysis provides decision-makers with the comparative data that will enable them to make a policy choice based on the relative feasibility of alternative courses of action.

card image—the equivalent amount of information on a magnetic computer tape that can be stored on one punched card. (see: *computer*)

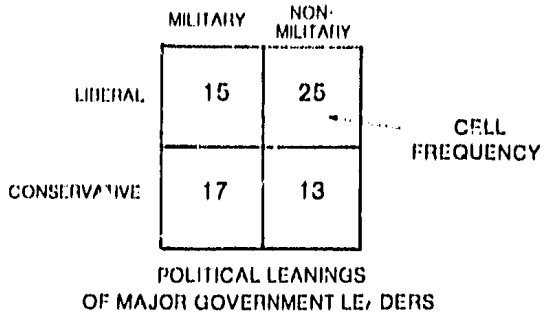
categorical distribution—a distribution in which the data are sorted into categories according to some qualitative description rather than by numerical value.

causal chain model—a large-scale economic model which looks at the changing pattern of the model's variables over time. (see: *model, variable*)

causal model—a simplified description of reality which isolates the causal relationships between a few variables—X, Y, and Z, for example. A causal model does not seek to prove that X necessarily causes Y, but merely that the variations in X can be used to explain the variation in Y. (see: *model, causality*)

causality—the relationship which links two or more political variables together in a cause-effect sequence to generate a particular event. Analysis of causality seeks to identify the antecedent action or change in one variable, called the independent variable, that produces or helps to produce a change in a second variable, the dependent variable. Causality can seldom be "proved" in the strictest sense, but it usually inferred from observed sequential relationships. (see: *independent and dependent variable*)

cell frequency—when a frequency distribution is classified into categories, the subcategories are called cells; the frequency with which observations fall into a particular cell is the cell frequency. (see: *distribution*)



central limit theorem—the theorem which gives the normal distribution its central place in the theory of probability and the theory of sampling. In its simplest form the theorem states that the variances of a given group of data will tend to be normally distributed—that is most of the cases will tend to cluster at a mid-point with fewer cases at either extreme—as the amount of data tends to infinity. (see: *distribution, normal distribution, variance*)

central processing unit—the unit of a computing system that includes the circuits controlling the interpretation and execution of instructions to the computer. (see: *computer, program*)

central tendency—a characterization of the "middle" value of a variable, or averages of various types which summarize the information contained in a set of data. (see: *measures of central tendency, variable*)

chain—a sequence of terms, values, or items, such that each item depends in some defined way upon the previous items in the series. The most common of these is the Markov chain. (see: *MARKOV ANALYSIS*)

chain index—an index number in which the value at a given period is related to a base in the previous period, i.e., Argentine political stability in 1972 compared with 1971 and 1973 compared with 1972, as distinct from an index number which is related to a fixed base, i.e., Argentine political instability in 1971, 1972, and 1973 related to 1968. The comparison of non-adjacent periods is usually made by multiplying consecutive values of the index numbers. For example, if the value of Argentine political instability, in 1973 based on 1971 is 500, and the value for 1971 based on 1968 is 400, the chain index for 1973 based on 1968 is $400 \times 500 = 2{,}000$ (divided by 100, since the index numbers were based on 100 as a standard). (see: *index number*)

chance—the quality of occurring randomly, or through unknown causes that makes a particular political event unpredictable. (see: *random*)

chance variation—variation in statistical observations due to the action of random, as distinct from systematic or predictable, factors. (see: *random*)

channel drift—(see: *distortion*)

Chapman-Kolmogorov equations—a set of equations used in the theory of stochastic (random) processes, giving the state of a system at a certain time in terms of the known states at previous times. (see: *stochastic, system*)

chi square—(pronounced "kye square") a test determining whether some observed distribution differs significantly from a purely random distribution. A chi square test involves algebraic computation followed by interpretation from a special chi square table found in the appendix to many statistics texts. Scores tell what percentage of the time the observed distribution could be reached purely by chance. If the data reaches a confidence level of .95, for example, we may be 95 percent certain that the observed distribution was caused by factors other than chance, because random events would have produced this type of distribution only 5 percent of the time. (see: *confidence level, significance level*)

circular formula, circular series—an operation which gives every item in a series an upper and lower boundary by making the first item also the last. For example, in a series of six terms there are only five differences between them. If, for analytical purposes it is desirable to have six differences, the first term is made a "pseudo" seventh term. This is equivalent to regarding the series as circular. Any resulting formula in the analysis is also said to be circular. The device is used in serial correlation analysis. (see: *CORRELATION*)

class, class boundaries, class interval, class frequency, class marks—the total number of items in any set of data may be grouped according to convenient divisions of the data to make subsequent analysis less tedious. Such a group is called a class. The highest and lowest values in each class are called the class boundaries, the difference between the highest and lowest values is the class interval, and the mid-point of the class intervals is the class mark. The number of items falling into each class is called the class frequency.

cluster sampling—a sampling technique which selects a specified number of random clusters, or groups of observations, for study. This sampling procedure is especially useful in cases where the cost of searching for samples is relatively high, but the cost of analyzing selected samples is relatively low. For example, in checking the condition of railroad ties, instead of sampling one tie in each of 1000 random locations, it may be easier and cheaper to select only 100 random locations and then inspect 10 ties in each. (see: *CLUSTER ANALYSIS, SURVEY RESEARCH*)

coding—the classification and sorting of political or social data into categories or ranges in order to simplify and improve analysis of the data. In content analysis, for example, the units of analysis (words, themes, whole message items) must be examined and assigned to categories such as friendly, moderate, or hostile. (see: *CONTENT ANALYSIS*)

coefficient—a multiplier. In 5Y (5 x Y), 5 is the coefficient and Y is the variable. (see: *continuous variable, discrete variable, variable*)

coefficient of determination—(see: R^2, *R-squared*)

coefficient of reproducibility—in the *GUTTMAN SCALING* technique, a way of measuring whether the scale will be analytically useful. A coefficient (correct responses divided by total responses) of .90 or better indicates that the scale has indicated a common attribute of all the items on the scale. This is regarded as an effective application of the Guttman technique. (see: *GUTTMAN SCALE)*

cognitive dissonance—a psychological term for the normal human tendency to reduce inconsistencies which may arise when new knowledge conflicts with values, environment, or behavior, by either changing these factors or by selectively ignoring the new facts. Cognitive dissonance theory is particularly useful in leadership and decision-making studies. Analysis of a leader's cognitive dissonance behavior patterns, for example, should reveal whether that leader would opt for negotiations or hostilities when confronted by a reviled traditional enemy with superior military capabilities.

cohort—an aggregate of individuals who have had some common experience during the same time period. (see: *COHORT ANALYSIS*)

common factor—in factor analysis, any factor which appears in two or more variables. If the factor appears in all variables, it is called a general factor. If it is common to a group of variables it is called a group factor. (see: *factor, FACTOR ANALYSIS, variable*)

complete system of equations—in econometrics, a term used to denote all the equations determining the behavior of an economic system or part of such a system. (see: *ECONOMETRICS*)

computer—an electronic/mechanical mechanism which can perform arithmetic operations, store information of various kinds, and compare information with other information. It performs exactly the instructions it is given, usually in the form of a program which is simply a list of those instructions. (see: *program*)

concept—a mental image or construct formed by generalizing from the characteristics of a class or category of things. A concept is an abstraction to which a descriptive label is attached, and the label may then be applied to individual members of the class to which the concept refers.

conceptual framework—a set of related concepts that provides an analytic scheme for political research and gives the analyst an overview of his research material, so that data can be systematically sorted, arranged, and examined. (see: *concept, methodology, approach*)

conditional probability—the probability of occurrence of one event, given that some other event has occurred or will occur. P(A/B) is read "the probability of A, given that B has occurred or will occur." For example, we may assess the probability that the Redskins will win this weekend (event A) as 80 percent, or P(A) = .8. But we may think that the probability of them winning if Larry Brown is unable to play (Brown's injury being event B) is lower, say, 60 percent or P(A/B) = .6. If, on the other hand, the events are independent, then they do not influence the probability of each other. In that case, the conditional probability of one, given the other, is still equal to the original probability of the first. For example, the probability that the Redskins will win (80%) is independent of whether or not I go to the game (my presence being event C.) So, P(A/C) = P(A) = 80% or .8. (see: *joint probability*)

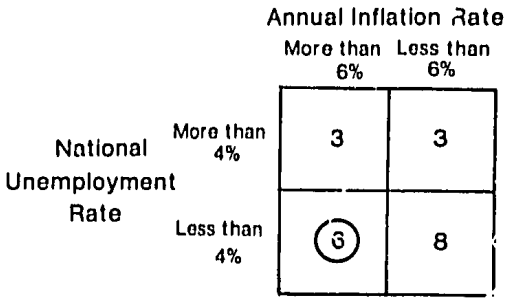
confidence level—the statistical probability derived from a significance test such as chi square. It tells how often the observed behavior would occur solely due to chance, and how certain we may be that non-random causes are operating. (see: *significance test*)

confluence analysis—a method of overcoming certain difficulties in *REGRESSION* analysis when errors of observation produce spurious linear relations in the observed independent variables. (see: *REGRESSION*, *independent variable*)

confounding—limiting the size of samples in large *FACTOR ANALYSIS* experiments by omitting some of the cases which relate to particular interactions deemed unimportant or of little practical significance from a policy point of view. (see: *FACTOR ANALYSIS*, *block*)

constraining equation—a mathematical expression of a resource limitation beyond which no conceivable solution is feasible. For example, in computing possible US GNP levels, one constraining equation might express the fact that the number of possible man-hours of labor input is limited by the population size. (see: *objective function*)

contingency table—an array of data in rows and columns. The simplest contingency table is a 2x2 matrix, illustrating two dichotomous variables and the frequencies of observation of each. For example, a 2x2 contingency table may show 20 observations of national unemployment and inflation rates. It shows, for example, that there were 6 cases when inflation of greater than 6 per cent was accompanied by unemployment of under 4 per cent. (see: *array*)



continuous variable—a variable which can assume fractional values as well as whole numbers. It describes quantities which may be infinitely divided, like distances or time, as opposed to discrete variables, such as men or machinery, which can not be fractions. (see: *discrete variable, integer, real number*)

continuum—a continuous range of related phenomena located along a line between two extreme values of the characteristic to which they are related. Events or behavior ranked on a continuum can be compared or quantified by assigning numerical values to each position along the continuum.

control group—a group representing the standard or normal level of behavior or activity which is not subjected to experimentation but is compared to the experimental group in order to assess the results of change.

covariance—the association between two variables, as shown on a scatterplot. *CORRELATION* is a measure of covariance, *REGRESSION* is a description of it. (see: *CORRELATION, REGRESSION*)

coverage—used in sampling in two senses: (1) to denote the scope of the material collected from the sample members; (2) to mean the extent or area of the population examined. (see: *sample, population, SURVEY RESEARCH*)

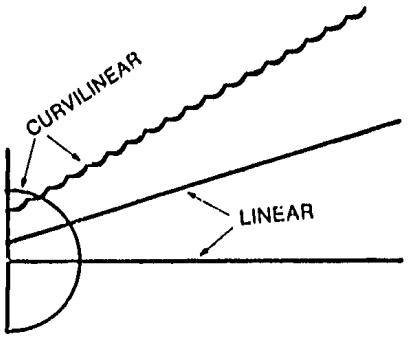
Cramer's V—a chi square based measure of association used to determine the strength of relationship between two tables of different sizes. Cramer's V varies between O, meaning no relationship, and 1, indicating maximum relationship. (see: *measure of association, chi square*)

critical region—one of two mutually exclusive regions in a sample space used in the testing of statistical hypotheses. If the sample point falls into one (the region of acceptance) the hypothesis is accepted; if it falls into the other, the hypothesis is rejected. The second region is customarily denoted as the "critical region." (see: *sample space*)

cross-correlations—correlations between series ordered in time, or space. Thus, if v_1, and v_2, and v_3 are variables in one series and w_1, and w_2, and w_3 are in another, correlations between v_1, and w_1 and between v_2 and w_2, etc., are cross-correlations.

culture—the aggregate of learned, socially transmitted behavior patterns characteristic of a group, which is developed and maintained through formal and informal learning, language, knowledge, folkways, beliefs, customs, traditions, and institutions. A political system is shaped by related cultural factors, and may in turn, promote cultural change by influencing the behavior patterns of the society.

curvilinear—a relationship between two variables displayed on a graph by curved lines. (see: *linear*)



cybernetics—originally a science dealing with complex electronic calculating machines, cybernetics has evolved into the study of information channels, decision-making processes, feedback mechanisms, and other methods by which a government or organization perceives, responds, and enforces. (see: *COMMUNICATIONS THEORY*)

D value—the largest absolute difference between the cumulative observed frequency of a variable and its cumulative expected frequency (or frequency if the variable were governed solely by chance), which is used in determining the significance of ordinal data. (see: *Kolmogorov-Smirnov D, data-ordinal, significance level, variable)*

data—the facts, statistics, and other forms of information that provide the raw material for analysis. There are four levels of data. *Ratio scale data* have a true zero point and equal intervals between data—for example, data on income levels from $0 to $100,000 with an equal interval between each dollar. *Interval data* lack a true zero point but have equal intervals. Temperatures are an example of interval data because the interval between each degree is the same but there is no true zero point. *Ordinal data* lack both a true zero point and equal intervals. The class rank of students is of

this type—in a class of 100, for example, the student ranked 100 is not necessarily twice the achiever of the student ranked 50. *Nominal data* do not even meet the requirement of order met by the three previous forms of data. The data may be "yes-no" answers to questions, or data arbitrarily assigned numbers for coding purposes (as localities in the US are assigned zip codes).

decision function—a rule of conduct which, at any stage of a sampling investigation, tells the statistician whether to take further observations or whether enough information has been collected to perform the statistical tests he wishes to make. (see: *sample, SURVEY RESEARCH*)

decomposition—the breaking down of a problem into its constituent parts. Used in time series analysis to mean the act of splitting a time series into four—a long-term movement or trend; oscillations of more or less regular period and amplitude about this trend; a seasonal component; and a random or irregular component—by the use of statistical methods. (see: *TIME SERIES ANALYSIS*)

deduction—the process of reasoning from the general to the specific and drawing conclusions from applicable premises. Deduction is used in political analysis when the primary objective is to explore and understand the implications of premises or to formulate hypotheses based on what is already known; the inductive method is used to test hypotheses once formed. (see: *induction, hypothesis testing*)

degrees of freedom—the number of elements in a set of data that can vary and still permit certain conditions to be met. Generally, the number of degrees of freedom is one less than the number of observations in the set. If four elements in a given set have a mean of 16, for example, the first three may take any values—e.g., 4, – 21, and 68—totalling 51—but then all three available degrees of freedom have been used. The fourth value must now be 13 in order for the mean of 16 to be met (51+13=64/4=16). The number of degrees of freedom is used in chi square and other statistical operations. (see: *chi square*)

delta (Δ)—an increment of change. If variable x goes from a value of 8 to a value of 14, the change in x (read, "delta x") is 6 ($x_1 - x_2 = \Delta x$). Δt often symbolizes a change in time. Delta is also shown as d, especially in the formula for the slope of a line; dy/dx, where the slope equals the change in y divided by the change in x. (see: *slope*)

demand—a measure of an individual's or a society's willingness and ability to purchase and consume a good at a given price. (see: *supply*)

dependent variable—the event or condition the analyst wishes to explain or predict in any research problem, as opposed to the independent variable, which is the predictor or explicator. The relationship between dependent and independent variables is not necessarily causal, but causality is usually inferred. (see: *variable, independent variable, intervening variable*)

derived statistics—obtained by an arithmetical operation from the primary observations. For example, population figures are primary and so are geographical areas, but population-per-square-mile is a derived statistic.

descriptive statistics—a branch of mathematics designed to summarize or describe the important features of a large amount of data without inferring anything that pertains to more than the data, as opposed to inferential statistics which involves generalizations, predictions, and subjective evaluation of the data. (see: *inferential statistics*)

descriptive survey—a sample survey where the principal objective is to estimate the basic statistical parameters (means, totals, ratios) of the population or its sub-divisions. (see: *SURVEY RESEARCH, parameter*)

deterministic model—an unchanging model which contains no random elements and for which the future course of the system is determined by its structure at some fixed point in time. (see: *model*)

development theory—the study of growth and change within political, social or economic systems, or change from one system to another, generally toward greater governmental capacity to cope with the demands made upon it. The terms traditional, transitional, and modern are commonly used to designate societies in different stages of development.

deviate—the value of a variable measured from some standard point of location, usually the mean. (see: *variable, mean*)

deviation—(see: *standard deviation*)

diachronic analysis—(see: *TIME SERIES ANALYSIS*)

dichotomous classification—the ordering of phenomena into two mutually exclusive groups (military/ non-military, for example). The purpose of such classification is to gain an understanding of political phenomena by relating them to overall descriptive categories for purposes of contrast and comparison.

digital computer—a computer which handles data in numeric form and processes data by operations based on counting. (see: *analog computer*)

direct sampling—a term used when the sample units are people and not some kind of record relating to them, such as census forms. (see: *SURVEY RESEARCH, sample*)

discrete variable—a variable which can take only integer or whole number values—the number of weapons in a cache, for example—as opposed to a continuous variable which can be either a whole number or a fraction such as the average age of enemy soldiers fighting each month. (see: *variable, continuous variable, integer, real number*)

discrimination—in multivariate analysis, a method for allocating items known to come from two or more populations to the correct population with a minimum amount of misclassification. (see: *multivariate analysis, sample, population*)

dispersion—the degree of heterogeneity of a set of observations. There are various numerical alternatives for measuring dispersion, including range, mean deviation, and standard deviation.

dispositions—personality traits that shape the individual's values, attitudes, and opinions toward political phenomena and his actions in the political arena. (see: *personality trait*)

distortion—occurs when information received by a system is changed or misinterpreted before it can be acted upon. *Amplification* of positive feedback is an example of distortion which involves an upward or downward spiral of self-reinforcing actions which lead to further actions of the same sort. *Negative-only feedback*, the opposite of amplification involves self-corrective actions which keep the system in equilibrium. *Channel drift* is a form of distortion where a political system attuned to one channel of information input may gradually shift toward another channel with unintended system consequences. Such consequences include, for example,

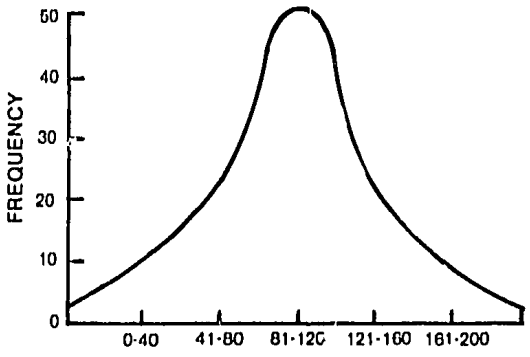
the societal and political changes which occur when newspapers begin to replace word-of-mouth communication in lesser developed countries, or when electronic media replace newspapers in developed countries. *Short-circuiting* is a distortion where data is not subject to all phases of the input-transformation-output system. (see: *SYSTEMS ANALYSIS, COMMUNICATIONS THEORY*)

distribution—the grouping of large sets of data into a number of classes by size, frequency, or some other variable. The information can be presented in a chart, table, or graph. When presented in graph form, it is known as a distribution curve (see: *frequency, frequency polygon, ogive, normal distribution*)

FREQUENCY DISTRIBUTION

I.Q. SCORES	FREQUENCY
0 - 40	5
41 - 80	20
81 - 120	50
121 - 160	20
161 - 200	5

FREQUENCY CURVE or POLYGON



Dollard-Doob hypothesis—(see: *frustration-aggression hypothesis, CONFLICT BEHAVIOR ANALYSIS*)

domain—the values that an independent variable may take.

dominating strategy—a strategy in game theory which is in no respect worse than all other strategies and in at least one respect better. (see: *GAME THEORY*)

down cross—the point where a time series, measured about its mean, changes in sign from positive to negative. Correspondingly, a point where it changes from negative to positive is called an *up cross*. (see: *TIME SERIES ANALYSIS, mean*)

dyad—a linked pair of entities—two states, leaders, political systems, interest groups, departments within the government, etc.—that exhibit in common some variable which is under analysis.

dynamic model—a mathematical computer model which is not derived statistically from time-series but which starts with a statement of system structure and by looking at multiple feedback relationships determines the dynamic consequences when the assumptions within the model interact with one another. A dynamic model does not primarily seek to optimize decision making or maximize system performance as in linear programming models but rather to identify significant causal relationships to help the analyst gain additional insight into system behavior. (see: *feedback loop, system, LINEAR PROGRAMMING*)

efficiency—the capacity to produce desired results with a minimum expenditure of resources; or the ability to combine given resources to maximum advantage.

eigen value—in *FACTOR ANALYSIS*, the sum of the factor loadings squared which measures the amount of variation in the variables accounted for by the factor. (see: *factor, loading, FACTOR ANALYSIS*)

elasticity—a measure of the sensitivity of one variable to change in another—commonly used in economics to demonstrate the change in demand or supply in response to price variation. The formula for computing elasticity of demand, for example, is the percentage change in quantity demanded divided by the percentage change in price. Thus, if a 20 percent price drop causes a 20 percent boost in demand, elasticity is equal to one, and the total amount spent on the good (equal to quantity times price) stays constant. If a 20 percent price cut causes a greater than 20 percent rise in demand (so that the total amount spent on the good rises) elasticity is greater than one, and demand is said to be "price elastic." Demand is inelastic when a 20 percent price cut causes a less than 20 percent rise (and total revenue falls).

empiricism—pursuit of knowledge by observation and experimentation. In political science, empirical theories (what is) can be contrasted with normative theories (what ought to be). In some political writings, empiricism is broadly equated with the use of scientific method or the behavioral study of politics. (see: *BEHAVIORALISM, method*)

endogenous variable—a variable whose value is dependent upon and determined by the process under study. For example, in an election model predicting party votes from unemployment and religion, the demographic factors are *exogenous*, and the voting results are *endogenous*. (see: *exogenous variable, dependent variable*)

engineering theory—often referred to as "policy science," engineering theory in political science tries to provide answers to political problems basically through the use of ends-means analysis. If policy makers are faced with the question of how to deal with rising unemployment, for example, they might develop an engineering statement that details the means for meeting that problem. Such a statement would set forth the variables relevant to achieving full employment and the manner in which those variables might be manipulated in order to bring it about. (see: *ENDS-MEANS ANALYSIS, COST-BENEFIT ANALYSIS*)

equilibrium—a state of balance ascribed to a political, economic, or social system. One form of equilibrium analysis assumes that environmental influences tend to affect the relationships within a system, constantly moving them away from, then back toward, a presumed preexisting point of stability. Equilibrium analysis sometimes assumes a dynamic equilibrium—that is, one in which no fixed point of stability exists but the elements of the system remain in some kind of balance with one another while the position of equilibrium fluctuates from one point to another. (see: *SYSTEMS ANALYSIS*)

events data—information concerning the behavior of individuals or nations. It is usually coded as input to a systematic data base or file and is dependent upon time. (see: *EVENTS ANALYSIS*)

exchange theory—seeks to explain political behavior in terms of market principles derived from economic models. Politics is, therefore, viewed as a process of resource distribution based on rational calculations of costs and benefits. Exchanges constitute a common form of political interaction. On the international level, for example, where each state may carry on interactions with more than 140 other states as well as numerous non-state systems, analysis of exchange levels—goods for goods, or deeds for deeds—can help bring economic insights to bear on a wide range of political behaviors and contribute to the development of typologies of state actions.

exogenous variable—a variable whose value is determined outside the model being studied, and is independent of it but which may influence variables in the model. A model of energy demand in Japan, for example, is affected by the exogenous political variable "oil decision" made by Middle Eastern exporting countries. (see: *endogenous variable, independent variable*)

expected value—a key concept in probability theory, statistics, and game theory. The expected value of a strategy is obtained by estimating the probability of each possible outcome, and multiplying each of these probabilities times the value (utility) associated with that outcome. The sum of these multiplications is the expected value of the strategy, and risk-adverse player will normally select the course of action with the highest expected value. For example: Strategy I has three possible outcomes:

Outcome	*Probability*	*Payoff*
A	.6	30
B	.2	80
C	.2	−20

Strategy II has two possible outcomes:

D	.5	60
E	.5	−10

The expected value of Strategy I is:

(.6 x 30)+(.2 x 80)+(.2 x −20)=30

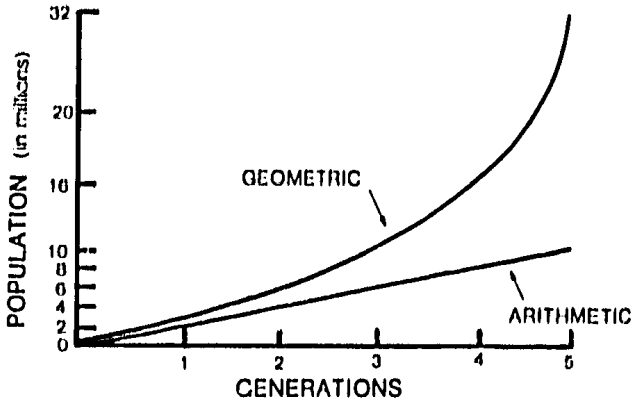
The expected value of Strategy II is:

(.5 x 60)+(.5 x −10)=25

exponent—the power to which a number is raised, that is, the number of times the number is multiplied by itself. In the equation, $2^3=2 \times 2 \times 2=8$, the 3 is the exponent.

exponential curve—a series of observations ordered in time, as displayed on a graph or chart, which increases geometrically at a constant, or approximately constant rate. In an arithmetic progression, each succeeding number is derived from the preceding one by the *addition* of a constant quantity. In a geometric or exponential progression, on the other hand, each succeeding number is derived from the preceding by *multiplying* it by a constant quantity. If the constant is 2, the series 2, 4, 6, 8, 10 would

form an arithmetic progression; the series 2, 4, 8, 16, 32 would form a geometric progression. Assuming the numbers represented population, the two series could be graphed as below. Typically population growth would follow the geometric (exponential) curve.



F test—a statistical test designed to tell whether two samples are significantly different from each other. The test measures the ratio of the variances of the two distributions, then uses an F table to tell how often such a difference would occur solely by chance. The process is similar to both the chi square and t tests, which tell how confident we can be that the observed changes are not random. (see: *confidence level, significance level*)

factor—in statistics, (1) a quantity under examination in an experiment as a possible cause of variation, or (2) an item in an average or an index number. (3) As adapted from psychology and used in multivariate analysis, it denotes an impact of the observed variables which may be regarded as part of those variables. (4) In *FACTOR ANALYSIS*, a common condition underlying a large number of variables or people is a factor. By studying a set of UN votes, for example, we may identify which votes belong together. Thus, industrial countries, former colonies, Arab states, or US allies may emerge as voting blocs or underlying factors in UN votes. (see: *multivariate analysis, variable, FACTOR ANALYSIS*)

factorial (!)—the product of successive numbers always ending with one. For example, 7! (read "seven factorial") equals 7 x 6 x 5 x 4 x 3 x 2 x 1 = 5040.

fair game—in the theory of games, a game consisting of a sequence of trials is deemed "fair" if the cost of each trial is equal to the expected value of the gain from each trial. In a contest between two adversaries with unequal resources, for example, it is deemed "fair" that the player A with the larger sum to stake has a better chance of ruining his opponent B because A can lose a larger sum from the game. (see: *GAME THEORY, expected value*)

fallacy of composition—the logical error of assuming that what is true for one individual is necessarily true for the whole group. For example, one individual at a football game can get a better view of the field by standing up, but if everyone tried this, they would all be likely to have the same, or even a worse, view than before. This fallacy slips easily into economics: in group savings and in national debt, for example, what is true on a micro level is not necessarily true on a macro level.

fallacy of division—the logical error that occurs when one erroneously assumes that characteristics of a group necessarily exist in each member of that group.

false dilemma—a logical fallacy which omits consideration of some available alternatives. Unfairly dichotomizing the possibilities to the two extremes, for example, may obscure moderate solutions and make a problem appear more perplexing than it really is.

feedback loop—in systems analysis (particularly systems dynamics) a circular, causal chain through which the past behavior of a system affects its future behavior. This may be contrasted with an open loop system, where outputs depend on inputs but inputs are not influenced by past behavior.

In the *positive feedback loop*, the causal chain acts to reinforce change, either up or down. Examples include the growth of deposited money via compound interest, arms races, the extinction of a species, and industrialization. Causal chains which resist rather than reinforce change are *negative feedback loops*. Given a disturbance pushing a system away from equilibrium, the negative loop responds in the opposite direction forcing the system back towards equilibrium. (see: *SYSTEMS ANALYSIS, SYSTEMS DYNAMICS*)

feedback system—a system in which the effects of the outputs are measured and the measurements are then reintroduced as inputs. This allows the system to constantly monitor its own performance, and adjust to the conditions of the environment even as it influences them. The classic feedback mechanism is the furnace thermostat which reads the room temperature as it adjusts it, in order to decide how much more heat is required.

field—in computer jargon, a column or set of adjacent columns on a computer punch card which are allocated to represent one piece of data or instruction. For example, a one column field may be punched 0 or 1 to indicate yes or no, and the year 1974 would fit on a four column field.

In field theory, the field" is a chart upon which actors are plotted according to their relationship to the variables under study and according to their relationship with other actors. (see: *actor, FIELD THEORY*)

Fisher's exact test—a test of significance for small samples (less than 20 observations) divided into a 2 by 2 table, which gives the exact probability of getting the same breakdown by chance as occurs in the table. (see: *significance level, chi square*)

fixed-sum game—a class of problems dealt with by game theory in which one player's gains are exactly equal to the other player's losses. This describes a perfectly competitive situation in which no cooperation is possible because the sums of the payoffs in each matrix cell are held to some constant level. (see: *variable-sum game, zero-sum game, GAME THEORY*)

fluctuation—any movement up or down between consecutive items of a series of numbers; the variation of a statistic from sample to sample.

folding—regarding all the values of a set of data as having the same sign—usually plus—in order to perform certain statistical operations which are complicated by the use of different signs in the data.

forecasting—prediction, in the customary sense of assessing the magnitude a quantity will assume at some future point in time. Distinct from "estimation," which attempts to assess the magnitude of an already existing quantity. For example, the final yield of a crop is forecast during the year and estimated at harvest.

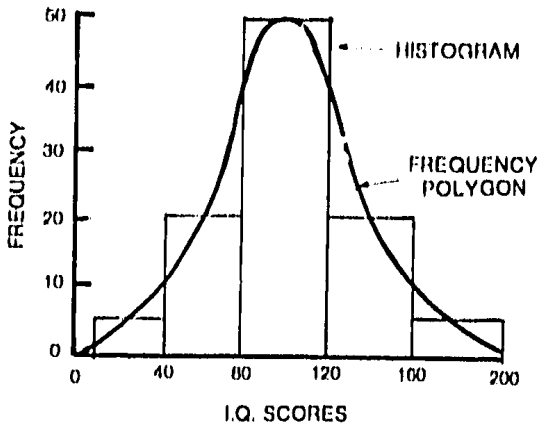
fractile—denotes the one value in a set of data above or below which lies a given fraction of the distribution. For example, quartiles divide the distribution into four parts, deciles divide it into 10 parts, and percentiles divide it into 100 parts.

frequency—the number of occurrences of a given type of event, or the number of members of a population falling into a specified class. Where the frequency is expressed as a proportion of the total number of occurrences or the total number of members, it is called the *relative* or *proportional frequency*. When the frequency of each occurrence is added to previous frequencies, the result is a *cumulative frequency*. (see: *distribution, frequency polygon, ogive, normal distribution*)

I.Q. Scores	*Frequency*	*Proportional Frequency*	*Cumulative Frequency*
0-40	5	1/20	5
40-80	20	1/5	25
80-120	50	1/2	75
120-160	20	1/5	95
160-200	5	1/20	100

frequency distribution—(see: *distribution*)

frequency polygon—the outline or silhouette of a histogram. It graphically displays a distribution of data by classes and by frequency of observation. Cumulative frequency, presented in graph form, is called an *ogive*. (see: *histogram, distribution, frequency, ogive*)



frustration-aggression hypothesis—the assumption that frustration always leads to some kind of aggressive reaction whether explicit or implicit. (see: *CONFLICT BEHAVIOR ANALYSIS*)

function—in statistics, a mathematical description of the relationship between two variables; a rule describing the value of one variable if given the value of the others. For example, the function $Y=6X+3$, tells how to derive the corresponding Y value if given any value of X.

In structural-functional analysis, *structures* are patterns which govern actions—including laws, constitutions or such concrete organizations as political parties and legislatures—while *functions* are the results or consequences of actions which satisfy the needs or demands of the system for its own maintenance. Defense of the state, stabilization of the economy, and execution of laws, for example, are political functions. (see: *structure, variable, STRUCTURAL-FUNCTIONAL ANALYSIS*)

fundamental probability set—a set of objects or events which are basic to a probabilistic situation, in the sense that all other objects or events under consideration are derived from them. It follows that all probabilities are expressible by the rules of addition, multiplication, etc., in terms of the fundamental set.

gain—in communications theory, gain measures the effective as opposed to the ineffective changes a system makes as a result of the communications it receives. (see: *distortion, lag, COMMUNICATIONS THEORY*)

gambler's ruin—the name given to one of the classical topics in probability theory, in which a gambler wins one predetermined sum of money for every success and loses a second predetermined sum for every failure. The play proceeds until the initial capital of one player is exhausted and he is ruined. The statistical problems involved are concerned with the probability of the ruin of a player given the stakes, initial capital and chance of success, and with such matters as the length of play. Variations of this problem can be used in war gaming. (see: *probability, GAME THEORY*)

game—(see: *GAME THEORY*)

gamma—(see: *Goodman-Kruskal gamma*)

general factor—(see: *common factor*)

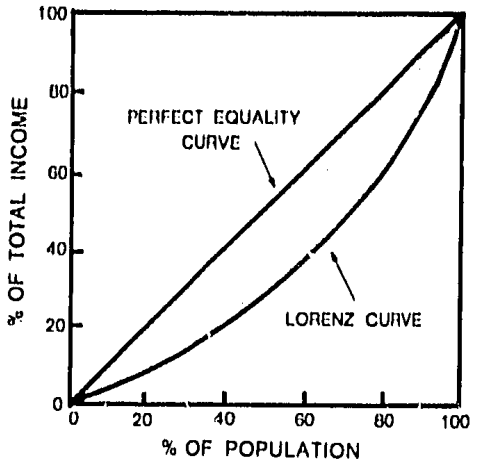
geometric mean—(see: *mean*)

geometric progression—(see: *exponential curve*)

Gestalt psychology—a psychological approach which basically argues that the whole is greater than the sum of its parts. That is, the examination of a given phenomenon in its totality yields different results or interpretations from those yielded by examination of the different parts that constitute it. More technically the Gestalt psychologist objects to treating the nervous system as a static, machinelike structure capable only of responding piecemeal to incoming stimuli. Rather the cerebral cortex is viewed as analogous to a force field which is in active equilibrium and in which each incoming stimulus affects the entire field.

Similarly in political science analysis the question is whether relationships and patterns can best be discovered by using an incremental Bayesian approach in which the analyst reacts to each individual piece of evidence or by a holistic approach in which the analyst gains insight about political reality by looking at and drawing conclusions from the totality of relevant evidence. (see: *FIELD THEORY*)

Gini coefficient—a measure of inequality, describing the uniformity of the distribution of some attribute (income, power) over some population (individuals, states). The index is equal to twice the area of difference between a Lorenz curve and the line of perfect equality. The coefficient (G) ranges from O (representing perfect equality in the distribution for everyone in the sample), to 1 (representing perfect inequality—one individual possessing everything and the rest of the population possessing nothing). (see: *coefficient, Lorenz curve*)



Goodman-Kruskal gamma—a measure of association for comparing ordinal level data in tables. Gamma varies between 0 and a maximum of 1 based on the probability of predicting one quality in the table given the other. (see: *ordinal*)

goodness-of-fit—in general the agreement between an observed set of values and a second set which are derived wholly or partly on a hypothetical basis, that is to say, derived from the "fitting" of a model to the data. In the process of validating models, for example, goodness-of-fit tests are used to compare the model's forecast with real world observations in order to infer the degree to which the model comforms to reality. (see: *chi square, Kolmogorov-Smirnov test*)

grade—the proportion of the total frequency that items occur in a continuous series with values less than or equal to the particular item being "graded." If we wish to grade item 494 in a series of 531 items, for example, its grade is 494/531, or 93%.

grid sampling—a form of cluster sampling, the clusters in this case being individual areas of a grid and hence consisting of groups of basic cells arranged in some standard geometric pattern. (see: *SURVEY RESEARCH, cluster sampling*)

group—a set of elements, individuals, or observations, distinguished by some common attribute or shared relationship. In group theory, the collectivity of individuals may be categorized in many ways. A *formal* or *organized* group has recognized goals and structures affecting group interaction. An *informal* group, such as like-minded politicians meeting at lunch to share views, lacks such explicit goals and organization structure.

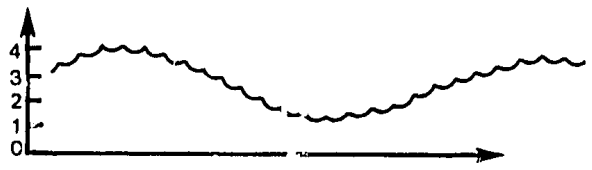
A *primary* group is a relatively small collectivity that engages in frequent and direct interactions (families, work groups, etc.) while a *secondary* group interacts infrequently and communicates mainly by indirect and impersonal means. A *categoric* group consists of individuals who do not necessarily interact but share some common attribute, such as income level, religion, etc.

The label *interest* or *pressure* group is applied to collectivity that seeks to influence governmental policy in favor of some goal or shared concern of the group. (see: *GROUP THEORY, interest group*)

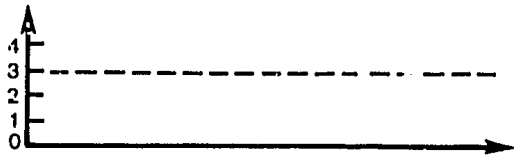
group factor—(see: *common factor*)

harmonic analysis—a regression analysis of data which, when graphed, has periodic, winding components. The harmonic analysis finds that straight line which can be drawn through and will best fit all the data. (see: *REGRESSION*)

harmonic mean—(see: *mean*)



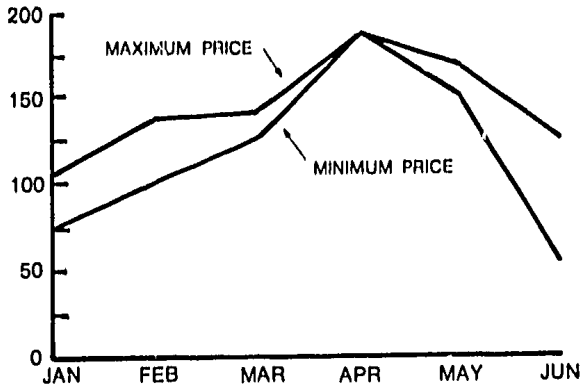
ORIGINAL DATA GRAPHED



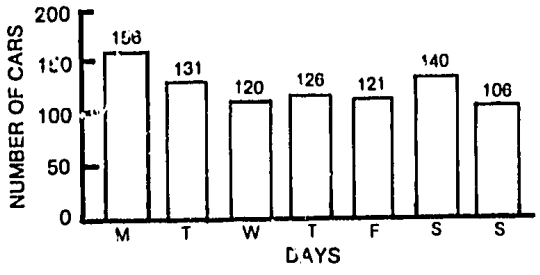
HARMONIC ANALYSIS GRAPH

heuristic—intellectual aid used in the search for new ideas and insights. The use of models as heuristic devices, for example, contributes to an understanding of politics because it permits the analyst to compare the model with real situations. (see: *model, SIMULATION*)

high-low graph—a form of graph used to depict ranges of variation in successive intervals of time. For example, daily price variation might be represented by taking time intervals of one month on the horizontal axis and, at each monthly point, showing the maximum and minimum price attained during the previous month. The highest prices can then be joined by a line drawn from high point to high point; similarly the low prices. Or, for each month, the high and low points may be joined by a vertical bar.



histogram—a graphic form of the frequency distribution in which the number of cases within each class is represented by a vertical bar whose height is proportional to the number of cases within the class. For example, a histogram describing how many cars passed a certain street on each day of a typical week, might look like this:



historical approach—the traditional approach to political analysis in which information is organized and interpreted chronologically. Continuities with earlier epochs may be determined and linkages established. The cumulative development of ideas during one time span, for example, may be studied to determine its impact on the institutions and societal changes that occurred during a subsequent period.

Hollerith card—sometimes called a key punch card, it is divided into 80 columns and information is punched on to it to be read by the computer. (see: *Hollerith code*)

Hollerith code—a punch card code in which the top three positions in a column are called zone" punches, (labeled R, X, and O, from the top down) and are combined with the remaining 9 digit punches to represent alphabetic, numeric, and special characters. The letter A, for example, is a combination of an R and a 1 punch in Hollerith code.

holism—the theory that group properties or characteristics are distinct from those of the individuals who comprise them. In political science, the division in approaches has generally been between those who regard the group as the proper level for the study of politics, and those who focus on the individual and his behavior as the basic unit of analysis. (see: *GROUP THEORY*)

homeostasis—the tendency toward maintenance of stability in a system through self-adjustments that provide compensating responses to disruptive or destabilizing influences. It has been argued that a system which loses the capacity to deal effectively with disruptive inputs faces the problem of disintegration or the possibility of being transformed into another kind of system. (see: *SYSTEMS ANALYSIS, feedback loop, feedback system*)

homogeneity—in statistics, a term used to indicate different populations, or samples from different populations, that are identical in at least one respect. For example, if two populations have the same mean (average) they are homogeneous even if they have different dispersions. (see: *population, sample*)

hypergeometric distribution—the distribution of a discrete variable (one which can take only whole number values) with a finite number of cases. As the number of cases tends to infinity, the distribution tends to the normal distribution form. (see: *discrete variable, normal distribution*)

hypothesis—a statement of an expected relationship between variables that may be tested empirically to determine its validity. A hypothesis may be derived from observation, deduced from a larger body of theory, or based simply on a hunch that the analyst is willing to use provisionally.

hypothesis testing—the attempt to confirm or disconfirm a factual proposition, or hypothesis, by gathering and analyzing relevant evidence. Testing refers both to the use of certain types of measuring instruments, such as intelligence tests or attitude scales, and to statistical tests of the level at which observed relationships in data are sufficiently different from chance relationships (level of significance). In the latter case, a null hypothesis, or opposite of the original hypothesis, is formed with the hope that it can be disproved. If the null hypothesis can be rejected, confidence in the original hypothesis is increased. (see: *null hypothesis, significance level, validation*)

hypothetical population—a statistical population which has no real existence but is postulated to be generated by repetitions of a given event (such as throws of a die, aircraft failures, elections, etc.). The hypothetical population is then used as a control group for comparisons with the actual population data.

ideal type—a concept whose characteristics are represented in so pure a form that examples of the concept are rarely if ever found in reality. Ideal types are used by political analysts in *GAME THEORY*, model building, SIMULATION, and the development of general theory and are used as a base for judging, explaining, or investigating reality. An ideal type of a supply-demand capitalist economy can be created, for example, in order to determine how far an actual system departs from the model. (see: *GAME THEORY, model, SIMULATION*)

idiographic explanation—an explanation based on specific propositions that explain unique or individual social or political phenomena as opposed to a *nomothetic* explanation which is based on general propositions applicable to classes of events or objects. History is often regarded as essentially an idiographic discipline, while political and other social sciences that emphasize explanation by theoretical generalization are more nomological.

illusory association or illusory correlation—an association between attributes which is statistically significant but involves no direct causal connection. If in Europe, for example, the possession of blonde hair was found to be positively correlated with an ability to ski, the association would be illusory in the sense that either attribute does not cause the other. While such an association might be statistically valid, it is due not to causality but to the accidental circumstances that the blonde nationalities inhabit the northern European countries where skiing is common. (see: *attribute, association, CORRELATION*)

image theory—concentrates on the analysis of a nation's perceptions of foreign societies as transmitted from generation to generation through the education system, folklore, news media, etc., to ascertain likely reactions to different cultures. The theory emphasizes that impressions of foreign nationalities are formed in childhood and are not imposed upon people by their leaders, although leaders can and do manipulate impressions, stereo-types, and prejudices.

impact panel—in panel studies, an experimental technique which interviews individuals before and after a planned event in order to ascertain the impact of the event upon attitudes.

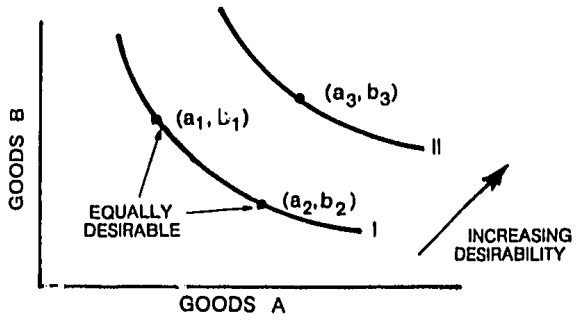
independent events—events which do not influence the probability of occurrence of each other. For example, the probability of obtaining heads on the second toss of a coin is independent of the result of the first toss. (see: *mutually exclusive events, conditional probability*)

independent variable—a variable whose changes are used to predict or explain changes in some other variable called the dependent variable. (see: *variable, dependent variable*)

index number—a figure which discloses the relative change, if any, of prices, costs, or similar statistical phenomenon from one period of time selected as the base period. The base period usually is assigned the index number 100 and any changes from it represent percentages. For example, if an index of cost of living with the year 1969=100 as a base rises to 150, the cost of living has increased by 50% since 1969.

indicator—an event, observation, or value used to measure an abstract concept. All abstractions must be measured indirectly—by indicators. For example, the number of riots or the crime rate might be used as indices of the abstract concept—civil order. (see: *validity, variable*)

indifference curve—(iso-preference line) a graphic description of various combinations of two goods which will leave the consumer equally satisfied. Used in economics, or in any analysis involving trade-offs: guns vs. butter, unemployment vs. inflation, or new car vs. home improvement; where limited resources mean that more of one type of goods can only come at the expense of less of the other type. An indifference curve shows how much more of one type of goods a customer would need in order to give up one unit of the other and still feel equally well off. In a set of indifference curves, each one represents a different level of overall utility or well-being, and individuals seek to attain the highest one, subject to budget constraints. Along any given curve, the individual is indifferent—he feels he would just about break even with any combination, and he has no preference between them. Here, points (a_1, b_1) and (a_2, b_2) are on the same indifference curve: the consumer will sacrifice some of goods B (measured by the quantity $b_1 - b_2$) in order to attain more of goods A (measured by the quantity $(a_1 - a_2)$ or vice versa—he is indifferent between the two combinations. Point (a_3, b_3), however, like all points on indifference curve II, is preferred to any point on curve I



indirect sampling—sampling from documents, or some record of the characteristics of a population, rather than the recording of information obtained first hand. A political analyst may find this method of sampling for survey research techniques best suited to his needs since he is usually unable to survey in the country he is studying. (see: *SURVEY RESEARCH, sample, population*)

Induction—a process of reasoning which attempts to reach logically valid generalizations from individual facts, cases, or events. Inductively reached generalizations are never quite certain to be true or valid; they may always be corroborated or disproven by subsequent observation. Induction contrasts with deduction, which is a process of logically reasoning from general statements or premises to conclusions about individual cases. (see: *deduction*)

inductive statistics—(see: *inferential statistics*)

inference—drawing a conclusion from facts and premises. In statistics, inference is the process of drawing causal implications from statistical data. An important distinction must be made between causality and correlation. Statistical techniques alone can establish that two variables are closely related and that they move together, but this does not necessarily prove a cause and effect relationship. If we find, for instance, that an increase in the number of teachers in an area has been paralleled by an increase in the number of liquor stores, we can not "prove" that one has caused the other. Rather, both may have been caused by a third factor: rising population. While statistical techniques alone can not firmly prove causality, the absence of any significant correlation between two variables can be powerful in disproving a false hypothesis of causality. (see: *CORRELATION*)

inferential statistics—inferential statistics (inductive statistics) the science of drawing causal implications from numerical data, and of making inferences about a large group of individuals when only a small sample has actually been observed. This is distinct from descriptive statistics, which provides methods to collect, organize, and present information on an entire group without drawing implications. (see: *descriptive statistics*)

influence—the capability of a political actor to affect the behavior of others in a manner favored by the actor. (see: *actor, power, authority*)

information system—in systems analysis, a set of interrelated structures that receives an input of new data, processes this data by comparing it to memory and previously established values, and then makes a decision about the data. This decision leads to storage of the data in memory and, if appropriate, to the implementation of actions which may in turn cause a feedback into the information system. (see: *SYSTEMS ANALYSIS, data, distortion, feedback system*)

Input—any influence that affects the functioning of a system. The use of "input" as a descriptive and analytical term in political science calls attention to the dynamic linkage between the system and its environment and between different elements of the system. Inputs relate directly to the making of decisions and to the outputs that follow. Thus, in a political system, inputs result in demands upon and support for the system which are converted into outputs in the form of authoritative policies and implementing actions.

Instrumental variable—a generic term to refer to any predetermined variable that can be used to help identify and estimate the parameters of a large system of equations. Instrumental variables are most frequently used in large econometric models in which the number of variables is greater than the number of available observations.

Integer—a value which can only be expressed as a whole number, as opposed to real values which can have fractions. The number of cabinet members in government—15, for example—is an integer, while the average age of the cabinet members—57.3 years—is a real number (see: *real number, variable*)

Integration—the process by which two or more political units increase their cooperative contact with one another; the political merger of two previously separate units. The condition of integration is also called *political community.*

Interactions—in events analysis and communications theory, interactions are activities and events between individuals or groups, usually across national boundaries, which have immediate political significance. Official visits, military forays, or aid and trade flows are examples of interactions. The routine exchanges across borders such as mail flows or tourists are termed *transactions,* and have less political significance. (see: *EVENTS ANALYSIS, COMMUNICATIONS THEORY*)

Intercoder reliability—a method for checking the reliability of the coding process in content analysis by having several coders analyze the same messages and by comparing the results. The amount of agreement is then computed and expressed numerically as the index of intercoder reliability. (see: *CONTENT ANALYSIS*)

intercorrelation—the relationship of a number of variables among themselves, as distinct from their relationships with an "outside" or dependent variable. (see: *variable, CORRELATION*)

interest group—(pressure group) a collectivity that seeks to influence governmental policy in favor of some goal or shared concern of the group. Interest groups can be subclassified as *associational* (organized interest groups), *institutional* (groups based on major social institutions such as bureaucracies, education, church), *non-associational* (groups lacking continuity in structure or regular procedures for articulating interests such as ethnic, regional, or class groups), and *anomic* (largely spontaneous, temporary groups that express discontent through direct actions such as demonstrations or riots). (see: *group, GROUP THEORY*)

interface—any mechanism or physical location where two different objects or kinds of objects meet. In international relations, the boundary between two countries is an interface. The term is often used to represent a border or boundary that affects the flow of communications between the entities involved. In computer terminology, the man-machine interface is the computer terminal.

interpenetrating samples—two or more samples taken from the same population by the same process of selection. The samples may or may not be drawn independently of each other. (see: *sample, population, SURVEY RESEARCH*)

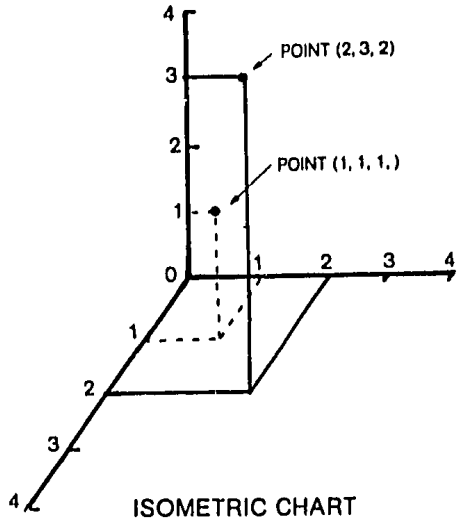
interval scale data—(see: *data—interval*)

intervening variable—an event or condition that affects the operation of the independent variable on the dependent variable. If, for example, we are interested in examining the effect of political propaganda on public attitudes, time may be introduced as the intervening variable by allowing some interval between the interviewing of respondents to ascertain the extent to which their attitudes are shaped by propaganda. (see: *dependent variable, independent variable*)

interviewer bias—bias in the responses or recorded information which is the direct result of the actions of the interviewer. This bias may be due, among other things, to failure to contact the right persons for survey; failure of the interviewer to establish proper rapport with the informant, with the result that imperfect or inaccurate information is offered; or to systematic errors in recording the answers received. (see: *SURVEY RESEARCH, sample, bias*)

inverse probability—the probability approach which endeavors to reason from observed events to the probabilities of the hypotheses which may explain them, as distinct from direct probability, which reasons deductively from given probabilities to the probabilities of contingent events. Bayes' theorem is an example of such reasoning from observation to probabilities. (see: *probability, hypothesis, BAYESIAN ANALYSIS*)

isometric chart—a chart which attempts to depict three-dimensional material on a plane. The distances on the three axes are measured on an equal scale.

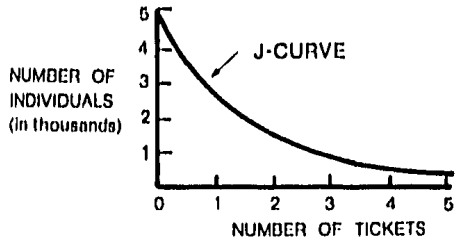


ISOMETRIC CHART

isomorphism—similarity in form or structure. In political analysis it generally refers to the similarity between a conceptual model and the institutional or behavioral attributes that it represents. If a model is to be useful it must be in some degree isomorphic to the system it represents. (see: *model, attribute, heuristic device*)

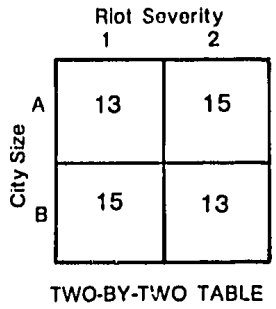
iso-preference line—(see: *indifference curve*)

J-curve—a strongly skewed frequency distribution curve in which the mode (the value observed most often) occurs at one extreme of the range. For example, in showing how many speeding tickets a number of individuals had been given during the past five years, a curve would show some dispersion, but most of the individuals would have recorded zero violations. (see: *skewed distribution, normal distribution*)



joint probability—the probability that both events A and B will occur. If A and B are independent, it is simply the product of their separate probabilities. For example, the probability of a flipped coin landing heads is .5; the joint probability of two heads in a row is .5 x .5=.25. (see: *conditional probability*)

Kendall's tau—a statistic for measuring the strength of association between ordinal level variables—that is, variables which lack a true zero point and equal intervals. It is most often used to measure the association between two variables as represented on a two-by-two table. When computed, usually by computer, the *tau* ranges from 0, indicating no association, to +1, indicating complete association. (see: *data-ordinal, measure of association, variable*)

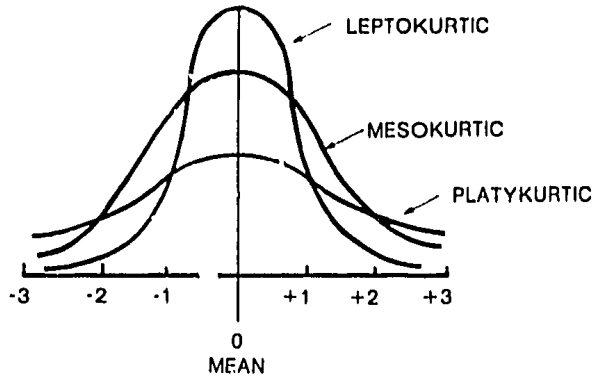


TWO-BY-TWO TABLE

key punch card—a card, sometimes called a Hollerith card, divided into 80 columns for carrying information from some source to a form that the computer can use—making the information "machine readable." (see: *Hollerith code*)

Kolmogorov-Smirnov D—a statistic which measures the significance (or probability that the observed data would occur solely by chance) of ordinal level variables (those which lack a true zero point and equal intervals). The *D value* is the largest absolute difference between the cumulative observed frequency of a variable, and the cumulative expected frequency if the variable were governed entirely by chance. Once the D value is computed, a table is consulted to find the corresponding level of significance of this particular data. (see: *data-ordinal, significance level*)

kurtosis—a measure of how much of the total area under a frequency distribution curve lies within plus or minus one, two and three standard deviations from the mean. When a distribution is perfectly normal, it is said to be *mesokurtic*, when it is peaked, it is *leptokurtic*, and when it is flat, it is *platykurtic*.



laboratory approach—the conduct of scientific experiments or tests under conditions arranged or controlled by the experimenter. In social science research, the analyst aims at holding certain variables constant so that he may systematically manipulate, observe, describe, and measure the remaining variables. Simulations and small group experiments are the principal applications of the laboratory approach in political analysis. (see: *variable, SIMULATION, GROUP THEORY*)

lag—in communications theory, the time interval between receipt of the communication and the reaction to it. (see: *COMMUNICATIONS THEORY*)

lambda—a statistic for measuring the strength of association between two variables by giving the "probable reduction of error" between the two. A lambda of .60, for example, means that knowing the independent variable will reduce by 60% the number of errors made in predicting the dependent variable. For example, knowing the size of a city where riots occurred might reduce by 60% the numbers of errors in predicting riot severity. (see: *measure of association, significance level*)

latent function—the relevant but unintended consequence of activities initiated to satisfy the needs or demands arising from individuals and groups within a political system. The manifest function of a political machine, for example, is to maintain and exercise political power, but its latent functions may include meeting welfare needs and providing channels for social mobility. (see: *STRUCTURAL-FUNCTIONAL ANALYSIS*)

latent variable—a variable which is unobservable but is supposed to enter the structure of a system under study, such as power in politics, or demand in economics. (see: *variable*)

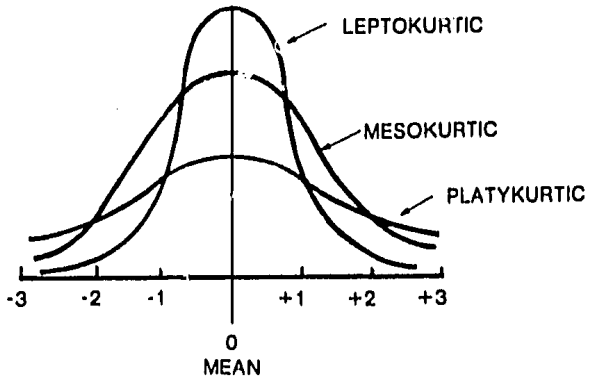
law of large numbers—(see: *Bernoulli theorem*)

least squares—the most common rule for drawing a *REGRESSION* line. The least squares method specifies that the regression line should be drawn as the "best fit" of the data on a scatterplot. It minimizes the sum of the squared vertical distances from each point in the group to the regression line.

If the line shown in the chart is the proper least squares regression line, then the sum of the squares of the distances represented by the dotted lines will be smaller than that achieved by any other possible line on the graph. The least squares method instructs a computer to derive this line, which describes the observed data better than any other line, and can be used to predict the dependent variable (number of riots, for example) if the independent variable (in this case, level of unemployment) is known.

legitimacy—the quality of being justified or willingly accepted by subordinates that converts political power into "rightful" authority. The consensus that provides the legitimizing factor in the exercise of power may be cultivated through the sanctity of tradition, by the devotion of people to a particular leader, or by the acceptance of the supremacy of law. (see: *power, authority*)

leptokurtic—a tall, thin frequency distribution curve. (see: kurtosis)



Likert scale—a simple scale for scoring the attitudes of respondents (called judges) according to their agreement or disagreement with statements which measure a test variable—such as hostility toward a neighboring country. Ordinarily there are five choices for each statement: strongly agree, agree, uncertain, disagree, strongly disagree. Each statement is then given a score, 1 to 5, equal to the average response of the judges. The judges are also scored according to their average response and then ranked according to their position on the test variable. (see: *SCALING, scale, variable*)

line sampling—a method of sampling in a geographical area. Lines are drawn across the area and all members of the population falling on the line, or intersected by it, are included in the sample. (see: *SURVEY RESEARCH*)

linear—involving a straight line; an equation without exponents. $2X+3Y=6$ is a linear equation; $X^2+3Y=14$ is non-linear. (see: *curvilinear, exponent*)

linear model—a model in which the equations connecting the variables are in linear form, that is, without exponents. (see: *model*)

linkage theory—studies the convergence of national and international systems, the relationships that link the stability, functioning, institutions, and goals of national political systems to variables in their extended environment. (see: *system, SYSTEMS ANALYSIS, variable*)

load—in communications theory, demands which are communicated to the communications system which challenge the efficacy of the system. The demand for more rapid transmission of ideas from the elite in the political system down to all other levels, for example, constitutes a load. (see: *COMMUNICATIONS THEORY*)

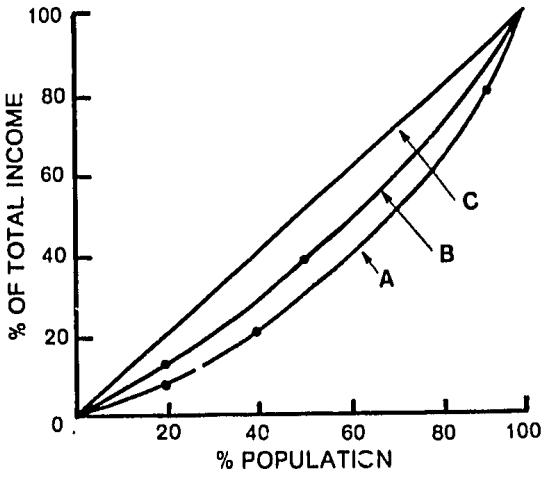
loading—in factor analysis, factors are made up of interrelated variables. Loading is the degree to which any given variable correlates with a factor as expressed on a scale from 0, indicating no correlation to +1 for positive and −1 for negative correlation. Factors are artificial constructs which have no inherent meaning; the analyst must infer what the factors represent from knowledge of which variables are most heavily loaded on them. If 3 of ten variables under consideration—age of party member, years in party, and highest office held—are "loaded" on a factor at the .75 or better level, the analyst may surmise from the variables included that the new factor represents "seniority." (see: *variable, factor, FACTOR ANALYSIS*)

logarithm—an exponent, used in a system especially helpful in simplifying computations involving very large or very small numbers. For example, if B is a positive number (other than one) and $B^L = A$, then the exponent L is the logarithm of A to the base B. In symbols, if $B^L = A$, then $L = \log^B A$. For example, $2^3 = 8$, so $\log_2 8 = 3$. Also, $\log_3 9 = 2$ because $3^2 = 9$. Instead of carrying out lengthy multiplications (12,650,000,000 x 683,000, for example) a quick and close approximation can be gained with the use of a logarithm table, found in the appendix of most statistics or mathematics books. The log to the base 10 of the first number is 7.1021 and the log of the second is 5.8357. We add these logarithms to multiply the bases, and their sum, 12.9378, corresponds to 8,667,000,000,000, which is close to the actual answer of 8,665,250,000,000.

long run—in economics, a time period long enough to permit a firm to adjust its fixed plant facilities to conditions of changing demand. Such time periods are conceptual rather than calendar measures—a small firm may adjust its plant size fairly quickly, but in heavy industries or utilities, lags of several years may occur before all the necessary adjustments may be made. (see: *short run*)

longitudinal analysis—(see: *TIME SERIES ANALYSIS)*

Lorenz curve—a graphic representation of the distribution of some attribute (income, power) over some population (people, states).



For example, Lorenz curve A shows that the lowest 20 percent of the population in this income group receives 5 percent of the total income of the society. Similarly, the lowest 40 percent gets 20 percent, and the highest paid 10 percent of the population gets 20 percent of the total income. Curve B represents a somewhat more equal distribution, with the bottom 20 percent earning 10 percent of the income, and the bottom 50 percent getting 35 percent. Line C is the line of perfect equality. We measure the area between this line and the Lorenz curve, then double it to derive the Gini index. This ranges from 0 to 1 and is a simple measure of inequality. (see: *Gini index*)

loss matrix—in decision theory a matrix specifying the economic loss or gain incurred according to the various decisions which can be taken and the various situations which can exist. (see: *matrix, DECISION THEORY*)

		DECISIONS			
		A	B	C	D
SITUATIONS	1	-5	+8	+3	+9
	2	-3	-4	+4	+3
	3	-2	+5	+1	-1
	4	+6	+5	+7	0

machine readable—information that can be processed directly by a computer with no transfer of information from one medium to another. Information on key punched cards or transmitted electrically and stored on tapes or discs is machine readable. (see: *computer, key punched card*)

macro/micro analysis—alternative levels of analysis, the micro level consisting of subunits of the higher level macro unit. In macro analysis the researcher directs his attention to the collective, systemic, institutional, composite, or group level. In micro analysis, he focuses attention on the parts, subsystems, components, or individuals that comprise the collectivity. A study made of a legislature might constitute macro-level analysis compared to the study of its committee system, but micro-level analysis compared to the study of the larger governmental system of which it is a part.

management by objective (MBO)—a system for establishing both the output requirements and measurement criteria of each managerial position and the periodic conversion of these requirements and criteria into measurable time-bounded objectives which are linked to future planning. The goal of MBO is to improve managerial effectiveness—that is, the extent to which a manager achieves the output requirements of his position.

management science—the use of scientific and particularly mathematical methods to assess the effectiveness of organizations from the perspectives of managers. It is of greatest interest to political scientists in the fields of public administration or urban affairs. (see: *SYSTEMS ANALYSIS*)

manifest functions—in structural-functional analysis, the purposeful or intended consequences of activities that satisfy needs or demands arising from individuals and groups within a political system. (see: *STRUCTURAL-FUNCTIONAL ANALYSIS, latent function*)

manifold classification—the final grouping of a population which has been divided into a number of mutually exclusive classes according to some characteristic and then each class subdivided by reference to some second, third, etc. characteristic. (see: *population)*

marginal (marginal cost, marginal revenue, marginal utility)—the term "marginal" refers to the increment added to the total cost, revenue, or utility by the last extra unit. This is "marginal" in the sense of "last," not in the sense of "unimportant." *Marginal cost* is the amount it costs to produce the last unit of a good—a perfectly competitive firm will produce until marginal cost equals price. *Marginal revenue* is the change in total revenue due to the last sale—a monopolistic firm will produce until marginal revenue equals marginal cost (usually at lower production and higher sales price than a perfectly competitive firm). *Marginal utility* is a psychological concept, not directly measurable. It refers to the change in total satisfaction or happiness resulting from the last unit of consumption. *"The law of diminishing marginal returns"* refers to the fact that marginal utility generally declines after a certain level of satisfaction has been reached. A second car, for example, may be valuable to a family, but not as valuable as the first, without which they would be immobile. A third car is somewhat less valuable and succeeding cars are of even less utility. *"The law of*

diminishing total returns" says that at some point, additional units may actually be bad, and subtract from total utility. For example, if a family gets too many cars, its driveway is always jammed, license fees rise, and they get in the way of each other. The family may then feel itself to be better off—i.e., to have a higher total utility—without these last cars, which are then said to be of negative marginal utility.

Markov chain—(see: *MARKOV ANALYSIS)*

matched samples—a pair or set of samples in which each member of a sample is matched with a corresponding member in every other sample by reference to qualities other than those immediately under investigation. The object of matching samples is to obtain better estimates of differences by removing the possible effects of extraneous variables. A survey of political attitudes, for example, might match individuals of several different samples according to age, income, education, race, and sex, but not residence, in order to ascertain regional differences in political beliefs. (see: *SURVEY RESEARCH, sample, variable)*

matrix—a table of numbers in which the row and column positions have some meaning, as well as the number entered in that row and column position. For example, a matrix might be a table whose rows are nations in descending order of wealth (as judged by GNP per capita) and whose columns are GNP, population, and amount of US aid

COUNTRY	GNP	POP	US AID
Argentina	$25,420 m.	24 m.	$22 m.
Chile	$ 7,385 m.	9 m.	$10 m.
Brazil	$34,600 m.	98 m.	$96 m.
Dominican Rep.	$ 1,395 m.	4 m.	$19 m.
Ecuador	$ 1,602 m.	6 m.	$10 m.

maximin, minimax—strategies derived from game theory in which players try either to maximize their minimum gain (maximin) or minimize their maximum loss (minimax).

If several different courses of action each offer varying arrays of payoffs, the maximin player will select the strategy which has the best minimum

guarantee—to make sure that he has a chance for at least some success. In the matrix below, player A is selecting his attack strategy in bombing player B's supply depots. The numbers in the cells tell how much of B's supplies A can destroy. If A is a conservative or a pessimist, he will select strategy III, his maximin, because the lowest value in that column (8) is still higher than the lowest value in any other column (6, 0, 2).

Player B's Defense Strategy	Player A's Attack I	II	III	IV
1	22	8	8	11
2	14	21	16	2
3	6	0	12	10

The minimax player, by comparison, will select the strategy which has the smallest worst value—to guarantee that if things go wrong it will still not be too bad. If B is designing defenses to protect his supply depots from A's attacks, his strategy 3 is a minimax because its highest possible value, 12, still hurts B less than the highest possible values from the other strategies (22, 21). Minimax is thus the defender's version of a maximin strategy. (see: *saddle point*)

McNemar test—a statistical test used to determine whether an observed difference between proportions, or percentages, derived from two related samples is significant or whether that difference could be a random accident. The McNemar method might be used, for example, to test whether a propaganda broadcast affected a group's political views. The two samples would be taken before and after the broadcast and tested to see if any observed differences in political views were significant. (see: *confidence level, significance level*)

mean—the arithmetic average, written $\bar{x}$ and pronounced "x-bar." The mean is equal to the sum of all observations divided by the total number of observations. For example, the mean of the series: 2, 4, 4, 8, 9, 10, and 12 is 7. The mean is one of three common measures of central tendency which try to use one number to describe a series of numbers.

The *geometric mean* is a special case of the mean found by taking the Nth root of the product of N numbers. For example, the geometric mean of 2, 5, 10, and 15 is the fourth root of 2 x 5 x 10 x 15 = 6.22. The geometric mean is used to average rates of change, percentages, and ratios, and for highly skewed distributions such as income distribution.

The *harmonic mean* is the reciprocal of the arithmetic mean of the numbers being averaged. For example, the harmonic mean of 2, 5, 10, and 15 is 4.62 since 4 ÷ (1/2 + 1/5 + 1/10 + 1/15) = 4 x 30/26 = 4.62. (see: *measure of central tendency, median, mode*)

mean deviation—the average difference between the values in a set of values and their mean or median. If 3, 4, 5, 8, and 10 are values in a set, for example, their mean is 6 (3 + 4 + 5 + 8+ 10 = 30 divided by 5 = 6). When the differences between each value and this mean (3, 2, 1, 2, 4) are added and divided by the number of values (5), the mean deviation of 2.4 is found. (see: *mean, median*)

mean range—the arithmetic mean of the ranges of a set of samples of the same size. The mean range in repeated sampling may be used as an estimator for the population standard deviation. If the ages of 20 people in three different samples are from 20 to 30 for the first sample, 20 to 45 for the second, and 20 to 60 for the third, the ranges of the three samples are 10, 25, and 40 respectively. The mean range for the three samples is 25, since 10 + 25 + 40 = 75/3 = 25. (see: *mean, range, population, sample, standard deviation*)

measure of association—a statistical test which shows the degree of dependence or independence which exists between two or more variables. Some of the more commonly used measures of association include: Cramer's V, Goodman and Kruskal's gamma, Kendall's tau, Lambda, Pearson's C, Phi-square, Somer's D, Tschuprow's T, and Yule's Q.

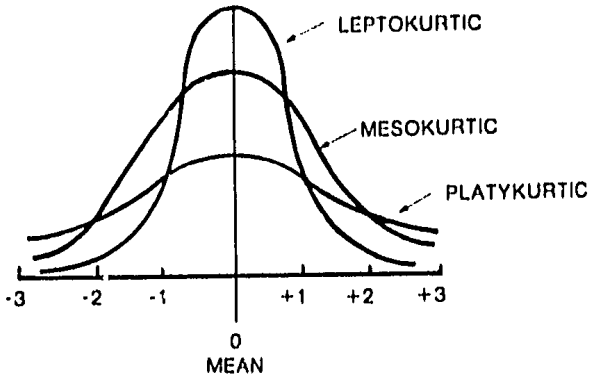
measure of central tendency—any of several statistics which summarizes the "middle" value of a set of data. The mean is the most common measure of central tendency, particularly for interval or ratio level data. If the data are ordinal, the median is used, and the mode is most often used to best summarize nominal level data. (see: *data, mean, median, mode*)

measure of location—(see: *measure of central tendency*)

measures of variability—any of several statistics which summarize the spread or variation of data around its central value. The most common are range, variance, and standard deviation.

median—the middle value of a population, denoted "M." The value which has half of the group above it and half the group below it. For example, in the series: 2, 4, 4, 8, 9, 10, 12 the median is 8. (see: *mean, mode*)

mesokurtic—a normal frequency distribution curve. (see: *kurtosis*)



methodology—a body of knowledge and technique relating to the process and assumptions of scholarly inquiry within a discipline. (see: *approach, technique*)

minimax—(see: *maximin*)

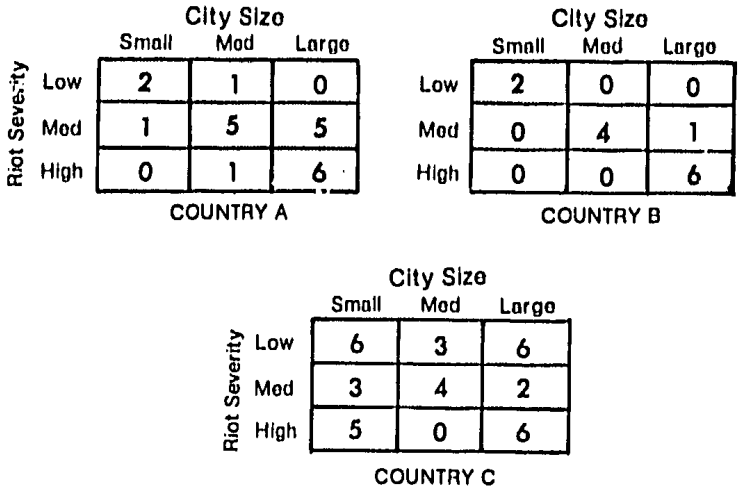
modal personality—a description of an entire group as revealed by the dominant personality traits of its adult members. Modal personality is essentially an average, and any discussion of it should include an estimate of the degree of variation from the mean. Complex industrial societies may be multi-modal, and personality patterns may cluster around two or more dominant types. Developing states are often uni-modal—that is the modal personality is also the "national character." (see: *personality trait*)

mode—the value occurring most often in a population. In the series: 2, 4, 4, 8, 9, 10, 12 the mode is 4. (see: *mean, median*)

model—a simplified description of reality which abstracts and generalizes about the complexity of the real world in order to accurately and clearly isolate the relatively few variables of particular interest to the analyst. (see: *variable, SIMULATION*)

moment—in statistics, the first moment is the mean of the distribution, the second moment is the variance, the third is the measure of skewness of a distribution, and the fourth is the kurtosis.

monotonic—a relationship which tends in the same direction when shown in a matrix. Both diagrams A and B indicate monotonic relationships. Diagram C does not.



morphogenesis—in systems analysis, a set of interrelated processes that tend to preserve a system's values in the face of a changing environment by evolving its structures. The international system after World War I and II attempted to preserve the system value—peace—in a changing environment by setting up first the League of Nations and then the UN. In contrast, homeostatic systems tend to preserve structures. Morphogenesis is said to be a better social model because a society in fact does much more than preserve its central structures; it is continually in the process of elaborating upon and changing those structures so that institutions will still serve the society's values even when conditions change. (see: *SYSTEMS ANALYSIS, homeostasis, model, structure*)

multicollinearity—in regression analysis, a situation in which the independent variables are more highly correlated among themselves then with the dependent variable. This situation makes the regression unreliable. (see: *REGRESSION, independent variable, dependent variable*)

multiple correlation and regression—techniques which measure how much of the variation in a dependent variable can be explained by two or more independent variables acting together. (see: *CORRELATION, REGRESSION, veriable*)

multiple correlation coefficient—(see: R^2)

multivariate analysis—a generic term for any analysis which takes into account more than two variables. (see: *variable)*

mutually exclusive events—events which cannot occur simultaneously, e.g., heads and tails are mutually exclusive on a single flip of a coin. (see: *independent events*)

N, n—the number of items in a series or population, or the last one of those items. Usually capitalized when referring to the entire population, lower case when referring to a sample. (see: *population, sample*)

national character—a modal personality which appears to be dominant for an entire state. (see: *personality trait, modal personality, NATIONAL CHARACTER STUDIES*)

natural experiment—research utilizing observation of human behavior or some other phenomenon in the normal field setting without injecting artificial stimuli. When a researcher studies the political statements made by businessmen at lunch without introducing questions to elicit answers, for example, he is conducting a natural experiment.

negative relationship (or inverse relationship)—a relationship between two variables such that when one rises in value, the other falls, and vice versa. (see: *positive relationship*)

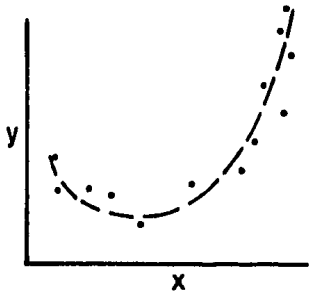
nested hypotheses—a sequences of hypotheses in which the hypothesis at any stage is contained in all hypotheses later in the sequence. (see: *hypothesis*)

noise—in communications theory, a series of random disturbances which increase the possibility that the information received will be different from the information sent. (see: *COMMUNICATIONS THEORY)*

nominal scale data—(see: *data)*

nomothetic explanation—explanation based on general propositions applicable to classes of events or objects rather than specific propositions that explain unique or individual social or political phenomena. The latter is called an *idiographic* explanation. Political and other social sciences that emphasize explanation by theoretical generalization are considered nomological, while history is often regarded as an essentially idiographic discipline.

nonlinear correlation—a correlation between two variables which cannot be graphed as a straight line. (see: *CORRELATION*)

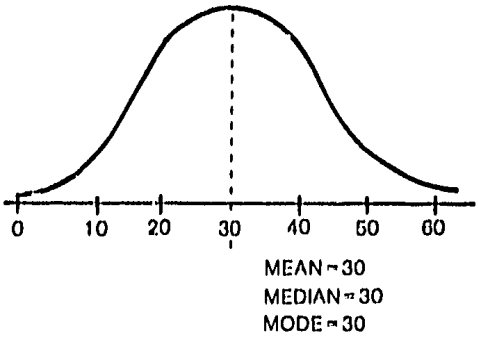


Non-Linear Correlation

nonparametric statistics—(see: *statistics*)

non-zero-sum games—(see: *GAME THEORY*)

normal distribution—a sample in which the cases tend to cluster about their average and progressively decrease in number toward either extreme of measurement. When graphed, the normal distribution appears as a symmetric bell shaped frequently distribution curve in which the mean, median and mode are identical. (see: *distribution, frequency, frequency polygon, mean, median, mode, statistics*)



STANDARD NORMAL DISTRIBUTION

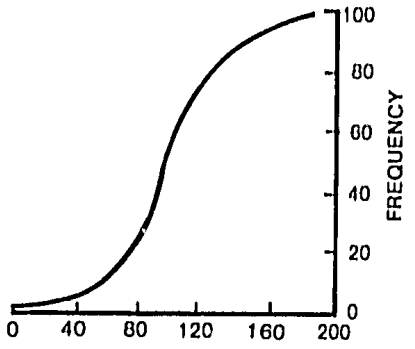
null hypothesis—the opposite of an original hypothesis, one which an analyst formulates with the expectation that it will be discredited by the data gathered. One using the scientific approach does not usually try to prove directly that a hypothesis is true but seeks to discredit its opposite with a given degree of probability. For example, it may be hard to prove directly that a certain strain of virus causes a disease. The theory may be supported, however, if diet, weather, insects, other viruses, etc., can be discarded from consideration. If the test of the null hypothesis shows it can be rejected with 99% confidence, the original hypothesis is strengthened, but not proven true. (see: *hypothesis testing*)

objective function—the criterion by which we select the best solution from a realm of alternatives. For example, in trying to build the cheapest battleworthy tank, the objective function would be cost, and the constraints would be speed, firepower, strength, etc. (see: *constraining equation*)

observation—an element of a population or sample, a specific case. (see: *population, sample*)

odds—the chance that a given event will occur, expressed as a ratio rather than a percentage as in probability. If the probability of an event occurring is 55%, for example, the odds for its occurrence are 55 to 45 (100 minus 55 = 45) and the odds against its occurrence are 45 to 55, or 11 to 9 for and 9 to 11 against, respectively. (see: *probability*)

ogive—a special case of a frequency polygon which graphically represents the cumulative frequency of a set of data. The ogive below, for example, shows that 20 out of a sample of 100 had IQ scores under 80 and 80 had scores less than 120. (see: *distribution, frequency, frequency polygon*)

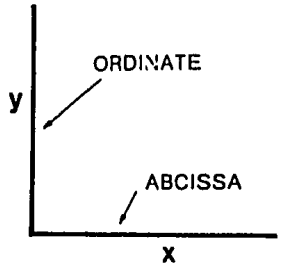


OGIVE

opportunity cost—the foregone benefit which could be reached by reassigning resources from their current use. The opportunity cost of investing in a ballistic missile system, for example, is measured by the amount of benefit that could have been gained by spending for some other program. Opportunity costs involve no cash outlay, but they are real in the sense of representing benefits or income which we deny ourselves by selecting the present policy. Opportunity cost, therefore is useful as a way of measuring the effectiveness or desirability of some investment or resource allocation problem, because it asks how much could be gained by alternative programs.

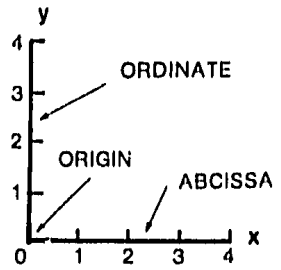
ordinal scale data—(see: *data)*

ordinate—the left vertical axis of a graph, designated Y. (see: *abscissa)*



organization—a goal-seeking collectivity of individuals having some kind of structure designed to help achieve the collective goals. In formal organizations, role relationships among the members are explicitly stated, while informal organizations evolve from repeated contacts among people without an express statement of goals or defining of roles. (see: *ORGANIZATION THEORY, GROUP THEORY*)

origin—the intersection of the x and y axes of a graph, where both equal zero. (see: *abscissa, ordinate)*



output—activities carried on by a political system in response to demands or stresses placed upon the system in the form of inputs. Outputs take the form of governmental policies, programs, decisions, and implementing actions. (see: *COMMUNICATIONS THEORY, SYSTEMS ANALYSIS, input)*

paradigm—a model, pattern, or example that helps organize thought and give direction to research. (see: *model)*

parameter—a condition or value that is given or held constant or assumed to be a limiting condition in a research problem. Parameters serve as boundaries or constraints upon the operation of the variables. Parameter values may be arbitrarily assigned for research purposes. In a decision-making study, for example, perfect information may not be available in fact but is assumed as a parameter of the situation.

In statistics, *a parameter* is a numerical characteristic of a population, such as a mean, median, upper limit, or lower limit. This is contrasted with a *statistic*, which is a numerical characteristic of a sample of the population.

Pareto optimal—a weak optimality condition which describes situations in the distribution of some desired good from which no one individual can be made better off without making someone else worse off. Essentially it is an efficiency criterion, eliminating waste, but not concerned about the equality of the distribution. Thus, if we are dividing $1000 among 3 people, a $700/$200/$100 split is a Pareto optimal combination, because all the money has been used up, and no one can be made better off (i.e., get more money) without making one or both of the others worse off. There are obviously several different possible distributions which satisfy the requirements of a Pareto optimality: $500/$500/$0; $1000/$0/$0; or $333.33/$333.33/$333.33. Pareto optimality is referred to as a weak optimality condition, because in some cases it may be that equality of distribution is more important than eliminating all waste. Thus, for a true socially optimal distribution, $300/$300/$300 may in some cases be preferable to $600/$300/$100, although the latter is a Pareto optimality and the former is not. But $333.33 each might be preferable to both.

partial correlation—(see: *CORRELATION*)

payoff matrix—a table used in game theory to list possible strategy options for the players, and to describe the outcomes for each.

Player B \ Player A	I	II	III
1	5 / -1	-6 / 16	-2 / 9
2	28 / 11	21 / 4	3 / 8
3	28 / 11	6 / 2	4 / -1
4	-6 / 12	9 / 0	-1 / 14

The player picking strategies on horizontal rows (player B) gets the return in the lower corner of the intersection box, and the column player receives the payoff in the upper right. Thus if player A picks his strategy I and player B selects strategy 4, the outcome will be worth −6 to player A and +12 to player B.

Pearsonian product-moment correlation (r)—a statistic which measures the degree of association between two variables which, if plotted on a scatter diagram, would form a straight rather than a curved line. The notation "r" is simply a shorthand way of expressing the coefficient which can range from 0 to +1 for positive correlation and 0 to −1 for negative correlation.

The correlation coefficient can be translated into a percentage by squaring it and multiplying the result by 100 ($100\ r^2$). If the correlation between the percentage of the vote for a party and the percentage of blue collar workers is +.42, for example, this means that 18% of the total vote for the party is accounted for by the percentage of blue-collar workers in the voting population ($100 \times .42^2 = 18\%$). (see: *measure of association, CORRELATION*)

Pearson's C—also called the contingency coefficient, Pearson's C tests the strength of relationship (measure of association) between variables as represented in tables of 2-by-2 size or larger. The C value ranges from .71 for 2-by-2 tables to 1.0 as the number of rows and columns increases. Some social scientists recommend computing C only for tables of at least 5-by-5, since only then does the maximum value of C approach 1.0. (see: *measure of association*)

perception—the act of becoming aware of things by means of the senses. It involves two related operations—receiving impressions through the senses and the assignment of meaning to the sensory impressions. Studies of social perception deal with differences in people's perceptions arising from differences in their cultural and social backgrounds, values, and beliefs. (see: *attitude, personality trait, PERCEPTION ANALYSIS, PSYCHOLINGUISTICS*)

perfect competition—a theoretical model of an economy, used by economists to characterize broad guidelines. This usually descriptive model includes the presence of a large number of independent, well-informed sellers and buyers, none of whom are large enough to affect the prices in a well-organized market. The perfect competition model is without advertising, product differentiation, or barriers to entry. No national economy has ever closely approximated this model, but some industries—notably agriculture and national stock exchanges—seem nearest, and the model is a valuable learning tool.

performance evaluation review technique (PERT)—a system aiding in the efficient control of research and development efforts by isolating job components, bottlenecks, uncertainties, etc., which are involved in achieving the desired output subject to relevant constraints. The technique requires three different estimates of the time needed to perform a particular task. One is the "most likely" time, the others are "optimistic" and "pessimistic" estimates. An expected value is derived from this probability distribution in order to efficiently allocate resources.

permutation—any one of the sequences or changes in position possible with a group of data. The permutations of A, B, and C are:

A B C	B A C	C A B
A C B	B C A	C B A

The formula for finding the number of permutations of N things taken R at a time is:

$$P = \frac{N!}{(N-R)!}$$ (N! is read "N factorial")

Thus the number of permutations of 5 things taken 3 at a time is:

$$P = \frac{5!}{(5-3)!} = \frac{5\times4\times3\times2\times1}{2\times1} = \frac{120}{2} = 60$$

(see: *factorial*)

personality trait—the persisting, organized disposition in a person that leads him to respond in characteristic ways to his environment. Personality traits are usually equated with the underlying dispositions that determine behavior rather than the behavior itself, but are sometimes regarded as acts, or characteristic modes of behaving. Personality traits are interrelated, although psychologists differ on how traits are organized within a personality structure. Psychologists generally hold that personality traits precede attitudes and therefore tend to shape attitudes. The traits that shape responses to political stimuli are called *political personality. Modal personality* refers to personality features most widely shared by a group of people, or the most prevalent personality type. (see: *attitude, behavior pattern, model personality, NATIONAL CHARACTER STUDIES*)

Phi-square—a statistic which measures the strength of an association between two dichotomous variables, such as whether sex makes any difference in an individual's voting Democrat or Republican. Phi-square is equal to chi-square divided by the number of observations and usually varies between 0, indicating no association and 1, indicating complete association. (see: *chi square, dichotomous*)

	Male	Female
Democrat	25	25
Republican	25	25

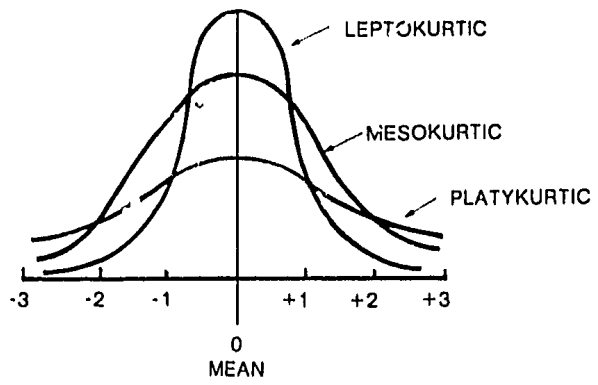
ϕ^2 (Phi-square) =0 (no association)

	Male	Female
Democrat	0	50
Republican	50	0

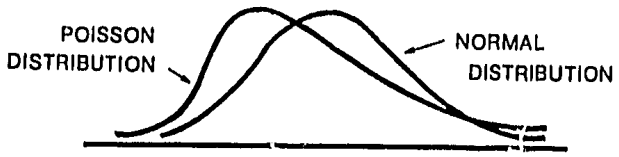
ϕ^2 (Phi-square) = 1 (complete association)

planning, programming, and budgeting system (PPBS)—a management system introduced to the executive branch in the early 1960's which includes three elements: *planning*—the study of objectives, alternative ways to achieve them, and contingencies and possible responses; *programming*—a system of evaluating the objectives of an activity and relating the costs or inputs needed to the benefits or outputs produced; *budgeting*—the process of distributing scarce available resources among competing claims. (see: *DECISION THEORY*)

platykurtic—a flat frequency distribution curve. (see: *kurtosis*)



Poisson distribution—a special type of probability distribution which often applies when there are a large number of independent events and for each of which there is only a small probability that a certain outcome will occur. The Poisson probability distribution has been applied the analysis of the number of accidents likely to befall a large number of workers per day, or the number of machines out of a large number which may break down in any one day. Its graphed plot is a single peaked curve with a steep ascent to the top and a more gradual descent to zero. (see: *normal distribution*)



political action—any observable, overt behavior of an individual or social group carried on within a political system. It may be unplanned, random behavior or part of a coherent decision process. A response to political actions undertaken by others is a *political reaction.* (see: *ACTION THEORY*)

political anthropology—the study of governmental institutions and practices among ethnic communities, particularly in primitive and tribal societies. Such study probes the relationship of political behavior to the broader group culture and examines the ways in which political institutions and practices evolve.

political communication—the transmission of meaning having relevance to the functioning of a political system. Such communication could include symbolic acts (burning a draft card), assassination, voting, speeches, communiques, or the simple sending and receiving of spoken messages. (see: *COMMUNICATION THEORY, CONTENT ANALYSIS, cybernetics*)

political community—(see: *integration*)

political culture—the pattern of orientations toward government and politics within a society. Political culture generally connotes the psychological dimension of political behavior—beliefs, feelings, and orientations—and is the product of the historical experience of the whole society as well as the personal experiences that contribute to the socialization of each individual. Within national political culture, elite and mass subcultures are often distinguishable. (see: *NATIONAL CHARACTER STUDIES, socialization*)

political development—growth and change within political systems, or change from one system to another, generally toward greater governmental capacity to cope with the demands made upon it. Commonly associated with increasing complexity, specialization, and differentiation of political institutions in a society, regardless of their democratic or authoritarian character. The terms traditional, transitional, and modern are commonly used to designate societies in different stages of political, economic and social development. (see: *development theory, political socialization*)

political ecology—the entire physical, cultural and social environment of a political system excluding other political systems. (see: *ECOLOGICAL ANALYSIS*)

political philosophy—the branch of intellectual inquiry dealing with ideas about politics, particularly ideas relating to political values, the nature of political reality, and the intellectual assumptions of political analysis. As *normative* theory, it seeks to clarify political values and to define what is desirable and moral. As speculation about what is rather than what ought to be, political philosophy is concerned with the basic nature of reality rather than particular observable manifestations of it. While not rejecting observation and experience, the political philosopher looks for rational interconnections between things as a means of describing and explaining political reality while the empirical scholar seeks scientifically verifiable hypotheses as the bases of explanation. As *analytic philosophy,* political philosophy deals with the meaning of words and concepts, the logical consistency of arguments, ways of knowing truth, and the grounds on which a proposition may be taken as true or false.

political pluralism—a society in which power is widely distributed among numerous groups arrayed in shifting patterns of conflict, competition and cooperation with one another.

political science—the systematic study of government and politics. The discipline has moved in the past century from an emphasis upon formal institutions and legal relationships to a concern for processes, the behavior of individuals and groups, and informal relationships. Methodologically, political science has supplemented the predominantly legal, historical and descriptive analysis of an earlier period with the methods and perspectives of modern behavioral science. (see: *BEHAVIORALISM*)

political socialization—the learning process by which individuals acquire beliefs, feelings, and values about government and politics. From the perspective of society, political socialization is the means by which political culture, that is the pattern of political orientations within the society, is maintained or changed. The family and educational system are generally regarded as the most important socializing agents, although political orientations may also be shaped by exposure to mass media, organized groups, informal groups, or any experience having political relevance. Political socialization is a key determinant of political behavior. (see: *attitude, political culture*)

political sociology—the study of political institutions and processes in their social setting. It is concerned with the ways in which political phenomena influence, and are influenced by, other aspects of society. The whole society or macro approach deals with such questions as the social foundations of power, the impact of social and economic class conflict on politics, and the reciprocal influence of political institutions upon social class. The micro approach to political sociology focuses on particular political institutions as social organizations, including their formal and informal structures, leadership patterns, methods of managing conflict and relationships with other organizations. (see: *GROUP THEORY, ELITE ANALYSIS, ORGANIZATION THEORY*)

political system—the pattern of human relationships through which authoritative decisions are made and carried out for a society. A political system is distinguished from other social systems by four characteristics: it is universal in its reach, extending to all members of a society; it claims

ultimate control over the use of physical coercion, it has the right to make binding decisions; and its decisions are authoritative, bearing the force of legitimacy and a substantial probability of compliance. (see: *legitimacy, system, SYSTEMS ANALYSIS, SYSTEMS DYNAMICS*)

political theory—a body of thought that seeks to evaluate, explain, and predict political phenomena. Political theory is also a subfield of political science concerned with political ideas, values, concepts and prediction of political behavior. In this broad sense it has two major branches—*political philosophy* with its value, analytic, historical, and moral concerns, and *empirical theory* devoted to an effort to explain, predict, guide research, and organize knowledge through the formulation of abstract models and scientifically testable propositions.

politics--human activity concerned with making and implementing decisions vested with the authority of the society for which the decisions are made. Politics connotes activity or process, whereas *political system* implies the existence of structures or patterned relationships.

polity—the political organization of a society. Polity may refer to the citizenry of a particular country or it may refer to the institutional forms and processes through which the country is governed. The concept may apply to highly organized states as well as very primitive societies in which recognized political authority has barely emerged. (see: *political system*)

population—any group of individuals or things which have at least one characteristic in common. (see: *sample, parameter*)

positive relationship—a relationship between two variables such that the values of each rise and fall together. (see: *negative relationship*)

post hoc ergo propter hoc—a logical fallacy which may occur when attempting to establish causality. It results from mistakenly assuming that merely because something occurred after some other event, it must have occurred because of it. Temporal sequence can disprove causality, but can not prove it.

post behavioralism—an intellectual movement in political science dating from the late 1960's, that asserts the obligation of political scientists to become more "relevant" and concerned with values and to use their special knowledge of politics and political systems to improve society. It represents an attitude toward the discipline and the profession of political science held by some political scientists rather than a well developed doctrine or school of thought.

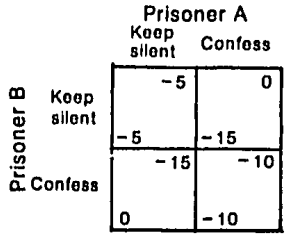
power—the capacity to affect the behavior of others in some desired way deriving from the threat of punishment, force, or coercion, as distinct from *authority* in which compliance is voluntary.

power function—in the theory of hypothesis testing, the probability that the null hypothesis will be rejected when in fact the alternate hypothesis is true. (see: *hypothesis testing, null hypothesis*)

premise—an idea, belief, or assumption one holds to be true. It is the starting point of a syllogism used in deductive reasoning. (see: *deductive reasoning, syllogism*)

present value—the current worth of future income. The formula for present value is given by the amount of money to be received divided by the sum of one plus the interest rate. Thus, if the interest rate is five percent, the present value of an expected $100 to be gained a year from now is only $100 ÷(1+.05) or about $95.24. This is the amount of money you would have to invest now at 5 percent in order to get back $100 one year from now.

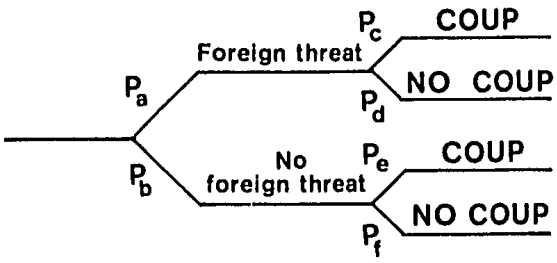
prisoner's dilemma—a class of variable sum matrix problems dealt with in game theory. This describes the situation in which the best alternative for both players is to cooperate, but due to a lack of communication or trust, they each have incentives to defect, or break their coalition, bringing results neither side wants. The game takes its name from a situation in which the police have apprehended two criminals suspected of a major crime. The police already have enough evidence to convict both for a related, less important crime, but unless one confesses, the major charge will not stick.



The police separate the two criminals, explain the situation, and offer each a reduced sentence in exchange for a confession implicating the accomplice in the major crime. The matrix below displays the choices. If each keeps silent, both get a light sentence for the minor charge (worth −5 to each). If one confesses and the other does not, the confessor goes free (worth 0) and the silent one gets the heavy sentence (worth −15). If both confess, both get a reduced sentence on the major crime (worth −10). In this situation, the best strategy for both is silence. But with no communication between them, and little trust, each may perceive it to be in his own interest to confess. Moreover, each knows that the other is facing a similarly tempting situation, and the game has its saddle point in the lower right, with mutual ruin. This type of game is often used by analogy to describe situations like superpower diplomacy, retail sales competition, etc. (see: *variable sum game, saddle point*)

probability—the chance that a given event will occur, expressed as a fraction from zero (signifying an impossible event) to one (signifying an event that will definitely occur), or the corresponding percentage from 0 to 100.

probability diagram—a graphical method for breaking a problem into its component parts, assessing the probability of each part and thereby finding the overall likelihood of an occurrence. If the analyst were assessing the probability of a coup in a foreign country, for example, and he felt the possibility of a foreign threat to be the major determining factor as to whether a coup would occur or not, he might construct the following probability diagram to assist his analysis. (see: *probability, DECISION THEORY*)



probability distribution—a prediction of the relative frequencies with which possible values of a variable will occur. For example, in predicting the number of Redskin football victories this season, we could pick just one number, as our prediction, but we might do better by defining a probability distribution such as: .1 chance of winning 10 or fewer games; .2 chance of winning 11 or 12 games; .3 chance of winning 13 or 14 games; .3 for 15 or 16 games; and .1 of winning more than 16 games—totaling 1.0 or 100% probability.

process—a sequence of related actions or operations. As a general concept, political process embraces all the activities by which people attempt to gain and wield legitimate influence within a society. (see: *political system, politics*)

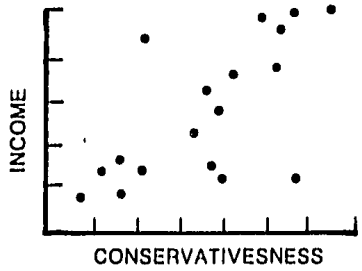
program—a set of instructions or steps that tell a computer how to handle a complete problem.

projective methods—a research technique used to measure psychological variables by inducing human subjects to project their inner attitudes, feelings, values, motives, and needs upon external objects. The most common projective methods are association, constructions, completion, and expression. In the use of association techniques, the subject is asked to respond to words or pictures with the first idea that occurs to him. The response will presumably be a projection of his attitudes and feelings. Construction requires the subject to produce a story or picture which is then interpreted. With the completion technique, the subject is asked to finish a sentence, story, or incomplete situation. Expressive techniques permit the attitudes and feelings of the subject to be deduced from his behavior in some observed activity—talking, playing games, and so on.

property space—one of two basic graphic forms of data representation. The *tabular property space* cross tabulates two variables, each divided into any number of subvariables, in a table. A *scatter diagram* represents the two variables as the two axes of a graph, and pinpoints each observation on the graph.

CONSERVATIVENESS	INCOME: Low	INCOME: Mod	INCOME: High
Low	5	2	1
High	1	4	5

TABULAR PROPERTY SPACE



SCATTER DIAGRAM

psychological scaling—the quantification and measurement of either certain personality traits or entire personality types. The analyst ranks individuals in particular categories based on his knowledge of their behavior. The categories are given values which can be statistically manipulated. "Miller's Scale Battery of International Patterns and Norms," for example, contains 15 rating scales to help ascertain important attitudes and beliefs within national cultures.

A number of scales and indexes have been designed to quantify and analyze individual behavior. A VIP behavior and performance study, for example, identifies five measurable factors—task orientation and functioning, emotional adjustment, interpersonal style, leadership style, and influencibility. (see: *SCALING, personality trait, behavior pattern*)

psychology—the science of human and animal behavior; the study of an organism in all its variety and complexity as it responds to the flux and flow of the physical and social events which make up the environment. These responses tend to coalesce into *behavior patterns,* which can be either overt or covert. Psychologists employ many methods for discovering and analyzing behavior patterns beyond simple observation. For example, when an individual is angry, subtle changes occur in his digestive tract, his vascular system, and breathing, all of which can be discovered through the use of special instruments. (see: *personality trait, behavior pattern*)

psychoanalysis—a branch of psychology directed toward the understanding, cure, and prevention of mental disorders. Generally the roots of human behavior are sought in unconscious motivation and conflict.

public opinion—beliefs and attitudes of people generally on a social or political issue. Public opinion may relate to matters of fact or of preference. (see: *SURVEY RESEARCH, attitude, behavior pattern*)

quartile or Q value—the three values in a set of data below which lie the bottom 25% of the data (Q_1), below which lie the bottom 50% of the data (Q_2), and below which lie the bottom 75% of the data (Q_3). Quartiles are a class of arbitrary divisions called fractiles which can be changed to suit particular analysis needs. *Deciles*, for example, divide the data into increments of 10%. (see: *fractile*)

MONTH	NUMBER OF RIOTS	CUMULATIVE PERCENT
1	0	.0
2	3	2.5
3	12 (Q_1 =riot 30)	12.5
4	21 (Q_2 =riot 60)	30.0
5	36 (Q_3 =riot 90)	60.0
6	24	80.0
7	12	90.0
8	9	97.5
9	3	100.0
10	0	100.0

R-squared (R^2)—a measure of the "goodness-of-fit" in regression analysis. If $R^2=0.78$, then 78% of the variation of the dependent variable is explained by or determined by the independent variables. R^2 is also called the *coefficient of determination.* The square root, r, is known as the *multiple* or *Pearsonian product-moment correlation coefficient.* (see: *REGRESSION*)

random—governed by chance. A random variable is one whose values are assigned by chance; a random sample is one whose elements are selected by chance. This means that each number or element in the population must have an identical probability of being selected.

range—the difference between the largest and smallest values in a set of data. (see: *variable, domain*)

rank order correlation—a technique of correlation which involves the ranking of subjects (for example, countries) in descending order of magnitude on each of two variables (for example, gross national product and size of armed forces) and determining the extent to which rank order on one variable is similar to the ranking on the other. This form of correlation takes only the difference in rank into account. (see: *CORRELATION, product moment correlation*)

rational actor model—traditionally the most widely used model in political analysis, it sees the nation as a unitary and rational actor. A single entity, such as the Soviet Union, Peking, or the Third World, is assumed to be capable, like an individual, of having ideals, perceiving interests, reacting to events, laying plans, having purposes, making decisions, taking actions, and pursuing policies. In his book, *Essence of Decision,* Graham Allison contrasts the rational actor model (which he calls Model I) with the organization (Model II) and bureaucratic politics (Model III) models. The organization model assumes that the actions of a government are not the actions of a national entity but those of the semi-autonomous organizations making up the government. The bureaucratic politics model conceives of government as an arena of competition among individuals within that government. (see: *BUREAUCRATIC POLITICS, ORGANIZATION THEORY*)

ratio scale data—(see: *data—ratio*)

raw score—the actual original numerical value of an item or individual according to one test or measurement. This value may be transformed into a standardized z-score to allow the comparison of values from entirely different sorts of tests. (see: *z score*)

real number—a value which can be expressed either as a whole number or a fraction—age or height, for example—as opposed to an integer, such as number of persons or weapons, which must be a whole number. (see: *integer, variable*)

recruitment—the selection of people to fill roles in a social or political system. Recruitment refers both to the filling of formal, legal positions, such as president, legislator, or civil servant, as well as less formal roles, such as lobbyist, party activist, or media representative. Specific techniques of recruitment may include co-optation, appointment, election, rotation, apprenticeship, examination, and purchase. Since recruitment is a function performed within every political system, structural-functional analysis treats it as a major category of inputs. (see: *STRUCTURAL-FUNCTIONAL ANALYSIS, POLITICAL SYSTEM*)

reductionism—a method of explanation in which the attributes or behaviors of the whole are reduced to, or explained by reference to, the attributes or behaviors of its constituent parts. A reductive approach, for example, might explain the functioning of a political system in terms of personality variables, or by reference to particular components of the system, such as actor characteristics, processes of political socialization and recruitment, political values, and so on. Reductive explanation may be contrasted with *holism*, which focuses on the functioning of the whole system rather than its parts. (see: *holism*)

reference group—any group of people from which a person adopts attitudes, beliefs, or values that affect his behavior. His behavior is therefore determined by reference to theirs. A *comparative reference group* provides a standard for comparative self-evaluation—for example, a group of co-workers, classmates, or persons of one's own social background. A *normative reference group* helps to shape the values, attitudes, and norms by which the individual evaluates the world around him. (see: *attitude, behavior pattern, personality trait*)

regressor—any independent variable in a regression relationship with another variable. The variable "unemployment," for example, would be a regressor in a regression study of the causes of riots. (see: *REGRESSION, independent variable*)

regret—in decision theory, the difference between the total loss as a result of a decision and the unavoidable loss regardless of the decision. It is this excess loss that has to be minimized. (see: *DECISION THEORY*)

rejection region—(see: *critical region*)

relative frequency (or proportional frequency)—a frequency which is expressed as a proportion of the total number of occurrences or the total number of members. (see: *frequency*)

I.Q. SCORES	FREQUENCY	RELATIVE FREQUENCY
0 - 40	5	1/20
41 - 80	20	1/5
81 - 120	50	1/2
121 - 160	20	1/5
161 - 200	5	1/20

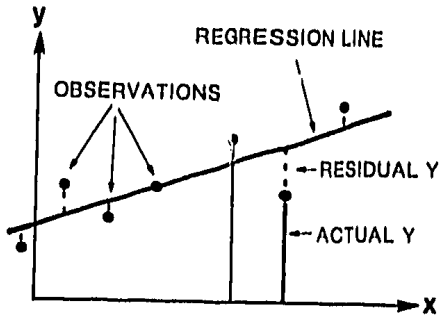
relaxed oscillation—in a time series model, when the value of a phenomenon, although oscillating, generally increases during a given period of time through the operation of its internal forces. These forces precipitate a "crisis" or "bursting" followed by a return to zero after which the process begins again. Models of some economic phenomena, such as inflation and depression, follow this pattern. (see: *model, TIME SERIES ANALYSIS*)

reliability--the extent to which a measurement technique produces the same results time after time, assuming the relevant conditions to be held constant. It implies constancy, dependability, and accuracy of measurement, but it does not imply that the measurement technique will accurately measure the variable the researcher is concerned with, in the way he wants it measured. That is a question of *validity*. (see: *validity*)

reproducibility—an experiment or survey is said to be reproducible if, on repetition of the experiment or survey under similar conditions, it gives the same results.

research design—the overall plan for a research project, specifying what the researcher proposes to do and how he proposes to do it.

residual—the difference between the actual value of an observation and the estimated value found from the regression formula. If the observed values were plotted on a graph with the best fitting regression line drawn through them, for example, the residual would be the sum of the differences between the points and the regression line.



In regression analysis, the residual represents the portion of the observed variation in a dependent variable which is not explained by the independent variables. For example, if we are attempting to predict government instability, we may find that a knowledge of demographic factors such as unemployment, urbanization, size of the armed forces, and literacy allows us to correctly predict instability 80% of the time. The residual in this case, 20%, tells how often we will be wrong because of the insufficient explanatory power of the data. If the residual is large, we have not specified all the important independent variables which influence the variable under examination. (see: *REGRESSION*)

return period—in *TIME SERIES ANALYSIS*, the interval of time taken by the series to return to some assigned value, usually an extreme value. For example, the "return period of flooding" in a river is the time it takes the river to recede to normal levels between floods. (see: *TIME SERIES ANALYSIS*)

Rice-Beyle cluster bloc—(see: *CLUSTER ANALYSIS*)

risk—generally synonymous with uncertainty. Sometimes risk refers to situations where the probabilities are known, and uncertainty to situations where they are not.

role—the set of behaviors expected of one who occupies a given position in society. Political roles are the expected behaviors associated with positions as legislator, party leader, voter, or revolutionary, which relate to making and implementing authoritative decisions for society. The occupant's actual behavior is called *role enactment*, which is affected by his perceptions of the role expectations of others, his own interpretation of his role, his sensitivity to demands for specific enactments generated by the

situation, and his skills and capacity to respond. A group of roles occupied by a single individual is a *role set.* Roles within a set that make contradictory demands upon the individual produce *role conflict.* (see: *behavior pattern*)

routine—a computer program or set of instructions arranged in proper sequence to cause a computer to perform a desired task.

run—a single continuous performance of a computer routine.

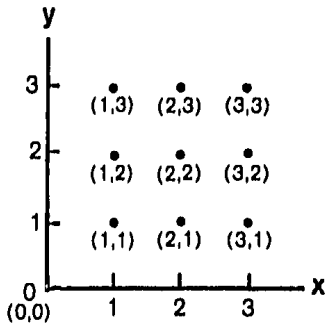
sample—a small part of a population, selected to represent the entire population. (see: *population, statistic*)

saddle point—a solution to a game theory matrix which represents both a maximin to one player and a minimax to the other. Any unilateral change from the saddle point results in a reduced payoff for the player making the change, thus the saddle point represents a stable situation. Not all matrices have saddle points, and in those which do the point is not always reached during the course of the game.

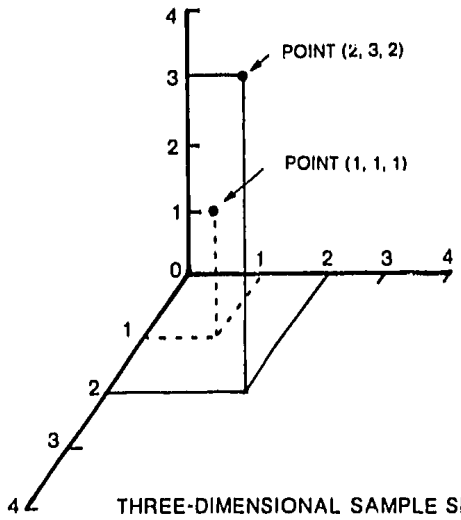
An example of a saddle point is in the matrix below. Player A is a maximin player, attacking the storage areas of player B, the minimax player. The number in each cell tells how much material is destroyed—A wants a high figure, B a low one. Player A selects strategy III because the lowest value in the column (70) is higher than the lowest value in either column I (30) or column II (50). Player B selects strategy 3 because the highest value in row 3 (70) is still lower than the highest value in either row 1 (90) or row 2 (80). The solution, the intersection of row 3 and column III is a saddle point, representing a stable equilibrium. It is stable because if either player suddenly tries to disrupt it by switching to another strategy, the outcome is worse for him. (see: *maximin, minimax*)

		Player A (MAXIMIN) ATTACK STRATEGIES		
		I	II	III
Player B (MINIMAX) DEFENSE STRATEGIES	1	30	90	80
	2	80	50	80
	3	50	60	70

sample space—a set of points plotted on a graph corresponding to all possible values of all possible samples.



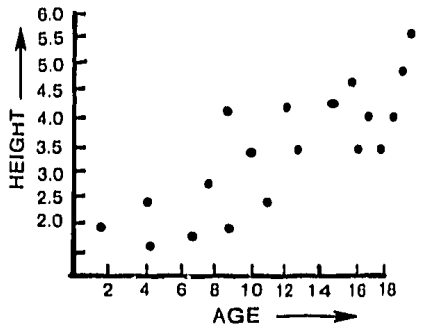
If in the two-dimensional sample space the X axis represents three issues in the UN and the Y axis represents votes (yes, no, and abstain), the sample space would show all the possibilities for those votes. If a third variable were added—level of development of the voting country (underdeveloped, developing, and modern)—as axis Z, a three-dimensional sample space would be needed. The point 2, 3, 2 whould then represent an abstention by a developing country on the second issue. The point 1, 1, 1 would represent a yes vote by an underdeveloped country on issue 1.



THREE-DIMENSIONAL SAMPLE SPACE

scale—a means of classifying, ranking, or otherwise measuring objects according to their possession of a given characteristic. A scale may be constructed to measure one of four levels of data: nominal, ordinal, interval, or ratio. A scale of political instability, for example, might rank instability in a country each month from 0 to 5, ranging from no overt acts of violence (0), through assassination (3), to civil war (5). (see: *data, SCALING, GUTTMAN SCALE*)

scatter diagram—a graph plotting observed points, used to visually inspect the relationship between the variables represented on the two axis. For example, a scatter plot with height on the vertical axis and age on the horizontal axis may show 19 observations like this:



set—any defined collection, category, group, or class of things. A set must be defined by listing all of the objects in it, or more commonly, by a rule stating the criteria for membership. A part or sample of a set is known as a *subset*. The *universal set* consists of all elements of whatever is being examined, and a set without any members in it is an *empty set* or *null set*. *Set theory* is a branch of mathematics devoted to the manipulation of numerical sets.

short-circuiting—(see: *distortion*)

short run—in economics, a period of time in which a firm may vary its utilization of some but not all resources. In the short run, output may be varied by hiring and firing, by purchasing more or fewer inputs, etc., but not by major changes in fixed plant facilities. Changes in physical plant take longer in most cases, and these adjustments are made only in the long run. The dividing line between the long and short run is flexible and will vary from industry to industry. (see: *long run*)

sigma or sum--the Greek letter sigma is the symbol used to indicate the operation of successive additions. The numbers above and below the Σ indicate the upper and lower limits of the operation. For example,

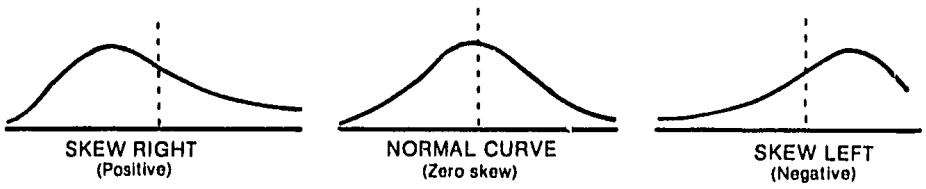
$$\sum_{1}^{3}(n+1)=9 \quad \text{because } (1+1)+(2+1)+(3+1)=9$$

$$\sum_{3}^{6}(n^2)=86 \quad \text{because } 3^2+4^2+5^2+6^2=86$$

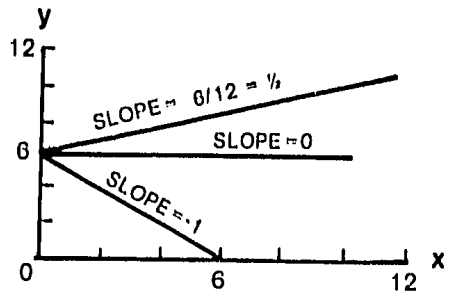
significance level—(also called confidence level) the statistical probability, derived from a significance test, that the observed data would occur solely by chance. (see: *significance test*)

significance test—a statistical test designed to determine the probability that the distribution of observed data could occur entirely by chance. These tests generally specify some low probability level, such as 1 or 5 percent, as the threshold for consideration of the data. This means that unless we are 99 or 95 percent certain that the observed distribution could not occur by chance, the data is considered insignificant. Some of the more common tests of significance include: chi square, F test, Fisher exact test, Kolmogorov-Smirnov D test, McNemar test, t test, and z test. (see: *confidence level*)

skewed distribution--a lopsided frequency distribution, in which the mean, median, and mode do not coincide as they would in a normal distribution. A statistic measuring skewness will equal zero when there is left-to-right symmetry; a positive value indicates a distribution skewed to the right; and a negative value indicates a distribution skewed to the left. (A distribution is said to be skewed to the direction in which the excess tail lies.) (see: *kurtosis*)



slope—a measure of the steepness of a line. It is computed by measuring the number of units the line rises (or falls) per unit of movement to the right. That is, if moving 12 units to the right involves a 6 unit rise in the graph of the line, then the slope is 6/12 or 1/2. A negative slope indicates that the line falls from the upper left of the graph to the lower right, and the absolute value of the slope then tells how sharp the fall is.



smoothing—any procedure that reduces the fluctuations or random scatter in data, or imposes a pattern upon seemingly random data. *Regression* is a type of smoothing. (see: *random, REGRESSION*)

socialization—(see: *political socialization*)

social matrix—the web of social relationships within which political behavior occurs. Political roles are examined in relation to other social roles and to the group and class structure of the society. An analyst studying the social matrix of an African government, for example, would begin by analyzing the social aspects of the society—tribal organization, class structure, clan system, traditions and mores—and would study the political leaders in this context. (see: *role, society*)

social mobility—(see: *social stratification*)

social science—the applications of scientific methods to the study of human behavior, the body of knowledge produced, or the disciplines engaged in such inquiry. *Behavioral science* is a newer term for essentially the same approach. Anthropology, sociology, psychology, economics, and political science, are generally regarded as the social sciences. History, geography, law, and public administration, are frequently added to the list.

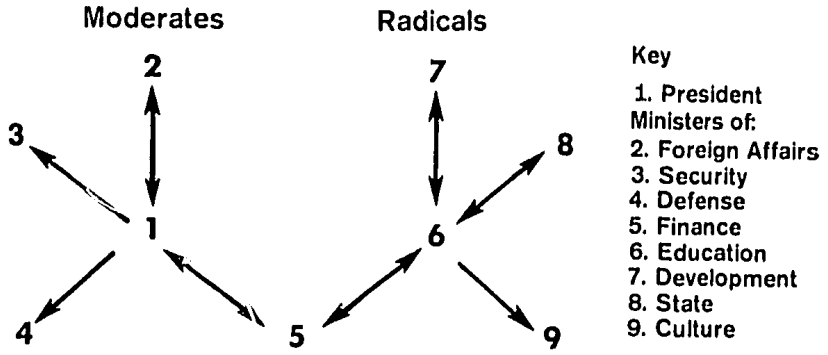
social stratification—the division of society into classes or strata, hierarchically ranked according to perceived differences in wealth, prestige, or other social characteristics. Movement of individuals from one level to another or the shift in relative position of a whole stratum or class is known as *social mobility*. Political analysts are interested in stratification as part of the broad area of political behavior, particularly in the study of political power, influence, and authority. (see: *political sociology*)

social system—an aggregation of two or more persons who interact with one another in some patterned way. A social system, consisting of interacting people, may be distinguished from a *cultural system*, which refers to learned modes of behavior characteristic of a society, and from a *personality system*, which refers to the organized dispositions within a single individual that lead to characteristic responses.

society—an aggregation of people who have certain common attributes that distinguish them as a group and who interact with one another in some characteristic way. In political science, society is treated as the larger system, usually a nation, within which political actors, or interactions, may constitute a subsystem.

sociogram—the graphic representation of relationships within small groups by means of circles and connecting lines, or some other geometric figures. (see: *SOCIOMETRY*)

Sociometrically-mapped diagram



Somer's d—a statistic which measures the degree of association between two variables. The statistic varies between 0 to indicate no relationship and +1.0 to indicate a monotonic relationship (one in which the relationship tends in the same direction when graphed on a matrix). (see: *measure of association*)

spillover—an integrative process by which cooperation between political units in one issue area leads to further cooperation in other areas. Spillover typically arises from the attempt by states to resolve problems created by their previous cooperative acts. The expansion of the European Coal and Steel Community into a wider European Economic Community is a classic case of spillover. (see: *INTEGRATION THEORY, FUNCTIONALISM*)

spurious—a false causal relationship. Correlation analysis, for example, may indicate a high association between variables A and B, but causal inferences would be unfounded if in reality variable C causes the observed movement in both. (see: *inference*)

stability—a condition of a system whose components tend to remain in, or return to, some constant relationship with one another. Stability is identified with the absence of basic or disruptive change in a political system, or the confining of change to acceptable or specified limits. (see: *system, SYSTEMS ANALYSIS*)

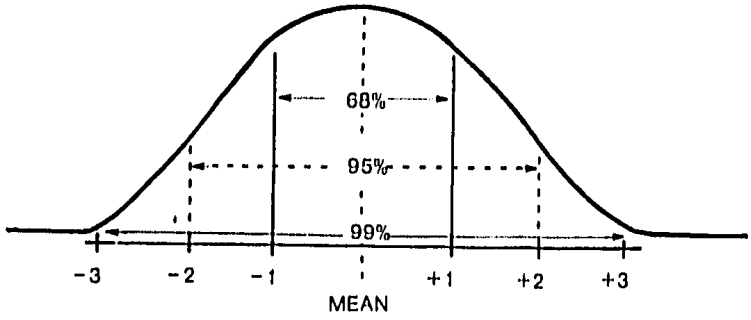
standard deviation—the most common measure of the dispersion around the central value of a group of numbers. Standard deviation, denoted "s," is equal to the square root of the *variance*, which itself is equal to the sum of all the differences between the mean value of a set of values, divided by the number of values. For example, the two sets of data—5, 6, 7, 8, 9 and 2, 3, 7, 8, 15—have the same mean and median—7—but very different standard deviations—1.4 and 4.6, respectively—reflecting that the second set is much more widely dispersed about its mean than is the first. The standard deviation of the second set is determined by the process:

$$\text{VARIANCE} = (2-7)^2 + (3-7)^2 + (7-7)^2 + (8-7)^2 (15-7)^2 =$$

$$106/5 = 21.2$$

$$\text{STANDARD DEVIATION} = \sqrt{21.2} = 4.6$$

In a normal distribution, about 68% of the observations will occur within one standard deviation above or below the mean. About 95% of the cases will be within 2 standard deviations of the mean, and about 99% will fall within 3 standard deviations above or below the mean value of the distribution. (see: *dispersion, normal distribution*)



standard error of estimate—the standard deviation of the distribution of points around a regression line. This tells how much of the total variability observed in the data remains unexplained by the regression analysis. (see: *standard deviation, REGRESSION*)

statistics—the science of dealing with the collection, classification, and presentation of numerical data. As applied to political research and analysis, statistical methods are used to summarize data and reduce it to manageable form, to locate and evaluate patterns in the data, and to assist the analyst in making reliable inferences from the data. Statistical procedures are *parametric* if they are designed to deal with samples of data having a normal distribution—that is, the cases making up the sample tend to cluster about their average and progressively decrease in number toward either extreme of measurement. *Nonparametric* statistical procedures—including Kendall's tau and Goodman and Kruskal's gamma—do not require normal distribution.

A statistic is also the name given to a numerical characteristic—such as the mean, median, or mode—of a *sample* as contrasted with a parameter, which is a characteristic of the larger *population* from which the sample is drawn. (see: *parameter*)

stochastic—a variable or process involving chance. A variable is stochastic if the values it assumes are governed by chance. The results from repeatedly flipping a coin, for example, would be a stochastic variable, as opposed to a variable such as a pricetag, which is calculated and assigned. (see: *random*)

Stouffer's H—a technique to give greater stability to Guttman scales by basing each scale item on three or more subitems, thus reducing the effect of random errors. Item 3 on a riot severity scale might be composed of subitems (a) vandalism, (b) looting, (c) arson. If at least two out of three of the subitems occurred, item 3 would be listed as occurring. (see: *GUTTMAN SCALING*)

strategy—a plan of action to defeat an opponent or to achieve some other goal. Associated with game theory where it is used to mean the plans of the players. (see: *GAME THEORY*)

structure—a pattern of related roles or established relationships among people. A political party organization is a formalized pattern of related roles, and is therefore a structure. The established relationship between the party organization and the voting public, although less formalized, could also be regarded as a structure. In *STRUCTURAL-FUNCTIONAL ANALYSIS*, structures are the patterns which govern actions, while *functions* are the results or consequences of actions. (see: *STRUCTURAL-FUNCTIONAL ANALYSIS, function*)

student's t—(see: *t test*)

supply—a measure of the willingness and ability of an individual or a society to produce and sell a good at a given price. (see: *demand*)

survey data—data which provides information on individual cases under observation but does not summarize the characteristics of a number of cases as does aggregate data. Import and export figures by transaction, or specific counts of people per household are examples of survey data. (see: *SURVEY RESEARCH*)

syllogism—an argument consisting of a major premise, a minor premise, and a conclusion. A syllogism is the major tool of deductive reasoning, which moves from general classes to specific cases. If the premises of a syllogism are true and the reasoning is logical, the conclusions must also be true.

Major premise: All Washington Redskins are overpaid.

Minor premise: Larry Brown is a Washington Redskin.

Conclusion: Larry Brown is overpaid.

(see: *deduction*)

system—any set of elements that exist in some patterned relationship with one another. The elements of a system may be specified by the analyst to meet his needs, for example, a group of people might be specified as the elements making up a political or social system, or the elements might be specified as political interactions rather than particular individuals.

The way in which the analyst defines the elements of the system and the relevant relationships among them will determine the *boundaries* of the system—the analytical line separating the system from its environment. An *open system* is one whose functioning is affected by inputs from the environment. A *closed system* is unaffected by its environment. (see: *SYSTEMS ANALYSIS*)

t test (or student's t)— a statistical test, designed to tell whether two samples are significantly different. The test compares the means of the two distributions to tell how often such a difference could be accounted for simply by chance. The t test is a special case of the F test, with wide uses in statistics to tell how much confidence we can place in the hypothesis that the two samples are significantly different. (see: *confidence level, significance test*)

tautology—a statement that is logically true by definition but can neither be confirmed or contradicted by facts. For example the statement, "The Soviet Union is or is not a dictatorship," is a tautology since it includes all logical possibilities and cannot be controverted. It also imparts no information about the subject. Propositions of mathematics or logic are also tautologies because they are necessarily true within their own logical systems.

taxonomy—a systematic classification scheme for the orderly arrangement of subject matter into categories or classes according to perceived similarities and differences. The classification of all information on a country into political, economic, military, social, biographic, geographic, and historic is a simple taxonomy.

technique—any systematic method of analyzing data in an attempt to answer theoretical questions about it. Technique is contrasted with an *approach* in that the latter is a strategy of analysis that provides a unique viewpoint and includes a set of tools or techniques for examining data. *CONTENT ANALYSIS, SIMULATION,* and *REGRESSION* are examples of techniques, while the use of social psychology in political analysis is an approach. (see: *methodology*)

tendency statement—the positing of a relationship between variables as a probability rather than a certainty. It usually suggests the direction but not the magnitude of the relationship and is subject to verification or disconfirmation, for example, the statement: Political participation tends to increase with the citizen's level of education." (see: *hypothesis, probability, variable*)

terminal—a point in a system or communication network at which data can either enter or leave.

Tetrachoric correlation—a form of correlation used when both variables have been condensed into dichotomies (high-low, large-small, etc.). The computation of the Tetrachoric coefficient is complicated, so tables have been developed as an aid. (see: *CORRELATION, coefficient, variable*)

theory—an idea or body of thought that purports to explain, predict, or prescribe in any field of inquiry. A single theoretical proposition or generalization is sometimes referred to as a theory, but the term usually implies the linking together of a number of related propositions in some coherent structure, such as a theory of power elites or voting behavior in international organizations.

Thurstone scale—a scale which ranks individuals or groups of individuals at approximately equal intervals according to some

criterion—UN members, for example, ranked according to the proportion of time they voted with the US. The equal intervals between the rankings permit the use of statistical techniques involving arithmetic manipulation of the data. (see: *SCALING*)

time-sharing—the alternate use of a device, such as a computer, by two or more users.

traditional approach—an approach to political research and analysis that emphasizes humanist, legal, and philosophical perspectives in contrast to the predominantly scientific and empirical approach called behavioralism.

transactions—a type of events data consisting of the routine exchanges across national boundaries such as mail flows or tourists. *Interactions* are the public, political flows of events such as official visits or government communiques. (see: *EVENTS ANALYSIS, COMMUNICATIONS THEORY*)

transformation—any mathematical operation performed upon a set of values to convert or transform these values into numbers more easily manipulated in statistical tests. The most common types of transformation include squaring the value, taking its square root, dividing the value into 1, and finding its log. If a set of data involves some positive and some negative values, for example, the normal procedure for finding their average (dividing the sum of the values by the number) would be inadequate since the negative values would be subtracted from the positive ones. In such a case, the values are first squared (each multiplied by itself) so that the resulting values are all positive, and then averaged.

triserial correlation—an extension of biserial correlation (correlation of a dichotomous variable such as "male-female" with a measurable variable) to be used when a trichotomous variable (such as low, medium, and high levels of riot severity) is correlated with a continuous interval-level variable (such as unemployment rates). (see: *CORRELATION*)

true values—mathematical values such as coefficients or predictions, based upon all possible observations (the whole population) rather than on a sample set of observations. (see: *population, sample*)

truth table—used in symbolic logic, a list of the truth values of certain logical statements. For example, if A represents any statement and "not A" represents its opposite, then one or the other, but not both, may be true. The table below shows another possibility: if statement A is true and statement B is false, then anything that is included in both A and B (read "A intersection B") is false, and anything in A or B but not both (read "A union B") is true.

A	Not A ($\bar{A}$)	B	A and B ($A \cap B$)	A or B ($A \cup B$)
T	F	F	F	T

Truth tables are used in symbolic logic for analysis and also have applications in statistical inference and computer work. Truth in symbolic logic is not necessarily truth in fact, however, if the hypothesis, in fact, is not true. (see: *Boolean algebra*)

Tschuprow's T—a measure of association which is used to show the degree of dependence or independence between two or more variables when these are arranged in table form. The T value varies from 0, indicating no association, to 1 indicating complete association. (see: *measure of association, chi-square*)

Type I error—(see: *alpha error*)

Type II error—(see: *beta error*)

uncertainty—generally synonomous with risk. Sometimes risk refers to situations in which the probabilities of relevant events are known, and uncertainty to situations in which they are not.

universe—in survey research, a population—that is, all members of any class of things, including people, that might be studied. A sample, by contrast, is a smaller group of things taken from the larger universe or population for more detailed scrutiny. (see: *SURVEY RESEARCH*)

up cross—the point where a time series, measured about its mean, changes in sign from negative to positive. Conversely, the point where it changes from positive to negative is called a down cross. (see: *TIME SERIES ANALYSIS*)

util—in the study of policy options, an artificial measure of utility.

utility—psychic satisfaction derived from consumption or gain. Utility is a subjective, personal value. It is not the same as "usefulness" in the functional sense—a painting, for example, may not be useful, but it does clearly afford satisfaction. The concept of utility in economics is key to the theory of consumer demand. In political analysis, utility is important to the study of policy options and decision-making. (see: *DECISION-MAKING ANALYSIS*)

validation—in hypothesis testing, a means of establishing the accuracy of the measuring instrument, or the validity of the test, in measuring the relevant variables. An instrument is considered valid if it measures the characteristics it is designed to measure. (see: *hypothesis testing, variable, validity*)

validity—the accuracy with which a measurement technique measures the variable an analyst is investigating in the way he wants it measured. Validity contrasts with reliability which is the extent to which a measurement technique produces the same results time after time. A public opinion poll that consistently finds 75% of its respondents favoring monarchy in the US might be very reliable, in that only monarchists were sampled, but would not be a valid measure of the actual state of opinion in the country.

value—a concept of what is desirable or good or, in measurement terminology, a difference in amount or kind of a variable. The variable I.Q., for example, may have values ranging from 0 to 200.

variable—any characteristic or an object or class of objects that may vary in quantity or quality for different members of the class or for the same individual at different points in time. In political science, variables do not usually refer to changes in the behavior of the same individual but to differences between individuals in the same class. Legislators, for example, may be compared with one another on the basis of such variables as party affiliation, age, geographic region, or education.

The event or condition the analyst wishes to explain or predict is called the *dependent variable.* The factors or conditions which may have explanatory or predictive power are the *independent variables.* The relationship between dependent and independent variables is not necessarily causal, although causality is often inferred.

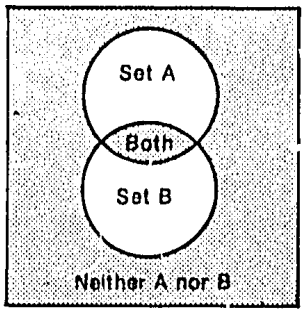
variable sum game—a class of problems dealt with by game theory which includes elements of both competition and cooperation. The sums of the payoffs to the two players in the different matrix cells are not equal, meaning that some combinations of payoffs are better for each player than are other payoff combinations. Unlike a fixed sum game, in which one player always loses as much as the other gains, a variable sum game also includes regions of cooperation. Some, but not all, interests are held in common by the players in a variable sum game. For example, in the game below, both players seek to avoid the outcome resulting from the combination of strategies II and 3 (both lose heavily in this situation). Beyond that, however, there is competition—A's best outcome (the intersection of strategies II and 1) is not good for B, and B's best outcome (strategies II and 2) is not good for A. (see: *fixed-sum game, zero-sum game*)

		Player A's Strategies I	II	III
Player B's Strategies	1	37 -2	40 6	21 25
	2	-1 12	-8 100	9 21
	3	35 2	-25 -40	16 23

variance—a measure of the dispersion of a set of numbers around their central value. Denoted s^2, variance is the sum of the squares of the differences between each value and the average of all the values, all divided by the total number of values. In the sample, 5, 8, 11, 16, 20, for example, the mean is 12 and the variance is $(12-5)^2+(12-8)^2+(12-11)^2+(12-16)^2+(12-20)^2=146/5=29.2=s^2$. (see: *standard deviation*)

vector—a quantity which has magnitude and direction and is most often represented by an arrow on a sample space or scattergram. The head of the arrow indicates the direction, and the length of the arrow represents the magnitude of the quantity.

Venn diagram—a drawing, used in Boolean algebra, to represent sets, their elements, and operations concerning them. For example, in the universe represented by the box below, there are some elements that fall in set A, some in set B, some in both, and some in neither.



verification—(see: *hypothesis testing*)

weighting—a ranking of observations according to their importance or some other criteria, before computation is performed. In multiple regression, for example, we may wish to compensate for possible errors in the data by weighting. If an observation has error E, then weight $W=1/E$ is used so that the larger the error, the less weight the observation will have in the total regression equation.

Yate's correction—an adjustment made in the computation of a chi square when the data base is small. If one or more of the expected cell frequencies in a 2 x 2 chi square matrix is less than 5, the normal chi square computation will overstate the true relationship. Yate's correction is designed to deflate this. (see: *chi square*)

Yule's Q—a quick statistical test which tells the approximate degree and direction of the association between two variables. It is a relatively unsophisticated test, confined to 2 x 2 contingency tables (arrays which display two dichotomous variables and the relationships between them). It ranges from -1 to $+1$, with 0 indicating no significant associations. For

example, the calculation of Yule's Q for sets of votes show political affiliation to be strongly associated with a yes or no vote on issues 1 and 3. On issue 2, however, there is no such marked relationship.

ISSUE I

	Lib	Cons
Yes	25	6
No	14	35

Q= .82

ISSUE II

	Lib	Cons
Yes	21	24
No	18	16

Q=-.13

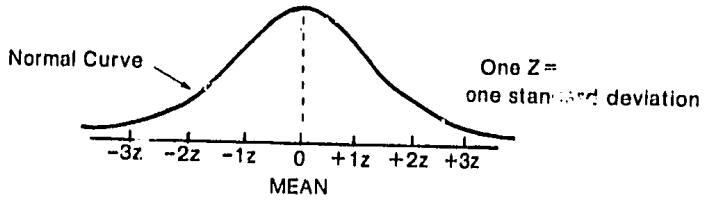
ISSUE III

	Lib	Cons
Yes	2	28
No	21	14

Q=-.91

In Guttman scaling, items having a Yule's Q value of .80 or higher with each other are scalable. (see: *measure of association, GUTTMAN SCALING*)

z score--a standardized measurement unit equal to one standard deviation of a normal distribution, which is used to compare the results of entirely different tests or measures. The z score is obtained by subtracting the original raw score from its mean value, and dividing that result by the standard deviation. Riot frequencies can be compared by unemployment levels, for example, by converting both to z scores. (see: *normal distribution, standard deviation*)



z test—a statistical test of the significance of data (or the probability that the same data would occur solely by chance) found by changing the data into standardized z scores and consulting a table of the probabilities of various z values occurring by chance. If an analyst was checking the hypothesis that the average age of world leaders was 55, for example, he might sample 20 leaders and find the average of the sample to be 61. But is 61 significantly different from 55 or could the difference be due to chance in selecting the sample? Changing both averages into z scores and comparing them on the table, he finds there is a 2.5% probability of getting a sample mean age of 61 or higher when the real mean age is 55. This chance is