# Chapter 1 Introduction



# 1.1 The Domain of Social Sciences

Social sciences correspond to a vast and rapidly growing area that encompasses investigations into diverse phenomena happening in the society, the economy, and the environment. In fact, social sciences deal with people—individuals, groups, or firms. As Bhattacharjee (2012) puts it, social sciences taken as a single branch of knowledge define the science of people or collection of people such as cultural groups, trading firms, learned societies, or market economies and their individual or collective behavior. That way social science embraces psychology (the science of human behavior), sociology (the science of social groups), political science (dealing with political groups), and economics (the science of firms, markets, and economies).

Some of the phenomena studied in social sciences are too complex to admit concrete statements; on some we cannot have direct observations or measurements; some are culture (or region) specific while others are generic and common. Data including laboratory measurements, survey observations, responses to questions, documents, artifacts, mission and vision statements and similar entities available in social sciences for scientific investigations into the 'behavior' phenomenon are so vague, uncertain, and error-prone that methods of investigation and techniques applied in physical sciences cannot be immediately used without necessary modifications. In fact, disagreements among observers or investigators on the same features of the same individuals are quite considerable, and it becomes difficult to generalize findings or conclusions based on a single set of data.

Measurements play an important role in any scientific investigation, to the extent that the quality and adequacy of pertinent measurements do affect the credibility of findings from the investigation. Measurement in the social sciences may be conceived as a process linking abstract concepts to empirical indicators. It transforms concepts into accounting indicators or schemes. The following phases in this transformation can be clearly identified.

1. The abstract definition of the phenomenon or concept that is to be studied.

<sup>©</sup> Springer Nature Singapore Pte Ltd. 2018

S. P. Mukherjee et al., *Statistical Methods in Social Science Research*, https://doi.org/10.1007/978-981-13-2146-7\_1

2. The breakdown of the original concept into 'constituent concepts' or 'dimensions.' The original concept corresponds, more often than not, to a complex set of phenomena rather than to a single directly observable phenomenon.

3. An indicator is assigned to each dimension.

4. Usually, an aggregate indicator is developed, unless characteristics of the phenomenon do not justify the construction of some synthetic indicator. In other cases, the aggregate indicator entails construction of an accounting scheme, as for instance a social accounting matrix or accounts of employment or of health.

All this implies that measurement in the context of social phenomena involves aspects of both a theoretical and an empirical character. Data are needed to construct and validate theories, at the same time theories are needed to generate and validate data.

The breakdown of a phenomenon into measurable dimensions is rarely unique, in terms of the number of dimensions—preferably non-overlapping or un-correlated and their identification in terms of data-based indicators. The problem becomes more complicated when the phenomenon is dynamic, and we can develop a reasonable breakdown at any point of time which may not be a reasonable representation of the phenomenon at a subsequent time point. In some cases, the dimensions are not really amenable to a direct enumeration or even identification. For example, when we have to deal with feelings, aptitudes, and perceptions, we construct scales by assuming certain continua and by noting the responses to some questions believe to reveal the chosen dimension.

# 1.2 Problems in Social Science Research

While scientific studies are invariably concerned with 'variations' in some features or characteristics across individuals and groups, over time and over space, in the context of social sciences many of these features which vary randomly are only 'latent' variables, unlike 'manifest' variables studied in physical or biological phenomena.

Let us consider a typical theme for research, viz. greater frustration among highly educated young persons about the prevailing employment situation than among people with lower levels of education and/or with lesser ambitions in life. To examine the applicability or validity of this proposition in a particular society or region or some suitably defined group, we need evidences bearing on entities like 'ambition,' 'levels of education,' 'frustration,' and 'perceived employment situation' in respect of some individuals in a 'sample' that adequately represents the group or population in relation to which the validity of the proposition was to be examined. And the first and the third features defy unique and objective definitions and, subsequently, measures. Evidently, any form of analysis based on some evidences collected on such latent variables will attract a lot of uncertainty. However, we cannot take our hands off and have to try out some reasonable surrogates or substitutes which are manifest and can be quantified. Of course, the choice of surrogates for 'ambitions' and 'frustration' is not unique, and the responses that are likely to arise to some questions carefully constructed to reflect on these latent variables cannot be scaled uniquely and cannot be subsequently summarized uniquely. We have to keep in mind this non-uniqueness associated with evidences that are most often necessitated in social science research.

In psychology, we talk of psychophysical experiments essentially dealing with responses to various stimuli. In education, we sometimes conduct an experiment to find out which of several alternative ways of teaching a new language is the most effective. In political science, we can think of an experiment to conduct an election in several alternative ways to identify the most preferred alternative. And rarely will experts or referees or judges will agree on the most effective or most preferred or most likely alternative. Such differences in assessment is just natural, and the confusion or inconsistency arising from such disagreement is unavoidable.

Dealing quite often with latent variables which are quantified in various equivocal terms and based on relatively small sample sizes, conclusions reached in many social science research studies are hardly 'reproducible' and hence are hardly 'scientific.' At the same time, we cannot drop all such latent variables or variables which defy unique quantification from our investigations and we deal with multiple variables in any study that make it difficult to determine the sample size that will be adequate to provide credible inferences regarding the many parameters that have to be estimated or all the hypotheses to be tested, except in terms of a number (of units) that will be too resource-intensive to really canvas.

Several so-called international agencies which have recently mushroomed and which attempt to rank different countries in terms of 'abstract' entities like 'charitygiving' only serve to dish out unscientific findings that cannot carry any conviction, but can be used wrongly by some interest groups to portray some countries poorly or in a lofty manner.

The choice of indicators based most often on some proxies or surrogates of the feature or characteristic understudy is not unique, and there is hardly any criterion to accept on in preference to another. Sometimes, a wrongly chosen indicator has led to lack of credibility of the final result based on an index that combines the various indicators. Earlier, the United Nations Development Programme took 'mean years of schooling' as an indicator of educational attainment of a country, to be taken along with the percentage of literates among adults. One should note that mean years of schooling for an individual as also for a group may increase as a consequence of stagnation and, that way, may be a negative indicator of educational attainment.

Evidences bearing on different social or cultural phenomena are mostly gathered through sample surveys, and an important decision to be taken in this regard is the choice of an appropriate sampling design to come up with an adequate sample size that can ensure credible estimates of the different parameters of interest and tests of different hypotheses with reasonable power. It is not uncommon to find a small sample used to come up with a general statement that can hardly beget any credibility.

The choice and use of an appropriate sampling design to suit the purpose of a sample survey throwing up adequate evidences of reasonable quality to make valid inferences is a bad necessity. And the inferences are to be valid in respect of a certain 'population' in which the investigator is interested and from which the sample has to be drawn. Thus, delineating the population of interest is a primary task, and in

social science investigations there could arise situations where this task is quite complicated. For example, if a national sample survey is to be conducted for getting a good estimate of the number of persons suffering from a certain disease which attracts some taboo, the problem of delineating the population of interest—which should not be the general population—poses serious problems.

Another big issue concerns the size and selection of the sample used in surveys to collect data on both measurable features of individual respondents as also on traits possessed by them that evade direct measurements. The sample must be large enough to make the findings reproducible, and the data must be collected with due care to secure proper evidences that can throw light on the underlying phenomenon or phenomena. Findings of many investigations fail to become reproducible because of shortcomings in such surveys.

### **1.3 Role of Statistics**

Statistics, being a scientific method—as distinct from a 'science' related to one type of phenomena—is called for to make inductive inferences regarding various phenomena like social tension, frustration among educated youths, exploitation and consequent feeling of alienation among neglected tribals, erosion of patriotic feelings among the young these days, religious fanaticism leading to tensions in the society, loyalty of middle-income customers to some brands of a consumer good, loss of credibility of democratic institutions over time, etc., based on evidences gathered.

In the context of a growing public demand for more credible and insightful view of distributive justice, and better and more comprehensive analysis of long-term and wide-area effects and outcomes of social expenditure by different agents, contemporary research has to come up with reasonable and defensible answers to such questions as: How does education affect employment? Does business development have an impact on crimes? To what extent are family formations and decisions are affected by economic prospects and employment security? What are the implications of a forward-looking prevention policy in health, long-term care, and the elderly?

It is true that social scientists are aware of the fact that answers to such questions are bound to be somewhat specific about time, space, culture, and other considerations. However, howsoever the group of interest may be defined, it will be surely larger—and, in some cases, much larger—than the 'sample' that can be conveniently canvassed in any research investigation. Thus, the need for inductive inferences based on evidences and some models is strongly felt.

Inductive inferences are made or have to be made in several distinct situations, viz.

(1) we have limited evidences available on a phenomenon, and we like to go from this sample of evidences to make a conclusion about the phenomenon itself (that really corresponds to an infinite population of evidences that can arise, at least in theory).

#### 1.3 Role of Statistics

(2) we observe the currently available in a damaged or an altered set of evidences pertaining to a phenomenon that occurred in the past and we like to infer about some aspect(s) of that past phenomenon.

(3) we have observations relating to a phenomenon revealed in the recent past or currently and we like to infer about how it unfolds in the future.

We must bear in mind the fact that in induction - unlike in deduction - premises provide some support to the conclusion or inference made on the basis of evidences available along with some 'model' for processing the evidences. In Deduction, the conclusion is warranted by the premises. This implies that with any inductive inference is assumed associated some amount of uncertainty, due both to uncertainties in the evidences made use of as also the uncertainty inherent in the use of statistical tools for processing the evidences.

This inferential uncertainty has to be quantified if alternative ways for processing of evidences or even if different sets of evidences bearing on the same phenomenon are to considered. And the concept of probability is brought in to quantify uncertainty involved in a given exercise in inductive inference. Evidential uncertainty is also handled in terms of fuzziness and related measures.

While statistical methods and techniques deal essentially with 'variations' in some features or characteristics across individuals and groups, over time and over space, to bring out a pattern behind such variations which can be taken further to offer an explanation of the observed variation, in the context of social sciences many of these features which vary randomly are only 'latent' variables, unlike 'manifest' variables studied in physical or biological phenomena and even those which are 'manifest' may be mostly 'categorical' or even 'nominal' to which standard statistical techniques cannot directly apply without some necessary modification. More often than not, social phenomena reveal interrelations among constructs or variables bearing on them which cannot be studied in terms of usual dependence analysis. Variables involved can be classified as endogenous and exogenous, after delineating the boundaries of the system in which the study is embedded, while the classification as dependent and independent is not pertinent.

Statistics—meaning both statistical data as also statistical reasoning—are becoming active partners in the world of social science research, promoting and supporting, using and questioning ongoing theoretical studies. Statistics not only provides valuable empirical evidence against which theoretical constructs can be tested, but also theoretical frameworks putting them to the test of the measurement process. Theories, in fact, are the main ingredients for developing the conceptual frameworks underlying the quantification of social phenomena. Their viability and effectiveness to cope with the dynamism and comprehensiveness of social change represents a crucial test of their validity. Theories are validated by empirical data and, therefore, the quality of data made use of in this context is a vital issue. Only close collaboration between social scientists and statisticians can bring about improvements in social statistics and, that way, in social science researches.

As is the case with researches in other domains, social science research generally if not necessarily—involves collection, aggregation, and analysis of multiple characteristics or features exhibited by the individuals or units in the group under investigation. In fact, factor analysis as an important technique for analyzing multivariate data was introduced in the context of psychological investigations to identify factors or traits which explain observed correlations between different pairs of subjects in which individuals have been tested. Methods of clustering and classification also had their initial applications in social investigations to identify homogeneous groups based on the different features of the individuals. Multi-dimensional scaling as an important tool for data visualization cropped up in connection with linguistic ability studies and related aspects of human behavior. Several techniques like conjoint analysis were developed during researches on consumer behavior.

In recent times, we quite often access data on multiple attributes based on which we like to compare several entities like different locations or institutions or societies or strategies or deployment plan, etc., and assign ranks to these entities so that we can identify and concentrate on the 'best' or the 'worst' situation needing 'urgent' or 'convenient' intervention. In fact, such multi-attribute decision problems are of great interest and importance in social science research. Indeed, before we can pool data on the same phenomena from different sources—and such data could be purely qualitative in character like opinions or judgements or ranks etc—we should examine the extent to which they agree among themselves. Similarly, meta-analysis or analysis of analyses carried out on the same phenomenon by different researchers possibly following different models and methods is required to ensure consolidation of analyses to enhance the substantial content of any social research study.

We also get data on social interactions among individuals or on decisions of individuals and groups to move from one place or one profession or one job to another. Such data can reveal important latent features about the individuals as also about groups on proper analysis by techniques which have emerged over the years.

### **1.4 Preview of this Book**

This book is not intended to be a standard textbook on the subject of statistics for social science research covering all types of phenomena studied in social sciences and the whole gamut of statistical techniques that are being used or are required to be used in that context. It is just a supplementary reading covering only some selected techniques which are widely applied and often warranted in some areas of social science research. In fact, some researches in social sciences have led to the development and use of some of the methods and techniques discussed in this book. Content Analysis is one such example. Several techniques in multivariate data analysis owe their origins to psychology, e.g., factor analysis. The same is true about scaling techniques originally used in the context of psychological tests. While Management Science may not be regarded as a component of social sciences, research in marketing has to deal with human behavior like preferences for certain brands or grades of a certain product when it comes to a purchase decision. And there should be no reservation to accept such studies as research in social science.

Thus, product scaling and multi-dimensional scaling are statistical techniques which are found useful in marketing research.

Techniques dealt with in this book range from those which relate to problems of data credibility in studies in which confidentiality is a major concern and responses are likely to be untrue to techniques involved in pooling data from different sources, from simple scoring of responses to items in a questionnaire used in an opinion survey to analysis of multivariate data. Some of these techniques are of relatively recent origin, while several others have found their way in social sciences as well as in other areas of research quite some time back.

Statistics—meaning both statistical data as also statistical reasoning—are becoming active partners in the world of social science research, promoting and supporting, using and questioning ongoing theoretical studies. Statistics not only provides valuable empirical evidence against which theoretical constructs can be tested, but also throws up theoretical frameworks putting them to the test of the measurement process. Social science research is primarily empirical in character and inferences made about a whole lot of social phenomena are inductive in nature, being based on data which are quite often subjective. Such data-based inferences, taking for granted some postulates and some model behavior of the data, do naturally use relevant statistical techniques and corresponding softwares.

Sometimes a distinction is made between qualitative and quantitative research. It is difficult to illustrate purely qualitative research, except to indicate that qualitative and logical thinking to draw conclusions from the data in hand coupled with qualitative interpretation of such conclusions in the context of the phenomenon or group of related phenomena also constitute useful research in social science. There has been a growing tendency among researchers these days to quantify many constructs and features (variables) that defy direct or unequivocal quantification. While it is true that statistical methods and techniques are involved in a quantitative analysis, it must be remembered that such methods and techniques should enhance the substantive content of research and not just the technical content. The latter objective may call for application of latest available statistical techniques, while the former focuses on a pragmatic and, may be, limited use of such techniques only to derive strength from whatever constitute the premises for making inferences about the phenomena under investigation.

Right from planning a data-gathering exercise, through making the data collected and documented amenable to quantitative analysis and carrying out necessary test for 'poolability' of data gathered from different sources, getting appropriate analysis done on the data as eventually accepted, to reaching evidence-based inferences and interpreting results in the context of the research project, we need to-day Statistics in every stage—imaginatively and effectively—to enhance not merely the technical content of the study but also its substantive content. Data visualization as being somewhat similar to and still distinct from dimension reduction is quite useful in exploratory research on certain types of social phenomena like disagreements among judges in assessing relative positions of certain objects or subjects in terms of some relevant attributes. There arise problems in getting correct or truthful responses to questions pertaining to sensitive or confidential items like consumption of drugs or income from dubious sources or unusual behavior, etc., and in such cases, randomised response techniques [RRTs] are used in some studies to extract reasonable estimates of some parameters of interest without asking direct questions on the underlying issue. RRT has been treated in Chap. 2 for both qualitative and quantitative data.

Before embarking on any quantitative analysis, qualitative analysis often helps to answer certain questions relating to disputed authorship of some piece of literature or to the trend in public opinion about some contemporary issue like limits to freedom of speech and the like. In such cases, data are scattered in some reports or recorded speeches or other artefacts like photographs or banners. We can think of a content analysis (taken up in Chap. 3) to come up with some sensible answers to some vexing questions that can evade a sophisticated approach to secure a 'correct' answer.

In many studies on opinions or skills or competencies and similar other attributes, we use tests or instruments to develop some measures in terms of scores assigned to responses to different items. And these scores in different subjects, in different environments and in different times, may not be comparable and we need to scale them properly before we can make use of the scores for any decision or action. In many socioeconomic enquiries, for example, an organizational climate survey to bring out employees' perceptions about the climate for work prevailing within the organization, we often try to seek responses from individual employees on a statement like 'I get full cooperation from my peers and colleagues in discharging my responsibilities.' Each respondent is to tick one of five possible categories to indicate his/her perception about this issue, viz. strongly disagree, disagree, undecided or indifferent, agree, and strongly agree. The number of response categories could be seven or nine or some other odd integer. Different scaling techniques have been discussed in Chap. 4.

Recent times see a wide variety of data streaming in from different sources to throw light on the same phenomenon may be dispersed over different locations or institutions or groups. In respect of each such location/institution/group, the data may not be all equally revealing about the nature and magnitude of the phenomenon under study. We are required to rank these different entities to identify the 'best' and the 'worst' situations, so that we can prioritize our interventions in them accordingly. There are a few techniques for this multi-attribute decision-making problem, and we focus on a widely used technique, viz. TOPSIS where the concept of an 'ideal' situation and distances of different situations from the 'ideal' are the components. The use of this technique for data integration has been explained with an illustration from environmental pollution data in Chap. 5.

Chapter 6 is devoted to an emerging topic of judging quantitatively agreement among different raters or experts or judges in situations like diagnosis by several medical men of a disease some patient is suffering from, or reliability of a test battery in a psychometric test as judged by a group of subject experts, or relative importance of a particular feature of a product or service in assessing the latter's quality in the eyes of several potential customers, or opinions expressed by several political analysts on the likely impact of an agreement signed between two countries on international trade, and the like. Any attempt to pool the assessments or ratings or diagnoses has to be preceded by a statistical assessment of agreement among the opinions or judgements. This is not the same as application of Delphi or similar techniques to make the ratings or opinions or judgements converge. All this has been explained with ample illustrations in Chap. 6.

Meta-analysis is another recent paradigm in social science research. Here the idea is to make use of all the available evidence which may be in the form of several pieces of information (derived from some data) from different sources, some of which may be in the form of expert opinions. Evaluation of each piece of information enables us to determine the weight to be attached to it in pooling information. However, pooling information demands that different pieces of information are not conflicting with each other. Finally, we have to choose an appropriate method to combine the different pieces of information and express the reliability of the final conclusion. This is the content of Chap.7.

Coming to data analysis, it must be admitted that such data are necessarily multivariate and, more often than not, the data set covers a large number of units or individuals which differ among themselves to different extents in respect of several observable or measurable features which are correlated to different extents one with the others. It will be desirable to group such units or individuals into homogeneous clusters before we analyze relations among the variable features within each cluster separately. And, we should even start looking at the variables themselves before we subject them to further analysis. Toward that reduction in dimensions, we may profitably use factor analysis as also principal component analysis, and we identify and extract artificial combinations or components that can be carefully interpreted in terms of the research objectives.

Whenever a research encompasses more than one periods of time or, say, generations of individuals, we may be interested in noting the changes or transitions of the individuals across social groups or occupations. Such mobility studies are also quite useful in market research to reveal customer loyalty to certain product or service brands using a 'mover–stayer model.' In such mobility studies, Markov Chains and related tools play an important role. In fact, Renewal-Reward Process models have been used in studies on occupational mobility. Chapter 12 is devoted to the subject of social and occupational mobility along with some related issues in manpower planning.

Coming to other aspects of data analysis, it must be admitted that such data are necessarily multivariate and, more often than not, the data set covers a large number of units or individuals which differ among themselves to different extents in respect of several observable or measurable features which are correlated to different extents one with the others. It will be desirable to group such units or individuals into homogeneous clusters before we analyze relations among the variable features within each cluster separately. And, we should even start looking at the variables themselves before we subject them to further analysis. Toward that reduction in dimensions, we may profitably use factor analysis as also principal component analysis to identify and extract artificial combinations or components that can be carefully interpreted in terms of the research objectives. In Chap. 8, we deal with clustering techniques and

Discriminant Analysis while in Chap. 9, we discuss principal component analysis, and in Chap. 10, we take up study of factor analysis.

Under multivariate analysis, two very important techniques are clustering and classification. Under the problem of clustering, we try to find out the unknown number of homogeneous inherent groups in a data set as well as the structure of the groups. But under classification, the basic problem is discrimination of objects into some known groups. One of the most basic abilities of living creatures involves the grouping of similar objects to produce a classification. Classification is fundamental to most branches of science. The information on which the derived classification is based is generally a set of variable values recorded for each object or individual in the investigation, and clusters are constructed so that individuals within clusters are similar with respect to their variable values and different from individuals in other clusters. The second set of statistical techniques concerned with grouping is known as discriminant or assignment methods. Here the classification scheme is known a priori and the problem is how to devise rules for allocating unclassified individuals to one or other of the known classes.

Principal component analysis (PCA) is a dimension reduction procedure. The method is useful when we have obtained data on a number of variables (possibly a large number of variables), and believe that there is some redundancy in those variables. In this case, redundancy means that some of the variables are highly correlated with one another, possibly because they are measuring the same phenomenon. Because of this redundancy, it should be possible to reduce the observed variables into a smaller number of principal components (artificial variables) that will account for most of the variance in the observed variables.

Factor analysis presented in Chap. 10 is a statistical method used to study the dimensionality of a set of variables. In factor analysis, latent variables represent unobserved constructs and are referred to as factors. Factor analysis attempts to identify underlying variables, or factors, that explain the pattern of correlations within a set of observed variables. It is often used in data reduction to identify a small number of factors that explain most of the variance that is observed in a much larger number of manifest variables. Its basic difference from principal component analysis (PCA) is that in PCA variables are replaced by a small number of linear combinations which are expected to explain a larger part of the variation, but it is usually not possible to correlate these linear combinations with some physical phenomena. But in factor analysis, the newly derived latent variables are extracted as factors representing some physical phenomena. Given a set of scores for a group of persons corresponding to aptitude tests in subjects like mathematics, physics, statistics and their performances in 100-m race, long jump, high jump, etc., one may extract two latent factors, viz. intellectual ability and physical ability.

There are situations where we like to compare entities like music, or dance or an object of art or just any product available in many variants or brands and we need to scale these entities (generally called products) to get an idea about the relative merits of the different entities or relative distances between them on a straight line or a two- or three-dimensional surface. We have one-dimensional scaling provided by Thurstone's Law of Comparative Judgment, further taken up by Mosteller and

others. To get a better visualization of the relative distances or dissimilarities among the entities, multi-dimensional scaling was introduced by Torgersen. In Chap. 11, we discuss about this aspect of data analysis.

Whenever a research encompasses more than one periods of time or, say, generations of individuals, we may be interested in noting the changes or transitions of the individuals across social groups or occupations. Such mobility studies are also quite useful in market research to reveal customer loyalty to certain product or service brands using a 'mover–stayer model.' In such mobility studies, Markov Chains and related tools play an important role. In fact, Renewal-Reward Process models have been used in studies on occupational mobility. Chapter 12 is devoted to the subject of social and occupational mobility along with some related issues in manpower planning.

Social network refers to the articulation of a social relationship, ascribed or achieved, among individuals, families, households, villages, communities, regions, etc. The study of social networks is a fast widening multidisciplinary area involving social, mathematical, statistical, and computer sciences. It has its own parameters and methodological issues and tools. Social network analysis (abbreviated SNA) means an analysis of various characteristic of the pattern of distribution of relational ties in a social group and drawing inferences about the network as a whole or about those belonging to it considered individually or in groups. Bandyopadhyay et al. (2009) have discussed in detail how graph-theoretical and statistical techniques can be used to study some important parameters of global social networks and illustrate their uses in social science studies with some examples derived from real-life surveys. In Chap. 13, we consider a few features or characteristics of a social network and explain how these features can be measured. Then we discuss the possibility of using sampling techniques in case of large networks.

#### **References and Suggested Readings**

- Bandyopadhyay, A., Rao, A. R., & Sinha, B. K. (2009). *Statistical methods for social networks with applications*. California, USA: Sage Publication.
- Bhattacharjee, A. (2012). Social science research: principles, methods and practices. Creative commons attribution - non-commercial-share alike.
- Garrona, P., & Triacca, U. (1999). Social change: Measurement and theory. *International Statistical Review*, 67(1), 49–62. Research Eds. H. M. Blalock, & A. B. Blalock (pp. 5–27). New York: McGraw Hill.
- Lazarsfeld, P. F. (1965). Le vocabulaire des siences sociales. In R. Boudon, & P. F. Lazarsfeld (Eds.), *Methodes de la sociologie* (Vol. 1). Paris-La Hayne: Mouton.