Journal of Official Statistics
Vol. 1, 1985, pp. 35–47

Chance, Statistics, Sampling[1]

Leslie Kish[2]

1. Preface

Statistics and statisticians deal with the effects of chance events on empirical data. Chance is present and relevant in all that we know, in all that we do, and in all conditions that affect us. Chance intrudes into fields as diverse as sports and games, into accidents and insurance, into physics and genetics, and into the social sciences. When dealing with empirical data from any field, the separation of chance effects from causal relations becomes the function of statistics. The best known tools for that separation are the many "tests of significance." Thus, statistics has a special function that penetrates all other fields; and its ubiquitous yet specialized nature makes it an unusual discipline. Two of the newest branches of statistics – experimental design and sample design – illustrate these interrelated, but specialized, functions. Both experimental and sample design evolved from the need to reduce, to control, and to measure more efficiently the effects of chance factors in scientific research and in other investigations that rely on empirical data.

Statistical science and its various branches arrived late in history and in academia, and they are products of the maturity of our intellectual development. The proper recognition and treatment of chance effects developed slowly, mostly in the twentieth century, because our natural prejudice does not easily accept the presence of chance everywhere in our lives, and its varied and complex relations to cause, to necessity, and to free will.

2. Synthesis of Chance and Necessity

All natural processes and all human activities display – to the practiced eye – diverse combinations of chance with the critical factors of skill, determinism, necessity. Sports and games like baseball, football, and bridge are intuitively designed to resemble our more serious efforts and to prepare us for them.

The structure and language of sports abound in creative ways for combining chance with skill. In American football the top team must emerge in short seasons of 10 to 15 games; hence stronger teams carry heavy odds even in single games, and we speak of an "upset" when the "underdog" wins. I savor the sophistication of the headline one Friday in our Michigan Daily: "Upset Unlikely." But in baseball the words "underdog" and "upset" are not used for single games, and the better teams emerge leisurely in a long season of 150 games. Sophisticated structures of skill and chance enable the stronger teams to prevail more often while still leaving the hope and

[1] The Henry Russel Lecture of 1981 at the University of Michigan. This article first appeared in Statistica, anno XLII n. 1, 1982, pp. 3–20. It is reprinted with the kind permission of Professor Italo Scardovi, Editor, Statistica. Only minor technical changes have been made in accordance with the style guide for the Journal of Official Statistics.

[2] Professor Emeritus, Institute for Social Research, University of Michigan, USA.

excitement of a "sporting chance" for the weaker teams. The end-of-season playoffs among the top teams provide excellent examples of this principle. The winner is the champion for the record, but we know, or suspect (and even hope) that the best team does not prevail always and necessarily. Baseball's World Series is won in seven games by the better teams in only 80 percent of the contests (Mosteller (1952))[3]. Also in tennis, it was in this spirit that Björn Borg said of winning Wimbledon in 1977: "I think I am number 1 for the moment."

The extremes near pure chance and pure determinism are rare either in games or in nature, and mixtures are far more common and far more interesting. "Pure" chance games are approximated with coin tossing and dice throwing, and with roulette wheels and lotteries. For these games the probabilities are designed to be constant, known and independent for each trial. The materials and operations, that is, the coin and the toss, are supposed to be "honest"; and "fair" odds for each trial are supposed to balance inequalities in the probabilities, such as odds of 5 to 1 against an ace in dice. (The odds in games are not entirely "*fair*" when they provide a "take for the house," but these are supposed to conform to announced percentages.) Most games, especially the most interesting ones, do not depend entirely on either chance or skill, but are mixtures of both. Blackjack, poker and bridge involve such mixtures, as do betting on races and on football pools. Even chess games reflect chance elements although chess lies near the skill end of the scale. When in sports the unequal skills of antagonists threaten to exclude the influence of chance, chance is reinforced with various devices, such as odds and points in betting, or handicaps, or the stratification of teams into leagues and divisions with more equal skills. On the other hand, the large components of chance that exist in individual games are usually shifted toward the skill side by aggregating games into longer matches, tournaments, and seasons. The popular wisdom about sports and games could well be used to spread understanding of the role of chance in all aspects of nature and life.

Humankind has paradoxical attitudes toward the concept of chance. After designing games of pure chance and demanding "fair" trials, "true" dice, wheels, and lotteries, humans still try somehow to improve their chances and to impose some control over them. They may try dishonest tricks, or may try to alter their chances with hunches, dreams and other private signals; or with charms, amulets, magic, or prayer. Also we tend to reject chance as explanation and to insist that "it cannot be mere coincidence" even for events that are mostly just that. Thus many comments on sports and on the stock market serve sage explanations for mere random phenomena. (Weaver (1963, Ch. 13).)

Humankind has been trying for a long time to synthesize chance with necessity, to combine luck and skill. We have been learning gradually to do it better, and also in more and more fields. We still find the task difficult, partly because the combinations in each field appear different in character as well as in proportions. Also partly because the task of combining

[3] "If we compare pairs of teams, products, drugs, or persons on the basis of a fixed number of binomial trials, and identify the member of the pair that wins the majority of trials as the better, we may be in error. For example, if the members of a pair are evenly matched, a decision on the basis of performance is equivalent to coin-flipping. On the other hand, if one of a pair is actually better, it is more likely that the better member will also be the winner. If we carry out such comparisons on many pairs under roughly comparable conditions, it is of interest to estimate the over-all effectiveness of the decision technique in the kinds of situations that have occurred in practice. Data from the World Series are available to illustrate many facets of this type of problem... The probability that the better team wins the World Series is estimated as 0.80, ..." (Mosteller (1952)).

chance and cause must be relearned each time by minds inclined to choose either one or the other.

Primitive people's fascination with chance, fate and fortune was early and widespread. They were first filled with wonder, with awe of hidden powers, magic, miracles, and with *ad hoc* explanations. Only *after* discovering rules for connecting cause with effect and after formulating natural laws, did humankind learn to *counterpose* chance effects to explain exceptions to and deviations from those rules. Much later came our *combined* views of causal laws with chance effects. Kendall (1968) said: "At the end of the seventeenth century the philosophical studies of cause and chance... began to move close together," and were developed in sciences like astronomy and demography only in the eighteenth and nineteenth centuries. "This development had an important impact on the theory of chance itself. Previously chance was a nuisance, at least for those who wished to foresee and control the future. Man now began to use it for other purposes, ... and to bring it under control, to measure its effects, and to make due allowance for it."[4] Thus statistics, which must combine chance and cause, developed largely in our twentieth century, although it needed only the human mind and not powerful material tools. Bork (1967) notes that "books of random numbers" could only and typically be inventions of the twentieth century[5].

[4] "During the eighteenth and nineteenth centuries the realization grew continually stronger that aggregates of events may obey laws even when individuals do not. Uncertain as is the duration of any particular human life, the solvency of a life insurance company is guaranteed; uncertain as may be the sex of an unborn child, the approximate equality of numbers of the two sexes is one of the most certain things in the world."

"This development had an important impact on the theory of chance itself. Previously chance was a nuisance, at least to those who wished to foresee and control the future. Man now began to use it for other purposes, or if not to use it, to bring it under control, to measure its effect, and to make due allowance for it. For example, errors of observation were found to follow a definite law, and it became possible to state limits of error in measurement in precise probabilistic terms. In the twentieth century we have seen similar ideas worked out to a high degree of precision: in the theory of sampling, where we are content to scrutinize only a subset of a population, relying on the laws of chance to give us a reasonably representative subset; or in the theory of experimental design, in which unwanted influences are distributed at random in such a way that chance destroys (or reduces to minimal risk) the possibility that they may distort the interpretation of the experiment. Man cannot remove chance effects, but he has learned to control them." (Kendall (1968)).

"The conscious *notion* of chance appears to be a late comer among the basic conceptual tools by which man gains mastery over his world of action and perception. Sigerist states categorically that primitive men always deny chance or randomness in the occurrence of disease. Illness is always precipitated either by natural causes or by the sick or somebody else's action."

"The idea of Fortune as a category of events appeared in Aristotle's works. But, beginning with him, the concept of Fortune has always been very close to its opposite pole, Fate. This ambivalence reappears in people's everyday philosophies relative to a large number of social situations; what seems to be determined by chance is at the same time decided in advance by fate." (Aubert (1959)).

[5] "I want to use this book as a beginning theme for this paper. The production of such a book is entirely of the twentieth century. It could not have been produced in any other era. I do not mean to stress that the mechanism for doing it was not available, although that is also true. What is of more interest is that before the twentieth century no one would even have thought of the possibility of producing a book like this; no one would have seen any use for it. A rational nineteenth century man would have thought it the height of folly to produce a book containing only random numbers. Yet such a book is important, even though it is not any of the usual lists of one hundred great books.

"That this book is strictly of the twentieth century is in itself of importance. I claim that it indicates a cardinal feature of our century: randomness, a feature permeating many different and apparently unrelated aspects of our culture. I do not claim that randomness is the only feature which characterizes and separates twentieth century thought from earlier thought, or even that it is dominant, but I will argue, admittedly on a speculative basis, that it is an important aspect of the twentieth century." (Bork (1967)).

The three-stage development of the idea of chance in the minds of children was described in 1951 by Piaget and Inhelder (1975): At first "...the child does not distinguish the possible from the necessary... and there is no differentiation between what is deducible and what is not...". After seven years of age... "there is differentiation and hence an antithesis between chance and operations, ... chance defines ... the unpredictable." However, after eleven years, "...on the contrary, there is a synthesis between chance and operations..."[6]

I propose that the child's three stage discovery of combining cause and chance is similar to the history of humankind, and that this similarity is no mere coincidence but a psychological analogy to the controversial biological theory that "ontogeny recapitulates phylogeny." In both systems the similar developments are imposed by the constraints of development. As Gould (1977b) said: "Given a small and simple starting point (the phyletic amoeba or the ontogenetic zygote), nature can only build complexity in a limited number of ways. Piaget uses the same style of argument, however different the content. The parallels are real, but phylogeny does not cause ontogeny. Again, two independent sequences follow similar paths under the influence of a common constraint – the structure of the human mind itself."

[6] "During the first period there is no differentiation between what is deducible and what is not because the intuitive anticipation remains halfway between operation and chance itself. During the second period there is differentiation and hence an antithesis between chance and operations, these latter determining the domain of the deducible while chance defines, therefore, the domain of the incomposable and the irreversible, in a word, the unpredictable... In the course of stage III, on the contrary, there is a synthesis between chance and operations, the latter allowing the field of fortuitous distributions to be structured in a system of probabilities by a sort of analogous assimilation of the fortuitous with the operative." (Piaget and Inhelder (1975)).

Combining chance and necessity seems to be the product of mature minds. The primitive mind does not even distinguish the two. Later, on the second level, we are tempted either to find determining causes or to accept mere chance effects as alternative explanations for observed phenomena. Only much later do we acquire the idea and capacity for combining the two explanations. This ability comes late but with great effects not only for statistics, but also for science and for philosophy. Consider the emergence of Darwin's theory of evolution and of genetics from Mendel to De Vries, Morgan and Watson. Scardovi (1976) writes: "Mendel's was the first explicit indeterministic paradigm in the history of science, the herald of the theory of quantum mechanics which was to make change the law of the Universe and to characterize the behaviour of a physical particle as unpredictable, like the destiny of a Mendelian unit. Mendel's paradigm was thirty years ahead of the philosophy and science of his time. The complete lack of recognition accorded him by his contemporaries can thus be explained. Mendel also surprises us with his statistical insight: he was capable of seeing in his numerical data random oscillations about a limiting value, and was thus able to extract the ideal proportion toward which the outcomes in the combined heredity of several traits tend, perceiving the important law of independence...". And Gould thus describes Darwin's theory of evolution: "First, variation must be random, or at least not preferentially inclined toward adaptation. For, if variation comes prepackaged in the right direction, then selection plays no creative role, but merely eliminates the unlucky individuals who do not vary in the appropriate way. Lamarckism, with its insistence that animals respond creatively to their needs and pass acquired traits to offspring, is a non-Darwinian theory on this account. Our understanding of genetic mutation suggests that Darwin was right in maintaining that variation is not predirected in favorable ways. Evolution is a mixture of chance and necessity – chance at the level of variation, necessity in the working of selection. Secondly, variation must be small relative to the extent of evolutionary change in the foundation of new species. For if new species arise all at once, then selection only has to remove former occupants to make way for an improvement that it did not manufacture. Again, our understanding of genetics encourages Darwin's view that small mutations are the stuff of evolutionary change." (Gould (1977 a).)

Another explosive discovery of the fundamental role of chance occurred in quantum mechanics. "Statistical ideas had been used in the nineteenth century physics, but then it was always assumed that the basic laws were completely deterministic. Statistical calculations were made when one lacked complete information or because of the complexity of the system involved. In the statistical interpretation of quantum mechanics I have just described, however, randomness is not accepted purely for calculational purposes. It is a fundamental aspect of the basic physical laws themselves. Although some physicists have resisted this randomness in atomic physics, it is very commonly maintained. A famous principle in contemporary quantum mechanics, the 'uncertainty principle', is closely related to this statistical view of the laws governing atomic systems." (Bork (1967).)

The fundamental role of randomness in evolution and in quantum mechanics became generally understood only in our own century. These great discoveries, and others as well, have not only led to acceptance of the ideas and of the roles of randomness, but without the growth of these ideas those discoveries would also have been impossible. But the ideas of the joint roles of chance and skill face resistance even today. Monod (1971), writes: "Even today a good many distinguished minds seem unable to accept or even understand that from a source of noise natural selection alone and unaided could have drawn all the music of the biosphere. In effect natural selection operates upon the products of chance and can feed nowhere else; but it operates in a domain of very demanding conditions, and from this domain chance is barred. It is not to chance but to these conditions that evolution owes its generally progressive course, its successive conquests, and the impression it gives of a smooth and steady unfolding." Even Einstein, the wisest mind of our century, is quoted (Hoffman (1972)), as having said: "God does not play dice with the universe." Yet Einstein-Bose statistics define a basic game of chance for elementary particles.

Lewis Thomas, one of our favorite science writers, laments (1980): "Today, an intellectually fashionable view of man's place in nature is that... The universe is meaningless for human beings: we bumbled our way into the place by a series of random and senseless biological accidents... I cannot make my peace with the randomness doctrine; I cannot abide the notion of purposelessness and blind chance in nature. And yet I do not know what to put in its place for the quieting of my mind."

The interaction of cause and chance in the basic aspects of sciences in general and of statistics in particular have profound philosophical implications, as Einstein, Monod, and Thomas note. The implications are revolutionary, but they need not be tragic. Listen to Schumacher's hopeful voice (1973): "When the Lord created the world and people to live in it – an enterprise which, according to modern science, took a very long time – I could well imagine that He reasoned with himself as follows: 'If I make everything predictable, these human beings, whom I have endowed with pretty good brains, will undoubtedly learn to predict everything, and they will thereupon have no motive to do anything at all, because they will recognise that the future is totally determined and cannot be influenced by any human action.

On the other hand, if I make everything unpredictable, they will gradually discover that there is no rational basis for any decision whatsoever and, as in the first case, they will thereupon have no motive to do anything at all. Neither scheme would make sense. I must therefore create a mixture of the two. Let some things be predictable and let others be unpredictable. They will then, amongst many other things, have the very important task of finding out which is which'." I only disagree with the phrasing of the last two sentences, because things do not come cleanly sorted into predictable and unpredictable; rather each thing, in and of itself, must be seen and understood as a mixture of both.

The complementary roles of chance and necessity appear both fundamental and ubiquitous. That the discoveries of diverse syntheses of the two occur so late to mature minds have several consequences. First, I believe that this vital outlook should belong to the philosophical, humanistic, and moral heritage of all scientists, and further of all citizens. As H.G. Wells said: "Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write." Second, I also believe that statisticians have natural responsibilities for teaching this outlook.[7]

3. Statistics and statisticians

Because chance, randomness, and error constitute the very core of statistics, we statisticians must include chance effects in all our patterns, plans, designs, and inferences. In all our endeavors we cannot avoid the basic philosophical problem of empirical science: to make inferences to large populations and to infinite universes from limited samples of data that are all subject to random errors. We cannot avoid the basic philosophical problems posed by David Hume. The diverse schools of statistical inference represented by R.A. Fisher, J. Neyman, and L.J. Savage are valiant efforts to capture within a statistical framework Hume's problem of inference. Similarly, the works of philosophers like Popper (1959), Salmon (1967), and Burks (1977) have profound meaning for statisticians.

Furthermore, related profound questions are sometimes asked about whether the concepts and definitions of probability should encompass *all* sources of uncertainty and embrace *all* doubtful propositions. Consider the uncertainties involved in some unique and momentous events. For example: the influence of great historical leaders, of wars, of religions and of great revolutions; the

[7] "Thus the complementary roles of chance and necessity appear basic and ubiquitous. That the discoveries of the diverse syntheses occur late to mature minds have several consequences. First, I believe that this vital outlook should belong to the philosophical, humanistic, and moral heritage of all scientists, and further of all citizens; also I believe that statisticians have natural responsibilities for teaching it. Consider the broad implications of several recent discoveries about the role of chance: the genetics of individual health, intelligence, and behaviour; the statistics of population genetics (for which Fisher has been called the foremost biologist of our century); the occurrence of cancer; the lengths of our lives."

"Appreciating the pervasiveness of chance should help us cope better with diverse aspects of life. About success in affairs, in business, even in science, we may say: 'Fortune favors the prepared mind.' The winners of millions of dollars or of Nobel prizes tend to accent the prepared mind, the losers emphasize fortune. This tendency for biased judgment, between winners and losers, is also found in national surveys of attitudes (and these self-protective biases may be psychologically healthy)."

"The tragic biblical Job might have been happier and wiser if he knew that his plagues were due to chance. The triumphs or the problems of your children may be due to chance, not only to your behaviour – despite what Freud may say; a statistical view may protect parents against false pride or against guilt and despair. But we are not mere helpless puppets of chance, and we can improve our chances – for example, by quitting smoking, with regular exercise, and by losing weight. Recognition of the interplay of chance with discernible causes may yet lead us to a better way of life and moral philosophy. Somebody may even start a new religion of 'Statisticology'." (Kish (1978).)

development of our species and of other species, wild and domesticated; the birth of our planet and our solar system; and the theory of the Big Bang. Such events have also been considered by Poincaré (1952), as a... "contrast between a very trifling cause that is unappreciable to the observer, and considerable effects, that are sometimes terrible disasters."[8] Victor Weisskopf's word is "amplification" for imperceptible chance effects growing into a catasrophe, such as the thermonuclear holocaust. Here I also suspect philosophical links to the new "catastrophe theory" in applied mathematics (Poston and Stewart, (1978)). Also to new models of "punctuated equilibria" for evolutionary change (Gould (1980)).[9] In both of these theories the smooth curves favored by deterministic models are replaced by abrupt discontinuities that suggest stochastic changes. However, to be brief and modest here, let us avoid these deep philosophical considerations. I also wish to avoid controversies around the word probability by confining myself to a narrower concept of chance.

Statistics is a peculiar kind of enterprise of contradictory character, because it is at the same time so special and yet so general. Statistics exists only at the interface between chance and empirical data. But it exists at *every* such interface, which I propose to be both necessary and sufficient for an activity to be properly called statistics. In scientific research of every kind, in government, commerce, industry, and agriculture; in medicine, engineering, education, sports, and insurance; statistics has a special function wherever and whenever empirical data are treated. This widely spread yet specialized character differentiates statistics from other disciplines, which tend to cover in depth all aspects of their own special domains.

Let us agree here that statistics alone does not embrace all kinds of uncertainties, especially unique events. Statistics deals with chance effects on empirical data concerned with classes of events, typically with large classes. I want to cover what statisticians do in practice, in teaching, in writing – and also what we should be doing. Statistics is a diffuse enterprise that at one end has vital roots in abstract mathematical theories of probability and there it shades into profound areas within mathematics proper. At the other end, it branches into practical procedures for collecting and analyzing data of all kinds and there it gets embedded into the methods and interpretations of all substantive fields of measurement.

Statistics and statisticians must remain in touch with both of these ends, but they can neither completely encompass nor reside exclusively in either end. Without data, we would remain solely within the deductive arena of mathematics. Without concern for chance events, we could not be distinguished

[8] "Chance, then, must be something more than the name we give to our ignorance... it may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous phenomenon... contrast between a very trifling cause that is inappreciable to the observer, and considerable effects, that are sometimes terrible disasters." (Poincaré (1952)).

[9] "Eldredge and I refer to this scheme as the model of *punctuated equilibria.* Lineages change little during most of their history, but events of rapid speciation occasionally punctuate this tranquility. Evolution is the differential survival and deployment of these punctuations. (In describing the speciation of peripheral isolates as very rapid, I speak as a geologist. The process may take hundreds, even thousands of years; you might see nothing if you stared at speciating bees on a tree for your entire lifetime. But a thousand years is a tiny fraction of one percent of the average duration for most fossil invertebrate species – 5 to 10 million years. Geologists can rarely resolve so short an interval at all; we tend to treat it as a moment.)

If gradualism is more a product of Western thought than a fact of nature, then we should consider alternate philosophies of change to enlarge our realm of constraining prejudices." (Gould (1980).)

from all those substantive areas where humankind has learned to observe, count, and measure. For counting and measuring in fields like accounting, banking, voting, and scoring in sports, there exist specialists who work at refining instruments and operations, and at eliminating mistakes and cheating. Statisticians bring a different and special view to the study of errors, which involves an explicit recognition of chance effects, a probabilistic view and treatment for them, and the incorporation of that view into research design and into the interpretation of observations. To err is human, to forgive is divine – but to include errors in the design is the statistician's contribution. It is the recognition of the effects of chance, of variability, of fluctuations that distinguishes the work of statisticians from others in just about every field.

The data of other scientists come chiefly from within their own disciplines – although they may occasionally also take side trips into other fields. But statisticians get *all* their data from other fields, and from *all other fields,* wherever data are gathered. Because we have no field of our own, we cannot work without others, but they also cannot do without us – or not very well, and not for very long. Statistics is a symbiotic way of life, a marginal and hyphenated existence. Still statisticians lead useful, rewarding, and interesting lives.

Where does the statistician's work come in? In any scientific enterprise we can distinguish a hierarchy of four problems and decisions. First comes the choice of variables to consider, and second the design of models for the relations between the variables. These two decisions belong to the scientist or expert in the substantive field. Third comes the estimation of parameters, and fourth the assessment and analysis of the variability and errors of these estimates. For these two later stages statistics are essential, and a specialized statistician may also be needed. The experts, the scientists, may do the statistics for themselves without statisticians, and they may do it well enough – but often they do it badly.

Consider, for example, the simple model $d = (1/2)\ gt^2$ for the distance covered in t seconds by an object falling freely to the earth's surface. The choice of the variables d, g, and t, and the constants 1/2 and 2 come from the physicist's model; and the 2 is not subject to doubt or error – it cannot be 3 or even 2.01. But the statistician gets involved in designs for measuring d, t, and especially g, the gravity parameter. To measure the distribution of g may pose a formidable challenge for the research design, requiring both a background in relevant variables and foresight about future uses of the model. This can be an interesting problem in experimental and in sampling design.

The situation is similar in more complex models, whether in the physical, biological, or social sciences. The model $d = (1/2)\ gt^2$ cheats reality because it excludes air resistance and wind from the *freely* falling body. I used this formula from freshman physics decades ago to estimate the height of the George Washington Bridge from the timed fall of my hiking stick. I lost my bet by a factor of about two or three and suffered a humiliating and lasting lesson in deceitful models. The model also omits a term e for experimental errors in the measurements, and separating e from g can be challenging statistical work. Furthermore, air resistance and wind must be brought into the equation if a vacuum cannot be used to exclude them.

We are now headed in the right direction: research often cannot be conducted in vacuum, *in vitro*, in the laboratory. Astronomy, meteorology, and acoustics are examples from the physical sciences of how multivariate things can get in the real world. Meteorology cannot cope with five dimensions: three for air movement plus heat plus moisture. But economics begins with perhaps seven variables and can easily get over 100 with modern computers. One econometric model of the USA was said to contain 1 200 variables – and that monstrous model was not yet saturated!

Explanatory models in the social sciences, in economics, sociology, or psychology, often have this linear form:

$$y = b_0 x_0 + b_1 x_1 + b_2 x_2 \ldots . b_p x_p + e \, .$$

This can have many terms, a large number p. The predictand y can represent stock prices, or crime rates, or IQ scores, etc. We can take an honest look at the differences between such equations in the social sciences and idealized simple models in the physical and natural sciences, like the gravity model. The differences between the two camps is not absolute, but only quantitative and relative in the proportions of complex and soft expressions versus simple and hard models. But those differences are most relevant to statistics, and statistics are relevant to them.

First and foremost, models in the social sciences tend to be more multivariate; the number p above is large more often, and often very large. The models are necessarily multivariate due to our ignorance of what variables can or cannot be excluded; also because it is not safe or not realistic or not useful to exclude them. I believe other differences could be reduced to this multivariate character, but I also feel that it is heuristic to list a few other major contrasts.

Second, the nature of the model may not be clear. For example, instead of the additive linear model as in the regression above, the predictors could be in a multiplicative (or logarithmic) model, as in the gravity model. Furthermore, the predictors x may represent complex variables like Z^2 och $Z^{1.7}$.

Third, the nature and the direction of causation are often not clear. Discovery may be hampered by long delays between cause and effect. Furthermore, the direction of causation may be confused by reversals due to anticipation, because humans anticipate; for example, they save money in order to buy a house later.

Fourth, the value of the error term e is often unfortunately and embarassingly large, because the models are both inaccurate and incomplete. Furthermore, the statistical distribution of the errors e may be complex, and this provides work for statisticians, and especially for samplers, as we shall soon see.

"Exact sciences" is a term I avoid – as any statistician should. True that chemists and physicists know of some exact numbers like mass numbers and atomic numbers for elements. They can also produce fairly precise constants for pure substances under highly standardized situations. But chemical and other engineers must work with much less pure raw materials in the real world. And the raw materials of biologists and especially social scientists tend to come dirtier still. But is astronomy an "exact science"? What about all the surprises from probes to our moon and to our sister planets and their moons? Each surface is different: rock or sulphuric acid or frozen gas or gorgonzola cheese for all we know! "Exact science" is a misnomer not only statistically but philosophically as well.

Note also that for complex and soft models the same hierarchy of four stages exists as for the simple gravity model. First the scientists choose the variables, and second they decide on the form of the model. Third, with statistical help, the values of the parameters must be estimated, and fourth comes the assessment of errors. In the design of observations, scientist and statistician must collaborate; though both may reside within one skull – as in R.A. Fisher's or Harold Jeffreys'. For the social scientists, stages 1 and 2 of model building tend to be more tenuous and complex, hence more remains for the statistical tasks in stages 3 and 4.

4. Sampling

The most prominent aspect of statistics is *statistical analysis,* which occupies most of the energy of the discipline, with prestige from and deep roots in difficult mathematics. But the discipline has four other aspects that, in my view, are badly neglected. First, methods for the *collection* and observation of data are essential to statistics, but these methods are also diffused into and intimately linked to fields of application. Second, statistical *computing* belongs to a giant field now, but severing its ties to statistics would entail grave danger. Third, statistical *inference* has deep roots in scientific philosophy, and we touched briefly on this aspect earlier. Fourth, statistical *designs* are a vital and creative aspect of statistical practice that I shall now discuss.

Modern statistical designs, although they have earlier roots, begin with two recent publications: R.A. Fisher's *The Design of Experiments* in 1935, and Jerzy Neyman's 1934 paper on sampling. The introduction of chance effects through randomization plays important roles in both of these two areas, but the two roles differ in significant ways.

Experiments and surveys can be counterposed as alternative methods for carrying out research. The choice between them should depend on the strategic goals and on limitations of resources. Empirical researchers typically must face three fundamental obstacles to inference: control of treatments, representation of target populations, and realism of research setting. With limited resources usually only one, seldom two, of these three aspects can be brought under firm statistical control. (Kish (1975).)

Experiments are strong on control over the predictor variables, because these predictors, the contrasting treatments, are assigned with randomization to the experimental subjects. This (ideally) removes any biases from disturbing extraneous variables, and controls the nature and direction of causation from predictor to predictand variables. Modern experimental design allows for great flexibility and efficiency, and for powerful statistical manipulation.

However, true experiments are often not feasible, especially in the social sciences, but also in geology, paleontology, astronomy, biology, and medicine. Second, experiments are often weak on representation of broad and dispersed populations. Third, experimental settings may provide only poor imitations of the realism of natural settings.

In contrast, probability selection in survey sampling yields *representation* through randomization of the sample subjects over designated populations. Also, the survey measurements can often resemble the natural settings for the predictor variables. To approach the realism of natural settings, survey sampling may be combined with controlled investigations as a substitute for true experiments. These have great variety and many names, such as quasi- or natural experiments or controlled observations, and they form the bases of many advances in medicine, education, and the social sciences.

The concept of sampling, like the concept of inference, pervades the entire field of statistics, yet the methods of survey sampling must comprise a special area within that field. Statistics are needed to separate causal structures from the "white noise" of chance effects. Chance effects, due to errors of sampling and of measurements, exist wherever the observations, that is the data, are results of empirical investigation, and not of complete determinism, nor of logical or mathematical deduction.

Survey sampling departs from classical statistics in its view and treatment of the structure of the chance errors *e* within populations. Classical statistical analysis is based on assuming that those errors are distributed in a pure random manner. In contrast, the methods of survey sampling are strategies designed to introduce randomness into the samples, regardless of how complex the actual distributions of errors happen to be within the target populations. Thus survey sampling deals with chance errors at a second level of complexity[10].

For distinguishing causal relations from chance errors it would be convenient and powerful to assume that the errors of observation are simply random, or "identically and independently distributed, IID" as mathematical statisticians love to intone the magical phrase. This assumption of uniformity and independence of errors greatly facilitates mathematical results and is the foundation of most classical statistical analysis. It assumes that the errors of observation *e* come from the same distribution, from the same well-shaken urn, no matter how and whence the data come. This resembles a "pure" game of chance played against nature, and is motivated by the desire for a simplicity similar to simple games of chance, like coin tossing and dice. However, it seems that nature usually – I say always – plays more complex and challenging games, with complex error structures.

Sample designs are essentially strategies to counter with randomization those complex structures of errors in populations. The strategies have three major components. First, some operation of probability selection must be designed to introduce controlled randomization over the distribution of errors in the designated frame population. Second, this operation must match the nature of available resources to the procedures of collection and analysis. Third, the input of limited costs and efforts must be applied with high efficiency toward desired objectives. Furthermore, those statistical objectives are typically highly multipurpose, involving many variables and also many domains – by geography, age, social class etc; and involving increasingly multivariate analyses nowadays. All these multiple and diverse statistics can be obtained from one and the same complex probability sample.

[10] Kish and Frankel (1974).

	STATISTICS		
SELECTION METHODS	1 Means and totals of entire samples	2 Subclass means and differences	3 Complex analytical statistics, e.g. coefficients in regression
A. Simple random selection of elements			
B. Stratified selection of elements		Available	Conjectured
C. Complex cluster sampling		Available	Difficult: *BRR, JRR, TAYLOR*

Fig. 1. The present status of sampling errors. Row 1 is the domain of standard statistical theory, and column 1 of survey sampling

The nature and size of the design must be fitted to available resources of time, money, and people; it must be efficient and economical; and, to be practical, the actual operations must be carried out in the field and in the office essentially as planned in the design.

The science of survey sampling deals mostly with the components of variances and with the cost factors involved in the many statistics of complex sample designs. The art and craft of survey sampling concern the practical applications of that science to the available resources, especially in the operations of data collection and of selection frames. Random numbers, selected mechanically from a table or from a computing program, identify listing numbers in a frame, and these numbers must identify the actual units to be selected and observed for the sample. (Kish (1965)).

The term "selection frame" denotes the complex procedures for associating, listing, and numbering the members of the population for eventual selection. For example, a sample of adults in the United States may require selection frames in five stages: first a sample of counties, then of area blocks from the selected counties, then segments from the blocks, then dwellings from the segments, and finally adults to be interviewed from the dwellings. The operations must be controlled over the five stages to obtain these adults with the desired and determined final probabilities. Similar procedures exist for selecting persons from frames of telephone numbers. Also for selecting organisations and business firms; and groups, institutions, schools and universities, and then classes, teachers, and students within them. Similar methods and procedures are applied for selecting samples for studies of agriculture and livestock, natural resources, etc.

The science of sampling, and even much of its art and craft to a great extent, possesses great generality and transportability across many fields of application and into all parts of the world. That is why we have been able to conduct a successful Sampling Program for Foreign Statisticians over 20 years, with graduates in 85 countries and in a great variety of fields of application – of which our University may be proud.

A recent example of the importance of sampling was the World Fertility Survey, which has obtained high quality samples and standardized data of great social content in about 60 countries. This is the largest social science project to date and an enviable achievement for united humankind. (Verma, Scott and O'Muircheartaigh (1980)).

Finally, sampling and experimental designs are only two of many new fields laid open with statistical recognition of chance effects. Time series and spectral analysis comprise another rich new area with several branches. Epidemiology has created many new applications of probability and statistics. Furthermore, McNeill (1976) recently used those ideas to write new interpretations of world history in his *Plagues and Peoples*. Industrial quality control, begun by Walter Shewhart, in combination with industrial relations, is remaking Japanese history. In Monte Carlo simulations we introduce chance effects with modern computers to solve problems that defy mathematical analysis. Other new fields of the effects of chance and of statistics still are awaiting discovery, and we must remember that "chance favors the prepared mind."

5. References

Aubert, V. (1959): Chance in Social Affairs. Inquiry, 2, pp. 1–24.

Bork, A.M. (1967): Randomness in the Twentieth Century. Antioch Review. Reprinted in Risk and Chance, ed. Dowie, J. and Lefrere, P., Milton Keynes: The Open University Press, 1980.

Burks, A.W. (1977): Chance, Cause, Reason. Chicago: The University of Chicago Press.

Gould, S.J. (1977 a): Ever Since Darwin. New York: W.W. Norton, p. 12.

Gould, S.J. (1977 b): Ontogeny and Phylogeny. Cambridge: Belknap Press of Harvard University.

Gould, S.J. (1980): The Panda's Thumb. New York: W.W. Norton, p. 184.

Hoffman, B. (1972): Albert Einstein, Creator and Rebel. New York: Viking Press, p. 193.

Kendall, M.G. (1968): Chance, in Dictionary of History of Ideas. New York: Charles Scribners & Sons, pp. 335–340.

Kish, L. (1965): Survey Sampling. New York: John Wiley and Sons, Section 1.

Kish, L. and Frankel, M.R. (1974): Inference from Complex Samples. Journal of the Royal Statistical Society (B), 36, pp. 1–37.

Kish, L. (1975): Representation, Randomization and Control. Ch. 9 in Blalock H.M. et al. (eds.), Quantitative Sociology, New York: Academic Press.

Kish, L. (1978): Chance, Statistics and Statisticians. Journal of the American Statistical Association, 73, pp. 1–6.

McNeill, W.H. (1976): Plagues and Peoples. Garden City, N.Y.: Anchor Press.

Monod, J. (1971): Chance and Necessity. New York: Vintage Books p. 118.

Mosteller, F. (1952): World Series Competition. Journal of the American Statistical Association, 47, pp. 355–380.

Piaget, J., and Inhelder, B. (1975): The Origin of the Idea of Chance in Children. New York: W.W. Norton, p. 215. French original, 1951.

Poincaré, H. (1952): Science and Method. New York: Dover, Chapter on Chance, pp. 67–69.

Popper, K.R. (1959): The Logic of Scientific Discovery. London: Hutchinson.

Poston, T. and Stewart, I. (1978): Catastrophe Theory and Its Applications. Pitman, London.

Salmon, W.C. (1967): The Foundation of Scientific Inference. Pittsburgh: The University of Pittsburgh Press.

Scardovi, I. (1976): Pagine dimenticate di storia della scienza. Genus, 32, pp. 1–44 (English summary).

Schumacher (1973): Small is Beautiful. New York: Harper & Row Perennial, pp. 224–225.

Thomas, L. (1980): On the Uncertainty of Science. The Key Reporter (ø BK), 46, pp. 1–4.

Verma, V., Scott, C. and O'Muircheartaigh (1980): Sample Designs and Sampling Errors for the World Fertility Survey. Journal of the Royal Statistical Society (A), 143, pp. 431–473.

Weaver, W. (1963): Lady Luck: The Theory of Probability. Garden City: Anchor Press, and London: Pelican Books (1977).

Received September 1984