Methodos Series 10

Daniel Courgeau

Probability and Social Science

Methodological Relationships between the two Approaches

Springer

Probability and Social Science

METHODOS SERIES

VOLUME 10

This Book Series is devoted to examining and solving the major methodological problems social sciences are facing. Take for example the gap between empirical and theoretical research, the explanatory power of models, the relevance of multilevel analysis, the weakness of cumulative knowledge, the role of ordinary knowledge in the research process, or the place which should be reserved to "time, change and history" when explaining social facts. These problems are well known and yet they are seldom treated in depth in scientific literature because of their general nature.
So that these problems may be examined and solutions found, the series prompts and fosters the setting-up of international multidisciplinary research teams, and it is work by these teams that appears in the Book Series. The series can also host books produced by a single author which follow the same objectives. Proposals for manuscripts and plans for collective books will be carefully examined.
The epistemological scope of these methodological problems is obvious and resorting to Philosophy of Science becomes a necessity. The main objective of the Series remains however the methodological solutions that can be applied to the problems in hand. Therefore the books of the Series are closely connected to the research practices.

For further volumes:
http://www.springer.com/series/6279

Daniel Courgeau

Probability and Social Science

Methodological Relationships between the two Approaches

Springer

Daniel Courgeau
Institut National d'Etudes
Démographiques (INED)
Boulevard Davout 133
75980 Paris Cedex 20
France

ISBN 978-94-007-2878-3 e-ISBN 978-94-007-2879-0
DOI 10.1007/978-94-007-2879-0
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2012932214

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my beloved wife, Hella

Acknowledgements

I would like to thank all the persons with whom I exchanged ideas and discussed different problems raised in this volume during the last 10 years: from the point of view of probability and statistics (first Marc BARBUT, unfortunately deceased just before the publication of this book, who opened the way to this volume, as editor of *Mathematics and Social Sciences -Mathématiques et Sciences Humaines-* and of the *Electronic Journal for History of Probability and Statistics* on publishing during these 10 years a number of papers I wrote on these themes, Xavier BRY, Henri CAUSSINUS and in the memory of Henry ROUANET who unfortunately died in 2008); from the point of view of social science (Alban BOUVIER, Eric BRIAN, Luc BUCHET, Tom BURCH, Ana DIEZ ROUX, Fatima JUAREZ, Alain GALLAY, Jean-Claude GARDIN, Harvey GOLDSTEIN, Nico KEILMAN, Denise PUMAIN, Jean-Marc ROHRBASSER, Isabelle SEGUY, Herbert SMITH, Eric VILQUIN, Bernard WALLISER, Guillaume WUNSCH); from the point of view of philosophy and methodology (Robert FRANCK, Pierre LIVET, Federica RUSSO); and many other researchers for discussion on these themes, particularly during the session on *Epistemology in demography and sociology*, I organized at the IUSSP (International Union for the Scientific Study of Population) General Conference in 2005.

Two anonymous referees must also be thanked for their very positive and constructive review on the outline of the manuscript.

Many thanks finally to Jonathan MANDELBAUM for his very clever translation of the whole volume.

Contents

General Introduction xiii
Links Between Probability Theory and Social Science at Their Inception xv
Statistical Inference, Induction, and Social Science xx
Concordance Between Basic Probability Concepts and Social Science xxi
Overview of Entire Volume xxiv
General Conclusion xxxiii

Part I From Probability to Social Science

Introduction to Part I 3

1 The Objectivist Approach 7
1.1 Objectivist Probability 8
1.1.1 First Step: A Classical Theory 8
1.1.2 Second Step: A Frequentist Theory 9
1.2 Paradigm and Axiomatics of Objective Probability 12
1.2.1 The Paradigm 14
1.2.2 The Axioms 17
1.3 Objectivist Statistical Inference 25
1.4 Application to Social Science 28
1.5 Problems Posed by the Objective Approach 38

2 The Epistemic Approach: Subjectivist Interpretation 43
2.1 Subjectivist Probability 44
2.2 Paradigm and Axiomatics of Subjective Probability 55
2.2.1 The Paradigm 56
2.2.2 The Axioms 57
2.3 Subjectivist Statistical Inference 67
2.4 Applications to Social Science 69
2.5 Problems Posed by the Subjective Approach 79

3 The Epistemic Approach: Logicist Interpretation 85
3.1 Logical Probability 86
3.2 Paradigm and Axiomatics of Logical Probability 98
3.2.1 The Paradigm 98
3.2.2 The Axioms 100
3.3 Logicist Statistical Inference 112
3.4 Application to Social Science 113
3.5 Problems Posed by the Logicist Approach 129

Conclusion of Part I 133
A Unique Notion or a Multi-faceted Notion? 133
Inference and Decision-Making Theory 138
Does Cumulativity Exist in Probability? 144

Part II From Population Sciences to Probability

Introduction to Part II 151

4 The Dispersion of Measures in Population Sciences 155
4.1 From Pascal and Fermat's Wagers to Graunt's Bets 156
4.2 Introduction of Epistemic Probability in Population Sciences 159
4.3 Toward an Objectivist Approach in Population Sciences 166
4.4 Return of Dispersion in the Event-History and Multilevel Approach 170
4.5 Estimating the Age Structure in Paleodemography 174
4.5.1 Methods Proposed Earlier 175
4.5.2 A New, Truly Epistemic Approach 182
4.5.3 Example of Archeological Application 186
4.6 Conclusion 189

5 Closer Links Between Population Sciences and Probability 191
5.1 The Framework: The Population and the Individual 193
5.2 The Object: Not Individual Behaviors but More Abstract Concepts 197
5.3 The Cross-Sectional Approach 200
5.3.1 Independence Between Phenomena 200
5.3.2 Dependence of Characteristics on Society 203
5.3.3 Problems Posed by the Cross-Sectional Approach 208
5.4 From a Longitudinal Vision to a Full-Fledged Event-History Approach 209
5.4.1 An Initially Longitudinal Approach 210
5.4.2 An Event-History Approach 213
5.4.3 Problems Posed by This Approach 219

5.5 From a Latent Hierarchical Vision to a Multilevel Approach 220
5.5.1 An Initially Hierarchical and Latent Vision 221
5.5.2 A Contextual Then Fully Multilevel Approach 227
5.5.3 Problems Posed by the Multilevel Approach 230

Conclusion to Part II 233
Intensity of Ties with Probability 233
Links Between Probability, Social Science, and Counterfactual Causality 235
Does Cumulativity Exist in the Population Sciences? 238

General Conclusion 243
Generality of the Use of Probability and Statistics in Social Science 243
Revisiting Causality in Social Science 249
Revisiting the Notions of Individual and Levels 252
Predicting Behavior in Social Science 256
Epilogue 259

Glossary 261

References 263

Author Index 295

Subject Index 303

General Introduction

The notion of chance has always been present in human culture from earliest antiquity, and all peoples have used a wide variety of games in which chance plays a relatively important role (David 1955). For example, traces of the game of knuckle-bones[1] can be found during the First Dynasty in Egypt as early as 3500 B.C.E., and Roman soldiers played it by betting on the sides that would turn up after a throw. There is evidence of dice games in Mesopotamia, Egypt and Babylon dating from the third millennium B.C.E. Games with sticks were played by the Mayas, the Greeks, the Romans, the ancient Bretons, and the Egyptians—along with card games, chess, and others. These games were disseminated either for religious purposes, such as Jewish Talmud (Hasofer 1967; Rabinovitch 1969, 1970), divining among the Greeks, the Romans, and Tibetan Buddhists, or for recreational purposes (David 1955). However, this application of chance was not formalized more rigorously until much later[2] (Kendall 1956) and in gradual stages. Examples include the Latin poem *De Vetula*, possibly written by Richard de Fournival between 1200 and 1250, *Liber de ludo aleae* by Cardano written in approximately 1564 but published only in 1663, a fragment by Galileo Galilei (ca. 1642), and the studies by Pascal and Fermat (1654, 1922). These texts set the stage for the emergence of probability theory as a full-fledged scientific discipline.

The same pattern applies to the study of population and the efforts to enumerate human beings. Population counts were already performed by the Egyptians around 3000 B.C.E., partly to meet labor requirements for the construction of the Pyramids; they were carried out in Mesopotamia during the same period for religious reasons, by Moses in Sinai at God's behest ('Take a census of the whole community of Israelites by clans and families, taking a count of the names of all the males, head

[1] Small bones of the tarsus connected to the tibia and fibula. Knuckle-bones of hoofed animals such as sheep and goats have been found in large quantities on archeological sites dating back to at least 40,000 years. The knuckle-bone in these animals is roughly symmetrical; in others such as cats and dogs, it is totally asymmetrical and thus unsuitable for games of chance.

[2] Italian authors from the early fourteenth century to the fifteenth century offered various partial formalizations: see the article by Meusnier (2004).

by head [...] 20 years of age and over [...]', *Numbers*, 1, 2), by the Chinese Emperor Yu or Yao in the Empire of the Center after a great flood in 2238 B.C.E., by the Greeks in sixteenth century B.C.E (Missiakoulis 2010), the Romans, the Incas of Peru, and others (Hecht 1977). Once it had developed an organized structure, a State manifestly needed to count not only its citizens but also its economic resources. Here as well, however, the analysis of censuses and registers occurred much later, when scientists succeeded in measuring and quantifying phenomena that were previously God's secret. The first such analysis was published by Graunt (1662), followed by Christiaan and Lodewijck Huygens (1669, see Huygens 1895; Véron and Rohrbasser 2000) and Leibniz (early 1680s, see Rohrbasser and Véron 2001). The social sciences—such as demography, economics, and epidemiology—could now enter the scene.

A large body of literature has addressed these two broad themes separately: first, the history, methodology, and epistemology of probability and statistics (Gouraud 1848; Todhunter 1865; Matalon 1967; Hacking 1975, 1990; Krüger et al. 1986; Stigler 1986; Porter 1986; Daston 1988; Gigerenzer et al. 1989; Desrosières 1993; Barbin and Lamarche 2004); second, the history, methodology, and epistemology of population and other social sciences (Durkheim 1895; Landry 1945; Granger 1967; Piaget 1967; Franck 1994, 2002; Berthelot 2001; Courgeau 2002, 2003; Martin 2003).

Our purpose here is entirely different. We want to examine the historical connections between those two broad sectors. Analysis and research projects were not carried out independently of one another but, on the contrary, in close interaction.

Pascal, for example, worked on mathematics (*Essay sur les coniques*, 1640), probability theory (*Traité du triangle arithmétique III*, 1654), physics (*Récit de la grande expérience de l'équilibre des liqueurs*, 1648), and philosophy (*Entretien avec Sacy sur la philosophie* and *Les pensées*, 1670). Leibniz worked alternatively on logic, mathematics, probability theory, history, linguistics, law, politics, philosophy, and other disciplines. All these subjects are addressed in his complete works (see the website: http://www.leibniz-edition.de/).

The same is true of many researchers since the seventeenth century, although a greater specialization developed over time. In the twentieth century, for instance, Fisher worked simultaneously on probability theory, statistics, and genetics throughout his life, Keynes on economics and probability theory, and so on. Our aim here is to describe the origin and development of the relationships that have always existed between these disciplines. That is what makes this volume different from its predecessors. In the first part of this General introduction, we illustrate the links that were established between probability theory and the social science at their very inception.

As noted above, the concept of probability arose from the examination of the outcomes of a wide variety of games such as dice and cards. It took shape through a theoretical and mathematical evaluation of the number of possible outcomes, assumed to be equally likely. This *geometry of chance* (*géométrie du hasard*), as Pascal called it, does not suffice in social science, where probabilities cannot be determined in advance. All we can do is perform a certain number of comparable tests and observe a posteriori the number of events occurring in the sample, such as the number of deaths in a population. How can we then use these figures to reconstruct an unknown probability? What significance should we assign to the principles of statistical inference and induction that we can use to infer a probability from an earlier observation of facts?

Moreover, the social sciences have not yet managed to define their specific ‘object’ and their present state ‘may be compared to that of natural sciences in pre-Galilean times’ (Granger 1994). The complex and changing life experience that constitutes a ‘human fact’ still needs to be conceptualized as a scientific object, and we shall try to make some progress toward that goal here. In the second part of this introduction, we address the issues raised by this statistical inference, the problems encountered in social science and possible ways to solve them.

Despite its flowering in the seventeenth century, probability theory was not axiomatized until the twentieth century, with the work of Kolmogorov (1933). However, while the role of axiomatization is to define mathematical beings in formal terms, it does not tell us what entities in nature can be represented by them. For instance Kolmogorov clearly conveys his belief that not every event has a probability (1951):

> Certainly not every event whose occurrence is not entirely determined under given conditions has a definite probability under these conditions

and he asserts his frequentist position (1933). Nevertheless, with slight alterations, his axioms can apply to other approaches to probability theory—for instance, the subjectivist or logicist approaches. It is therefore important to realize that ‘probability theory formalizes something that, in a manner of speaking, ‘exists’ independently; the divergences concern the nature of that ‘something’ which, according to this approach, is represented by the mathematician’s probability’ (Matalon 1967).

In the third part of this introduction, we shall examine this axiomatization and the problems encountered in applying it to a universe of experience—in social, biological, or physical science.

The fourth and final part will outline the path followed in this book, so that the readers can locate their position in the overall plan at all times.

Links Between Probability Theory and Social Science at Their Inception

While the investigations by Greek philosophers and mathematicians did not lead them to probability theory or to social science (Granger 1976), their work did enable them to raise the issue of chance and introduce the notion of *justice*, a crucial factor in the establishment of links between probability and social science.

For instance, Aristotle already made a clear distinction between things that ‘always occur identically and others [that occur] frequently.’(Physics, 196b). In the Nichomachean Ethics, he writes (III:3):

> And in the case of exact and self-contained sciences there is no deliberation […]; but the things that are brought about by our own efforts, but not always in the same way, are the things about which we deliberate […]. Deliberation is concerned with things that happen in

> a certain way for the most part, but in which the event is obscure, and with things in which it is indeterminate.[3]

While he does not succeed in formalizing this probable outcome correctly, the introduction of the notion of justice (fairness) and its formalization led to rules that preceded probability theory by centuries and made it possible. Aristotle defines justice as ‘that kind of state of character which makes people disposed to do what is just and makes them act justly and wish for what is just’ (Nichomachean Ethics, V:1). He goes on to formalize the notion as follows:

> The just, therefore, involves at least four terms; for the persons for whom it is in fact just are two, and the things in which it is manifested, the objects distributed, are two. And the same equality will exist between the persons and between the things concerned; for as the latter—the things concerned—are related, so are the former; if they are not equal, they will not have what is equal, but this is the origin of quarrels and complaints—when either equals have and are awarded unequal shares, or unequals equal shares.

Aristotle views justice as a critical element in, for example, concepts such as markets and money, which allow contracts between different and unequal persons. He therefore extends the argument by stating:

> This is why all things that are exchanged must be somehow comparable. It is for this end that money has been introduced, and it becomes in a sense an intermediate; for it measures all things, and therefore the excess and the defect—how many shoes are equal to a house or to a given amount of food. (Nichomachean Ethics, V:5)

What ultimately gives justice its full importance in the genesis of the notion of probability is the random contract (Daston 1988).

In his book *Liber de ludo aleae*, written in the mid-sixteenth century but not published until 1663, Cardano invokes Aristotle to define a fair wager:

> Other questions must be examined in a subtler manner, for mathematicians too can err, but differently. I did not want this issue to be set aside, for many people, who have not understood Aristotle, have erred, incurring losses. Thus there is a general rule that requires us to consider the total circuit,[4] and the number of outcomes representing all the ways in which a favorable result can occur, then to compare this number to the rest of the circuit, and lastly to examine the proportion to be used in reciprocal wagers so that they apply to equal terms.[5]

We shall see later how to formalize such a line of argument, which enables us to compute the probability of an event using the notion of circuit.

---

[3] Translation by W.D. Ross, http://classics.mit.edu/Aristotle/nicomachean.html.

[4] Cardano uses the term ‘circuit’ to denote the set of throws of different dice that can be examined in a given game.

[5] Reliqua ergo subtiliter consideranda; cum etiam in Mathematicis deceptio contigat, sed alia ratione. Volui hoc non latere, quia multi non intellegentes Aristotelem, decipiuntur, & cum iactura. Vna is ergo ratio generalis, vt consideremus totum circuitum, & ictus illos, quot modis contingere possunt, eorumque numerum, & ad residuum circuitus, eum numerum comparentur, & iuxta proportionem erit commutatio pignorum, vt equali conditione certent.

For Pascal as well, fairness is the concept that enabled him to develop the 'geometry of chance' Indeed, he presented his treatise in the following terms (Pascal 1654):

> …an entirely new treatise, on a subject hitherto utterly unexplored, namely: the distribution of chance in games that are governed by chance—what is known in French as *faire les partis des jeux* [setting the odds of the game]; the uncertain outcome is so well controlled by the fairness of the computation that each player always receives exactly the amount consistent with justice.[6]

Pascal goes on to show how reasoning allows progress in this area, where experience seems of little use to him:

> And it is there, surely, that we must seek by means of reasoning all the more so as we are less likely to be informed by experience. Indeed, the results of ambiguous chance are rightly attributed to fortuitous contingency rather than to natural necessity. That is why the issue has drifted uncertainly until today. But now, having remained impervious to experience, it has failed to escape the empire of reason. And thanks to geometry, we have reduced it so effectively to an exact art that it partakes of geometry's certainty and has already made bold progress. Thus, by combining the rigor of scientific demonstration with the uncertainty of chance, and reconciling these apparent opposites, it can, drawing its name from both, rightfully claim this astonishing title: *The Geometry of chance*.[7]

In the third section of his *Traité du triangle arithmétique* [Treatise on the arithmetical triangle] (1654), Pascal spells out the prerequisites for reasoning on chance:

> …the money that players have wagered no longer belongs to them, for they have relinquished their property of it; but, in exchange, they have received the right to expect the share of that money which chance can give them, under the terms they have agreed upon at the outset.

In the third section, Pascal also formulates the two principles that he views as the prerequisites for computing probability:

> The first principle, which is designed to determine how shares should be divided, is this.
>
> If one of the players finds himself in such a situation that, whatever the outcome, a certain sum accrues to him in the event of loss and gain, without chance being able to deprive him of it, he must not wager it, but take it in its entirety as guaranteed. This is because the wager must be proportional to the chances, and since there is no risk of loss, he must withdraw the entire amount undivided.
>
> The second principle is this. If two players find themselves in such a situation that, if a player wins, he is entitled to a certain sum, and if he loses, the sum will go to the other

[6] Novissima autem ac penitus intentatae materiae tractatio, scilicet de compositione aleae in ludis ipsi subjecti, quod gallico nostro idiomate dicitur *faire les partis des jeux*, ubi anceps fortuna aequitate rationis ita reprimitur ut utrique lusorum quod jure competit exacté semper assignetur.

[7] Quod quidem eô fortius ratiocinando quaerendum, quò minus ten tando investigari possit. Ambiguae enim sortis eventus fortuitae contingentiae potius quam naturali necessitati meritò tribuuntur. Ideò res hactenus erravit incerta; nunc autem quae experimento rebellis fuit rationis dominium effugerenon potuit. Eam quippè tantâ securitate in artem per Geometriam reduximus, ut certitudinis ejus particeps facta, jam audacter prodeat; & sic matheseos demonstrationes cum aleae incertitudine jungendo, ab utraque nominatinem suam accipiens, stupendum hunc titulum jure sibi arrogat: *aleae Geometria*.

> player; if the game is of pure chance and if the chances of winning are equal for both players and therefore the chances of winning are no greater for one player than for the other, if they want to part ways without playing, and reclaim their legitimate shares, they should divide the sum at stake in half, and each should take his half.

Pascal clearly indicates that this is a game of pure chance, i.e., for example, that the dice are not loaded. Using the arithmetical triangle, he generalizes this result to the broader case in which the players break up the game at a time when the first player is missing *m* shares and the second player *n* shares. Interestingly, Fermat, who discussed his approach in his correspondence with Pascal on this subject (Pascal 1922), reached the same result by means of a purely combinational method, and this enabled Pascal to conclude:

> I admire your method for wagers, all the more so as I comprehend it very well; it is entirely your own, and has nothing in common with mine, and reaches the same goal but easily. Our [mutual] understanding is thus restored.

In this exchange, Pascal and Fermat were addressing objective probability, for the chances of winning are determined by the fact that the playing tokens have not been tampered with. But Pascal's wager takes the reasoning further and introduces epistemic probability, for unique events, such as the existence of God. In a section of the *Pensées* entitled *Infini rien* [*Infinite nothingness*] (1670), he shows how an examination of chance can lead to a decision of a theological nature. Let us summarize his approach briefly here; we can return to it in greater detail in later sections of this book, when needed. Pascal argues as follows. Consider an individual who hesitates between faith and unbelief, but does not want to rely on the testimony of believers, doctors of the Church or miracles. Pascal begins by stating how the question is formulated absent experimental data:

> And let us say: God is or is not; but to which side shall we lean? Reason is of no avail here. An infinite chaos separates us. A game is being played at the far end of this infinite distance, where heads or tails will turn up. What will you wager?

Pascal shows that we must wager the existence of God, and that a probabilistic approach is possible here:

> Let us weight the gain and loss, wagering tails that God exists. Let us estimate the two outcomes: if you win, you win all, and if you lose, you lose nothing: therefore, without hesitation, wager that God exists. That is admirable.

Here, Pascal examines a hypothesis—the existence of God—and shows that the previous probabilistic argument, which concerned the occurrence of events that could reoccur in identical conditions, remains possible. While we can criticize its premises, this reasoning closely resembles that of game theory, but is based on entirely different arguments.

Let us now examine the situation in social science at the time. The first experiment in social science was, in fact, provided by John Graunt (1662), who submitted his findings to John Lord Roberts, Lord Privy Seal, in these terms:

> Now having (I know not by what accident) engaged my thoughts upon the Bills of Mortality, and so far succeeded therein, as to have reduced several great confused

> Volumes into a few perspicuous Tables, and abridged such Observations as naturally flowed from them, into a few succinct Paragraphs, without any long series of *multiloquious Deductions*…

Graunt's approach effectively summarizes many observations by means of clear statistical tables. He uses these mortality statistics to deduce, through probabilistic reasoning, the population of London, formerly estimated at six million by worthy persons:

> Next considering, That it is esteemed an even Lay, whether any man lives ten years longer, I supposed it was the same, that one of any 10 might die within one year. But when I considered, that of the 15000 afore-mentioned about 5000 were *Abortive*, and *Still-born*, or died of *Teeth, Convulsion*, *Rickets*, or as *Infants* are *Chrysoms*, and *Aged*. I concluded, that of men, and women, between ten and sixty, there scarce died 10000 per Annum in London, which number being multiplied by 10, there must be 100000 in all, that is not the 1/60 part of what the *Alderman* imagined.

We shall see later on the errors committed in this reasoning. Suffice it to say here that Graunt's method is still highly approximative and his hypotheses extremely crude. Compiling a true life table would require, at the very least, a series of age-specific probabilities of dying, which are far from constant. This was achieved decades later by Edmond Halley (1693), who set out to estimate the 'Degrees of the Mortality of Mankind' from the bills of mortality and birth of Breslau, a town whose population was less affected by migration than that of Graunt's London.

William Petty (1690) generalized the approach—which he designated as *Political Arithmetic*—not only to demographic issues but also to economic, political, epidemiological, administrative, and other issues in social science:

> The Method I take to do this is not yet very usual: for instead of using only comparative and superlative Words, and intellectual Arguments, I have taken the course (as a Specimen of the Political Arithmetic I have long aimed at) to express myself in Terms of Number, Weight, or Measure; to use only Arguments of Sense, and to consider only such Causes as visible Foundations in Nature; leaving those that depend upon the mutable Minds, Opinions, Appetites of particular Men, to the Consideration of others. (Petty 1690)

It is under the label of political arithmetic that the social sciences developed during the seventeenth, eighteenth, and early nineteenth centuries. The rise of political economics began with Petty and de Boisguilbert (1695), followed by Cantillon (1755), Quesnay (1758), and Adam Smith (1776). After Graunt and Petty, demography and epidemiology progressed thanks to Halley (1693), Süßmilch (1741, 1761–1762), and Deparcieux (1746). But the term *demography* did not appear until much later—in its French form of *démographie*—in the title of Guillard's book *Eléments de Statistique Humaine ou Démographie Comparée* (1855). Epidemiology followed a similar path. The term was initially used to denote a medical discipline devoted to large-scale outbreaks of infectious diseases. But it did not emerge as a scientific discipline until the nineteenth century, most notably with the founding of the London Epidemiological Society in 1850.

Statistical Inference, Induction, and Social Science

If the probabilities of successive plays in a game of pure chance can be computed by strictly rational means, in other cases—particularly in social science—they cannot be *a priori* probabilities but only be determined *a posteriori*. Unlike Pascal, who was working on results that he could regard as equiprobable, Graunt had to use empirical observations to deduce probabilities of dying. It is after these observations on human mortality that Pascal's successors tried to generalize the notion of probability.

This involves going back from effects to causes—from empirical observations to the factors that generate them—in order to achieve greater certainty and, above all, greater generality in the analysis. This is known as the problem of statistical inference. After the initial efforts by Jacob Bernoulli (1713) to solve it, the solution was eventually proposed by Bayes (1763). Condorcet and Laplace developed it as the mathematical instrument perfectly suited to social science, where the *a priori* probabilities of causes were always unknown. Let us briefly review the issues raised and solved, which we shall examine in greater detail in the first section of this volume.

Jacob Bernoulli died in 1705, but his book, published by his nephew Nicolas, did not appear until 1713. In it, the author clearly stated the problem of a priori and a posteriori probabilities:

> But, in truth, another path is open to us in our quest for what we are seeking. What we cannot obtain *a priori* can at least be determined *a posteriori*, i.e., we shall be able to extract it by observing the outcomes of many similar examples; for we must assume that, later on, each fact can occur or not occur in the same number of cases as it was previously observed to occur or not occur in similar circumstances.[8]

The problem that Bernoulli is trying to solve is thus indeed complementary to the one raised by Pascal: when we do not know the a priori probability, we must obtain it a posteriori, from the observation of many similar outcomes. However, we are dealing here with objectivist probabilities, where the law of large numbers enables us to confer an objective, non-equivocal status upon the notion of probability. In the process, Bernoulli demonstrates a theorem still known in probability theory as the weak law of large numbers:

> Thus it is this problem that I now propose to solve, after having reflected on it for twenty years: its novelty and great usefulness, combined with its great difficulty, may exceed in weight and value all the other chapters of this thesis.[9]

[8] Verum enimverò alia hîc nobis via suppetit, quâ qæsitum obtineamus; & quod *à priori* elicere non datur, saltem *à posteriori*, hoc is, ex eventu in similibus exemplis multoties observato eruere licebit; quandoquidem præsumi debet, tot casibus unumquodque posthac contingere & non contingere posse, quoties id antehac in simili rerum statu contigisse & non contigisse fuerit deprehensum.

[9] Hoc igitur is illud problema, quod evulgandum hoc loco proposui, postquam jam per vicennium pressi, and cujus tum novitas, tum summa utilitas cum pari conjuncta difficultate omnibus reliquis hujus doctrinæ capitibus pondus and pretium superaddere potest.

His demonstration of the theorem is perfectly correct, but he was expecting a fuller result from his investigations. His finding applies to objectivist probabilities whereas his work was intended to apply to subjectivist probabilities, as he clearly states in his treatise.

Jacob Bernoulli accordingly demonstrates that if we know the probability of a phenomenon (assumed to be constant in successive observations), then, when we increase the number of observations, the observed frequency will diverge from its probability by a given quantity, which we can determine with the aid of that number and can set to as low a value as we want.

Bayes managed to go further by proving the opposite theorem, at least in a simple given case. He begins his article (1763) by clearly announcing the problem he intends to solve:

> *Given* the number of times in which an unknown event has happened and failed: *Required* the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named.

That is indeed the principle of statistical inference. The approach consists in using the observation of occurrences of an event to draw an inference on the probabilistic distribution responsible for the phenomenon—i.e., to provide an analysis of a past phenomenon, or a prediction of a similar future phenomenon. Throughout this volume, we shall see the various meanings that have been assigned to statistical inference and their links to social science.

Concordance Between Basic Probability Concepts and Social Science

As noted earlier, the notions of chance and of counting populations as well as some of the events they experience have been present in human thought since earliest antiquity. However, the concepts were not refined and initially mathematized until around the seventeenth century—by Pascal and Fermat (1654) for probability and Graunt (1662) for social science. This mathematization should logically lead to a more precise search for the bases on which to build a more robust theory of probability and social science.

Our introduction attempts to outline some of these bases in order to show the concordance or discordance between probability and social science. We shall elaborate on the bases in growing detail throughout the rest of the volume.

As Pascal and Fermat showed, probability could be mathematized, paving the way for probability theory. However, the research on probability focused on concepts that did not fit into the mathematics or the logic of the period: events, proof, randomness, chance, likelihood of an event, expected winnings, and so on—none of these concepts entered into the formalization of social science. Likewise, some of the chosen examples drawn from social science since the very inception of probability theory clearly showed the theory's potential use in fields other than games. Hence the need to define those concepts more precisely in order to use them with greater confidence and ensure that everyone was referring to the same things when applying them.

From the outset, Cardano clearly enunciated the main precondition of equiprobable outcomes, without which there could be no fair wager:

> The most fundamental principle of all in gambling is simply equal conditions, e.g., of opponents, of bystanders, of money, of situation, of the dice box, and of the die itself. To the extent to which you depart from that equality, if it is in your opponent's favor, you are a fool, and if in your own, you are unjust.[10]

This broad notion of equality is thus indeed the bedrock of probability theory, without which it would be meaningless.

Similarly, when Huygens sought to axiomatize probability—in the first true handbook on the subject published in 1657 under the title *On ratiocination in dice games*—he clearly showed that one cannot determine the fair amount of a wager except in a game with even chances:

> … I start from the hypothesis that in a game, the chance of winning something has a value such that if we possess that value we can obtain the same chance by means of a fair game, i.e., a game that seeks to deprive no-one.

Again, this is a fundamental notion that allowed a reasoned investigation of probability. The notion was taken up by many later authors to serve as a basis for probability theory. Hence the classic definition of probability as: the ratio of the number of positive outcomes to the total number of outcomes provided that these are all equally possible. However, this definition contains a circular element, *equally possible* being an exact synonym of *equally probable*.

For instance, in a game of heads or tails, the only way to determine that the coin is as likely to land face up as face down is to toss it an infinite number of times. In social science, the problem is even trickier, for we must assume that the probability of an event is identical for all individuals in a given population.

This question was directly addressed by Henry (1959) in his discussion of a fundamental issue in demographic analysis:

> A homogeneous cohort may be viewed as consisting of identical individuals whose life histories differ only by chance. We can classify their histories according to the events that characterize them and the dates of their occurrence. This yields a statistical history of the cohort: a given proportion of individuals has experienced a given type of history. Let us now imagine that each individual in the cohort can repeat his or her history indefinitely; the infinite set of histories of each individual could, in turn, be classified according to the same criteria as before; we would obtain a statistical history of the individual. For a homogeneous cohort, the statistical history of the individuals who compose it is identical to the statistical history of the cohort.

[10] Is autem, omnium in Alea principalissimum, aequalitas, ut pote colusoris, astantium, pecunarium, loci, fritilli, Aleae ipsius. And quantumcumque declinaueris ab ea aequalitatae aduersum te, stultus es, & pro te iniustus.

However, Henry is then forced to admit that actual cohorts do not consist of identical individuals and that no human group is homogeneous. This finding undermines the analytical methods commonly used in demography, which assume cohort homogeneity or do not address that homogeneity. The author examines the equally theoretical case of a heterogeneous cohort formed by the amalgamation of infinitely large homogeneous cohorts. Once again, we are faced with difficulties similar to those encountered in probability theory when analyzing equally probable outcomes. Henry shows that error can be null only when the cohort is, in fact, homogeneous with respect to the topic studied. We shall return to these issues later.

Another basic notion of probability theory was defined somewhat later by Jacob Bernoulli (1713) and elaborated by Cournot (1843). It involves the case where the possibility of an event may be so close to zero that we may regard it as *physically impossible* or, on the contrary, so close to unity that we may regard it as *physically certain*. In Chap. IV of Part IV of *Ars Conjectandi* (1713), before demonstrating his theorem on the law of large numbers, Jacob Bernoulli clearly states:

> Some new points must be examined here, which may never have occurred to anyone before. We certainly still need to ask ourselves why, after the number of observations increases, there is a greater probability of reaching the true ratio between the number of cases where a given event can occur and the number of cases in which it cannot, so that the probability ultimately exceeds all given degree of certainty…[11]

This notion of certainty or 'moral' impossibility opposed with mathematical impossibility was widely discussed throughout the eighteenth century and in the early nineteenth. It was then revisited more thoroughly by Cournot (1843), who introduced continuity in the measurement of probability. This enabled him to discuss the notions of physical or moral possibility and impossibility:

> *The physically impossible event is therefore the one whose mathematical probability is infinitely small*; and this single statement imparts substance—an objective and phenomenal value—to the theory of mathematical probability.

Let us take the example of a jar containing a single white ball and an infinity of black ones. The probability that a blind agent will extract the white ball is mathematically possible but in fact so small as to be physically impossible. However, the only way to demonstrate this physical impossibility by means of Bernoulli's theorem is to draw an infinity of balls from the jar.

Do we find a similar notion in social science? Again, we can refer to Cournot (1843), who tells us:

> The acts of living, intelligent and moral beings have no explanation, in the present state of our knowledge, and we can boldly proclaim that they can never be explained by the mechanics of geometricians.

[11] Ulterius aliquid hic contemplandum superest, quod nemini fortassis vel cogitando adhucdum incidit. Inquirendum nimirum restat, an aucto sic observationum numero ita continuò augeatur probabilitas assequendæ genuinæ rationis inter numeros casuum, quibus eventus aliquis contigere & quibus non contigere potest, ut probabilitas hæc tandem datum quemvis certitudinis gradum superet …

The notion of probability is therefore the only one applicable to social science, for this second notion of physically impossible event is perfectly suited to human acts. The two disciplines—probability theory and social science—set out to measure and quantify phenomena regarded as secrets of the gods before the seventeenth century: games of chance (such as dice, and cards) and games of life (births, diseases, deaths, migrations, and so on). The means used for these measurements and quantifications will, of course, form the basic theme of this book.

We now reach the twentieth century, in which the axiomatization of probability reached its broadest extension and in which the social sciences sought firmer foundations on which to address human affairs in rational terms.

After a series of more or less fruitful attempts to axiomatize probability (Laemmel (1904), Broggi (1907), Bernstein (1917), von Mises (1919), Slutsky (1922), Łomnicki (1923), Steinhaus (1923), Ulam (1932), Cantelli (1932), etc.), the work of Kolmogorov (1933) is now regarded by most probability theorists as the most consummate foundation for the science. We shall examine its basic principles in greater detail throughout this book, and point out the links between that axiomatization and the way in which we can interpret that formalization. Despite near-general acceptance of the axioms, controversies over the nature of this calculation and its possible interpretation persist in barely muted form. We shall therefore need to examine in greater detail how the different approaches view social science, in order to assess their validity in that field.

In social science, we are still a long way from axiomatization, and 'the transformation of the complex and changing life experience that constitutes the human fact into a scientific object—even in those of its aspects that are commonly recognized as public—remains problematic' (Granger 1994). We shall therefore need to examine in detail the multiplicity of viewpoints adopted on human facts over time in order to identify the operation that may enable us to reconstruct them in all their complexity. For this reconstruction, probability may prove essential.

Overview of Entire Volume

This volume will be structured as follows:

Part I From Probability to Social Science

Introduction to Part I

Depending on the historical period examined and the authors, the number of alternatives theories of probability is very variable and ultimately leads us to distinguish three broad types: objective probability, subjective probability, and logical probability. The last two categories can be grouped under the heading of epistemic probability.

Chapter 1 The Objectivist Approach

Classical probability theory relied from the outset—as early as Aristotle—on the notion of fairness. In 1654, Pascal referred to it for the purpose of defining a fair wager. But the notion, fully applicable to games of chance, did not hold up when transposed to the social sciences. These needed to assume that the now unknown probability of a demographic event—or, more generally, a social event—nevertheless existed, and remained the same throughout the period observed. This led to the notion of frequentist probability. The nineteenth-century debates over its validity showed that it cannot be applied to all feelings of uncertainty. The approach is suited to only a small number of social phenomena—particularly demographic ones.

The paradigm of objective probability had to reconcile the notions of equipossibility and physical impossibility. While the first was not specific to objective probability, the second proved indispensable, contrary to what later happened for epistemic probability. Objective probability is confined to events that can repeat themselves in identical conditions. Therefore, we cannot speak of the probability that a proposition, unique by nature, is true.

A proper search for axioms, however, did not become possible until after the establishment of set theory and axiomatics in the late nineteenth century. Setting aside many other attempts, we describe in greater detail two main types of axiomatization of probability, which were to formalize the two notions of paradigm. The first, introduced by von Mises in 1919, defined the notion of *collective* as the origin of probability. But many authors questioned the notion's consistency, undermining von Mises's axiomatics. In the end, it was the second type, introduced by Kolmogorov in 1933, that won the acceptance of most authors working on objective probability.

At this point, it is important to see how to apply objective probability to the statistics supplied by the physical and social sciences: this is known as the problem of statistical inference. The aim is to make the best use of the incomplete information available in order to move from data on a given phenomenon to the prediction of a similar phenomenon in the future. But, as the notion of 'an objective probability that a proposition is true' is meaningless, all we can estimate here is the probability of obtaining the observed sample if the hypothesis underlying the prediction is met.

We give some examples of applications of this approach to the social sciences. In developing political arithmetic, Graunt and Arbuthnott still used the notion clumsily. Another application concerns epidemiology, with the analysis of the effects of inoculation to prevent smallpox. Likewise, in sociology, Durkheim sought to identify social phenomena stripped of all extraneous elements by using the method of concomitant variations, i.e., a regression method.

This approach raises various problems. For example, while it allows a proper analysis of the outcomes of games with no cheating, it cannot determine whether a player is cheating or not. Similarly, the statistical inference made possible by objective probability is imperfectly suited to the study of decision-making. And it is suitable for analyzing only a small proportion of social phenomena.

Chapter 2 The Epistemic Approach: Subjectivist Interpretation

To apply probability calculus to the greatest possible number of feelings of uncertainty, however subjective they may be, we must abandon the notion of frequency—the foundation of objective probability—and hence the notion of physical impossibility. In 1713, Jacob Bernoulli envisaged what is now called a direct approach, which actually takes the probability of the studied event as a given. In 1763, Bayes solved the problem of the inverse approach, which assumes not only that the probability is unknown, but that its very existence is hypothetical. This leads to the notion of epistemic probability, which becomes fully subjective when one takes the view that it can be defined only for a specific individual, and not for an event as in the objective approach. As a result, the scope of application is substantially enlarged. For instance, we no longer need to assume the lack of cheating, for this probability is also defined in situations where players cheat, and the probability that a proposition is true now has a clear meaning.

The subjective-probability paradigm must rely on notions that differ from those underlying the objective approach. The notion of *coherence* in individual behavior must be reconciled with the notion of utility of winning for the individual. Coherence means that the reasoning of individuals must not contain any intrinsic contradiction, even as they are free to adopt any probability value that they prefer for an event. The notion of *utility*, introduced by Daniel Bernoulli in 1738, represents the subjective value of the stakes and will depend on each individual's condition. We can complete this paradigm by introducing the notion of *belief*, which is not probabilistic but allows the formalization of a psychological level outside the forecasting domain, and that of *plausibility* in order to reintroduce probability.

We must now apply a set of axioms to characterize the choice made by a rational individual faced with an uncertainty situation. Here as well, many axiomatizations have been proposed and we shall describe only the main ones. In 1931, de Finetti showed that a set of personal opinions, if it satisfied certain axioms, could be represented by a numerical measure. His axioms specified the notion of coherence. Savage completed them in 1954 by introducing the notion of utility, which arithmetizes the preference relationship between actions. Interestingly, the resulting quantitative probability satisfies Kolmogorov's axioms. Some criticisms of the axioms led to modifications introducing the notion of belief, which exists independently of the notion of probability examined in this volume. We shall therefore give only a brief presentation of it: Suppes in 1974 and Shafer in 1985 proposed axiomatizations incorporating two probabilities; Smets, in 1990, proposed an axiomatization that did not even include the concept of probability.

The objectivist approach offered only a partial solution to the problem of inference by twisting its meaning. By contrast, the subjectivist approach provided a perfectly clear answer. Using a *prior distribution*[12] and a data set, it allows an

[12] We need to distinguish here the term *prior*, which denotes any information beyond the immediate data and even used to express our ignorance, from the term *a priori*, which denotes a proposition, whose truth can be known independently of experience (Jeffreys 1939; Jaynes 2003).

estimation—under certain conditions—of a *posterior distribution* that predicts a future phenomenon. To ensure this outcome, the notion of *exchangeable* events, introduced by de Finetti, becomes indispensable.

We give examples of applications. The first concerns the combination of testimonies and is applicable in jurisprudence, artificial intelligence, and other areas. This problem has been addressed by many researchers over several centuries: the earliest solution used results found by Hooper in 1699; the latest uses Smets's theory of 1990. In our second example, the notion of exchangeability is applied to educational-science data for the purpose of drawing a correct statistical inference.

The approach is open to several criticisms. Psychological experiments have shown that, depending on how events are described, the subjective probabilities actually chosen by individuals do not necessarily meet the coherence principle. Although subjectivists reply that they study rational choices, the psychological problems posed by actual choices remain a fundamental issue. Moreover, an individual cannot always make choices transitively or even decide which choices to make: in such cases, his or her feelings of uncertainty cannot be represented by subjective probability. We also examine the criticisms of Savage's axiomatics by Allais in 1953 and show that the attempted modifications of his axioms cannot adequately explain all the phenomena connected to the choice paradox. The subjectivist approach seems too closely tied to individual psychology. Could a more logical yet still epistemic approach offer a means to avoid such criticisms?

Chapter 3 The Epistemic Approach: Logicist Interpretation

While a subjective probability is defined only for a given individual, a logical probability must be definable in the same manner for all individuals. For this, rather than start from the notion of personal odds for each individual, we must return to Pascal's notion of fair odds: when an individual wagers on a random event, fair odds yield a zero loss or zero expected gain. Yet fair odds will always reflect a degree of belief and are therefore applicable to all situations involving uncertain events, such as subjective probabilities.

The logical-probability paradigm introduced the logical notion of *consistency*, which specifies the required relationship between a proposition and the information available. Subjective probability depends on the individual. By contrast, logical probability, when obtainable in different ways, must yield the same result. It must also use all the information available for defining it. To this end, it incorporates the notion of *entropy* proposed by Shannon in 1948. Lastly, its focus is not on repetitive events, as in objective probability, or a single event, as in subjective probability, but on propositions made about events.

At this point it is useful to provide an axiomatics of the logic of propositions, introduced by Boole in 1854. It forms a basis for describing the main axiomatics of logical probability. The axiomatics proposed by Jeffreys in 1939 was initially rejected by most probabilists, philosophers, and statisticians of the time, but came to be recognized as highly innovative. However, without the notion of entropy, introduced later, Jeffreys was led to question the uniqueness of the choice of the prior probability. In 1961, Richard Cox showed that it was possible to derive the rules of probability from two axioms independent of the notion of set. One could thus use Kolmogorov's axioms, applying them now not to sets but to propositions. However, these axioms contain implicit conditions that van Horn later spelled out in order to make them more comprehensive. Similarly, while Cox effectively introduced the notion of entropy, it is Jaynes (2003) who showed more clearly how to use it to estimate a distribution of prior probabilities under different information scenarios.

For its application to social science, the epistemic approach concentrated on the incomplete information available on a phenomenon in order to draw inferences on the outcome of future experiments, using the consistency condition. It would no longer consider the personal probability that different individuals may choose, but those that they should choose on the basis of information shared by all. Statistical inference and probability would then form an inseparable whole.

We give examples of the use of logical probability in social science. The first example, from demography, is Laplace's application to the masculinity proportion at birth in 1781. The second is an application to legal science, which we illustrate with a wide-ranging review of results from 1785 to 2003, from Condorcet to Jaynes, via Laplace, Quetelet, Poisson, and others.

We conclude with a discussion of problems posed by this approach. The first is that impossibility and logical necessity are incompatible with the notion of zero probability for certain events when they can actually occur. But this criticism, which would be valid for an Aristotelian deductive logic, does not apply to a logic of plausible reasoning. The second problem is the difficulty, in certain cases, of finding a single prior distribution, although in many other cases we can deduce a non-informative distribution directly from the distribution of observations. This leads to a more general problem of dependency between the language used to pose a problem and the prior probability that can be deduced from it. We offer some solutions, but it is important to realize that the problem is inherent in all forms of epistemic probability, whether logical or subjective.

Conclusion to Part I

We begin by setting the three different approaches described in the preceding chapters in the context of the history of probability. The classical theory of probability that prevailed from the mid-seventeenth century to the first half of the nineteenth century was a unified theory in which the three aspects were closely linked: the probability of an event was simultaneously objective (considering its long-term frequency when it could be measured), subjective (considering the degree of our belief in its occurrence), and logical (considering the notion of fair odds). This type of probability was used in all fields, particularly the social sciences. In the first half of the nineteenth century, many criticisms led specialists to prefer the objective approach, which soon established its dominance for reasons that we discuss. In the

1930s, Kolmogorov's axiomatization of objective probability was swiftly followed by an in-depth examination of subjective and logical probabilities, although this did not result in their immediate adoption. They did not regain a stronger position until the second half of the twentieth century. However, they did not loosen the grip of objective probability—particularly in the social sciences, where it prevails to this day. We conclude with a methodological reflection on this revival of subjective and logical approaches, which leads us to examine if there is some cumulativity in probability.

Part II From Population Sciences to Probability

Introduction to Part II

We now examine the development of population sciences to show their methodological ties with probability throughout their history. While we cannot discuss all the social sciences—our work is not an encyclopedia—we show, when possible, that some methods used in this field are also suited to many other social sciences. We can thus extend the conclusions of these chapters beyond the specific field of population sciences.

Chapter 4 The Dispersion of Measures in Population Sciences

The aspect of probability that played a crucial role in the history of population sciences pertains to the *dispersion* of measures, either around their mean value, called *rate* (first sense of 'dispersion'), or as a function of other characteristics of the population studied (second sense). We devote particular attention to the use of statistical regression methods.

From the outset, Graunt's wager on the probability of dying is based on other hypotheses than Pascal's wager on the outcome of a game. Whereas Pascal can assume without too much difficulty that the odds are fair, it is far harder for Graunt to assume that the probability of dying is identical for all members of a population. Although the only information available to him was the number of observed deaths, he nevertheless chose that course in order to establish political arithmetic by positing an identical probability for all persons between ages 10 and 60. We show his errors, and how other researchers with access to fuller data improved his estimate by demonstrating that one should regard mortality as a function of age. Moving in the other direction, the introduction of the law of large numbers allowed Nicolas Bernoulli to refute Arbuthnott's argument on the distribution by sex of births in London from 1629 to 1710.

At the beginning of the nineteenth century, Laplace's application of the multiplier method, which allows a transition from observed births to the total population, supplied an estimate of the French population in 1782 within precise limits,