

Bruno de Finetti

Philosophical Lectures on Probability

*collected, edited, and annotated
by Alberto Mura*



Springer

PHILOSOPHICAL LECTURES ON PROBABILITY

SYNTHESE LIBRARY

STUDIES IN EPISTEMOLOGY,

LOGIC, METHODOLOGY, AND PHILOSOPHY OF SCIENCE

Editor-in-Chief:

VINCENT F. HENDRICKS, *Roskilde University, Roskilde, Denmark*
JOHN SYMONS, *University of Texas at El Paso, U.S.A.*

Honorary Editor:

JAAKKO HINTIKKA, *Boston University, U.S.A.*

Editors:

DIRK VAN DALEN, *University of Utrecht, The Netherlands*
THEO A.F. KUIPERS, *University of Groningen, The Netherlands*
TEDDY SEIDENFELD, *Carnegie Mellon University, U.S.A.*
PATRICK SUPPES, *Stanford University, California, U.S.A.*
JAN WOLEŃSKI, *Jagiellonian University, Kraków, Poland*

VOLUME 340

Bruno de Finetti

PHILOSOPHICAL LECTURES ON PROBABILITY

COLLECTED, EDITED,
AND ANNOTATED
BY

Alberto Mura
University of Sassari, Italy

With an Introductory Essay by
Maria Carla Galavotti

Translated by
Hykel Hosni

 Springer

Author

Bruno de Finetti[†]

Editor

Alberto Mura
Università degli Studi di Sassari
Dipto. di Teorie e ricerche dei sistemi
culturali
Piazza Conte di Moriana, 8
07100 Sassari
Italy
ammura@uniss.it

This translation was supported by the Alexander von Humboldt Foundation, the German Federal Ministry of Education and Research and the Program for the Investment in the Future (ZIP) of the German Government through a Sofja Kovalevskaja Award, as well as the Italian universities of Pisa and Sassari.

ISBN: 978-1-4020-8201-6

e-ISBN: 978-1-4020-8202-3

Library of Congress Control Number: 2007943065

© 2008 Springer Science+Business Media B.V.

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Contents

Preface	ix
Editor's Notice	xiii
De Finetti's Philosophy of Probability	xv
1 Introductory Lecture	1
Against the Axiomatic Approach	1
Subjectivism	3
Defining Probability	4
Proper Scoring Rules	4
2 Decisions and Proper Scoring Rules	15
Why Proper Scoring Rules Are Proper	15
Probability Depends on the Subject's State of Information	18
Sequential Decisions	19
Subjectivism versus Objectivism	21
For an Omniscient Being Probability Would Not Exist	23
3 Geometric Representation of Brier's Rule	27
Envelope Formed by Straight Lines	27
Operational Definition of Probability	28
4 Bayes' Theorem	31
Bayes' Theorem and Linearity	31
Statistics and Initial Probabilities	34
Bayesian Updating is Not a Corrective Revision	35
"Adhockereries"	36
Bayes' Theorem for Random Quantities	37
Inexpressible Evidence	39
5 Physical Probability and Complexity	47
"Perfect" Dice	47
The Lottery Paradox	49

Probability as Frequency	50
Probability and Physical Laws	51
Probabilistic Theories as Instruments	53
Random Sequences	54
6 Stochastic Independence and Random Sequences	59
Logical and Stochastic Independence	59
Propensities	60
Independence and Frequentism	61
Von Mises Collectives	62
7 Superstition and Frequentism	69
The Frequentist Fallacy	69
Idealized Frameworks	71
The Fallacy of Hypothesis Testing	72
8 Exchangeability	75
Urn Drawings with Replacement but Without Independence	75
Induction and “Unknown” Probabilities	78
Exchangeable Random Quantities	80
Alleged Objectivity and Convergence of Subjective Probabilities	81
9 Distributions	87
Introductory Concepts	87
Cumulative Distributions	89
Continuous Distributions Without Density	90
The General Case	93
Characteristic Functions	93
About Means	95
10 The Concept of Mean	97
Chisini’s Serendipity	97
Γ -Means and the Nagumo-Kolmogorov Theorem	100
Statistical Theory of Extremes and Associative Means	102
Inequalities Among Associative Means	104
Concluding Remarks	105
11 Induction and Sample Randomization	109
Exchangeability and Convergence to the Observed Frequency	109
Bayesian Statistics and Sample Randomization	110
12 Complete Additivity and Zero Probabilities	113
The Betting Framework and Its Limits	113
Finite and Countable Additivity	114
‘Strict’ Coherence	118

Conditioning on Events of Zero Probability	120
Allais' Paradox	123
13 The Definitions of Probability	127
Axiomatic, Classical, and Frequentistic Approaches	127
Indistinguishable Events and Equal Probability	130
Frequentism and Exchangeability	131
Von Mises' "Regellosigkeitsaxiom"	134
14 The Gambler's Fallacy	137
Against the Measure-Theoretic Approach	137
Gambler's Fallacy and Frequentist Fallacy	140
Events and Propositions	144
15 "Facts" and "Events"	149
A Pragmatic View of Events	149
On Elementary Facts	152
Events and "Phenomena"	155
16 "Facts" and "Events": An Example	159
A Sequence of Coin Tosses	159
A Graphical Representation	161
17 Prevision, Random Quantities, and Trievents	165
Probability as a Special Case of Prevision	165
The Conglomerative Property	167
Trievents	169
18 Désiré André's Argument	177
Heads and Tails: The Gambler's Ruin	177
The Wiener-Lévy Process	180
Again on Gambler's Ruin	180
The Ballot Problem	182
The Power of Désiré André's Argumentative Strategy	183
19 Characteristic Functions	185
Prevision and Linearity	185
Rotations in the Complex Plane	186
Some Important Characteristic Functions	188
Concluding Remarks	189
Cited Literature	191
Name Index	209

Preface

This book is a substantially revised and updated version of the volume *Filosofia della probabilità (Philosophy of Probability)* published in Italian in 1995, ten years after the death of Bruno de Finetti. It is not a work left in *written* form by de Finetti but comes from the transcriptions of the tapes recording a course for graduate students held by de Finetti in 1979 at the “Istituto nazionale di alta matematica” (National Institute for Advanced Mathematics) in Rome. To properly reflect the origin of the published material, a new title has been chosen for the present edition.

The lectures — which were held orally, without the support of any written text — were recorded with the aim of having them published. For that reason de Finetti and I had agreed on the recording of the lectures and his later editing of the resulting text. With this project in mind, along with the recording on tape, I took numerous notes and, above all, I noted carefully everything de Finetti wrote on the blackboard during the lectures.

Unfortunately, the original project could not be realized: other commitments forced me to put off the transcription of the tapes. Later, the poor health of de Finetti made this impossible. Only many years later, on the occasion of the tenth anniversary of the death of de Finetti did I have the idea of transcribing those old tapes and publishing them with comments and notes. Given the rather informal character of de Finetti’s lectures, the transcription required some editing in order to be published as a written text. This work was undertaken scrupulously in order to avoid altering de Finetti’s ideas.

Although originally it was aimed at an audience composed mainly of mathematicians, the course does not have, except to a small degree, the character of a mathematics course: instead it is rather a series of lectures in *philosophy of probability*, in which de Finetti deals in detail with almost all the themes of his subject.

This work therefore presents an exposition, also accessible to readers not trained in mathematics, of the ideas of the father of one of the most important philosophical and scientific schools of thought, which emerged in the past century and are becoming more and more popular in the present one: the neo-Bayesian school.

The intended spirit of the course is well illustrated in the “notice” which de Finetti himself wrote to be displayed on the Institute’s notice-board:

The course, with a deliberately generic title ["On Probability"] will deal with the conceptual and controversial questions on the subject of probability: questions which it is necessary to resolve, one way or another, so that the development of reasoning is not reduced to a mere formalistic game of mathematical expressions or to vacuous and simplistic pseudophilosophical statements or allegedly practical claims.

The notice made it very clear what the three targets of de Finetti's criticism were: (1) the formalism of those who see in the theory of probability a branch of an abstract systems of axioms; (2) the "vacuous and simplistic pseudophilosophical statements": an allusion to the various "definitions" of probability (such as frequency-limit, or "ratio between the number of favorable cases and the number of equally possible cases," and so forth) which he judged as superficial and without content; (3) "allegedly practical" behavior of those who consider as irrelevant the philosophical questions (or "conceptual" as de Finetti preferred to say) for the application of the calculus of probability in concrete situations.

In the first few days of the course there were about ten people attending. Most of them were graduate students of the ordinary courses of the Institute, but there were also other students. The number reduced with time. The presence of a limited audience meant that the course took on the character of an informal and interactive seminar. Usually, as the reader can realize, the seminars were a series of discussions about various aspects of the philosophy of probability.

The lectures contain numerous observations and comments, which are interesting in defining de Finetti's philosophy of probability. For example, note the interesting observations on inductive dynamics (Chapter 4) or physical probability (Chapter 5, the distinction — new and important in several respects — between *facts* and *events* (Chapters 15 and 16), the thesis according to which the meaning of a proposition can depend on empirical conditions (Chapter 17).

Although I added or modified numerous notes to the text with respect to the first Italian edition, the text from de Finetti's lectures has been translated into English without further editing.

When de Finetti ran the course I had known him for about a year. Around the end of 1977, Professor Francesco Barone (1923–2001), whose student I was, had the good idea of putting me in touch with de Finetti. When I went to visit him at his home in Rome, de Finetti welcomed me warmly and I had a conversation with him which lasted for more than five hours. That was the first of a long series of meetings, thanks to which I had the privilege of talking to him about every aspect of my research. At one of these meetings, which took place at the beginning of 1979, he told me he had accepted the invitation to hold the course presented in this volume. On that occasion he warmly asked me to take part, "also because — he said — the presence of a person already introduced to the philosophical problems of probability could be interesting."

I would like to take this occasion to express the debt I feel I have towards Bruno de Finetti. The impact that the meeting with him had on my education was decisive, his idea of mathematics as an instrument to be used in conceptual clarification and — above all — the concrete example with which he gave substance to this idea

constituted and continue to constitute the paradigm on which my scientific activity is based.

I would like to testify to his extraordinary human quality. I like to remember his ability in conversation, without ever tiring, even with a student — as I was at the time — as he would with a colleague, without a thought to the difference in competence, position, age, “importance” which there was between us. Rather than a manifestation of modesty there was a spontaneous manifestation of deep respect, which he lived profoundly, for every human being. I was also impressed by the extraordinary nobility and purity of his spirit, qualities which made him, in my eyes, so superior not only from the intellectual, but also from the moral point of view.

Special thanks go to Maria Carla Galavotti, not only for the new introductory essay which enriches this volume, but also for having encouraged and done everything possible in order that the book should see the light of day. I also owe particular gratitude to the translator Dr. Hykel Hosni for his numerous remarks, which were helpful in the final preparation of this edition. Moreover, I am indebted to Professor Nicholas H. Bingham who read a draft of the work and provided useful observations that I treasured in preparing the final version of the editor’s notes. I feel a special gratitude to Fulvia de Finetti, who kindly donated to me the copyrights on de Finetti’s recorded lectures. In accepting the donation I am pleased to commit myself to employing the royalties from this book to support studies on Bruno de Finetti’s scientific and philosophical thought. Finally, I am grateful to an anonymous referee who provided extensive comments and valuable suggestions.

Editor's Notice

The transcription of the text was undertaken respecting the needs of the written word and the conceptual content. Editing was used only in the case of excessive repetition or of comments not relevant to the course. Chapter and section titles were chosen by the editor according to the subject under discussion. The summaries were, however, reported with minimum modifications from the registers of the lectures compiled by de Finetti himself, and which were given (in copy) to the students of the course. Endnotes are by the editor. The editor has attempted to reproduce the various figures according to de Finetti as they were on blackboard and as they were copied into his note-book.

The five participants in the course, with their comments and questions have been named (after a suggestion by Marco Mondadori) 'Alpha,' 'Beta,' 'Gamma,' 'Delta,' and 'Epsilon' according to the order in which they intervened for the first time. Only for the first two has it been possible to verify their identity; the first was the editor and the second Anna Martellotti, nowadays a professor of Mathematical Analysis at the University of Perugia, who at that time was a graduate student from the Institute for Advanced Mathematics, where the course was held.

The editor has made available to those who wish to study them, the tapes on which the material published in this volume is based. A copy of them has been donated to the *Archives of Scientific Philosophy* at the Hillman Library of the University of Pittsburgh (USA), which already had other material by Bruno de Finetti (see Arden, 2004).

Part of the fourth lecture (Chapter 3) was lost because of the poor quality of the recording and the poor state of conservation of the tape. Three lectures of the course were not recorded at all and can thus be considered as lost. These were the third, fifth and nineteenth lectures. The relative summaries were, however, conserved in the register of the lectures: These are given below. The text in brackets is added by the editor.

LECTURE III (Thursday 15 March, 1979)

Applications of the above method [Brier's rule] to the case of three alternatives (instead of two, mention of the possible general case of whichever n).

Scheme of equilateral triangle; the vertexes represent the 3 possible cases (as for example, the three outcomes 1-X-2 in a game of football [where 1 denotes a home win, X a draw and 2 an away win]).

Information on prognostic gambling based on such concepts and their use in evaluating the probability to apply to facts of this kind.

LECTURE V (Wednesday 21 March, 1979)

The symbol '**P**' for 'probability' of an event; '**P**' extended (as "prevision," a better expression than 'speranza matematica' ['mathematical hope']); in both cases, **P** can also be interpreted as "price." Thanks to this interpretation (equivalent to that based on *proper scoring rules*) all the versions are unified and made intuitive.

Note: to speak of prevision $\mathbf{P}(E)$ of an event, or of a random quantity X is incorrect if we do not specify (or at least we must presuppose, not ignore) the circumstances H_0 , and write ' $\mathbf{P}(E \mid H_0)$.'

LECTURE XIX (Thursday 10 May, 1979)

Analysis and development of subjects from the previous lecture.

(Examination of when, after several successive "trials," it is possible to conclude that certain desiderata are satisfied or cannot be or the opposing hypotheses are both possible.)

De Finetti's Philosophy of Probability

One of the leading probabilists of the last Century, Bruno de Finetti not only made a fundamental contribution to statistical inference, but forged an original philosophy of probability that underpins his technical results and is the hallmark of his work. Throughout his life, de Finetti maintained that probability's wealth and depth of philosophical insight should not be disjoined from its mathematical aspects. His conviction that conceptual questions concerning probability have priority over both formal and practical issues inspired the pages of this book, which, as Alberto Mura clarifies in his Preface, originates from a series of lectures given by de Finetti in 1979. The following pages will first sketch de Finetti's life and personality, then outline the main traits of his philosophy of probability.

Bruno de Finetti: The Man

Bruno de Finetti was born in Innsbruck on June 13, 1906 and died in Rome on July 20, 1985. At the time Bruno was born, his father — an engineer — was director of the construction of the alpine electric railroad from Innsbruck to Fulpmes. Until the collapse of the Austro-Hungarian empire in 1918, the de Finetti family were Austrian subjects, but of "irredentist" tradition: his paternal grandmother Anna Radaelli was the niece of General Radaelli, commander of the defence of Venice against Austria in 1848–49. After the sudden death of his father in 1912, Bruno and his mother, who was pregnant with his sister, moved to Trento — his mother's home town — where he attended primary and secondary school. In 1923 he entered the Polytechnic in Milan, to continue the family tradition and train as an engineer. In his third year as a student, Bruno moved to the newly opened faculty of mathematics, from which he graduated in 1927, only twenty-one years old, already having some publications in scientific journals. In the autobiographical sketch "Probability and my life"¹ de Finetti writes that he encountered probability for the first time in his student years, through an article on Mendel's laws by the biologist Carlo Foà. After reading it, de Finetti wrote a paper on the subject and sent the manuscript to Foà, who was so impressed he showed it to various colleagues, including Corrado Gini, president of the newly created "ISTAT: Istituto Centrale di Statistica" (National

Census Bureau) and editor in chief of the journal *Metron*, which published the paper in 1926. Gini's appreciation of de Finetti was such that he offered him a position at ISTAT on graduation. De Finetti worked at ISTAT in Rome from the end of 1927 until 1931, when he accepted a post at the "Generali" insurance company in Trieste, the town of his father's family. In Trieste de Finetti was then appointed lecturer at the local university and subsequently became professor of financial mathematics. In 1954 he was nominated professor at the University of Rome and kept that position until he retired in 1976. By the time of his death, de Finetti was renowned for his contribution to probability and statistics and had received many honours, including a honorary degree in economics from the LUISS University in Rome in 1982. Decisive in spreading his ideas were de Finetti's lectures at the Institute Poincaré in Paris in 1935, later expanded in his famous article "La prévision, ses lois logiques, ses sources subjectives,"² and his subsequent encounter with Leonard Jimmie Savage, during a trip to the United States in 1950. The two entertained a fruitful exchange and it was largely through Savage's writings that de Finetti's ideas were brought to the attention of the English-speaking scientific community. Throughout his life de Finetti interacted with leading probabilists all over the world by taking part in international conferences and exchanging correspondence.³

Bruno de Finetti was a man of strong feelings. He was probably driven by the irredentist tradition of his family towards radicalism in politics. As a young man, he adhered to fascism, welcoming the nationalistic character of the movement, as well as its collectivistic tendency. He opposed the liberal idea that equilibrium can be obtained through individual profit and heralded collective economy as a way of achieving social justice. At the same time, he was a libertarian, who later in life favoured the introduction of divorce and abortion in Italy, struggled for peace and took part in the anti-militarist movement. De Finetti's participation in Italy's intellectual and political life is also attested by the numerous articles and letters on controversial issues he wrote for leading newspapers.

De Finetti had a natural inclination for mathematics. A letter to his mother, written in 1925 by the then 19-year-old Bruno, seeking her permission to quit engineering to study mathematics, gives an idea of his dedication to mathematics. There he writes: "mathematics . . . is always progressing, getting richer and sharper, a lively and vital creature, in full development, for these reasons I love it, I study it, and I wish to devote my entire life to it. I wish to understand what the professors teach and I want to teach it to others and I swear, swear that I will proceed further, because I know that I can, and I want to do it" (de Finetti, 2000, p. 733). He then describes with touching expressions his highly emotional state after attending a few lectures delivered at the newly opened faculty of mathematics: "Singlets, multiplets, covariant and countervariant tensors, variational calculus . . . I felt like the poor boy who, standing on tiptoe, discovers an enchanted garden beyond an unsurmountable fence. . . every formula, at least for someone who feels mathematics as I feel it, is a spark of a higher universe, which man conquers and absorbs with divine voluptuousness" (ibid., p. 734). Indeed, de Finetti's contribution to the field he so enthusiastically embraced is outstanding. From his first paper, written in 1926 at the

age of twenty, to his death at the age of seventy-nine, he published more than 290 works, including books, articles and reviews.

Always very outspoken and straightforward both in his professional and everyday life, de Finetti never hesitated to criticize the ideas he did not share. His vivid and incisive prose reflects his strong personality; his habit of inventing new expressions, as a rhetorical device to clarify his position has become proverbial. Some examples: “‘Press of deformation’ (instead of ‘press of information’) to highlight the peculiarity of most Italian journalists to “interpret” the facts as they like them instead of reporting as they are. ‘Bureau-phrenia’ to mean the gigantism and deformation of our national bureaucracy. ‘Idiot-crazy’ to mean the government of idiot politicians and managers” (Rossi, 2001, p. 3655). De Finetti offers a lucid motivation for such a habit by writing that “things said in a too colourless, painless, discrete, polite way, have no influence at all when they collide with preconceived ideas and prejudices. By contrast, a ‘strong’ word can be useful for destroying the mental defences of fossilized and adulterated ideas, which constitute the background of the so-called ‘established order’ (more properly called ‘established disorder’ by genuine right-thinking people). Thus, I think that creating mocking and funny words or sentences is a powerful, necessary and praiseworthy weapon for discrediting, fighting and pulling down the despicable conformity, which privileges maintenance of the worst raging institutions and ideas now in force” (quoted from Rossi, 2001, p. 3656).

Being the result of a transcription from a taped recording, this book forcefully depicts de Finetti's personality and conveys to the reader the flavour of his *vis polemica*, suggesting at the same time how deeply involved he was with the topic on his mind throughout his life: probability.

Probabilism

De Finetti is unanimously considered the founder of the subjective interpretation of probability, together with the English philosopher and mathematician Frank Plumpton Ramsey. Working independently at about the same time, both these authors suggested a view of probability as degree of belief obeying the sole requirement of coherence. To the definition of probability as degree of belief, de Finetti added the notion of exchangeability, which in combination with Bayes's rule gives rise to the inferential method at the core of neo-Bayesianism. The notion of probability as degree of belief developed by de Finetti between 1926 and 1928⁴ is the kernel of his philosophy of probability, centred on the idea that probability cannot but be conceived in a subjective fashion. A remark made right at the beginning of Chapter 5 of this book is noteworthy: there de Finetti claims he became convinced that probability is subjective in his student years, after reading Czuber's book *Wahrscheinlichkeitsrechnung*, which contains a short survey of the various notions of probability. This persuaded him that there cannot be an “objective” notion of probability, attributing to probability some sort of existence independent of those evaluating it, as if it belonged to “Plato's world of ideas” (p. 25).

De Finetti's philosophy of probability can be described as a blend of pragmatism, operationalism and a form of empiricism close to what we would today call "anti-realism."⁵ Such a philosophical position — labelled "radical probabilism" by Richard Jeffrey⁶ — portrays a conception of scientific knowledge as a product of human activity ruled by (subjective) probability, rather than truth. There is no doubt that the so-called Italian pragmatists, including Giovanni Vailati, Antonio Aliotta and Mario Calderoni exercised some influence on de Finetti's thought. In the article "Probabilismo," published in 1931 (de Finetti [1931] 1989), which he regarded as his philosophical manifesto, de Finetti mentions Mach's philosophy as a source of inspiration. This is reaffirmed in a letter to Bruno Leoni dated 1957, where de Finetti writes that his own attitude is close to the point of view of "scientific criticism (for instance Mach; 'operational definitions' in modern physics, 'behaviourism,' and so on), tracing back to the perspective of Hume, Berkeley, etc."⁷ He adds that, in this spirit, he considers it "meaningless to start with postulating an external reality and some laws governing it, instead of explaining the first in terms of a mathematical representation scheme useful to fix attention on part of our impressions . . . and the second as modes of thinking (based on our conscious and unconscious experiences together with the learning of 'science' forged by earlier generations) currently apt to forecast possible future impressions" (ibid.). The refusal of the idea that there are true, "immutable and necessary" laws deeply imprints de Finetti's output. In his words: "no science will permit us to say: this fact will come about, it will be thus and so because it follows from a certain law and that law is an absolute truth. Still less will it lead us to conclude sceptically: the absolute truth does not exist, and so this fact might or might not come about, it may go like this or in a totally different way, I know nothing about it. What we can say is this: *I foresee* that such a fact will come about and that it will happen in such and such a way because past experience and its scientific elaboration by human thought make this forecast seem reasonable to me" (de Finetti [1931] 1989, p. 170). Probability is what makes forecasts possible and the subjective notion of probability reflects in a natural way the fact that a forecast is necessarily referred to a subject, being the product of his experience and convictions.

Accordingly, probability "means degree of belief (as actually held by someone, on the ground of his whole knowledge, experience, information) regarding the truth of a *sentence*, or *event E* (a fully specified 'single' event or sentence, whose truth or falsity is, for whatever reason, unknown to the person)" (de Finetti, 1968b, p. 45). This notion, which he regards as the only non contradictory one, de Finetti believes will cover all uses of probability in science and everyday life. He therefore refused to attach an objective meaning to probability, and maintained that "probability does not exist" — a phrase that he wanted to be printed in capital letters in the *Preface* to the English edition of his *Theory of Probability*.⁸ This claim is rooted in his radical empiricism, which deems objective probability a metaphysical notion of which there is no need, and which is an obstacle to a correct understanding of probability.

A distinctive feature of the subjective approach developed by de Finetti is that it does not contemplate "correct," or "rational" probability assignments. In his words: "The subjective theory . . . does not contend that the opinions about probability are uniquely determined and justifiable. Probability does not correspond to

a self-proclaimed 'rational' belief but to the effective personal belief of anyone" (de Finetti 1951, p. 218). As a consequence of this position, it is admitted that two people, on the basis of the same information, make different probability evaluations. De Finetti's attitude in this connection differs sharply from the logicist interpretation of probability upheld among others by Rudolf Carnap and Harold Jeffreys, who believe in "correct" probability evaluations.⁹

Subjective probability needs an operational definition to gain both applicability and admissibility. Such a definition is usually given in terms of betting quotients, namely the degree of probability assigned to a given event by a given individual is to be identified with the betting quotient at which he would be ready to bet a certain sum on its occurrence. The individual in question should be thought of as one in a condition to bet whatever sum against any gambler whatsoever, free to choose the betting conditions, like someone holding the bank at a casino. Probability can be defined as the fair betting quotient he would attach to his bets. Coherence is a sufficient condition for the fairness of a betting system and behaviour conforming to coherence can be shown to satisfy the principles of probability calculus. These can be derived from the notion of coherence defined in terms of betting quotients, and all of its results are derivable from such a notion.¹⁰ Consequently, coherence becomes the fundamental and unique criterion ruling probability as degree of belief.

The operational definition of probability by means of fair betting quotients is by no means the only possible one. Already in 1931, de Finetti contemplated an alternative definition of probability, based on the qualitative relation "at least as probable as."¹¹ Later on, in the sixties and seventies, he turned to a definition of probability based on the so-called "penalty methods," like scoring rules of the kind of the well known "Brier's rule." Scoring rules are based on the idea that the device in question should oblige those who make probability evaluations to be as accurate as they can, and, if they have to compete with others, to be honest. This is the option taken by de Finetti in this book, where the issue is discussed at length in Chapters 2 and 3.

Having opted for the subjective approach to probability, de Finetti vigorously opposes the other interpretations of probability. Already in the introductory lecture de Finetti declares his aversion to the axiomatic approach on the one hand, and to the classical and frequentist interpretations of probability on the other. He regards the classical interpretation, based on the notion of "equally probable events" as rooted in a misunderstanding, because the very notions of equiprobability and independence are subjective, being the product of judgment, to be made by those who evaluate probability.

Things are even worse with frequentism, which according to de Finetti originates from the contradiction of assuming a link between probability and frequency, which should instead be taken as a matter for demonstration. "From a frequentist point of view, the frequency converges to the probability by definition. However there is no guarantee at all that this will happen. At the bottom of frequentism lies the mystification of making probabilistic laws certain" (p. 63). The idea that frequentism is based on superstition imprints the title of Chapter 7: "Superstition and frequentism," but Chapter 6 also contains a sharp criticism of the basic notions of frequentism, such

as those characterizing von Mises' sequences, namely randomness and infiniteness. De Finetti deems the idea of an infinite sequence of experiments "meaningless," because "even supposing that the universe will last a billion times longer than we currently think it is going to last, it is nonetheless of finite duration" (p. 56). Similarly, the notion of an irregular sequence is deemed a "pseudonotion," a conceptual outgrowth originating "from superstitions like the one according to which the numbers which have failed to come up in past lotto draws have a higher probability of coming out in the next draw, or similar cabbalistic beliefs" (p. 65).

De Finetti also opposes Reichenbach's idea that updating probability judgments is a self-correcting procedure. For him updating one's mind in view of new evidence does not mean changing opinion, but updating one's opinion in the face of new evidence. The updating process is made possible by the use of Bayes' theorem, discussed in Chapter 4, which for de Finetti represents not just the best, but indeed the only acceptable way of addressing probabilistic inference. De Finetti's Bayesianism should be construed in connection with the notion of exchangeability. A probabilistic property weaker than independence, exchangeability makes probability evaluation depend on the number of observed positive and negative instances, regardless of the order in which they have been experienced. Used in connection with Bayes' theorem, exchangeability allows subjective probability judgments to be updated by taking into account observed frequencies. In fact, the method devised by de Finetti bridges the gap between degrees of belief and observed frequencies and explains "why expected future frequencies should be guessed according to past observed frequencies" (de Finetti 1970, p. 143).

Therefore, after criticizing the frequentist claim to define probability in terms of frequency, de Finetti ascribes frequencies a pivotal role in the evaluation of probabilities. Their role is reflected by the representation theorem, which shows how information about frequencies interacts with subjective elements in making probability assignments. According to de Finetti, a clear-cut distinction between the definition and the evaluation of probability is needed and it is a decisive advantage of the subjective viewpoint to keep them separate. Having clarified that probability is subjective and should be defined in terms of degrees of belief, subjectivists admit that all sorts of information can be useful in order to evaluate probability: namely frequencies, but also symmetries, when applicable. Unlike subjectivists, that do not confuse the definition and the evaluation of probability, upholders of the other interpretations do; they embrace a unique criterion — be it frequency or symmetry — and use it as grounds for both the definition and the evaluation of probability. They take a "rigid" attitude towards probability, which implies a univocal choice of the probability function. On the contrary, the subjectivist approach is "elastic;" it does not commit the choice of one particular function to a single rule or method, but regards all coherent functions as admissible.

For de Finetti, the evaluation of probability is the result of a complex procedure involving both subjective and objective elements. Indeed, for him the explicit recognition of the role played by subjective elements within the formation of probability judgments is a prerequisite for the appraisal of objective elements. He always warned that the objective component of probability judgments, namely factual

evidence, is in many respects context-dependent: evidence must be collected carefully and skilfully, its exploitation depending on the judgment on what elements are relevant to the problem under consideration and enter into the evaluation of probabilities. Furthermore, the collection and exploitation of evidence depend on economic considerations that vary according to the context. For all of these reasons, the exploitation of factual evidence involves subjective elements of various kinds. Equally subjective, according to de Finetti, is the decision on how to let belief be influenced by objective elements like frequencies through the exchangeability assumption. As suggested by the preceding considerations, the issue of "objectivity," rephrased as the problem of how to obtain good probability evaluations, was taken very seriously by de Finetti, who addressed it in a number of writings, partly fruit of his collaboration with Savage.¹²

As a consequence of his overarching subjectivism and pragmatism, according to which science is a continuation of everyday life and subjective probability suits all situations where probability evaluations occur, de Finetti did not feel the need to ascribe a special meaning to the use of probability made within science. His refusal of "objective probability" goes hand in hand with his lack of consideration for such notions as "chance" and "physical probability." Interestingly, Chapter 5 of this book contains a few remarks to the effect that probability distributions belonging to statistical mechanics could be taken as more solid grounds for subjective opinions. This suggests that late in life de Finetti must have entertained the idea that a somewhat robust meaning can be attached to those probability assignments deriving from accepted scientific theories. However, such evaluations depend "on the theorizing that we make of phenomena, which is subject to change" (p. 52). As to the interpretation of theories, de Finetti takes a pragmatist and even instrumentalist attitude: "If one takes science seriously, then one always considers it also as an instrument. Otherwise what would it amount to? Building up houses of cards, empty of any application whatsoever!" (p. 53). These hints seem to point to a form of subjectivism apt to accommodate notions like "chance" and "probability in physics."

A step in that direction was taken by Ramsey, who developed it in conjunction with a pragmatist approach to theories and truth.¹³ To put it briefly, for Ramsey the distinctive character of the probabilities encountered in physics lies with the fact that they are derived from physical theories and their objective character descends from the objectivity ascribed to theories that are commonly taken to hold.

In a similar vein, the logicist Harold Jeffreys claimed that within an epistemic conception of probability there can be room for the notion of "physical probability," endowed with a peculiar meaning deriving from scientific theories, whose objectivity is also made to rest ultimately in our experience, from which it is obtained through the application of statistical methodology. Put briefly, Jeffreys' idea is that the notion of "objectivity" is produced by inferential procedures rooted in experience and can play a useful role in connection with some scientific hypotheses when the data yield a probability which is so high that one can draw inferences, whose probability is practically the same as if the hypotheses in question were true.¹⁴

There is a widely felt need to incorporate into subjectivism a notion of probability endowed with some kind of robustness, in view of its application within "hard"

sciences, like physics. While de Finetti's writings do not contemplate any notion of physical probability on a par with those entertained by Ramsey and Jeffreys, the claims contained in Chapter 5 of this book are evidence that his view is nonetheless compatible.

Maria Carla Galavotti

Notes

1. See de Finetti (1982a).
2. See de Finetti ([1937] 1980).
3. The bulk of such correspondence together with a large number of drafts, notes and documents of various nature form the "De Finetti Collection," belonging to the "Archives of Scientific Philosophy" at the Hillman Library of the University of Pittsburgh.
4. See de Finetti ([1976] 1980, p. 195).
5. See Galavotti (1989).
6. See Jeffrey (1992a,c).
7. See document BD 10-13-27 of the "Bruno de Finetti Collection", quoted by permission of the University of Pittsburgh. All rights reserved.
8. See de Finetti ([1970] 1990a); de Finetti ([1970] 1990b).
9. On the differences between Carnap's logicism and de Finetti's subjectivism see Galavotti (2003b, 2005).
10. The crucial link between coherence and the fundamental properties of probability was grasped by Ramsey but was actually proved by de Finetti.
11. See de Finetti ([1931] 1992).
12. See de Finetti and Savage (1962).
13. On Ramsey's notion of chance see Galavotti (1995, 1999). For a comparison between Ramsey's and de Finetti's philosophy of probability see Galavotti (1991).
14. For a detailed account of Jeffreys' position see Galavotti (2003a).

Chapter 1

Introductory Lecture*

The various conflicting conceptions of probability; preliminary objections against the so-called objectivist definitions; attitudes towards uncertainty, probability evaluations and criteria for obtaining sincere indications of the latter ("proper scoring rules"[†] — in particular Brier's rule and its mechanical interpretation: minimum moment of inertia if considered with respect to the center of gravity).

The present series of lectures takes place at a time when I find myself particularly weighed down by work. As a consequence, I would have preferred to postpone its beginning for some time. However, from another point of view, it is perhaps more useful to start it right now. Given that in fact I am engaged in the development of a conceptual rather than mathematical kind of work, a commitment like the present one, if time and energy consuming, might turn out to be of some help.

As to the course, I plan to conduct it by making use of some materials from the work I am currently writing.¹ This will also give me an opportunity to discuss such a work in the present occasion. However, I would also like to touch upon other themes in the perspective of a unitary treatment of probability, a treatment which, again, I would like to develop from a conceptual rather than a mathematical point of view. Indeed, in my own opinion, mathematics is interesting as an instrument for condensing concepts which, by their own right, without mathematics, would still be what they are.

Against the Axiomatic Approach

And I have to say that I am increasingly adverse to any sort of axiomatic foundation — especially in the field of probability. The reason for such an aversion is aptly illustrated by a famous phrase which — I believe — is due to Russell (I say "I believe" because I find myself always in doubt about the paternity of the sentences engraved on my memory). Russell (or whoever for him) used to say that an axiomatic treatment — and in general, in his opinion, all of the

* Lecture I (Tuesday 13 March, 1979).

[†] (Translator's note:) In English in the original text.

mathematics — “is such that we do not know what we are talking about nor do we know whether what we are saying is true or false.”² According to the axiomatic point of view, some axioms for probability are introduced (for example through certain linear equations over a certain field), without one being able to figure out, unless one knows it already, what “probability” means. One could, for instance, interpret probability as a measure and believe that the words “measure” and “probability” are synonymous.

On this subject I wrote, though that was an occasion on which I gave vent to a momentary feeling of anger, on “*Tuttoscuola*” de Finetti (1978b). I do not know if you are familiar with it: it is an illustrated journal for high-school, edited by Vinciguerra, who is a mathematics teacher in a secondary school. I wrote a letter to the editor of the periodical under the title *Rischi di una “matematica” di base “assiomatica”* where both “*matematica*” and “*assiomatica*” are spelt with a double “t”³

Unless more meaningful content, which however cannot follow from the axioms, is added to them, as they abstractly talk about relations among undefined objects. This holds in particular in the field of probability which, despite being partly abstract, is also, in a certain sense, concrete. It is concrete for we all utter sentences like: “It was more or less probable that the colleague I was expecting would come.”⁴ I would not really know how to translate the latter sentence in set-theoretic terms. Perhaps appeal could be made to the set of all “possible worlds,” so as to distinguish those worlds in which the colleague does come from those worlds in which the colleague does not. In this way, the measure of all the worlds of the former kind could be considered as the probability value. In my own opinion, this way of looking at probability contributes only to make things absurdly complicated.

I said at the beginning that in this period, I am very busy with various projects and the fact that I have to lecture the present course as well worries me considerably. The main commitment to which I have to devote much of my time is the drafting of the entry *Probabilità* for the *Einaudi* encyclopedia. The editor asked me for a hundred pages, many more than usually required for the other entries in the same encyclopedia. I have made numerous attempts to commence that work, yet it is only now that I have turned to a path which I hope will prove to be good enough to deserve pursuing. It is a similar kind of path that I will try to follow in the present course.

From that notice,⁵ where as a title I have written “On probability” (generically: neither “theory” nor “calculus” nor other things) it was already implicit that the points I shall insist upon are, for instance, the way in which the concept of probability must be thought of; the reason why probability can be translated into calculus;

* (Translator’s note:) De Finetti plays here with the consonance of three Italian words: “*matematica*” (“mathematics”), “*assiomatica*” (“axiomatic”) and “*matto*” (“mad” “insane”). Doubling a “t” in “*matematica*” and “*assiomatica*” allows de Finetti to ridicule the axiomatic foundations of mathematics by dubbing it as a foolish — “*mattica*” — activity (though no such adjective exists in Italian). Hence, the “risks” discussed in de Finetti’s letter are those arising from endorsing a “Mathemadics” based on “Axiomadics” to suggest that it is something foolish, that is to say, airy-fairy and utterly useless.

why a piece of evidence is such that an event has a certain probability; and so on. We will begin with general and introductory notions and then gradually build up the specific notions that are to be applied to the various cases. In this way, eventually, the meaning of probability should emerge clearly.

Subjectivism

In my own view — as many will already know, if they are familiar with this subject — probability has only a *subjective* meaning. That is, I think it is senseless to ask what probability an event has per se, abstractly. By an “event,” on the other hand, I mean a single well-defined fact. It is an event, for instance, the fact that today the express train coming from Milan arrived with a delay of between thirty and thirty-five minutes. And if I am talking about the probability of the latter event, I refer to the individual fact mentioned and not, for instance, to the “probability of delays in general on all the lines” or “on each day.”⁶ Hence, by “event” I mean a circumstance such that it is always possible to know whether it is verified or not. Of course, there could be situations in which this distinction between the two possible cases does not hold because a necessary condition for it to make sense has failed.⁷ Moreover, it must be borne in mind that whenever we say that an event is something that surely turns out to be either true or false, we are making a limiting assertion, as it is not always possible to tell the two cases apart so sharply. Suppose, for instance, that a certain baby was born exactly at midnight on 31 December 1978 and we had to say which was the year of his birth. How can it be decided in which year the baby was born, if his birth started in 1978 and ended in 1979? It will have to be decided by convention. These are the things that need to be stressed in order to avoid building up the entire framework of probability calculus on top of distinctions that are supposed to be determined and unquestionable. There is always a margin of approximation, which we can either take into account if we say “1978” or “1979.”

If one adopts the subjectivistic point of view, none of the traditional definitions whereby an objective meaning is allegedly attached to probability judgments can be accepted. At this point one naturally wonders: how can we talk about probability if we reject every proposal that has been made in order to define it? This is a legitimate question. Before giving you the answer, however, let me state in advance that the point of view I maintain is based on the thesis that *it is senseless to speak of the probability of an event unless we do so in relation to the body of knowledge possessed by a given person*. And if one really wants to talk about an objective probability, then one should say that every event has probability 0, if it is not verified and probability 1, if it is. Therefore, since we do not know whether the objective probability is going to be 0 or 1, we estimate it according to whether we are inclined towards favouring one or the other alternative. This is the most radical way to remove all those conundrums, those rather weird ways of asking if one probability evaluation is more or less true than another. “Everyone holds true that probability which she assigns on the basis of all she knows.” This sentence

is due — though I’m not entirely sure about this — to Borel⁸ (surely, as far as I remember, it was a Frenchman. I say “surely” despite being only *almost* sure, not being in a position to warrant it: given that I struggle quite a lot to recall the substantive things, I might well mix up the accidental ones — this however, with no intention to diminish anyone’s importance).

Defining Probability

Let us go back now to our question: how can probability be defined? How can we attach a meaning to the word “probability” if neither of the two most common ways of defining it can be accepted? One of them consists in the so-called *classical definition*, according to which the probability of an event E is the ratio of the cases in which E is verified to the number of all equally probable cases. This is an attempt to define probability through a tautologically true statement. Indeed, should the same probability be assigned to n cases and should it be known that an event E is verified whenever one of certain m subcases is verified and is not verified when one of the remaining $n - m$ subcases is verified, then, almost by definition (I say “almost by definition” because the definition which I regard as such is another, more neatly subjectivistic one and does not mention equally probable events), the probability would be m/n . This definition is circular, for it is based on the expression “equally probable” for which an independent definition is not provided. It must be said, however, that under the premise of equiprobability, the distinction between favourable and unfavourable cases is plausible.

The other definition mentions a frequency in a set of “repeated trials.” This one is a lot more vague as it is not clear at all what “repeated trials” are. Strictly speaking, repeated trials do not exist. That amounts in fact to a linguistic abuse, which could be explained in terms of the need for matching ordinary parlance. Although I do not want to adopt the typical aprioristic point of view of those who use this definition, I observe that it is based on the idea that some *identical* events exist which, as such, should also be equally probable. However, identical events do not exist: at most there are events with respect to which we might fail, naturally, to have preferences as a consequence of the equality of the conditions that are known to us. Yet there could be so many other conditions, which might be unknown to us and with respect to which those events might differ widely.

Proper Scoring Rules

Therefore, these two ways of defining probability (as a ratio of the number of favourable and unfavourable cases and as a frequency) are airy-fairy, unless one states beforehand what “equally probable” means. On the other hand, there is a way, and I believe it is the only one, which allows us to say exactly what we mean

by “probability.” Such is the method of *proper scoring rules*,* which consists in asking a person (let us call her “A”) what is the probability she assigns to an event E , A being warned that she will receive a score depending on the answer she will provide and on the value of the “objective probability” of E (in the sense specified above: 0% if E is false, 100% if E is true⁹).

Let us suppose that A answered by saying, e.g., “40%.” Had there been no penalization, A could have said whatever she liked, without her being subject to any negative consequence. In the presence of a proper scoring rule instead, A would be driven by her own interest — which consists in minimizing the prevision¹⁰ of the penalization — to give the exact answer.¹¹ How does it turn out to be the case? Let us consider the answer “40%.” Should A say so, what would the score be? It would be — according to the simplest scoring rule — exactly the square of the difference between the probability stated and the objective one. Obviously, there are two cases: either E is verified or E is not verified. In the former case, the objective probability is 100% and since we have supposed that she said 40%, the score would be $(40\% - 100\%)^2 = (-60\%)^2$. Instead, should E be disproved, the difference would be $(40\% - 0\%)^2 = (40\%)^2$. If A, on thinking that the probability value is 40%, said 35%, or 45% (or whatever other value), the prevision of penalization would have been greater. Why? Although the little calculation can be done effortlessly, I prefer, given that I am not fond at all of ready-made[†] algebraic formulae, to explain this through elementary and fundamental mechanical concepts.¹² It is in fact enough to think of two masses p and $100\% - p$ hanging at the end-points of a uniformly dense bar: $(100\% - p)^2$ would then be the moment of inertia, which is always minimum at the barycentre (Fig. 1.1).

Therefore, under this scoring rule, were a person willing to express her probability evaluations so as to minimize, from her own point of view, the prevision of the penalization, then she should state the probability value that she is actually assigning. And should she say p' instead of p , then the expected penalization would increase by $(p' - p)^2$. Similarly, in mechanics, if a body — suppose, for the sake of simplicity, it is again a bar — is rotated about a point p' distinct from the centre of mass p , what would then be the moment of inertia? It would be the central moment

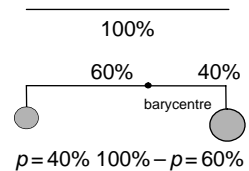


Fig. 1.1 Representation of probability as a barycentre

* (Translator's note:) In English in the original text.

[†] (Translator's note:) De Finetti refers to such formulae as *formulette*, literally “little formulae.” The diminutive emphasises here, as it will in the next chapter with respect to certain “little rules” — *regolette* — the lack of fertility of some cut and dried formulations of a mathematical problem and in particular, the meaning of probability.

(that is, relative to the centre of mass, which is the minimum) plus the moment of inertia of the whole mass relative to p' , were the latter concentrated at the centre of mass p .

Such a scoring rule is called *Brier's rule*.[†] The latter is used, for instance, by American meteorologists to estimate the probability of rainfall. And in the morning — so I am told — those estimates are published in newspapers and broadcasted by radio and television stations so that everybody can act accordingly.¹³ To say that the probability of rain is 60% is neither to say that “it will rain” nor that “it will not rain”: however, in this way, everyone is in a better position to act than they would be had meteorologists sharply answered either “yes” or “no.”

Indeed, by giving similar answers, one would pretend to be a diviner, when in fact nobody can seriously foresee (as if he were a prophet) that “it will rain” or that “it will not rain.” Of course, meteorologists refrain from giving such sharp responses and say instead, given all they know about the situation (which now, thanks to the charts, can be somewhat directly ascertained by anyone who enjoys having a glance at it) that in a certain situation, it is either not much, or much, or very much probable that it will rain. And yet, when a meteorologist uses expressions like “not much,” “much” or “very much,” the listener might interpret them in different ways: for instance, he may intend “much” either as “90%” or as “60%.” This is why I believe that this is not a serious way of expressing oneself. On the other hand, were a numerical evaluation given, one could indicate unambiguously what is the degree, between yes and no, of such a belief. The day after, of course, everyone would know whether it did or did not rain, yet since nobody can claim to be an infallible prophet, it would be silly to ask meteorologists to behave insincerely, pretending to be infallible, and to tell us whether “it will rain” or “it will not rain.” Such a sharp prediction would be utterly useless. For this reason I believe that the only sensible way to tackle the problem of prevision is to resort to the numerical appraisal of probability through the method of proper scoring rules.

ALPHA: Why is the adjective “proper” being used for scoring rules? And who introduced such rules?

DE FINETTI: I believe that Mitio Nagumo¹⁴ and the American McCarthy¹⁵ (he is not the politician McCarthy, of course) were the first to introduce them. Afterwards, many others have put forward the same proposal. Some years later, though ignoring the fact that someone had preceded us, we did here at the University of Rome, precisely at the Faculty of Economics, a forecast competition on the football championship¹⁶, where a probability evaluation had to be given to the results 1, X, 2 of every match.* In such a situation, somebody who ignored the teams' strength on

[†] (Translator's note:) In English in the original text.

* (Translator's note:) De Finetti's experiment closely mirrored the so-called *Totocalcio*, a popular gambling game in Italy, which is based on the outcome of football matches, mostly from *Serie A*. In a nutshell, the game consists in indicating, for each of the thirteen matches listed on a lotto-like card, whether the home team will win, draw or lose by marking 1, X or 2, respectively. There is an important difference between *Totocalcio* and de Finetti's competition: whereas in the former

the football pitch might choose the values 50%, 30%, 20%, respectively. Such a choice would be warranted by the fact that, statistically, 50% of the matches end in a home win, 30% in a draw and 20% in an away win. Yet those who play the *totocalcio* are hardly ever so incompetent as to ground their choice on this evidence only: they will also take into consideration whether the home team is among the league's strongest ones, or *vice versa*. As a consequence, those values are just *mean* values. Should a person base herself on them only, she would neglect other relevant information. It is exactly at this point that the difference between the concept of subjective probability and the alleged concept of statistical probability emerges. Let us make it clear: I do not intend to say that the concept of frequency is to be rejected. What I want to stress is that it must not be confused with the concept of probability. Since frequencies in certain cases are 90%, 9% and 1%, respectively, and in other cases are 20%, 30% and 50% (depending on whether the home team is much stronger or much weaker), a distinction should be made between the concept of *prevision of frequency* and that of *probability*: a distinction which the theories based on the statistical definition (taking the probability of the single case for a mean) usually fail to mark. It cannot be said that that mean is *the* probability, for if we averaged out the probabilities of winning on each match, we would obtain a value that is empty of practical significance with respect to the single case. Of course, it can rightly be said that it is a mean. But were it taken as *probability*, then it would be useful only if one had an obligation to give the same probability to every match (but then betting would make no sense whatsoever).

I still possess the records of the competitions that we have carried out. We could have a look at them tomorrow, so we can make some calculations and discuss some actual data: the data of an experiment conducted with Brier's rule. You will have a chance to see, with the hope that this will help in clarifying the matter, how the calculations of the penalizations were worked out; the evolution of the rankings as the competition moved on; and so on.

ALPHA: Is the use of the square, in the penalization, due to the fact that in this way one obtains the above mechanical analogy or (if one likes) representation? I have the impression that if one took the simple difference in place of the quadratic difference, one would get the same results.

DE FINETTI: There is certainly an analogy between mechanics and the technique of these penalizations. However, the use of the quadratic difference in place of the simple difference is essentially due to another reason. This is the requirement that the scoring rules be *proper*, that is, that they be such that the maximum benefit is obtained by stating one's own probability evaluation sincerely. This would not happen if we took the simple difference instead of the quadratic difference for the calculation of the penalization. In the case of proper rules, supposing that we could speak of objective probabilities and that they were 50%, 30%, 20% (I'm saying this against my opinion, although this might help to understand this case better),

gamblers can make multiple choices for any match (although this obviously raises the cost of the bet), no such option was available to those who took part in the competition set up by de Finetti.

then those are exactly the values that it would be advantageous for us to state, as the penalization would turn out to be smaller than it would be if the evaluations were instead 90%, 1%, 20% or 20%, 30%, 50%. The point here is not to make a comparison with “reality,” but rather to compare one’s own sincere evaluations with those one indicates. If one wanted to make an accurate prediction, one should either indicate 100%, 0%, 0% or 0%, 100%, 0% or 0%, 0%, 100%. Out of three people who chose to play respectively in those three ways, one would surely win. However such a strategy would not be ‘globally’ advantageous. It would be more convenient, instead, to indicate the evaluations exactly, without distorting them.

Tomorrow, I shall discuss another example to show that Brier’s rule is not just a game. The case of the game is useful as it is the one everybody grasps (and then, should it go wrong, never mind: it is just a game). In order to discuss the application of this method I will avail myself of a volume on oilfield search (written by a certain Grayson): *Drilling Decisions by Gas Operators* (1960). Among the data taken into account in the search for oilfields in a given area, the opinion of the experts is always taken into account in order to evaluate the potential profitability of drilling the terrain. But the answers provided by the experts, when formulated in words, tend to be rather elusive. Typically experts give evaluations like: “It is almost certain, but not sure; indeed it might not be the case.” and so on. They say and unsay in the one breath, not to risk too much. The result is that it is never quite clear what their answers mean. The expert tries to speak in such a way as to secure himself that, whatever happens, one cannot tell him that he was wrong. On the contrary, if questions are asked in such a way as to obtain a probability value as an answer, the ambiguity disappears. Hence, if one asks the experts whether there are signs of a salt dome (I am not an expert on the subject of oilfields but I know that a salt dome is a finding that suggests that the presence of an oilfield is likely) or other information about the pattern of the layers of earth crust, instead of the usual ambiguous answers (like, “it might be, but I do not know if that’s the case” and so on) they must provide answers like: “in the light all the evidence I possess, I estimate it to be 35%.” In such a case, it is possible to carry out the actual calculation of the cost, taking into account that there is a 35% probability that we will eventually find what we are looking for. Subjective probability is therefore an aid to give a reliable measure of what cannot be measured objectively. Unlike guessing whether there is oil (or a dome or something else) or not, we can do the probabilistic calculation of the expected cost and of the expected gain (if we are successful), to decide whether it is worth drilling or not. Otherwise, an analogous calculation might be done to decide whether it is not appropriate — to be surer — to perform a seismic prospection first (for instance, by setting off a mine to create some sort of artificial earthquake, so that useful information could be extracted from the way in which seismic waves propagate). It is clear, however, that all those pieces of information mirror the opinion of the expert (geologist, engineer and so forth) in a better way when they are expressed in the form of probability values rather than by words, which have a much more ambiguous meaning. Such words do not give a value between 0% and 100%. All they say is that it is ‘more 0% than 100%’ or ‘more 100% than 0%’.

Tomorrow, we shall insist on those issues through more concrete examples. Let me sum up, to conclude, the content of the lecture. I have asserted that the only serious device to talk about probability is the one based on proper scoring rules. As to probability, it is nothing but a state of mind, more precisely it is the degree of belief that an individual holds concerning the occurrence of a certain event. And proper scoring rules are the only adequate instrument by means of which that degree of belief can be measured. The procedure to obtain such a measure is as follows. A person X is asked to indicate her own probability evaluations. Such a person is warned that she will receive a score depending on the numbers she has stated. Scores, on the other hand, are devised *ad hoc* so that it is advantageous for X to indicate exactly those numbers which correspond to her own degrees of belief, as this would minimize the prevision of the penalization.

As to the other ways to define probability, I have said that those which are formulated in terms of frequency or proportion might, in some sense, find an application as well. However, they can only be applied as special cases of this general framework: otherwise one would inevitably end up talking of equal probability without knowing what probability is. The use of the word “probability” — even though certain formal constraints are met — is not enough: we have to make it meaningful. And, as far as I know, the only meaning that is independent of vicious circles is the subjective one.

Well, you have 23 hours to raise, boost or alleviate doubts and to finally express them tomorrow.

Editor's Notes

1. Such a work is the entry “*Probabilità*” (“Probability”) for the “*Enciclopedia Einaudi*” (1980).
2. Cf. Russell ([1901] 1957, p. 71): “We start, in pure mathematics, from certain rules of inference, by which we can infer that if one proposition is true, then so is some other proposition. These rules of inference constitute the major part of the principles of formal logic. We then take any hypothesis that seems amusing, and deduce its consequences. If our hypothesis is about anything, and not about some one or more particular things, then our deductions constitute mathematics. Thus *mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true*. People who have been puzzled by the beginnings of mathematics will, I hope, find comfort in this definition, and will probably agree that it is accurate.” (emphasis added).

It should be stressed that the formalistic account of pure mathematics that emerges from this passage is confined to represent Russell's thought about mathematics at the moment in which it was written (1900–1901), influenced by Peano's views that strongly impressed him at the *Congress of Philosophy* in Paris: “The most important year in my intellectual life was the year 1900, and the most important event in this year was my visit to the International Congress of Philosophy in Paris. . . . In Paris in 1900, I was impressed by the fact that, in all discussions, Peano and his pupils had a precision that was not possessed by others. I therefore asked him to give me his works, which he did. As soon as I had mastered his notation, I saw that it extended the region of mathematical precision backwards towards regions which had been given over to philosophical vagueness. Basing myself on him, I invented a notation for relations.” (Russell, 1944, p. 12).

Among the pupils of Peano there was in Paris (appointed to the permanent international commission) Giovanni Vailati (see note 17 on page 58), who had been his assistant at the

university of Turin between 1892 and 1895. As pointed out by Maria Carla Galavotti in the Introduction to this book (p. xviii), Vailati (who combined Peano's formalism with a pragmatist view of science and mathematics inspired by Peirce) had some influence on de Finetti — who, in any case, was acquainted with his work (cf. Chapter 10 page 100) and mentioned him often (see also Parrini, 2004, p. 49). Now, Vailati wrote a paper commenting on and defending Russell's definition of mathematics (Vailati [1904] 1987). It is not unlikely that de Finetti had notice about Russell's definition through Vailati's paper.

On the other hand, it should be noted that Russell, under the influence of Frege's work, later developed a logicist philosophy that departs from his early purely axiomatic viewpoint, and criticized Peano's approach to arithmetics: "It might be suggested that, instead of setting up '0' and 'number' and 'successor' as terms of which we know the meaning although we cannot define them, we might let them stand for *any* three terms that verify Peano's five axioms. . . . If we adopt this plan, our theorems will not be proved concerning an ascertained set of terms called 'the natural numbers,' but concerning all sets of terms having certain properties. Such a procedure is not fallacious; indeed for certain purposes it represents a valuable generalization. But from two points of view it fails to give an adequate basis for arithmetic. In the first place, it does not enable us to know whether there are any sets of terms verifying Peano's axioms; it does not even give the faintest suggestion of any way of discovering whether there are such sets. In the second place . . . we want our numbers to be such as can be used for counting common objects, and this requires that our numbers should have a *definite* meaning, not merely that they should have certain formal properties. This definite meaning is defined by the logical theory of arithmetic." (Russell [1919] 2002, pp. 9–10).

3. Indeed, the letter was published apart from the letters to the editor as a two-column article (whose sub-head contained: *A letter by Bruno de Finetti*). For the reader's convenience I copy it below:

Dear Vinciguerra,

I would be grateful to you if you could publish in "*Tuttoscuola*" some thoughts to which I have been frequently driven and whose clarification and publication appeared to me, as a consequence of a recent experience, more necessary than ever. I have had, in fact, the opportunity of comparing various people's (colleagues as well as young scholars) positions in regard to the role played by axiomatics as a basis to "define" the notions which constitute the object of mathematical investigation and as a basis to fix, formally, the ways of operating on them.

Long ago, when the initial statements of a theory used to be called "postulates" and constituted, accordingly, a somewhat idealized expression of circumstances which truly existed in "reality" (or were supposed to), the "mistake" (so to say) consisted in regarding such empirical properties as being logically necessary, universally sound in whatever conception, even in the abstract mathematical one. Space, for instance, even as a "mathematical object," couldn't be other than Euclidean: to doubt it was to expose oneself to the "shrieks of the Boeotians" (which even Gauss, the great "princeps mathematicorum," dreaded to such an extent that he relinquished to disclose his own reflections about non-Euclidean geometries).

Nowadays, on the contrary, since we have started talking of "axioms" instead of "postulates," that risk has vanished but we are now left, as a counterbalance, with its opposite which, in my own view, is much more substantive and dangerous. By definition, axioms are arbitrary sentences which are freely chosen as initial "truths," with the only obvious proviso that they must not be contradictory. This means that every conclusion that can be drawn from them must be considered sound, true. Sound . . . , but in what sense? In the sense of the abstract initial environment. Yet, it is often pretended (or it is presumed that we are allowed to concede) that such conclusions must be considered sound also in those concrete ambits in which, at first sight, it would appear "convenient" to adopt some abstract model, like a high fashion "prêt-à-porter" suit.

So has been done, for instance, by two English analysts who (as many, very many others) have assumed "complete additivity" among the "axioms" of probability theory, that is to say

the extension of the “additive property” to the case of infinitely (needless to say: countably) many incompatible events.

There is nothing really new in this; on the contrary, such an idea enjoys the greatest favour at the moment.

What is peculiar in their approach, however, is the fact that, as a justification for their choice, they explicitly point to “mathematical convenience”! [translator's note: in English in the original text] for without it “many results could not be obtained.” And on this basis they conclude that “The definition is therefore justified ultimately by the elegance and usefulness of the theory which results from it” [translator's note: In English in the original text]. (That is to say: the elegance of the resulting theory justifies the definition . . . independently of whether it is essentially sound or not in the actual reasoning about probability!)

A simple counterexample shows how complete additivity cannot be presumed to be always “sound.” If the “possible outcomes” are the rational points in the $(0, 1)$ interval, and the probability density is uniform (the total probability of the rational points of a subinterval of length p is p), the probability “1” of the sure event appears (so to say) to be the sum of infinitely many zero probabilities relative to the single rational points.

Besides this example (its persuasiveness would appear to be stronger or weaker according to whether one is familiar or not with probability and the various perspectives on it) I am anxious to point out how distorted is the pretension to extend a definition or a procedure relative to a concrete field of application to wider domains, without paying due attention to see whether the resulting concrete conclusions turn out to be reasonable or absurd, whether they satisfy common sense or violate common sense in their applications to reality. It almost looks as if any qualms against such a bad habit is considered to be an act of sabotage!

I took the liberty to double the “t” in “*matematica*” and “*assiomatica*” in the title as an admonition to similar trends and I suggested that the corresponding syllable should be put in bold face to urge that, throughout a bad and excessive deployment of axiomatics, mathematics is not turned into sheer madness (“*roba da matti*”).

Actually, the title appears entirely in bold face in the journal.

4. Presumably this is a reference to an episode that took place a few minutes earlier. Before the beginning of the lecture, de Finetti informed the attendants that a colleague of his expressed an intention to participate to the lecture, although he added that he was not sure whether this man had come. The colleague did not attend the lecture which, as a consequence of expecting him, began after some minutes of delay.
5. It is the notice posted up on the notice board of the *Istituto nazionale di alta matematica* (Translator's note: National Institute for Advanced Mathematics) where the course was held. The original text is reproduced in the preface to the present volume.
6. Allusion to the frequentist view according to which probability is always relative to some appropriate “reference class.”
7. For instance, were it known that the express train coming from Milan did not leave, it would make no sense to ask whether it arrived on time or not. On various occasions as well as in Chapter 17 pp. 169–171, de Finetti recommended resorting to a three-valued logic in order deal with situations like this. See also Chapter 4 note 2 on page 41.
8. “Mr Keynes thinks that important progress is made in observing that the notion of probability should be attached to a judgment, to a proposition, rather than to a fact or to an event. If I were to choose a card in a game of 52 cards, it is incorrect language to speak of the probability that it is a king of diamonds. One can legitimately speak of the probability of the judgment that I enunciate in affirming that it is a king of diamonds. *This probability is relative to my judgment and to the information that I possess about the experience which has just occurred; it would not be the same necessarily for another observer who possessed different information.*” (Borel [1924] 1964, p. 49, emphasis added).
9. According to de Finetti the percentage $x\%$ can be identified with the function $f(x) = x/100$. Hence, 50% and $1/2$ are equivalent for de Finetti. The advantage of this way of intending percentage lies in the fact that it allows us to express probabilities as percentages (by saying,

more easily, 30%, 50%, and so forth, instead of 0.3, 0.5, and so forth) without this resulting in any alteration of the scale: probability is still included in the $[0, 1]$ interval, which would coincide with the $[0\%, 100\%]$ one.

10. By the term ‘prevision,’ de Finetti means — as he himself will say later on in the second chapter — what in the Italian tradition is called ‘speranza matematica’ and in the English language one ‘expectation’.
11. To put the matter in a more general setting, let F be a person (called “the forecaster”) who is asked to tell her subjective probabilities about certain events E_1, \dots, E_n (considered as random quantities whose possible values are 0 and 1) that form a *partition* (that is such that one and only one of them will obtain, so that $\sum E_i = 1$) gives the answer q_1, \dots, q_n . Of course, the forecaster may not be honest if her real subjective probabilities p_1, \dots, p_n of the E_1, \dots, E_n differ from the respective declared numbers q_1, \dots, q_n . A *proper scoring rule* is a random quantity $f(q_1, \dots, q_n, E_1, \dots, E_n)$ such that

$$\mathbf{P}(f(q_1, \dots, q_n, E_1, \dots, E_n)) \leq \mathbf{P}(f(p_1, \dots, p_n, E_1, \dots, E_n)).$$

A proper scoring rule is said to be a *strictly proper scoring rule* if

$$\mathbf{P}(f(q_1, \dots, q_n, E_1, \dots, E_n)) = \mathbf{P}(f(p_1, \dots, p_n, E_1, \dots, E_n))$$

only if $p_1 = q_1, \dots, p_n = q_n$. Throughout the book de Finetti refers only to strictly proper scoring rules, even if he omits the adverb ‘strictly.’

The history of proper scoring rules begins in 1950 with an article by the meteorologist Glenn Wilson Brier (1950), who introduced the so-called Brier’s rule to be applied to meteorological forecasts. It is obtained by putting:

$$f(q_1, \dots, q_n, E_1, \dots, E_n) = k \sum (E_i - q_i)^2$$

A similar rule (the mean square error rule) was also independently proposed by Rudolf Carnap (1952) as a tool for measuring the success of his “inductive methods.” Carnap’s ideas have been later fruitfully developed by Roberto Festa (1993). Carnap’s rule turns out not to be proper in the sense defined above (Bröcker and Smith, 2006), but it would be proper if the class of admissible probability evaluations were so restricted that only those evaluations that satisfy some symmetry constraints (actually required by Carnap) were allowed. This suggests a generalization of the notion of proper (and strictly proper) scoring rule. A scoring rule may be called proper (respectively, strictly proper) *relatively to a constraint \mathbf{K}* , if it assumed that the numbers p_1, \dots, p_n and q_1, \dots, q_n satisfy \mathbf{K} . This generalization may be useful when the fulfillment of an admissibility criterion (e.g., exchangeability) is required for the probability evaluations. Of course, every [strictly] proper scoring rule is [strictly] proper relative to any constraint.

Independently of Brier and Carnap, Irving John Good ([1952] 1983) found another strictly proper scoring rule, commonly labelled as “logarithmic”, which is obtained by putting:

$$f(q_1, \dots, q_n, E_1, \dots, E_n) = k \sum E_i \log(q_i)$$

The computer scientist John McCarthy (see note 15), citing Good’s paper, found that the class of strictly proper scoring rules is determined by the class of strictly convex functions of q_1, \dots, q_n (McCarthy, 1956). More exactly, it is the class of those functions $f(q_1, \dots, q_n, E_1, \dots, E_n)$ whose prevision is an $n - 1$ hyperplane supporting a strictly convex function g at q_1, \dots, q_n . In the binary case, such a hyperplane is a straight line and with respect to the Brier’s rule, this property is illustrated by Fig. 3.1.

Independently of these works, proper scoring rules and McCarthy's theorem were rediscovered by de Finetti and Savage, who began to write a paper under joint authorship in the spring of 1960. This paper was completed by Savage and published posthumously in 1971 without de Finetti's authorship "though it owes so much to him at every stage, including the final draft" (Savage, 1971, p. 783).

In the present book, de Finetti presents the proper scoring rules mainly as a device to elicit personal probabilities so that both coherent and "sincere" assessments are encouraged. This allows him to replace the schema of bets to provide a new operational definition of subjective probability free of the difficulties afflicting the former (cf. 4 note 12 at page 43) one. For a detailed exposition of such difficulties, see also Mura, (1995b).

However, proper scoring rules are also, as de Finetti pointed out in the sixties, a means to assess the "goodness" of probability evaluators (de Finetti, 1962). Indeed, if in a large number of probability evaluations the score obtained by an assessor A tends to be to an important degree better than the score of a second assessor B on the same set of events, it seems reasonable to infer that B is a better assessor than A . To speak about "good" assessments beyond coherence is *prima facie* embarrassing from the subjectivistic viewpoint. In fact, if such a goodness is taken as *objective*, it seems to challenge the idea of subjective probability. However, this is not the case because, according to de Finetti, (1) the "goodness" is not relative to a supposed objective probability but to truth-values of single events, whose objectivity de Finetti never questioned and (2) the judgment of goodness is itself genuinely *subjective* in character, being nothing but our subjective prevision of the future performance of the appraisers. Of course, such a prevision may be based on past performances and standard Bayesian statistical methods may apply to this situation.

Two problems arise in this context. The first is to find the extent to which the "goodness" measured by different proper scoring rules are comparatively equivalent. In other terms: is it possible to define an order relation such that given two assessors A and B , all the proper scoring rules agree in attributing to A a better score than to B (or viceversa)? The answer to this question is partially positive. A total order in terms of goodness, common to all the proper scoring rules, does not exist, so that to some extent the "goodness" depends on the chosen rule. By contrast, several *partial* orders (defining some property as a kind of "goodness *ceteris paribus*") have been defined. Coherence and sincerity are just two such properties.

However, there are other virtues of good assessors beyond coherence and sincerity that all proper scoring rules to some extent reward. Among them two appear to be especially important and have been intensively studied: *calibration* and *sharpness*. From a Bayesian viewpoint we must use the probability values provided by the forecaster F via Bayes' theorem, so calculating our subjective conditional probability q , given that the subjective probability of F is p . This kind of calculation has, however, little theoretical interest unless we make special assumptions. The simplest one is to assume that we *rely* on the assessments made by F so that the probability that we assign to an event E given that F assigns to E the probability p is p . This appears to be a frequent situation (albeit to some extent idealized), especially if we suppose that F is an expert asked in order to form our subjective probability values in a field in which we have no expertise. Now, under such an assumption, a desirable property of F 's assessments is that for every $p \in [0, 1]$ the expected frequency (according to our evaluations) by which the events to which F assigns probability p is p . In this case we consider F *calibrated*. Calibration of F has a clear intuitive appeal if we have to rely on F 's assessments. The intuition is confirmed by analysis. As proved by A. Philip Dawid (1982a), every coherent assessor must consider herself as calibrated, so that we cannot completely rely on the assessments of another person unless we consider her calibrated. On the other hand, in the light of this result, it follows from the very definition of a strictly proper scoring rule that, *ceteris paribus*, the expected score of a calibrated forecaster is better than the score of a non-calibrated one.

Consider now forecasters F and G that we take to be perfectly calibrated. A partial order among calibrated forecasters that is rewarded by all strictly proper scoring rules is provided

by the notion of *refinement*, introduced by Morris H. DeGroot and Stephen E. Fienberg (1982, 1983, 1986). F is said to be at least as refined as G if the distribution of probability values D_G provided by G may be obtained from the distribution D_F provided by F by means of a stochastic transformation so that D_F stochastically dominates D_G at the second degree (DeGroot and Eriksson, 1985). Roughly speaking, in this case D_F is “sharper” than D_G . DeGroot and Fienberg proved that for every strictly proper scoring rule, a more refined forecaster receives a better score. Refinement under calibration may be considered as an *explicatum* for expertise. Unfortunately, however, refinement only induces a *partial* ordering, so that when two appraisers cannot be ordered by refinement, then the order in terms of the respective expected scores may depend on the adopted scoring rule.

The problem thus arises of choosing the best among the strictly proper scoring rules. A definitive answer to this question seems to be meaningless, because the optimal scoring rule ultimately depends on the preferences of the client of the forecaster and on the context. In spite of this, Reinhard Selten (1998) has provided a set of desiderata that appear to be reasonable in most situations and characterize Brier’s rule up to positive linear transformations. This result vindicates de Finetti’s and Savage’s preference for this simple rule.

On the topic of proper scoring rules an enormous amount of literature, either theoretically or experimentally oriented, has grown in several sciences (economics, psychology, statistics, meteorology and so forth). I recommend in addition to the papers cited above, for their relevance, the following works:

Kadane (1982); Lindley (1982); Dawid (1982b); Blattenberger and Lad (1985); Seidenfeld (1985); Schervish (1989); Seillier-Moisewitsch and Dawid (1993); Winkler (1994); Cervera (1996); Lad (1996); Clemen (2002); Cooke and Goossens (2004); Dawid (2004, 2007); Manski (2004); Gneiting, Balabdaoui, and Raftery (2005); Bröcker and Smith (2006); Krämer (2006); Gneiting and Raftery (2007).

12. The algebraic calculation can be found in Chapter 2.
13. In more recent times, with the rise of the so-called *artificial intelligence*, certain programs have been developed (called *expert systems*), which are capable of providing probabilistic weather forecasts. Such forecasts are subjective, for they mirror the habits of the human experts that the program tries to emulate. Many meteorological services, available on the press or on the Internet, provide automatically generated probability for various events like rain, storms and so on.
14. The reference to Nagumo is actually mistaken. Nagumo should rather be credited — together with Andrei Kolmogorov, who obtained the same result independently — with the theorem of associative means (hence called the Nagumo-Kolmogorov Theorem), which de Finetti will discuss extensively in Chapter 10.
See note 11 for details about the history of proper scoring rules.
15. John McCarthy (1927–), is a well-known American computer scientist who received the Turing award in 1971 and introduced the term “artificial intelligence.” To him is also due the invention and design of the LISP recursive programming language.
16. This experiment is so described in detail in de Finetti ([1970] 1990a, p. 195):
“The participants have to hand in, each week, previsions for the forthcoming matches, giving, for each match, the probabilities (expressed in percentages) of the three possible outcomes (in the order: win, draw, defeat); writing, for instance, 50-30-20, 82-13-05, 32-36-32, etc. Given the results, one can evaluate, game by game, the losses and the total losses for the day (and, possibly, a prize for the day), as well as the cumulative sum needed for the final classification. This final classification must be seen as the primary objective. If there are prizes, the largest should be reserved for the final placements, and, in order to conform to the spirit of the competition, the prizes must complement losses; i.e., they should depend on them in a linearly decreasing fashion. Some of the results of the competition actually organized by de Finetti are published in de Finetti (1982b, p. 6).”

Chapter 2

Decisions and Proper Scoring Rules*

Examples of important applications of proper probability evaluations, from the simple case of weather forecasts to the important and complex ones. As a typical example, the chain of provisions concerning the expected costs and benefits, and hence, the potential appropriateness of drilling an oil well or not — according to the volume Decisions under uncertainty. Drilling Decisions by Oil and Gas Operators by C. J. Grayson Jr (1960).

Since it makes no sense, in a context like this, to lecture as one would lecture to schoolchildren, participation should be as spontaneous as possible on both sides. I would like to hear from you if there are ideas which you have found obscure or if you are unclear about something. I will try — as far as possible — to alleviate your doubts, even though, clearly, they can vanish completely only with the progressive development of the course. It is, in fact, normal and inevitable that there are things which at the beginning of a course one understands only approximately, or which one learns without being able to see how do they relate to one another. In spite of this, your reactions are very important to me, since I can gather from them suggestions about how to proceed with the course. I hope that you have clarifications to ask. Indeed, those who do not have doubts, usually, rather than having it all clear, have swallowed the lecture as if it were castor oil — without relishing it at all.

Why Proper Scoring Rules Are Proper

BETA: I have to admit, I have grasped proper scoring rules very vaguely. I have understood that there is a certain penalization, but I have not been able to understand why is it appropriate to indicate the probability which one deems more reasonable. Why is it inappropriate to “cheat”?

DE FINETTI: It is not “inappropriate”; this has nothing to do with cheating. One is allowed to indicate whatever number she likes.¹

* Lecture II (Wednesday 14 March, 1979).

Suppose that a certain person X , under a proper scoring rule, assigns probability p and $1 - p$ to an event E and to its negation \bar{E} , respectively. Suppose further that X believes (for some reasons) that it is more appropriate for her to indicate, instead of p , the number π (which obviously has nothing to do with the constant π). Since she is subject to penalization, X will suffer a loss in any case; her best interest is to minimize the prevision of it.

Let us now ask what is the prevision of such a penalization. And let us ask whether such a prevision would turn out to be smaller or greater by taking π to be distinct from p . In other words, let us ask ourselves whether it is appropriate or not for X to indicate a value other than the value X feels is the one corresponding to her degree of belief. Should the event E turn out to be true, its value will be 1, whereas should the event E turn out to be false, its value will be 0. Therefore, E is a variable that can take either value 1 or 0 depending on whether E turns out to be true or false. The actual penalization depends, according to the formula $(E - \pi)^2$, both on the value of E (as it will turn out to be) and on the value of π stated by the person. However, in order to compute the *Prevision* of the penalization (and hence to compute the relative appropriateness) the value p (which corresponds to X 's actual degree of belief in the occurrence of E) must be used as the probability of E .

Our problem is to see X , should she “try to be smart” by indicating a value π other than p , whether she is really smart or not. Let us do some calculations. If X indicates π , the prevision of penalization is:

$$\begin{aligned}\pi^2(1 - p) + (1 - \pi)^2 p &= \pi^2 - \pi^2 p + (1 + \pi^2 - 2\pi)p \\ &= \pi^2 - \pi^2 p + p + \pi^2 p - 2\pi p \\ &= \pi^2 + p - 2\pi p.\end{aligned}$$

Now, had the subject said “ p ,” it would have held $\pi = p$ and hence, substituting in the last line π by p , the prevision of penalization would be:

$$p^2 + p - 2p^2 = p - p^2.$$

The difference between the penalization that one would face by saying “ π ” and the one she would face by saying “ p ” is:

$$\begin{aligned}(\pi^2 + p - 2\pi p) - (p - p^2) &= \pi^2 + p - 2\pi p - p + p^2 \\ &= \pi^2 + p^2 - 2\pi p \\ &= (\pi - p)^2.\end{aligned}$$

Therefore by saying “ π ” the penalization would increase by $(\pi - p)^2$ if compared to the case in which X said “ p .”

ALPHA: More precisely: the prevision of penalization increases, for the actual penalization does not depend on p being either π^2 or $(1 - \pi)^2$.

DE FINETTI: Yes, of course.

ALPHA: Prevision of additional penalization.

DE FINETTI: Yes. It is an additional penalization relative to the penalization which she would have incurred had X have said p . This result can be extended to any random quantity whatsoever. Even in the general case, in fact, the person who is subject to a scoring rule will have an interest in stating her prevision sincerely, as the penalization that she would consequently undergo would be smaller than the one she would undergo had she indicated a different value. More precisely, once again, it is not the real penalization that is increasing but the prevision of the penalization given the probability evaluations that the person elaborates at the moment she decides what to state.

I am asking you now something more substantive than arithmetical calculations: what is the aim of such penalizations? Why have I called into question this rule to define probability, instead of the cut and dried rules* like “the ratio of favourable cases to the possible ones, if equally probable” or “the frequency in the long run,” and so forth? What distinguishes this definition from the others?

BETA: By defining probability through such rules[†] we avoid entering a vicious circle.

DE FINETTI: That is to say?

BETA: Both the classical and the frequentist definition start from the idea of n equiprobable events. Since they are based on the notion of equal probability, they define probability in terms of itself. On the other hand, no circularity is involved in defining probability through proper scoring rules.

DE FINETTI: Undoubtedly: this formulation does not resort to the idea of equally probable events, or even equal under all respects. But what else does it take into account? Could probability be equal for every individual (even leaving out of consideration the fact that there might be individual differences according to which someone is a pessimist and someone else is an optimist)?

BETA: It seems a subjectivistic definition to me.

DE FINETTI: Probability, thus defined, mirrors a state of mind. But besides being subjective in that it pertains to a person, it also depends implicitly and evidently (evident things are those on which we have to insist harder, for they are so evident they become neglected) on . . . ? What does it depend on?

ALPHA: First of all, there is an objective probability measurement in terms of an individual's actual behaviour. That is: subjective probability can be measured without resorting to introspection.

DE FINETTI: In which way?

ALPHA: By observing one's actual behaviour in situations of possible gain or loss.

DE FINETTI: Under this respect, the operational procedure based on scoring rules is slightly idealized, as it does not take fully into account the distortions arising from the fact that many people like to risk for risk's sake. Gamblers, for instance (if they are more than just dreamers, in that they overestimate their possibility of

* (Translator's note:) See footnote at page 5.

[†] (Translator's note:) Beta is clearly referring to proper scoring rules here.

winning — especially if, as might well be the case, they happened to win the first time they played), play because they take pleasure in risking and they are often driven to make choices which they themselves deem unprofitable from the economic point of view. The behaviour of such players need not necessarily be incoherent. Yet if applied to those individuals, scoring rules would not provide the correct answer, for those individuals do not have the minimization of the prevision of penalization as their only interest.

Probability Depends on the Subject's State of Information

BETA: Moreover, the probability assessment depends on the knowledge possessed by the gambler.

DE FINETTI: This is the essential thing! When writing " $P(E)$ " we are using an incomplete notation, as we fail to mention the state of information of the individual who is making the evaluation of probability. By the expression "state of information" I mean *all of* the previous experience, everything the person has seen, heard, read and so forth. What we take into account is a person's behaviour on the basis of her probability evaluations and our aim is to judge whether that person is coherent or not.

Even if an individual were self-coherent, I might still find his opinions absurd and I might say: "Is he crazy?" Likewise he might say that I am crazy because my evaluations differ from his own. Of course, a third person might have reasons, on the grounds of what happened subsequently, to judge either my evaluations or the other person's ones more reasonable. But this is that sort of "hindsight" of which we are told "ditches are full." * I have never seen ditches full of hindsight, yet there is probably a rough kernel of truth in this sentence.

Hence, the state of knowledge of the individual who is making the prevision should always be taken into account. I introduced the notation (which seems handy to me, although there's nothing special about it) " H_0 " to mention generically, without making any further assumptions, the state of information. We usually write " $P(E | H)$ " to denote the probability of E conditional on H , where H is an event whose truth value is unknown to us. The probability is conditional on H in the sense that, if H is not verified, the gamble is called off (i.e., the sum that has been possibly paid is returned to the gambler). But instead of writing " $P(E | H)$ " it would be more correct to write " $P(E | HH_0)$." Actually, it would be more natural to prefix the " H_0 " to the " H ," but this notation is typographically preferable: otherwise "0" might look like an operator between the two " H "s. Of course, one could keep writing " $P(E | H)$," without the " H_0 ," as it is not very practical to write " H_0 " every time. However, the presence of " H_0 " must always be understood.


*(Translator's note:) De Finetti paraphrases here the proverbial expression *del senno di poi son piene le fosse*.

It is commonly said that an event E has a certain probability $P(E)$, meaning by the term ‘event’ every event of a certain type (for instance, obtaining Heads by tossing a coin) rather than a single well-defined fact. This habit is not erroneous in itself: after all, it is just a matter of terminology. Nonetheless, it has the disadvantage of suggesting that one actually wants to refer to the “trials of the same event,” in the sense intended by the objectivists. Probability, on the other hand, always concerns single events, even under the hypothesis that the trial is equally probable. Each event is what it is individually. And whenever probability is to be judged, such a probability should be thought of as conditional on a particular state of information.

Strictly speaking, the latter should be made explicit in full: yet this is practically impossible, for it would mean condensing into one proposition the entire experience that the bearer of the probability judgment has accumulated since birth, or even since he was in his mother’s womb.

Sequential Decisions

In any case, the probability of E is conditional on an acquired state of knowledge and possibly on an additional condition H . And it must always be intended as the probability of a single, well-defined, fact. I wanted to bring in here the book in which the problem of searching for oil fields is discussed (Grayson Jr., 1960). Unfortunately, I did not manage to find it. It develops the topic in a way I believe to be extremely appropriate to illustrate these concepts. The reason why the case of searching for oil fields provides an interesting example is that in order to ascertain whether it is profitable or not to drill a well (be it an exploratory well or an exploitation well or something else), many analyses need to be done. However, I can still report verbally to you how the argument is developed there, something which is in my opinion very useful.

Figure 2.1 shows a diagram representing the various steps of the decision process involved in deciding whether oil drilling should be performed or not. The diagram has – like all decision diagrams – a tree-like structure. This is because the decision-making process can be represented as a succession of steps. Such steps can be of two types: ‘decision nodes’ (denoted by *empty squares*) and ‘risk nodes’ (denoted by *empty circles*). The branches that originate from a decision node represent various alternative decisions. These are called ‘decision branches.’ On the other hand, the branches that originate from a risk node (called ‘risk branches’), represent a partition of ‘relevant circumstances.’ These are relevant in that the circumstances of the subsequent decisions depend on which circumstance of the partition is present at that stage of the process. The sign ‘’ indicates the branch that would be chosen in case the process passed through the node from which that branch originated. The *bold-face circles* indicate the terminal nodes (at which the process finishes). The numbers along the risk branches indicate the cost of the corresponding decisions. The percentage figure along the risk branches indicates the probability that the process continued along that branch in case the process passed through the node from which that branch originated.

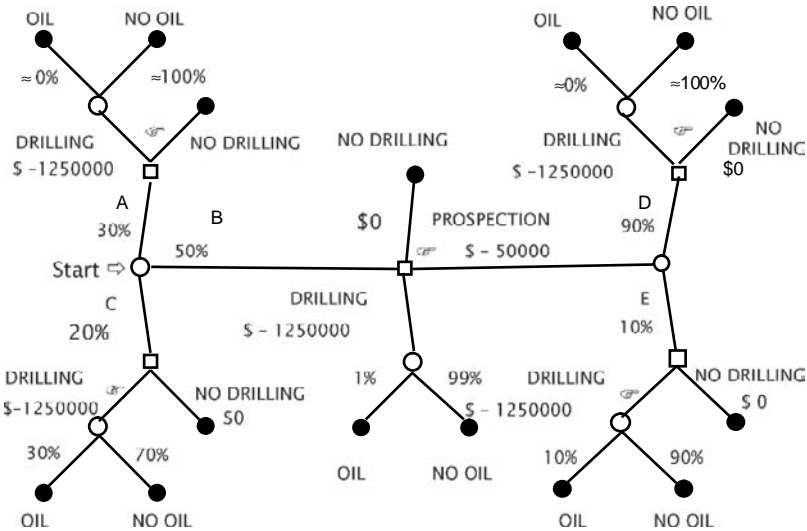


Fig. 2.1 Decision diagram for an oil well search

This is just an example, with no claim of corresponding to any real situation. Moreover, the quantities have been placed at random. Suppose that before deciding whether to perform the drilling or not, a geologist is asked her opinion concerning the presence of certain possible circumstances A, B and C (which represent the various geologic situations that are somewhat indicative of the presence of oil). Suppose that the geologist says that the circumstances A, B and C (which we assume to be exhaustive and pairwise exclusive) have the probability 30%, 50% and 20%, respectively. For all those circumstances it can be scrutinized as to what would be appropriate to do under the hypothesis that it happened. Suppose we knew that if circumstance A occurred, any drilling would surely turn out to be hopeless. On the other hand, suppose that if circumstance C occurred, we surely would have to go on with the drilling, for in that case there would be a 30% probability of finding oil. Suppose, finally, that if circumstance B occurred, a decision would have to be made between (a) going on with the drilling (which we suppose would cost 1,250,000), or (b) taking the interlocutory decision of performing a seismic prospection in order to decide, in the light of such an experiment, whether to continue with the drilling or not.

A seismic prospection consists in setting off a bomb in the subsoil in order to observe the propagation of the resulting artificial earthquake (this is a relevant piece of information as far as the knowledge of geologic layers conditions and so on is concerned). For each of the three options, the potential gain (in the event oil is found) will have to be taken into account and compared with the respective expected costs. In short, the problem here is to calculate the prevision of gain for each possible decision and then to make the decision that yields the maximum prevision of gain. Suppose that a decision was taken to go ahead with a seismic prospection. In this case the experiment could lead to — say — either the outcome D or the outcome

E, with probabilities 90% and 10%, respectively. In either case, a new decision will have to be taken about whether, in the light of the new state of information, the drilling is to be done or not, and so on.

I believe that this outline is sufficient to give a general idea of the approach (even though I have chosen some quantities at random). The main problem is to decide whether — in the presence of circumstance *B* — it would be profitable to perform the seismic prospection. Every path involves a certain cost and yields a certain probability that the oil search will be successful.

This example seems instructive to me in regard to showing that probability always refers to single events and that it is not always related to a frequency. Of course, it might well be that for many problems a mean is sufficient.² Suppose, for instance, that on average 5% of the items manufactured by a certain automatic machine turn out to be faulty. In such a case, this piece of statistical information is all that the buyer needs to know in order to judge the reliability of the product relative to its price. And, on the basis of the sales, the producer will have to decide whether to place all the items straight on the market or carry out a selection of the rejects first (an operation that would have a certain cost).

The fact that the probabilities are subjective does not rule out that two individuals who have to make a common decision might end up agreeing. Not only can it happen that two individuals agree on some probability assessments, but even failing this, it could still happen that the conclusions about the decisions to be taken turn out to be alike. And if the conclusions were alike, this would be an indication of their stability and therefore no further experiment would be necessary (unless a considerable change of opinion could be expected as a consequence of their outcome). Otherwise, it is possible to try to reach an agreement by repeating some experiment or by contriving more precise ones. It is not appropriate, however, to discuss the details of these problems now.

What has been said is just an example to illustrate that in order to make the correct use of probability, we must try to make a step-by-step evaluation of the chain of all the circumstances faced. This contrasts with the way these things are usually treated in handbooks. These tend to favour statistical data (percentages, observed frequencies and so forth): indeed, more schematic and artful examples are usually preferred, as they allow one to speak of probability as if it were something beyond dispute. In those examples, equally probable cases are such by definition: it is therefore forbidden to think that perhaps one of those cases could be more probable than the others. I do not mean to say that, in certain cases, it is not natural to judge them all equally probable. But — and this is the most important point — we cannot make an unquestionable dogma out of this evaluation, which is an opinion just like any other.

Subjectivism versus Objectivism

Maybe there is also another point upon which I have not yet lingered (though maybe I have vaguely pointed to it yesterday) and which I believe is worth clarifying. It is an elucidation about the various attitudes held by mathematicians, statisticians,

philosophers and so forth, towards probability. In this respect, I find some writings in the *Enciclopedia Einaudi* interesting, and in particular two articles by the Russian *Stefan Amsterdamski*.³ They are the entries *Caso/Probabilità* (Amsterdamski, 1977a, pp. 668–87) and *Causa/Effetto* (Amsterdamski, 1977b, pp. 823–45).^{*} The double title, in both cases, stresses the connection intervening between the two associated concepts. I will not discuss a third article by Amsterdamski included in the same encyclopedia, as I have not yet read it. Of those two articles I have written in the entry *Probabilità* (on which I am still working) for the same encyclopedia:

[They] open up the sight of a vast as well as complex topic, one which is very analogous, if only as a background, to the one planned for the present, more technical discussion which therefore will result enriched and more precisely informed by a pointed comparison, not a contraposition as the problem here is to propose and try to justify a univocal and precise choice among the wide spectrum of the options which have been put forward, or at least have not been excluded, by the above-mentioned pair of articles. In a schematic form and taking advantage of the possibility of referring to the broad and thorough survey by Amsterdamski, I can already make precise, in a few words, my position by saying that it corresponds to a version, very radicalized indeed, of point of view 2, vol. 2, p. 675 (de Finetti, 1980, pp. 1147–8).

The article continues with my formulation. Before reading it to you, I would rather read out loud the passage of the entry *Caso/Probabilità*, where Amsterdamski expounds, among many, the two stances that are the nearest to my point of view. Amsterdamski writes:

The existing interpretations of the concept of probability can be divided into two groups: the interpretations according to which the notion of probability is a characteristic of judgments and the interpretations according to which probability concerns the relations between classes of events (Amsterdamski, 1977a, p. 672).

Later on, with regard to the former group, he writes:

However, things change if one assumes that the concept of probability is always a characteristic of judgments and never concerns events. This opinion corresponds to one of the following points of view:

1. it either means that every probabilistic statement concerns logical relations among events and is, consequently, analytical;
2. or it means that probability, despite being always a characteristic of judgments, is not always a logical concept and that not every probabilistic statement is analytic, but some of them express the degree of belief in the truth of the statement, i.e. characterize the knowing subject's attitude towards it (Amsterdamski, 1977a, pp. 674–5).

If events are considered only in themselves, probability becomes a function of them only. This is exactly the *objectivistic* meaning of probability judgments. It is not said, in fact, “in my opinion,” nor is it said in whose opinion. It is as if the probability judgment were heaven-sent.

At this point you should be in a position to understand the reformulation of Amsterdamski's formulation 2 that I proposed in the entry *Probabilità*:

^{*} (Translator's note:) “Chance/Probability” and “Cause/Effect,” respectively.

Probability, despite being always a characteristic of judgments, is never a logical concept. Probabilistic statements are never analytic,⁴ but they always and only express the degree of belief that the judging subject assigns, given her current state of information, to the object of the statement. That is to say, concisely, that probability characterizes the knowing subject's attitude with regard to a given statement (de Finetti, 1980, p. 1148).

This sentence contains two theses. The former is that probability is not a characteristic of events; the latter is that probabilistic statements are not analytic. The former thesis is clear: probability is not a property of events. Rather, we can all, for the very fact that we are expecting an event, attach a probability to it. As to the latter thesis, it excludes that probability judgments can be logically warranted. I thus comment, in my article, this reformulation:

To clarify the situation in a more explicit form, it is enough to ask ourselves which answers an interrogated person might give with regard to an event, that is a given statement (which has an unambiguous and, to her, understandable meaning). Evidently, the possible answers among which anybody can choose the one corresponding to her own actual state of relevant knowledge are, in an objective sense, three: "yes," "no," "I do not know." The essential difference between the three answers lies in the fact that (in whatever version) the two extreme ones: "yes" (or "true," or "sure") and "no" (or "false," or "impossible") do have a univocal sense, a definite and categorical character, whereas the intermediate one "I do not know" (or "doubt" or "uncertain") has instead a provisional character as it only expresses the persistence of a current ignorance or indecision between "yes" or "no," which are the only two definitively conclusive answers (de Finetti, 1980, p. 1148).

Therefore, when the answer is either "yes" or "no" the question is settled. Otherwise, if it is "I do not know," then one has the whole spectrum of probabilities between 0 and 1 to choose from.⁵ In this case, the question will be to try to translate verbal expressions like "very much probable" or "much probable" or "little probable" or "very little probable," and so forth, into numerical terms. All these concepts are very general and despite my efforts to make them concrete, they might not yet be entirely clear to you. And, since we still have five minutes at our disposal, I would like to use this time to answer your questions.

For an Omniscient Being Probability Would Not Exist

GAMMA: The objectivist definition, which gives to an event a probability that is independent of the individual's own judgment, is a consequence of the fact that the body of pre-existent knowledge H_0 is often shared by the individuals, so that (almost) everyone agrees on giving the same probability values. For example, two individuals can be put in front of a coin they have never seen before. Those people, reasonably, will give probability $1/2$ both to the event Heads and to the event Tails. But if one person knew that the coin was unfair, his opinion would be radically different. Perhaps the essential point is that H_0 could be a *common* body of knowledge. DE FINETTI: Even if H_0 were common to the individuals, there still could be differences of opinion.⁶ When I say that probability is subjective, in the sense that it must be clear who is evaluating it, I do not mean to exclude that the evaluations

could coincide in many cases. There is no reason to exclude this. On the contrary, whenever the state of knowledge is alike, or the habits according to which certain things are thought of are alike, it is natural that the probability evaluations of distinct individuals, despite being subjective, would coincide. What must be avoided is turning this coincidence into a logical fact. We must not say that probability is unique, nor that in many cases the majority of people can indicate the probability “correctly,” whilst the others are wrong. In some cases one could even think that there is a mistake in the evaluation of probability: one could think, for instance, that somebody who has given another value, has done so because of a mistaken calculation rather than because she has a different opinion. Exactly like a shopkeeper who, by mistake, said that 50 cents and 20 cents were 60 or 80 cents (excluding the case where the shopkeeper deliberately gets the calculation wrong, to cheat). I hope I have not misinterpreted your question.

GAMMA: I was only trying to understand the objectivists’ reasons. It wasn’t really a question.

DE FINETTI: Nothing prevents us from studying those cases in which it seems somewhat natural to admit that certain things are equally probable, and perhaps even independent. In those cases, everyone would agree on evaluating the probabilities that follow from this premiss, unless they got the calculations wrong. And there is no objection to be raised against this, even though it should be borne in mind that it is practically impossible that this can happen, except, possibly, in the special cases of coins, urns, dice and so forth. And even in those cases, a complete coincidence is not guaranteed at all. Probably, no die is so “perfect” as to necessarily lead everyone to be *a priori* certain that the various sides will come up with the same frequency. At any rate, even though this appears to be reasonably confirmed by past experiences, it could never be so in a perfect way. And whenever the number of observed trials fails to be a multiple of six, it is logically impossible that equal frequencies can occur for each side.⁷ In fact, if probability were a frequency, we should demand that each of the sides 1, 2, 3, 4, 5, 6, always had to come up the same number of times. All excessively apodictical statements can only count as a vague, general idea.

A probability evaluation depends on many circumstances that we have not yet mentioned. For example, if we made withdrawals without replacement, step by step, we would extract every ball from the urn and it is certain that the frequency will eventually coincide with the actual proportion. If this is known from the outset, what will happen can be predicted with certainty. Otherwise, if this is not known, one could make up one’s mind about the composition of the urn only as the sequence of withdrawals approached its end. And only after drawing the last ball, would one know exactly the composition.

In my own opinion, in order to avoid falling into a self-trap, we should think of the whole issue in a practical rather than abstract way. If we started off with the presupposition that both equiprobability and independence were satisfied, then all the calculations could be done accordingly and no problem would arise. But both equiprobability and independence are subjective. And it could well happen that in special cases (like the typical textbook ones), everyone agreed on adopting a particularly simple and symmetric opinion. In cases of this sort, it would be hard to make

some other similarly reasonable choice. Still, probability must not be mythicized. We must not think that probability values are entities living in some Empyrean, like Plato's, world of ideas. We must not think that probability values are exact values to which we substitute approximated estimates.

For an omniscient being, probability would not exist. For this entity each proposition would be either true or false. I might doubt the fact that if I threw the little piece of chalk that I am now holding in my hands, it would eventually reach the blackboard. On the contrary, an omniscient being would have known, since the beginning of time, that that little piece of chalk would stop there, exactly at that point and that this table would be made of wood and that there would be this glass round-shaped object here.

Likewise, we should not forget that everything is relative to the world that an individual can know (or imagine) concretely at the moment or at the epoch in which she is living. If an individual from the thirteenth century were driven here through space-time (for instance, Dante — undoubtedly a supreme mind) he would probably fail to understand what we are talking about. Everything depends on the way we represent the things we want to deal with. Each individual judges the probability of those events that matter to him: because they catch her attention or because they concern her personally or even because they concern others or the future of the country or of the world. Still, it makes no sense to say that those probabilities exist in themselves.

Similarly, we could try to foresee the percentage of votes for each of the various parties running for the next elections. We could evaluate the *prevision* (or as it is traditionally referred to with a rather funny expression, the *mathematical expectation*)* of each percentage. But this would be a subjective prevision: everyone can suppose whatever she likes; one cannot say that a prevision is more or less right. A posteriori, it is surely possible to tell whether a prevision was even close to the actual value.

I have a tremendous dread of the mystification of concepts.

Editor's Notes

1. This is probably a misunderstanding. By saying "inappropriate," Beta meant "unfavourable" rather than "indecorous," as de Finetti appears to have understood. Similarly, the term "cheat," which we have put in inverted commas in the question asked by the student, was most likely intended kind-heartedly as a synonym for "trying to be smart" (an expression which de Finetti himself uses later on) rather than in its proper meaning of "unfairly violating the rules of the game."

However, de Finetti's answer is important for it brings to the foreground the fact that in the application of proper scoring rules, it is immaterial whether the subject is asked to indicate her degrees of belief or simply whatever numbers she likes. Under proper scoring rules, indeed, the

* (Translator's note:) Note that in consonance with the Italian traditional expression, the original text reads *speranza matematica*, literally *mathematical hope*.

subject will always have an interest in stating those numbers that correspond to her own degrees of belief.

2. It should be borne in mind that the frequency of favourable cases can be thought of as the mean truth-value of the events.
3. As a matter of fact, Stefan Amsterdamski (1929–2005) was a Polish (rather than a Russian) philosopher and sociologist. Born in Warsaw he actually lived a few years in the former USSR, first in Vilnius and then in Siberia, being deported there in 1941 with his family. Between 1970 and 1989 he was employed by the Institute of Philosophy and Sociology of the Polish Academy of Sciences. He was one of the founders of the so-called “Second Flying University” (1977–1981), which flourished just around the time Amsterdamski wrote the articles for the *Enciclopedia Einaudi* quoted in the text. This unofficial institution, where each class was held in a different private apartment, was established on the model of the so-called “First Flying University”, going back to 1883.

The initiative was strongly opposed by the Polish government. To support the Flying University, a letter, signed by world-wide leading intellectuals like Alfred Ayer, Michael Dummett, Michel Foucault, Robert Nozick, Hilary Putnam, Willard Quine, René Thom, Franco Venturi and many others was published in *The New York Review of Books* in 1980 (Ayer, Dummett, Foucault et al., 1980). As reported by an *Editor's Note* accompanying the Letter “Soon after this appeal was received, The New York Times of November 30 reported that a ‘police campaign’ was threatening the Flying University. Police have recently arrested and fined a number of lecturers and harassed and photographed students.”

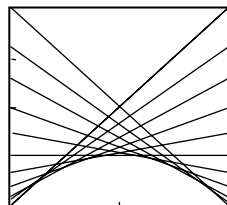
As a result, between December 1981 and November 1982 Amsterdamski was interned in a camp at Jawor in Darlowko. After his release, he taught at the *Collège de France* and worked in several foreign Institutions, like the *Institut des Hautes Études* in Paris (1984–1985), *Stanford University* (1991) and the *Institute for Human Sciences* in Vienna (1992–1993).

4. This standpoint seems to clash with the Definetian conception of the calculus of probability as “logic of the uncertain.” Actually, what de Finetti intends to say here (despite the use of the adverb “never”) is that those probability judgments that do not follow logically from the fundamental axioms of probability are empty of logical warranty. Those judgments that do follow from the axioms are in fact, for him, *tautologies* which do not express a genuine opinion. For the logicians, instead, *every* probability judgment is tautological. It is not, however, clear in which sense de Finetti’s reformulation of thesis 2 differs from Amsterdamski’s original one. On these topics, see Mura (1995b).
5. Here it should be added: “provided that other probability evaluations already fixed do not confine the interval of the possible probability values to a sub-interval of $[0, 1]$.” Indeed, in virtue of a theorem, proved by de Finetti in the thirties and called by him the *fundamental theorem of probability* ([1970] 1990a, p. 111), in the presence of the probabilities $P(E_i)$ ($i = 1, 2, \dots, n$) of a finite numbers of events, the probability of a further event E can be assigned, coherently, any value in a closed interval $[p', p'']$, where $0 \leq p' \leq p'' \leq 1$. See Chapter 12 page 117 and note 9 page 124.
6. This statement contains the main difference between subjectivism and logicism: for the latter, probability values, as rational degrees of belief, only depend on the event taken into consideration and the state of information H_0 . As a consequence, the states of information being equal, two rational people would be bearers of the same probability evaluations.
7. The following clause should be added: “on a finite number of trials.” In Chapter 6 below, de Finetti will also criticize the idea that probability can be a frequency with respect to an *infinite* number of trials.

Chapter 3

Geometric Representation of Brier's Rule*

Generalization and applications of proper scoring rules to the evaluation of probability: penalization $L_x = (E - x)^2$. Geometric diagram with straight lines and parabola: equivalence of approach between the formulation based on penalizations and the one which takes us back to classical schemes of equiprobability (where reference to the latter however does not lead to any vicious circle).



I would like to go back briefly to what I said last week in regard to proper scoring rules. It seems to me, in fact, that I might have introduced too many details without making things sufficiently clear. Since I am afraid that you might miss the most important point, I would like to focus more closely on it.

I shall illustrate in a way other than the algebraic one (which I find trivial) the definition of the probability of an event E through proper scoring rules. And since Brier's is the simplest among scoring rules, I shall refer to this latter.

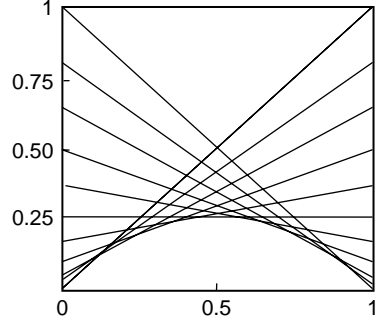
Envelope Formed by Straight Lines

Brier's rule consists in this: a person X is asked to evaluate the probability of an event E and she is told that she will receive a penalization L , which depends both on the value of E (1 if E is true, 0 if E is false) and on the number π stated by her. More precisely, the penalization will be $(E - \pi)^2$. It will therefore equal $(1 - \pi)^2$ if E is verified, and it will equal π^2 if E is not verified. The question we are interested in is: what reason do we have to believe that the number stated by X really corresponds to X 's subjective probability? I will try to explain this through Fig. 3.1.

The x -axis represents the ordered set of the possible values of p (that is the closed interval $[0,1]$). The prevision of penalization according to Brier's rule is displayed on the y -axis. For each pair of values of π and p , the prevision of penalization $P_p(L_\pi)$ is given by

* Lecture IV (Wednesday 14 March, 1979)

Fig. 3.1 Envelope formed by straight lines each representing a different elicited answer according to Brier's rule



$$\begin{aligned}
 P_p(L_\pi) &= p(1 - \pi)^2 + (1 - p)\pi^2 \\
 &= p(1 + \pi^2 - 2\pi) + \pi^2 - \pi^2 p \\
 &= p + \pi^2 - 2\pi p \\
 &= p(1 - 2\pi) + \pi^2.
 \end{aligned} \tag{3.1}$$

$P_p(L_\pi)$ is linear in p . This means that the curve representing the prevision of penalization as a function of p is, once the value π has been fixed, a straight line. To each value of π corresponds a distinct straight line. Figure 3.1 represents the straight lines relative to some values of π — more precisely the values $i/10$ ($0 \leq i \leq 10$). It will be noted, by observing the figure, that every straight line is a tangent to the parabola $p - p^2$ (which is therefore an envelope of them). It follows from this that no straight lines lie below the parabola. For each p , the optimal choice (among those X can make) consists in indicating that value of π which corresponds to the straight line of minimal ordinate at point p . But the straight line of minimal ordinate at point p is the straight line tangential to the parabola at p . But which value of π corresponds to this straight line? Let us observe that the derivative of $p - p^2$ is $1 - 2p$. Hence, the straight line tangent to the parabola is the straight line of slope $1 - 2p$. By equation (3.1) the slope of all penalization lines is $1 - 2\pi$. It follows that whenever $\pi = p$, the slope of all penalization lines is $1 - 2p$ and hence, the straight line will be the tangential to the parabola. Any other straight line will be of greater ordinate at point p . Hence, the best choice X can make consists in indicating p (that is, the value of π which coincides with her probability opinion).

Operational Definition of Probability

We have shown that Brier's rule has the property of forcing a coherent person to indicate her opinion, for if this person indicated a different value she would have suffered — as a prevision — an additional penalization. The importance of this result should be clear. Indeed, if we simply asked someone: “what is, in your opinion, the probability that the event E will take place?”, that person would not have any interest in saying what she really thinks.¹ But scoring rules possess yet another virtue. It is the fact that by operationally defining probability through such rules,

the difficulties arising from the definition given in terms of bets can be overcome.² In the betting framework the individual was put in a position to indicate the betting rate, leaving the faculty of inverting the sides of the bet (giving in that case the role of the bank to the bettor) to a competitor. Theoretically, in this situation, the bettor would have an interest in indicating a betting rate that does not favour either of the two sides, for otherwise the competitor could take advantage of this simply by choosing for herself the more favourable side. Unfortunately however, this schema works in a very imperfect way. In fact, it presupposes that the bettor is certain that her competitor shares her own probability evaluations. It is hard to satisfy this condition except in very special cases. Should the bettor foster the suspicion that the competitor had richer information or only that she was the bearer of a different probability opinion, she would no longer have an interest in indicating probability sincerely. On the contrary, the procedure based on scoring rules is sheltered from this difficulty, for under a scoring rule, the person who is making the probability evaluations does not have to come to terms with other people's evaluations.³

Editor's Notes

1. A strictly proper scoring rule forces a coherent person to declare her actual degrees of belief being such a declaration the optimal choice in a decision problem under uncertainty. This is exactly what in the literature about proper scoring rules is called 'elicitation'. Now, since every choice requires reflective deliberation, the strictly proper scoring rules not only discourage deliberately insincere declarations of degrees of belief, but also stimulate careful reflection, aimed at determining the numerical values of personal probabilities. The exercise consisting in the effort to give the best probability evaluations under a proper scoring rule should, according to de Finetti, gradually develop in the individual the ability to express accurately her deep-felt belief.

To test this hypothesis, de Finetti organized the football forecast competition about which he talked above (see Chapter 1, p. 6). While showing us the data of that research he would insist on how, at the end of the season, the average penalizations were smaller and the evaluations much less biased towards extreme values than they were at the beginning: this was, according to him, a consequence of the apprehension obtained by means of practicing with scoring rules.

2. The problem arises of explaining how de Finetti's operational elicitation of the subjective probability that a person X bears, really provides a not circular definition of probability, given that the suggested procedure is based on asking X to declare her subjective probability, which seems to be presupposed. It should be noted that in the light of Bayesian decision theory, as axiomatized by L. J. Savage, it is inessential to ask the subject to declare numbers that coincide with her subjective probability values. The only assumptions required are (i) that the preferences among acts of X satisfy the axioms of Bayesian decision theory, (ii) that the utility of a given score s is a positive transformation of its numerical value. If these two requirements are satisfied, X must declare numbers that satisfy the axioms of probability. As a matter of fact, de Finetti proved that if the numbers declared by X do not satisfy those axioms, they cannot be optimal, since a different set of numbers exists that would lead *with certainty* to a better score. This result is just the counterpart, with respect to the proper scoring rules, of the *Dutch book* argument (see Chapter 4, note 12 on page 43). So de Finetti may say, without circularity, that subjective probability is just the quantity operationally measured under the assumptions (i) and (ii).
3. For a more detailed explanation of the difficulties faced by the method of bets, I refer to Section 2 of Mura (1995b)

Chapter 4

Bayes' Theorem*

Subordinate (or conditional) probabilities: $\mathbf{P}(E \mid H H_0)$ (the probability of the event E in the state of information H_0 and conditional on H being verified.)

Bayes' theorem, inductive reasoning according to this framework; outline and criticisms of alternative methods (appropriate denomination introduced by I. J. Good = "adhockeyies").

The fundamental statement of Bayesian reasoning:

$$\mathbf{P}(E \mid H) = \mathbf{P}(E) \frac{\mathbf{P}(H \mid E)}{\mathbf{P}(H)}.$$

$\mathbf{P}(E \mid H)$ is obtained from $\mathbf{P}(E)$ by changing it by the same proportion as $\mathbf{P}(H)$ is modified conditional to E .

Bayes' Theorem and Linearity

We have already pointed to the logical operations on events and random quantities. I would now like to develop this topic more systematically. For the sake of simplicity, I will prescind from referring to proper scoring rules. Instead I will be referring to the interpretation of probability and prevision as *price*.¹

The probability of an event E , in fact, can be considered as the fair price (according to the evaluation of the person who is appraising it) of the right to receive 1 should E be verified and nothing should E fail to be verified.

$$\mathbf{P}(E) = \begin{cases} 1 & \text{if } E \\ 0 & \text{if } \tilde{E}. \end{cases} \quad (4.1)$$

Considering E as a random quantity, which takes the value 1 or 0 depending on whether E turns out to be true or false, the probability of E can be considered as the price of an aleatory offer amounting to E . By saying *price* I mean a linear quantity. E , on the other hand, can be identified with the random quantity $1 - E$. Linearity implies:

* Lecture VI (Thursday 22 March, 1979).

$$\mathbf{P}(\widetilde{E}) = \mathbf{P}(1 - E) = 1 - \mathbf{P}(E).$$

In words: the probability of the complement $1 - E$ of E (that is, of \widetilde{E}) equals the complement of the probability of E , that is, $1 - \mathbf{P}(E)$.

Let us now consider the arithmetic operation $A + B$. It does *not* give rise to an event. It does instead give rise to a random quantity, which equals 0 if A and B are both false, 1 if one of the two events is true and the other is false, 2 if both A and B are true:

$$A + B = \begin{cases} 0 & \text{if } \widetilde{A}\widetilde{B}, \\ 1 & \text{if } \widetilde{A}B \text{ or } A\widetilde{B}, \\ 2 & \text{if } AB. \end{cases}$$

The arithmetic sum must be distinguished from the *union of two events* which — unlike the former — is always an event. More precisely, the union $A \vee B$ of two events A and B is that event which is false only if A and B are both false and is true in all the remaining cases. The union has the following in common with addition: that it takes the value 0 if and only if both events are false. The reason according to which the union cannot, in general, be represented as an addition is that when both A and B are true, the sum takes the value 2 instead of 1.

If we have n events A_1, A_2, \dots, A_n , the union E of A_1, A_2, \dots, A_n is true if at least one of the events A_i ($1 \leq i \leq n$) is true. Then, it is enough to say that E is the *maximum* of these n numbers. Since they are all either 0 or 1, then the maximum will either be 0 or 1 depending on whether the A_i 's are all 0 or at least one of the A_i 's is 1. And in order to say that we take the greatest of those numbers, we write:

$$E = \bigvee A_i.$$

Let us now consider intersection. The intersection E of A_1, A_2, \dots, A_n could also be written in this way:

$$E = \bigwedge A_i.$$

The intersection E of E_1, E_2, \dots, E_n is true when all the E_i 's are true. That is, when there is no false one (hence there is no 0). Therefore, the intersection amounts to the product in the usual sense, for to say that the intersection is true, that is that these events are all true, is to say that they all equal 1 and there is no 0. And the product of many 1's, without any 0, is 1, whereas a product is 0 if among the factors there is even a single 0. In the case of union, on the other hand, the analogue (with addition in place of multiplication) does not hold, for if we write:

$$E_1 + E_2 + \dots + E_n$$

we get the random quantity X , that is the number of true events: a number which varies between 0 and n . But there is no connection between addition and union. $E = \bigvee E_i$ does not equal the sum $x = \sum E_i$, but equals the event $X > 0$. If X is 0, then they are all false and the union is false as well. If there is even a single true one, then the union is likewise true. What must be stressed is that every logical operation can be represented in arithmetic terms. Depending on the circumstances, the logical notation or the arithmetic one will be preferred.

Of course, all those events must be considered in relation to a state of information H_0 . That is, we should write:

$$(A \mid H_0), (B \mid H_0), (C \mid H_0), (D \mid H_0), \text{ and so on.}$$

The reference to H_0 is important, though in practice, in order to simplify the notation, it may be omitted.²

I would now like to illustrate the relation between the probability of an event E , $\mathbf{P}(E \mid H_0)$ and the probability of E subject to the obtaining of a hypothetical condition H , which we shall denote by $\mathbf{P}(E \mid HH_0)$. What does $\mathbf{P}(E \mid HH_0)$ or (omitting the state of information H_0) $\mathbf{P}(E \mid H)$ mean? It is the price that one judges fair to pay in order to receive 1 if both E and H turn out to be true (it is already known that H_0 is true) and nothing otherwise, with the proviso that should H fail to occur (independently of the occurrence of E) no action whatsoever would take place, that is, nothing would be paid and nothing would be received: if one had already paid, then one would be reimbursed. Hence we have

$$\mathbf{P}(E \mid H) = \begin{cases} 1 & \text{if } EH, \\ 0 & \text{if } \tilde{E}H, \\ \mathbf{P}(E \mid H) & \text{if } \tilde{H}. \end{cases} \quad (4.2)$$

Let us now turn to the relation between $\mathbf{P}(E)$ and $\mathbf{P}(E \mid H)$. As a consequence of (4.1) it holds that

$$\mathbf{P}(H) = \begin{cases} 1 & \text{if } EH, \\ 1 & \text{if } \tilde{E}H, \\ 0 & \text{if } \tilde{H}. \end{cases} \quad (4.3)$$

By virtue of the assumption of the linearity of prices, we can multiply member-wise (4.3) and (4.2). In this way we obtain

$$\mathbf{P}(H)\mathbf{P}(E \mid H) = \begin{cases} 1 & \text{if } EH, \\ 0 & \text{if } \tilde{E}H, \\ 0 & \text{if } \tilde{H}. \end{cases} \quad (4.4)$$

From (4.1) (given that $EH = HE$) it follows that

$$\mathbf{P}(HE) = \begin{cases} 1 & \text{if } EH, \\ 0 & \text{if } \tilde{E}H, \\ 0 & \text{if } \tilde{H}. \end{cases} \quad (4.5)$$

By comparing (4.4) and (4.5) we obtain

$$\mathbf{P}(HE) = \mathbf{P}(H)\mathbf{P}(E | H). \quad (4.6)$$

Similarly it holds that

$$\mathbf{P}(EH) = \mathbf{P}(E)\mathbf{P}(H | E) \quad (4.7)$$

and since $EH = HE$, from relations (4.6) and (4.7) it is immediately obtained that

$$\mathbf{P}(H)\mathbf{P}(E | H) = \mathbf{P}(E)\mathbf{P}(H | E) \quad (4.8)$$

and therefore, dividing both members of the identity (4.8) by $\mathbf{P}(H)$, we finally obtain

$$\boxed{\mathbf{P}(E | H) = \mathbf{P}(E) \frac{\mathbf{P}(H | E)}{\mathbf{P}(H)}} \quad (4.9)$$

Equation (4.9) provides the simplest formulation of *Bayes' theorem*. Here $\mathbf{P}(E)$ is the probability attributed to E in the state of initial knowledge H_0 . If, on the other hand, we suppose that H will occur (or we learn that H has occurred), then $\mathbf{P}(E)$ — the initial probability of E — varies by the same proportion at which $\mathbf{P}(H)$ changes conditional on E , giving rise to the new probability $\mathbf{P}(E | H)$.

Statistics and Initial Probabilities

The whole of subjectivistic statistics is based on this simple theorem of the calculus of probability. This provides subjectivistic statistics with a very simple and general foundation. Moreover, by grounding itself only on the basic probability axioms, subjectivistic statistics does not depend on those definitions of probability that would restrict its field of application (like, e.g., those based on the idea of equally probable events). Nor, for the characterization of inductive reasoning, is there any need — if we accept this framework — to resort to empirical formulae. Objectivistic statisticians, on the other hand, make copious use of empirical formulae. The necessity to resort to them only derives from their refusal to allow the use of the initial probability $\mathbf{P}(E)$. And they reject the use of the initial probability because they reject the idea that probability depends on a state of information. However, by doing so, they distort everything: not only as they turn probability into an objective thing (though this would not be that serious, for the distinction between “objective”

and “subjective” could well be — up to a certain point — a subjective one), but they go so far as to turn it into a *theological* entity: they pretend that the “true” probability exists, outside ourselves, independently of a person’s own judgement.

Bayes’ theorem provides statistics with a unitary framework, tracing the hundred empirical formulae of objectivistic statistics back to a unique principle. Rules of the form “whenever knowledge changes in such and such a way, let the probability change in such and such a way” might well be meaningful as approximations. But they cannot be preferable to Bayes’ theorem, which is such a simple truth that Cornfield (one of the subjectivistic statisticians, one of the best in the Bayesian field)³ once said that given how natural and unquestionable the result is, it is exceedingly solemn to call it a theorem.

Bayesian Updating is Not a Corrective Revision

ALPHA: The conditional probability is the new probability that we ought to give whenever we acquire new information. Do you consider this to be an actual change of opinion?

DE FINETTI: If we reason according to Bayes’ theorem, we *do not* change our opinion. We keep the same opinion, yet updated to the new situation. If yesterday I was saying “It is Wednesday,” today I would say “It is Thursday.” However, I have not changed my mind, for the day after Wednesday is indeed Thursday.⁴ Bayesian updating can be illustrated by means of a simple diagram (Fig. 4.1). Let us let H_0 represent everything that we initially know to be true. Suppose that the diagram is to scale: in this way, taking the area of H_0 as a unit, the area of the other events represents the respective probabilities. If we added a new hypothesis H , the probability of E conditional on H would cease to be equal to the ratio between the area of E and the area of H_0 , but would equal the ratio between the area of $EH \mid H_0$ and the area of $H \mid H_0$. If this ratio remained unchanged, then we would say that the events are independent. In such a case, the probability of $EH \mid H_0$ relative to the probability of $H \mid H_0$ would be the same as the one of E relative to H_0 . When a new piece of evidence H is acquired, H rules out all the parts that lie outside itself (that is, logically incompatible with H). Hence, the part of H_0 that is compatible with

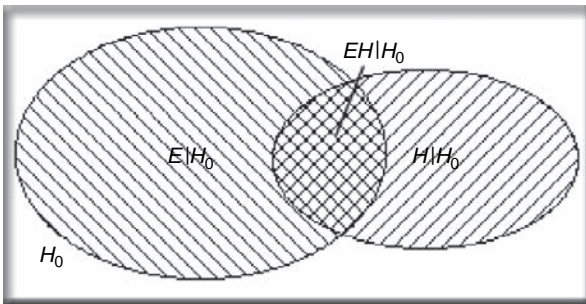


Fig. 4.1 Venn diagram illustrating Bayes’ theorem

the new state of information is HH_0 . By normalizing the measure (that is, putting the area of HH_0 equal to 1) and by preserving the proportions, we obtain the new probabilities. Thus, the probability of $E \mid HH_0$, will equal the area of $EH \mid H_0$ divided by the area of $H \mid H_0$.

“Adhockeries”

This explanation could be easily grasped even by schoolboys. Nonetheless there are many philosophical points of view that in an attempt at turning into objective what is subjective, arrive at complicated alternative formulae without verifying whether the latter have a foundation or not. Since 1764 — the year of the publication of Bayes' essay (Bayes [1764] 1970) — there has been an ongoing lively debate on this point. Among the major subsequent probabilists there have been divergences of opinion on some alternative empirical formulae, which were, at first sight, quite reasonable. I. J. Good calls those *regollette** (some of which are artful and some simplistic) “adhockeries.”⁵ I cannot deny that sometimes they might even have an approximative meaning: yet none of them is grounded on reasoning that is directly derived from the concept of conditional probability.

The defendants of those alternative rules argue that they are, in practice, more useful than Bayes' theorem and try to show that they better account for inductive behaviour. I do not mean to show any disrespect towards other scholars but I believe that those attempts lack any foundation. Although I am respectful of others' work, I cannot help but notice that every formula alternative to Bayes' theorem is grounded on empirical and qualitative arguments and therefore lacks a robust foundation.

I have mentioned those things to bring attention to the fact that to speak of the probability of an event *tout court*, without any qualification, does not have any concrete meaning. Rather, it must be kept in mind that probability is always relative to the state of knowledge of the person who is making the judgement. For instance, what is the probability that a certain person will recover from illness? The physician examines her, considers a variety of circumstances and then, if he does not manage to reach a sufficiently certain conclusion, prescribes some tests or some radiography. As soon as the radiographies are available he examines them, modifying continuously his initial opinion. He values this source of information that has provided some new elements, of which some could be favourable, others unfavourable. In such a case, it is hard to enumerate one by one all the various pieces of evidence, attaching a numerical value to them. Nonetheless, the pattern of reasoning is always the same, even though this is mentally developed in a somewhat intuitive way. The physician who is to make a prognosis reasons like this: “if that circumstance which I noticed during the examination were absent, the probability that the patient could recover would be this much. However, in the light of the signs that I have found (of which

* (Translator's note:) See footnote * on page 5.

one betters the clinical situation whilst the other — more important one — worsens it), the probability of recovering is, if only to a small extent, diminished.”

I am insisting more on the non-mathematical part: mathematics is not, after all, the most adequate instrument to put the concrete aspect of probability under the spotlight. From the mathematical — that is formal — point of view, there is no difference between the union of two events and the union of a bag of potatoes and a piece of meat. I could represent both unions in the same way. The mathematical part is not as decisive as is the understanding of the function that each single notion has in everyone's reasoning, even in the reasoning of those who are not familiar with the calculus of probability. The calculus, however, is useful to avoid slips, for this is a field in which slips are far from unusual.⁶

Bayes' Theorem for Random Quantities

What holds for events, holds for random quantities as well. Let us ask ourselves: what is the prevision $\mathbf{P}(X \mid H)$ of a random quantity X conditional on the occurrence of an event H ? To provide an answer to this question it is useful to recall formula (4.6):

$$\mathbf{P}(HE) = \mathbf{P}(H)\mathbf{P}(E \mid H),$$

from which we immediately obtain:

$$\mathbf{P}(E \mid H) = \frac{\mathbf{P}(H \mid E)}{\mathbf{P}(H)}. \quad (4.10)$$

Well, the formula about random quantities is just a simple generalization of formula (4.10):

$$\mathbf{P}(X \mid H) = \frac{\mathbf{P}(H \mid X)}{\mathbf{P}(H)}. \quad (4.11)$$

In many statistics textbooks, $\mathbf{P}(X)$ is denoted by the symbol \bar{X} or the symbol \bar{X} . If we consider the single values that X can take, we must distinguish between the *discrete* and the *continuous* case. In the discrete case, the set of values that X can take may have either a countable infinity or a finite number of elements. In this lecture, I will confine myself to considering the discrete case with a finite number of possible values only.

In this case, X can take one (and only one) of the n values x_1, x_2, \dots, x_n with probabilities p_1, p_2, \dots, p_n , respectively (in the continuous case, X can vary between $-\infty$ and $+\infty$). The sum $\sum p_i$ of those probabilities must be 1. And it holds that

$$\mathbf{P}(X) = x_1 p_1 + x_2 p_2 + \dots + x_n p_n. \quad (4.12)$$

Fig. 4.2 Representation of prevision as the barycenter among masses placed on a bar of uniform density



To put it in more familiar terms: $\mathbf{P}(X)$ is nothing but the arithmetic mean of the values that X can take, weighted by the values of the respective probabilities. In this case too it turns out to be helpful to resort to a mechanical representation, with masses x_1, x_2, \dots, x_n arranged on a bar (Fig. 4.2), whose distance from the center of mass is p_1, p_2, \dots, p_n , respectively. The prevision $\mathbf{P}(X)$ corresponds to the abscissa of the center of mass. The center of mass, in turn, is the point at which the bar must hang in order to be in equilibrium. This mechanical interpretation provides the most elementary way of explaining probability and prevision. I have already said on another occasion⁷ that in many texts — especially in the more dated ones — prevision was called “mathematical expectation.” Whether it is a positive or a negative expectation, this depends on the particular cases. But when it is negative, the term “expectation” seems inappropriate: for instance, the expression “mathematical expectation of death” is as funny as it is out of place.* I therefore prefer the word “prevision.”

The centre of mass has another property, which is useful in statistics and in the theory of probability: the centre of mass, in fact, besides being the point of equilibrium among the masses (that is the sum of the products of masses divided by the distance from the fulcrum) has yet another meaning. Is anyone among you able to recall it? Even if you should fail to recall it from your rational mechanics classes, it should follow from what I have said on proper scoring rules and in particular, on Brier’s rule. Brier’s rule, in fact, allowed us to find the centre of mass as . . .? What was minimum in the centre of mass?

AUDIENCE (chorally:) The moment of inertia.

DE FINETTI: Correct. Thus, there is a two-fold way to define prevision. It might seem that the one based on the moment of inertia is much more artful, but this is true only if we think of prevision without bearing in mind its more interesting properties. Brier’s rule can be used to find the center of mass without other information being available besides the penalizations relative to the various points. About them we know that they are proportional to the square of the distance from the centre of mass. If we do not know the centre of mass, we could think of determining it by rotating the bar about a point; by repeating the experiment with distinct points, we can determine the point at which the moment of inertia becomes minimum: that point is the centre of mass. In this way we have the advantage of determining the centre of mass directly, without confronting weights and distances. And it is instructive to realize that prevision can be represented as a centre of mass. And it is also important to be able to understand (by means of the probabilistic translation

* (Translator’s note:) As already mentioned above, the Italian expression for “mathematical expectation” is *speranza matematica*, literally: “mathematical hope.” This makes de Finetti’s remark particularly effective.

that can be made of it) that to ask what is the point at which a bar has to be hung so that it can stay in equilibrium is equivalent to asking which point it should be rotated about in order to have the maximum live force at the same velocity.

Inexpressible Evidence

ALPHA: I would like to ask a question concerning Bayes' theorem. A common objection raised against the thesis that this theorem provides the explanation for any probability updating in the light of new information is based on the observation that experience is so complex that it cannot always be expressed in the form of a proposition.⁸ Hence, not every bit of experience can be considered a well-defined event. Yet Bayes' theorem is only applicable to those cases in which the new evidence is a well-defined event. Therefore — the objection concludes — there are cases in which one's opinion is subject to change in ways which, despite not being incoherent, cannot be described by means of Bayes' theorem. In other words, our own experience is an extremely complex phenomenon and we can only have an undetermined and inexpressible experience of certain things whose relevance cannot, therefore, be captured in terms of Bayes' theorem.

DE FINETTI: That sort of experience can be taken into account as well, provided that the axioms — if we want to talk in axiomatic terms — remain unchanged. Otherwise we go off the track.

ALPHA: And what if, in time, we change our probability function? Is it possible to change it?

DE FINETTI: Of course, actually in general we ought to, for we continuously receive new information.

ALPHA: No, suppose that we changed it essentially, not simply by normalizing it according to the new state of information. That is, let us suppose that a person said: "If I were to go back to the previous state of information, I would no longer have the probability I had before." Is this possible?

DE FINETTI: Yes, in theory But if we say "state of information" I would not know

GAMMA: It seems to me that what ALPHA is saying can be interpreted in the following way. Bayesian statistics formalize the process of revising one's own state of information but with the proviso that what is initially considered to be true and absolutely certain must remain so throughout the entire process. Yet a situation could arise (and indeed this often happens in reality) in which the person, as a consequence of an experiment, changes her mind completely.

ALPHA: Let us suppose that we are to make a bet conditional on a hypothesis H . Let us suppose further that H will be verified and that no other information is available. Is it possible to "repent" of this bet after learning that H was verified?

DE FINETTI: There is nothing of which one cannot repent! Moreover, it must be said that a situation might change only by virtue of the fact that time goes by. Even if a high probability was attached to the possibility that a given person X would die before 1 January 1979, by the time we get to the eve of that date, it is likely that

the probability will be very small, no longer close to 1. And should X be still alive on 2 January 1979, the event at issue would no longer be possible. Some seeming violations of Bayes' theorem arise from the fact that some elements that are relevant to the situation are neglected. If they were made explicit, the accordance with the theorem would be restored. But it would be unrealistic to require that the entire state of information had to be made explicit, for, in order to do that, more time would be needed than what has elapsed since the beginning of the geological era. An ideal person should accurately take into account all the information she possesses. A person in the flesh makes her own evaluations according to what she remembers and is in a position to use. Furthermore, there are so many things that one knows, but which nonetheless one considers irrelevant in determining a certain probability. And there are things that a person deems known to her with insufficient accuracy (and hence she hesitates to give an evaluation of them).⁹ I believe that we should introduce the following pre-axiom: *all that is required of a probability evaluation is its coherence at the particular moment in which the evaluation is made, in a particular state of information.*¹⁰ But the state of information does not change the way a computer stores the data that are provided for it! I am an example of the opposite, for I forget many things just a moment later and then I become angry and say: "what was I thinking of? I have found something that appeared to be right and now I do not have a clue where I should look for it!" We cannot omit to consider the state of information of the person who evaluates the probability. And this state varies from one moment to another. At a given moment, one could feel more optimistic, in another more pessimistic. A tiny obstacle with no importance could be enough to change our mood. This is something that pertains to everything and, to an even greater extent, to probabilistic reasoning. If one did not accept to reason with the very limited resources that an individual has (in terms of memory, analytic ability or the capability of connecting all those things that can have a connection, etc.), then nobody could say anything. There are some debatable issues but they are situated at a less immediate level. At the immediate level, nothing is exact.

ALPHA: On this point, your thought is similar to Carnap's, who said that we should think of the project of a robot as a purely ideal case: men will get close to such a project to the extent that their limited resources will allow them.¹¹

DE FINETTI: On Carnap I should say that many fluctuations are observed in his writings (not in the logical part, but) in the way he looks at practical things. For instance, he puts forward the consideration (it is not only his own but rather a general one) that there are many subjective factors intervening in the evaluation of probability. It can happen, for instance, that a person who learns that a flat (of an acquaintance of hers, or say in the opposite building) has been destroyed by fire, even though she previously never thought of obtaining insurance against fires, decides to do so. Those facts are extremely important but I do not think that any formula helps to explain them.

ALPHA: Carnap has modified his positions during his lifetime. In the fifties he was deeply influenced by the position of Keynes after reading his *A Treatise on Probability* (Keynes [1921] 2004). Keynes denied that a unique probability function, adequate for any context, could be determined. Carnap, on the other hand, believed

that this was possible. At that time, Carnap was in a certain sense an objectivist, since he believed that probability values were logically determined. But, later on, when he realized that his programme of identifying a unique probability function could not get any further, he ended up admitting the presence of subjective factors. GAMMA: Carnap used to distinguish between probability₁ and probability₂.

ALPHA: I'm speaking of probability₁, which is probability intended as a rational degree of belief and not as a frequency. Carnap attached a dignity of its own also to the other concept: this is, however, another question altogether. Latterly, Carnap insisted significantly on the idea that probability is the "guide to life," in connection with decision theory. Hence his revival of those theorems on the theory of bets (his own and Ramsey's).¹² Carnap believed that it was possible to justify the axioms of probability theory by restricting the field of probability evaluations by means of certain principles of rationality. Among those is the principle of *coherence*, which forbids the disposition to lose with mathematical certainty in a determined betting framework. Carnap took those results with enthusiasm and hoped to go on in that direction. The goal was for him to put more constraints on subjective probability evaluations (besides the one of coherence). Through certain axioms — though some fail to be entirely convincing — he managed to restrict the domain of admissible evaluations up to confining it to a unidimensional continuum of inductive methods (characterized by the famous λ parameter). But Carnap, at a certain point, said (I quote from memory) "right now I cannot say whether it is possible to restrict the domain of probability functions beyond a certain limit or whether we shall always need to admit uncountably many of them."¹³ The divergence from his position seems to me to be essentially this: that even though you maintain a similar normative position, you only ask for the fulfilment of the condition of coherence. Carnap, on the other hand, believed that further rationality requirements existed.

DE FINETTI: Do such principles have a mathematical nature?

ALPHA: It is always possible to provide a mathematical translation of principles of this sort. For instance, the so-called symmetry axiom is the equivalent of exchangeability (although it is formulated in such a way as to refer to the totality of information).

Editor's Notes

1. See Chapter 17 for more on this theme.
2. Notice that de Finetti uses ' \mid ' as a logical connective. This involves a departure from classical bivalent logic with the introduction of a third truth-value ' \emptyset ' meaning "null." An event of the form $E \mid H$ (called 'trivalent' by de Finetti) is considered true if it is true that EH , false if it is true that $\bar{E}H$ and null if it is true that \bar{H} .

This *prima facie* seems to be at odds with the celebrated *Lewis's triviality results* (Lewis [1976] 1986, p. 139), according to which no iterable connective (let alone a truth functional one) is compatible (except in trivial cases) with the behaviour of ' \mid ' as determined by the axioms of finite probability. There are several variants of Lewis argument. In a nutshell Lewis remarked that the two following formulas (a) $(X \mid Y) \mid Z = X \mid YZ$ and (b) $\mathbf{P}(H) = \mathbf{P}(H \mid K)\mathbf{P}(K) + \mathbf{P}(H \mid \bar{K})\mathbf{P}(\bar{K})$ can be satisfied simultaneously only in trivial cases of no interest

(cf. Jeffrey, 2004, p. 15–16). Now, while (a) is undoubtedly a necessary property of ‘|’ as a connective, the property (b) holds for ordinary events but not in general for trievents (ordinary events being a special subclass of trievents). In the case of trievents a generalized formula holds instead of (b), which coincides with (b) when both H and K are ordinary events. For further details on de Finetti’s trievents, see Chapter 17, on pages 207–210 and Mura (2008).

3. Jerome Cornfield (1912–1979) was an important statistician who made relevant contributions to both theoretical and applied statistics. The unpublished studies that he developed in 1941 on the optimization of diets — the “diet problem” as he labelled it — foreran the development of linear programming.

Beginning with the early 1960s, Cornfield concentrated his efforts on developing Bayesian statistical techniques (to be applied especially in the biomedical field), strongly influenced by de Finetti, Savage, Jeffreys and Lindley. He wrote a landmark paper (1969) devoted to defending the basic Bayesian tenet according to which “any inferential or decision process that does not follow from some likelihood function and some set of priors has objectively verifiable deficiencies” (ibid., p. 617).

In that paper — which turned out to be extremely influential in promoting the Bayesian viewpoint among the biomedical statistical community — Cornfield rehearses — among the other considerations — the Dutch book argument and related theorems along the lines of de Finetti [1937] 1980. For further details about Cornfield’s contribution to Bayesianism see Zelen (1982).

4. This interpretation of Bayesian conditionalization is connected with the conception of probability calculus as the “logic of the uncertain.” According to this conception, probability laws are logical laws and hence, every probability evaluation obtained from other probability evaluations by the calculus of probability is nothing but the result of an analytic transformation, by means of which the initial opinions are simply put into a new form. In Mura (1995b) I have tried to defend, on new grounds, this point of view. See note 14 on page 125.
5. See Good (1965). It should be noted that for I. J. Good the term ‘ad hoc’ does not always have a derogatory import. Commenting on de Finetti’s paper (1969), Good (1969, p. 23) writes: “de Finetti is right when he says that I reject ‘ad hoc’ theories,” but only when perfect rationality is aimed at. I distinguish between rationality of Type 1, which involves complete consistency and rationality of Type 2 in which an attempt is made to maximize expected utility allowing for the cost of theorizing. . . . In practice it is usually necessary to use rationality of Type 2: the best we can usually hope for is consistency of judgments and discernments as far as we know at a given moment of time.”

De Finetti acknowledged this. In fact, in note 2 of de Finetti (1969) (added in proof), he wrote “Good’s position, as I gathered it from his talk at the Salzburg Colloquium, is less radical than I supposed. According to it, ‘ad hoc’ theories ought not to be rejected outright; their use may sometimes be an acceptable substitute for a more systematic approach.”

6. Cf. Chapter 1, p. 1.
7. In Chapter 2.
8. Allusion to the positions maintained by Richard C. Jeffrey ([1965] 1983). See also note 9 below.
9. In the context of this discussion, two widely debated problems emerge: the problem of uncertain evidence and the problem of indeterminate probabilities.

The former problem derives from the fact that the conditionalization process does not seem to take into account the relevance of clues of the form: “it seemed to me that . . .” For instance, let us suppose that a witness X , during the judicial trial concerning a certain murder, should state to be 0.9 sure that she can recognize the photographic picture of a certain person Y . Let H be the event that Y is actually the photographed person and let E be the event that Y is the murderer. Suppose that if H is true, then the judge’s subjective probability about E would be $\mathbf{P}(E | H) = 0.7$ and $\mathbf{P}(E | \bar{H}) = 0.1$ if the other way around. Suppose also that the judge relies on the probability assessments made by X , so that after X ’s testimony, the judge’s degree of belief about H is $\mathbf{P}_{\text{new}}(H) = 0.9$ (regardless of her previous subjective probability

borne before the testimony). How should the judge conditionalize on such uncertain evidence in order to evaluate her update probability $\mathbf{P}_{new}(E)$ that Y is the murderer? This problem has been solved by Richard C. Jeffrey ([1965] 1983), who proposed a generalized version of Bayes theorem, which is also capable of reconstructing this kind of change in probability evaluations like Bayesian conditionalization.

Jeffrey's solution (called *Probability Kinematics*) is given by the following formula:

$$\mathbf{P}_{new}(E) = \mathbf{P}(E \mid H)\mathbf{P}_{new}(H) + \mathbf{P}(E \mid \tilde{H})\mathbf{P}_{new}(\tilde{H})$$

If, instead of two alternatives H and \tilde{H} , there is a partition of n alternatives H_1, \dots, H_n , the foregoing formula easily generalizes to the following one:

$$\mathbf{P}_{new}(E) = \sum_{i=1}^n \mathbf{P}(E \mid H_i)\mathbf{P}_{new}(H_i)$$

Probability kinematics reduces to conventional Bayesian conditionalization when, for some i ($1 \leq i \leq n$), $\mathbf{P}_{new}(H_i) = 1$. It may also be further generalized to more complicated situations (for example, in the previous example, supposing that the judge does not completely rely on the witness' probabilities after her testimony, or supposing that there are multiple witnesses and so on (Jeffrey, 2004, pp. 53–61)).

The latter problem derives from the fact that sometimes people do not seem to be in a position to provide pointed probability evaluations, since they are rather undecided about the probability that has to be attached to events. In order to represent this behaviour mathematically, many have resorted to the idea of representing degrees of belief through intervals rather than pointed values (I can mention, among the others, Good ([1962] 1983); Koopman (1940); Smith (1961); Dempster (1967); Levi (1974); Suppes and Zanotti (1989)). De Finetti has opposed this framework by dubbing it a useless and illusory attempt to eliminate the inevitable idealizations embedded in any mathematical scheme (de Finetti and Savage, 1962).

10. According to de Finetti thus, Bayes' theorem, *qua* device for updating degrees of belief on the basis of the evidence collected, may be applied only in canonical contexts (like Bayesian statistics), where a conditional bet before the piece of evidence is acquired may be made in principle. In real life, the Bayesian updating model has a limited application, while the coherence constraints apply normatively without exception to the probability assessments made at a given time.
11. See Carnap (1971b, p. 508). Indeed, de Finetti explicitly criticized the Carnapian reference to the idea of a "perfectly rational" robot (de Finetti and Savage, 1962, p. 89n).
12. Allusion to the celebrated *Dutch book* argument, according to which, roughly speaking, assessments of probabilities that violate the axioms of finitely additive probability are "incoherent," since they entail the disposition to accept a set of bets that leads to a sure loss. In this schema it is assumed that the subjective probability $\mathbf{P}(E \mid H)$ of an event E conditional on H borne by a person X is measured by the so-called *fair betting quotient*. A conditional bet with respect to E conditional on H is a situation under uncertainty in which X has to choose a real number p and, after making this choice, will have a gain $s(1 - p)$ ($s \in \mathbb{R} - \{0\}$) if both E and H obtain, a gain $-sp$ if H obtains but E does not and a null gain if H does not obtain. Formally, it may be considered a 5-tuple $B = \langle E, H, X, p, s \rangle$. The amount s is called the *stake* of the bet. If s is a positive number, B is called a bet *on* ($E \mid H$). If s is a negative number, B is called a bet *against* ($E \mid H$). When H is the necessary event T , the bet is called *absolute* and its quotient may be regarded as a measure of the absolute subjective probability $\mathbf{P}(E)$. The number p is called the *betting quotient* of the bet. A bet $B = \langle E, H, X, p, s \rangle$ is said to be *fair* according to X if X is indifferent between B and the bet $B' = \langle E, H, X, p, -s \rangle$. If $s > 0$ then B' is a bet against ($E \mid H$) with the same quotient as B . If B and B' are fair bets, then p represents the conditional subjective probability $\mathbf{P}(E \mid H)$ borne by X about E supposing that H obtains. It

is assumed that the fair betting quotients are independent of the stakes s . This last assumption is only loosely correct (and requires the further proviso that the stakes are small with respect to the total fortune of X) when the gains are expressed by monetary amounts not linear in utility.

Suppose now that X assigns her fair betting quotients q_1, \dots, q_n to certain events E_1, \dots, E_n conditional, respectively to H_1, \dots, H_n . It is assumed that X is prepared to accept any simultaneous set of fair bets. A *book* is just a finite set $\mathbf{B} = \{B_1, \dots, B_k\}$ of bets. The gain \mathbf{g} of \mathbf{B} is the sum of the gains $\sum g_i$ of the bets B_1, \dots, B_k . In general, \mathbf{g} depends on the truth-values of E_1, \dots, E_n and H_1, \dots, H_n . When for every assignment of truth-values to all the events $E_1, \dots, E_n, H_1, \dots, H_n$, \mathbf{g} is strictly negative, \mathbf{B} is called a *Dutch book*.

The Dutch book theorem, as proved by de Finetti, asserts (a) that if the fair quotients q_1, \dots, q_n violate the laws of finitely additive probability, then there exists a Dutch book and (b) that no Dutch Book exists if q_1, \dots, q_n are in accordance with the laws of finitely additive probability. According to de Finetti, (a) shows that the violation of the laws of finitely additive probability entails an inconsistency, since the certainty of a negative balance involved by the existence of a Dutch book is at odds with the alleged fairness of all the quotients q_1, \dots, q_n . By a different interpretation of the same result, also endorsed — albeit secondarily — by de Finetti, the theorem proves that a behaviour based on subjective degrees of belief violating the laws of finitely additive probability is *irrational*, involving a ruinous disposition that may be avoided by complying with the laws. On the other hand, (b) shows that no strengthening of the formal axiom system of finitely additive probability may be justified by a Dutch book argument. And since de Finetti maintained that only probability assessments ruled out by such argument are unacceptable as inconsistent or “incoherent,” every additional axiom would exclude possible subjective degrees of belief and, therefore, should be rejected.

It should be stressed that for de Finetti, the Dutch book argument is entangled with an operationalistic stance. Degrees of belief borne by a person X are measured by X 's betting behaviour. Such a measure is based on bet quotients declared by X . However, according to de Finetti, a valid operationalistic procedure must not simply rely on the intellectual honesty of X . It is perfectly possible that X , asked to tell what is the betting quotient that she considers as fair about a certain event, does not provide the right answer. It is assumed that only when X has to face a real bet, her answer (which in this case assumes a performative character) may be reliably elicited. The experimental set-up should be such that the optimal decision for X (assuming that X aims at maximizing her gain) is to declare the actual fair betting quotients borne by her. Now, proposing to X a bet *on* ($E \mid H$) at the quotient declared by X herself does not work: in such a situation X would prefer to declare a lower quotient than the fair one. The same may be repeated for proposing to X a bet *against* ($E \mid H$): in such a case X would prefer to declare a number greater than the actual fair quotient. In an attempt to avoid such inconveniences, de Finetti invented a measurement schema where after X has declared her fair betting quotients, an opponent Y chooses the stakes. In particular, Y decides, for every i , whether X would bet *on* ($E_i \mid H_i$) or *against* ($E_i \mid H_i$). Unfortunately, also with this caution, de Finetti's procedure is not reliable, even if we assume that X and Y make their decisions on the basis of the same evidence. The reason is that X may bear expectations about Y 's degrees of belief, which may induce her to declare quotients that are not actually considered as fair. There is no way to escape this difficulty. De Finetti's betting experimental set-up put the bettor in a game-theoretic situation that may prevent the “sincere” elicitation of subjective probabilities. This is the main reason why de Finetti eventually gave up the betting framework and resorted to proper scoring rules (where the presence of a troubling opponent is missing). On this point see Chapter 12 on page 114 and note 3 on page 123.

The Dutch book theorem, in its original formulation, is about sets of conditional bets and provides an argument for requiring that conditional degrees of belief comply with the axioms of finitely additive probability. To apply this result to Bayesian conditionalization, we have in addition to assume that if at time t_0 the subjective probability of an event is $\mathbf{P}(E)$ and in the time interval $(t_0, t_1]$ the total evidence available to X passes from K_0 to $K_0 \wedge H$, X 's subjective probability on E at time t_1 must equal $\mathbf{P}(E \mid H)$. To prove this crucial assumption, a *dynamic*

version of the Dutch book was presented by Paul Teller, elaborating an idea originally due to Davis Lewis (Teller, 1973, pp. 222–25). In the Teller-Lewis schema the opponent of X may, in addition to deciding the stakes of every bet, plan sets of bets at different times (called, after van Fraassen (1984, p. 240) *betting strategies*). Violation of conditionalization entails with certainty a negative final balance for some betting strategy. A similar result, due to Armendt (1980), applies to Jeffrey's Probability Kinematics (cf. the foregoing note 9 on page 42).

The meanings of these results have to be taken with caution (see Christensen, 1991; Howson, 1993). To maintain the focus on conventional conditionalization, it should be recalled that it requires the proviso that $\mathbf{P}(E \mid H)$ at time t_0 equals $\mathbf{P}(E \mid H)$ at time t_1 (cf. Jeffrey, 2004, p. 52). This assumption seems to be straightforward if (a) the only novelty in the state of information of X at time t_1 is the passage from a total evidence K_0 to a total evidence $K_0 \wedge H$ and (b) at time t_0 $\mathbf{P}(H) > 0$.

It should be recalled that unlike other Bayesians, de Finetti did not require that a person that changes her mind about the probability of certain events is rational only if that change is the result of conditionalization. The change may occur on second thoughts even if the state of information remains the same (see above in this chapter on pages 39–40 and note 10 at page 43), so that de Finetti required dynamic coherence only under the provisos we have made explicit above. Moreover, according to de Finetti, a person may revise her probability evaluations by “repenting” of her evaluations held at time t_0 , so that her evaluations at time t_1 are derivable by Bayesian conditionalization from those priors that, at time t_1 , X maintains she ought to have borne in mind at time t_0 on the basis of the total evidence K_0 . Notice that this regret may be occasioned by the evidence collected in the time interval $(t_0, t_1]$ (see Chapter 8 on pages 78–79).

The origin of the expression ‘Dutch book’ is unclear (see Wakker, 2001a). I conjecture that it is related to the history of the Lotto game and, in particular, to the introduction in the Low Countries, at the beginning of the 16th century, of the so-called “Dutch Lotto,” in which in contrast with the so-called “Genova Lotto” — see note 8 on page 147 — the number and price of tickets as well as the number and amounts of prizes were fixed in advance, so that the organizer had, in any event, a positive gain (see Willmann, 1999).

For further details on the Dutch book argument see the Introduction to this volume by Maria Carla Galavotti on page xix. For a detailed analysis of the argument and different versions of it see Mura (1995b). The literature on the Dutch book argument is extremely extensive. The following list is limited to some of the most significant contributions: de Finetti ([1931] 1992); Ramsey ([1926] 1990); Carnap ([1962] 1971); Howson and Urbach [1989] 2005; Kemeny (1955); Lehman (1955); Shimony (1955); Adams (1961); Baillie (1973); Davidson and Pargetter (1985); Kennedy and Chihara (1979); Armendt (1980); Milne (1990); Berti, Regazzini, and Rigo (1991); Jeffrey (1992b); Howson (1993); Christensen (1996); Maher (1997); Titiev (1997); Vineberg (1997); Waidacher (1997); Hild (1998); Schervish, Seidenfeld, and Kadane (1998); McGee (1999); Silber (1999); Schervish, Seidenfeld, and Kadane (2000); Lindley (2000); Döring (2000); Cubitt and Sugden (2001); Corfield and Williamson (2001); Christensen (2001); Paris (2001); Border and Segal (2002); Coletti and Scozzafava (2002); Diecidue and Wakker (2002); Eaton and Freedman (2004); Seidenfeld (2004); Borkar, Konda, Sachs, and Mitter (2004); Hájek (2005); Bradley and Leitgeb (2006); Hosni and Paris (2005); Greaves and Wallace (2006); Diecidue (2006)

13. “We do not know today, whether in this future development the number of admissible functions \mathcal{M} will always remain infinite or will become finite or possibly even be reduced to one.” (Carnap, 1971a, p. 27).

Chapter 5

Physical Probability and Complexity*

Discussion (arisen on the initiative and with lively participation of the attendees of the course) of the point of view of some philosophically oriented authors (mentioned, in particular; Carnap, von Mises, Popper, Jeffreys, Cantelli, Reichenbach, and so on).¹

“Perfect” Dice

ALPHA: When did you understand that probability is subjective?

DE FINETTI: When I was an undergraduate student, probably two years before my graduation, while reading a book by Czuber, *Wahrscheinlichkeitsrechnung* ([1903] 1914)). Czuber was a Czechoslovak-born mathematician, though of Austrian nationality. If I am not wrong, he used to be a professor in Vienna.² That book briefly pointed to the various conceptions of probability, which were very hastily introduced in the initial paragraphs. I cannot now recall well the content of the book, either in general or with respect to the various conceptions of probability. It seems to me that it mentioned De Morgan as the representative of the subjectivistic point of view.

ALPHA: Augustus De Morgan?

DE FINETTI: Yes. If I remember well, De Morgan’s work, to which Czuber referred, was published in 1847 (in any case around the half of the Nineteenth Century).³ Confronting the various positions, I seemed to realise that every other definition was meaningless. In particular, the definition based on the so-called “equiprobable cases” seemed unacceptable to me. Although I found it natural to give to the distinct sides of an apparently “perfect” die the same probability, I could not see how one could give an objective meaning to probability on these grounds. The physical symmetry of the die looked to me as a circumstance that could explain why each individual attributes the same probability to the various sides of the die. But — I thought — there could be a thousand reasons to make an exception. For example, if one discovered that there was an imperfection in the die, or if one were influenced by the fact that after casting the die, frequencies distant from 1/6 have been recorded, ascribing this to some alleged imperfection of the die rather than to chance. Examples of this sort can be subjectively interpreted as situations in which one tries to give an objective justification for subjective opinions. This does not

* Lecture VII (Tuesday 27 March, 1979).

suffice to make the concept of probability objective: in fact, it can well happen that the distinct sides of a “perfect” die will show up with very disparate frequencies in a series of trials, without one being able to discover the reason for it being so. I therefore realized that a subjective component is always present, which consists in thinking of there being or not a cause — to make a concession to a philosophical terminology I distrust — to which such an unbalanced result needs to be ascribed. And it seemed to me that they meant to make something more than objective out of those subjective judgments by interpreting them as judgments that do not depend on the opinions of the individuals (opinions that moreover can be reasonable and, as such, shared more or less by everyone). It was taken for granted that those judgments, even though nobody would have ever shared them, would have been true *per se*.

As to the other points of view, I intended to say something about them today. Clearly each objectivistic conception consists in appealing to some objective data. So far, nothing wrong: it is something everyone should do, especially subjectivists, to avoid talking nonsense. But objectivists think that objective data, instead of being a piece of evidence of a circumstance that helps one forming an opinion, constitutes just the essence of probability.

ALPHA: It is not necessarily the case that an opinion is “true” just because it has been argued for by means of objective data.

DE FINETTI: Yes indeed. And there is no big difference if we look at the issue from an empirical point of view. One is not allowed, in fact, to say: “these are the mathematician’s or the philosopher’s quibbles, there is no need to be so subtle.” The reason is that there is a profound difference between those who see in objective circumstances reasons to argue for some probability values and those who pretend to define probability in terms of those circumstances instead. It must be added, however, that it is not true that on the one side of this watershed everything is objective whilst on the other everything is subjective. But if we take as objective what — if grounded on considerations which may well be objective and reasonable — is but a subjective judgment, we lose the possibility to examine, for any single case, all its concurrent circumstances. Moreover, in such a case, a person who makes a probabilistic judgment, would not be responsible for it.

Some try to turn the judgement of probability into something objective through the concept of frequency. For instance, instead of saying: “this is my evaluation,” they say: “this is the result of a series of a thousand tosses of a coin.” But this result is a *fact* that happened only once: a series of heads only or of tails only could have likewise happened. It is reasonable to attach a very small probability to such an outcome: but likewise reasonable would be to attach a small probability to any other succession of a thousand tosses, the one which actually occurred included. The identification of probability and frequency is an attempt to give an objective meaning to probability by means of useless complications. This is, however, a path which has been followed by a number of authors, among whom we must mention von Mises, Reichenbach and Borel (though the latter has not written many genuinely philosophical things).

The Lottery Paradox

ALPHA: Is there only that article on the work of Keynes, by Borel, which is published in the anthology edited by Kyburg?⁴

DE FINETTI: Borel wrote various things. One upon which he insisted was (once I had an argument on this point with him at the *Institute Poincaré*) that an event of very small probability is an event that *cannot* take place. He used to say (and his arguments were right in themselves): let us consider the probabilities 10^{-3} , 10^{-10} , 10^{-100} and 10^{-1000} . A probability of 10^{-1000} is roughly equal to the probability of picking by chance a particular atom in the entire universe. Of course it is such a small probability that . . .

ALPHA: This argument by Borel is similar to Popper's. Popper maintains the thesis according to which scientific theories are such only when they can be falsified. Indeed, according to him, if there is no possibility for a theory to be controlled empirically, then it is metaphysics (and this he said without any sort of contempt towards metaphysics). Herein, however, arises a difficulty. Those theories which are under a probabilistic form (Thermodynamics, Statistics, etc.) can never be refuted (even from the frequentist point of view, which Popper has defended in the first phase of his thought): after all, even an event of probability 0 can well occur. According to Popper we then have to take a methodological decision and consider a probabilistic theory *T* refuted if a certain event occurred which, as far as *T* is concerned, would otherwise be extremely unlikely. It is the scientific community that fixes, by convention, the refutation threshold. Let us look at some examples. We know that in the light of statistical mechanics it is very unlikely that the totality of the air present in this room will gather under the table. And we know that in the light of the second law of thermodynamics, it is extremely unlikely that if we filled-in with warm and cold water two adjacent compartments separated by a conductive wall, the temperature of the warm water would increase whereas that of the cold water would decrease. Nonetheless those are events that statistical mechanics does not rule out. Strictly speaking, thus, phenomena of this sort are compatible with statistical mechanics. If one day one of them should occur, it could rightly be said: "it happened by chance, which — after all — could have happened." Therefore, this would not make a case for rejecting statistical mechanics. However, suppose that the phenomenon occurred repeatedly during an attempt to test the validity of statistical mechanics. Suppose that ten experiments have been conducted and each time one of those phenomena occurred. A case like that, Popper says, is so unlikely in the light of statistical mechanics that we should consider that theory refuted. According to him, we cannot say: "unlikely as it may be, in the light of the theory, that event was compatible with the theory and hence, the theory is in accordance with experience."⁵

DE FINETTI: There is no logical necessity to take one position or the other. We could simply say: "in my opinion, this event occurred by chance." But when an extremely unlikely — according to the accepted theory — event is considered to be random, we simply mean that we are not considering this fact as such a serious anomaly that one should expect it to occur quite frequently in the future.⁶

ALPHA: A possible phenomenon can always repeat itself, infinitely many times. Even if a strange phenomenon occurred in the past for a million times in a row, it *could* always have happened just by chance.

DE FINETTI: It seems to me that the difference between facts that are certain to happen and facts that are uncertain to happen — and likewise, the difference between possible facts (though of very small probability, even though of probability 0) and impossible facts — is enormous, since it is a qualitative difference. To use a situation familiar to those who set things in terms of measure theory — as many do as a matter of principle, though doing so does not seem to me reasonable — a 0 probability is equal to the probability of hitting exactly a point with an arrow whose extremity has zero thickness — provided that it makes sense to speak of hitting a point in the mathematical sense. But if one said: “since each point has 0 probability of being hit, the probability of any point being hit by an arrow is in turn 0,” one would say something absurd.

ALPHA: This is the lottery paradox . . .

DE FINETTI: Yes indeed. It is a fictitious paradox, for it is a consequence of an attempt to transform probability into certainty. To say, with Borel, that since the probability of picking a certain particle in the universe is 10^{-1000} it is *impossible* to pick exactly that one, is to use an argument that could be repeated for any other point. Thus it would be impossible to hit any of the points. But suppose that I am just about to fire an arrow: what would happen? Indeed, if it is impossible to hit any point, because the probability that it will be hit is null . . .

ALPHA: The probability of hitting at least one point is 1.

DE FINETTI: We cannot, thus, confuse the vacuum with the minimum that we can think of. We cannot confuse a geometric point with a bunch of circumstances, however complicated they might be.

Probability as Frequency

There is usually another way to make probability objective, which consists in resorting to the concept of frequency. The latter is often also taken as a *definition* of probability. By saying that probability and frequency are one and the same, one makes a very ambiguous statement. The following example will make the reason for this clear. Suppose that a person enjoys herself by counting all the times that a certain number k is drawn in the *lotto*. Suppose that after a considerable number of drawings, that person finds out that k came out in exactly 1% of cases. If probability and frequency were the same thing, that person should say that the probability of drawing the number k is 1%. On the other hand, a person who attached probability $1/90$ to every single drawing should predict that the frequency will be $1/90$. But it would be very unlikely if the frequency turned out to be, for each number, exactly $1/90$. If we said that *on average* it is $1/90$ (by averaging out all the numbers), this would certainly be true but in a trivial way for, the numbers being 90, the mean of the frequencies would always be $1/90$.

There are many variants of frequentism, none of which seems to me in a position to overcome these difficulties. If one carried out a great number of trials and

arranged them in a sequence according to whatever ordering — be it temporal or not — or if one took the sub-sequences (for it might well happen that there are two individuals of whom one assisted in every trial whilst the other did not), one should expect that almost all of those sequences agreed on the fixed probability. Nonetheless, even if this happened, we could hypothetically think of an individual who, by chance, looks at the die only when 2 shows up.

ALPHA: In this context the problem of reference classes emerges, which gives rise to enormous complications.

DE FINETTI: Yes, those are complications that just contribute to moving the difficulties of frequentism from one step to the next one.

Probability and Physical Laws

ALPHA: There is another objection against probability as a degree of belief. Carnap, who also accepted the frequentist concept, said: if probability was always subjective then physical concepts like entropy would become psychological, in which case laws such as the law of entropy would become psychological laws, a conclusion that looked unacceptable to him.

DE FINETTI: The interpretation that a person gives to phenomena does not influence them. That certain phenomena, for instance that the motion of gas molecules, follows (if only statistically) the law of entropy, is a fact that can be observed. We could say (as actually many of those who do not split hairs say) that it is *sure* that entropy will increase. Even without entering into the subject of Physics (after all I am not a physicist), let us suppose that we mix up two dusts (for instance to make a dye) by blending some red dust with some brown dust. After mixing well, the result will be a mixture of a colour that is at first sight uniform. However, if we then looked at the mixture with a magnifying glass, we would see grains (or even small pieces) of distinct colours. In my opinion what seems to be the point here is just this, *there is nothing in Nature that specifically tries to realize a mixture that looks macroscopically uniform*.

ALPHA: But the law of entropy comes in a probabilistic form. It involves probability. What sort of probability is that? Can a physical law refer to subjective probabilities? How should the law of entropy be interpreted?

DE FINETTI: Probability should not be considered as the reason why the fact takes place in a certain way.⁷ Rather we must say: the fact takes place in such a way as to suggest — for the way in which it takes place — the attribution of a higher probability to certain events. If we believe that there is no reason to the contrary (for instance difference in the density between parts or grains of different colour) then every configuration is possible and almost every possible configuration (except for a small number) is such that the two colours appear more or less uniformly mixed. It is therefore reasonable to give a high probability to the fact that a uniform colour will be observed.

ALPHA: Hence, the probability distributions brought to us by statistical mechanics are suggestions about how to obtain certain subjective probability opinions.

DE FINETTI: I would say they provide more solid grounds for subjective opinions. By looking at the outcome of a phenomenon we could be driven to formulate a rule by virtue of which, in each case, things would blend in that way, as if it were a necessary law of nature. On the other hand, by means of those theories, we can show that the occurrence of certain macroscopic circumstances is rather natural, as they correspond to a great quantity of microscopic “equally probable” cases (this is a dangerous expression, which is to be understood in the sense that among those cases there are no privileged ones, so that a uniform distribution of them looks reasonable).

ALPHA: You say: from a deterministic theory like classical mechanics, one can obtain, by means of certain reasonable assumptions, laws which are formulated in terms of subjective probabilities. Therefore, those laws are not objective *laws* of nature. Rather, they are *rules* to give, in the light of a deterministic theory, reasonable probability evaluations.

DE FINETTI: This is, more or less, my position.

ALPHA: Let us now consider an essentially non-deterministic theory like quantum mechanics. In this case there is no deterministic theory from which statistical conclusions can be derived. Quantum mechanics is a theory that has probability at its core, in a way that cannot be related to reasonable distributions. Do you think that the probability of quantum mechanics is also subjective?

DE FINETTI: What would we mean by saying that it is objective? At the end of the day, what we call, for instance, ‘the probability of an impact’ is nothing but the degree of expectation we have, in the light of the many experiments that lead to the construction of a certain theory, of a certain impact taking place. Those experiments may either consist in the observation of individual cases, or they may be constituted of statistical observations: in either case there is no reason to believe that the observed regularities (even though an adequate explanation of the latter is not available to us) will not occur again. It must, however, be kept in mind that all this depends on the theorizing we do about the phenomena, which is subject to change. Although the theory speaks of “non-deterministic facts,” or of facts that are called ‘impacts’ just by convention, it is all about mathematical operators in a field.

GAMMA: We should distinguish better between ‘interpretation’ and ‘evaluation’ of probability. When Alpha says that the probability occurring in the formulation of a statistical law is not subjective (as his question seems to suggest), he confounds these two concepts.

ALPHA: I did not mean to express an opinion. I only wanted to know the professor’s reply to a frequent objection.

GAMMA: Fair enough, yet it corresponds to the opinion of someone (even if they are not present in this room). Some say that the probability of Physics is not subjective. Why? Because the way in which probability is evaluated in Physics is objective. The subjective interpretation of probability does not involve rejecting the evaluation based on the observation of frequencies: in this case too, the point is to apply Bayes’ theorem to certain initial distributions. It is well known that after a great number of trials, the observed frequency can provide an excellent approximation of the final probability. Though evaluation is one thing, the meaning is another, which is the

interpretation of probability. It cannot be said that a statistical probability is really involved in those typically paraded examples (of statistical mechanics, etc.). Rather, what happens is that in such contexts probabilities are evaluated on the basis of objective data.

ALPHA: However, if we have a theory that provides the initial distributions, what is the nature of those distributions? I am speaking of a theory that tells one how to evaluate probability before making observations. If we take (a typical problem in Physics) a closed circle with a material point moving inside, the theory will tell us the probability of finding the point at a given time in a given location.⁸

DE FINETTI: Initial conditions, the laws of motion, etc., are given.

Probabilistic Theories as Instruments

ALPHA: Those probabilities can be derived by making certain assumptions. Nonetheless, they are *intersubjective*, for they are considered reasonable by whoever accepts the theory. I would also like to observe that your point of view implies an instrumental view of science: a scientific theory would not be other than an instrument to make forecasts, though probabilistic forecasts. In such a case, it would not make any sense to say, of a probabilistic theory, that it is true. A probabilistic theory would not be other than a set of rules for making forecasts. That is, it would only be an instrument.

DE FINETTI: If one takes science seriously, then one always considers it also as an instrument.⁹ Otherwise, what would it amount to? Building up houses of cards, empty of any application whatsoever!

ALPHA: Science can be understood as a set of assertions that at the very least have a pretention to be true. The most refined ones say “verisimilar.” Here verisimilitude is not to be understood as *likelihood**, but as *truthlikeness*†. That is, it must be intended as closeness to truth.¹⁰

DE FINETTI: If we knew where truth was, we would not get close to it, we would go straight there. And if we do not know where it is, we cannot even know how far it is from us.

ALPHA: It is a normative idea. We may have good reasons to believe that a new theory is more verisimilar than the previous one, if the new theory survives certain empirical tests, exactly where the previous one failed.

DE FINETTI: Speaking of empirical tests in the field of probability is a contradiction in terms: what we can say is that every possible result is possible; there is no possible result that would belie the theory.

ALPHA: This is what I said before: probabilistic theories are not refutable.

DE FINETTI: Yes, probabilistic theories are not refutable. Nonetheless, if there are phenomena that fall short of being explicable by means of a certain probabilistic

* (Translator’s note:) In English in the original text.

† (Translator’s note:) In English in the original text.

theory (or are such that, in the light of this theory, it seems reasonable to consider a very small probability) which occur “a bit too often,” then we could take into consideration another theory, according to which it is reasonable to expect those phenomena to occur a bit more often. Of course, when we say that the probability of certain events is very small, we do not mean to assert that they *cannot* happen or that they can only happen once. Yet, if during an experiment in which we expected a certain phenomenon to happen just once, or even ten times, it should occur a thousand times, then a commonsense person would say: “let us see if this does not make a case for changing the framework.” This is the problem of choosing an entire theoretical construction. Even if we put probability to one side and considered say Einstein’s theory in relation to classical physics, we would see that it presents many differences (suffice it to think of the concepts of space and time, theorized in relation to the speed of light and to Minkowski’s geometric interpretation, and so on): it is a different framework in which to arrange everything we can see in the universe. Even a probabilistic theory can be very slightly modified with respect to the observable previsions, yet with major changes in the theoretical formulation, in the substrate of the whole approach. But, going deeper and deeper, we can see phenomena which, with this initial approximation, would escape. And the question is to see whether, by slightly modifying the approach, we obtain something reasonable rather than something artful, which would have the sole aim of accommodating some strange result that has been observed (but which could be a consequence, say, of a malfunctioning photographic slab). A theory frames phenomenon into a general theoretical scheme. Sometimes, something catastrophic, unforeseen, happens and then everything can change radically. In this way, for instance, we can move from pre-Einsteinian theorization to Einsteinian theorization.

Random Sequences

I would now like to move on to the topic I was meaning to address. This is very similar to the one we have developed so far (which constitutes a more general premiss, which I do not dislike at all). More precisely, I would like to examine the idea of a sequence of independent, equally probable events. If someone were to cast a die under the hypothesis that it was a fair die, she would expect that the sides from 1 to 6 showed up pretty much with the same frequency, and in an irregular fashion (not — for instance — in the order 1,2,3,4,5,6,1,2,3,4,5,6, and so on). However, if very disparate frequencies were observed instead of the predictable ones, and some regularity were noted, one would be led to say that the hypotheses that they were equiprobable and independent events had been refuted. Well, these reflections are particularly important from the point of view of inductive reasoning, for they can be considered relevant to certain hypotheses on the outcome of successive casts. However, to say that they refute or confirm (depending on the occurrence of either of those outcomes) the hypothesis that it is a sequence of independent, equiprobable events, is plain nonsense. Whatever the observed sequence, there is no reason to draw from it the information that objectively confirms a probabilistic hypothesis. If,

in fact, we were to suppose that every sequence is equally probable, then no matter what sequence we observe, from an objective point of view, it would tell us nothing about the outcomes of the subsequent casts. Nonetheless, from the subjective point of view, the outcome of the casts can reasonably be considered relevant to the prevision of successive casts. This is not to belittle the value of those inferences, yet it would be a problem if one wanted to draw a conclusion that is stronger than what is just reasonable. Moreover, the word “reasonable” is only useful to characterize an attitude. And also the notion of irregularity (*Regellosigkeit*) is a notion to which only a *psychological* meaning can be attached. However, from an objective point of view, every sequence is at the same level as the others.

ALPHA: But there is a relatively recent (going it back to 1965) attempt by Kolmogorov¹¹ and also (independently) by another scholar (whose name I cannot recall now).¹² It is a way of defining the concept (that goes back to von Mises) of irregular sequence or random sequence.¹³ An earlier analogous attempt was made by Church.¹⁴ Kolmogorov’s idea is as follows. Every 0-1 finite sequence can be trivially generated by enumerating its elements. But whenever the sequence is regular, it can be condensed in a rule that can generate it and the more regular the sequence, the less complex such a rule will turn out to be. According to Kolmogorov, a binary sequence is irregular whenever it cannot be generated by means of a recursive rule which, once translated into binary code, is irreducibly shorter than the sequence itself.

GAMMA: How is it possible to decide that there is no such rule?

ALPHA: I did not say that there exists any effective criterion to decide it. Actually, there are some theorems to the effect that there is no recursive method for deciding whether a sequence is random or not in Kolmogorov’s sense. In probabilistic terms, we can show that the longer the sequences are, the more probable it is that they are random. A theory has been developed, which however I do not know well, as I did not have a chance to study it in full detail.

DE FINETTI: I am not familiar with this theory.

ALPHA: It was discussed in the last philosophy of science congress held in Ontario.¹⁵

GAMMA: There is, on this topic, an observation by de Finetti that I take the liberty to quote. If we took the first triplet of the *Divina Commedia* and we translated it into binary code, it would, probably, according to Kolmogorov’s definition, turn out to be random. Maybe, the ordering of the digits “0” and “1” would not allow us to recognize the triplet. Nonetheless, one who knew that it was the transcription of the Dantesque triplet into binary code, would not consider it a random sequence.

ALPHA: It seems to me that we are shifting from the syntactic point of view to the semantic one. One thing is the *sequence* of letters in the triplet of the *Divina Commedia*, which can well be random, another thing altogether is the *content*, what the triplet says. Indeed, to the very same words distinct contents can be associated or even no content at all: nonetheless, the sequence of letters, as such, can well be random.

GAMMA: Even without content, there is already a non-random aspect, a determined sequence of words within a system.

ALPHA: If there is a regularity, this is always recursively characterizable. Every regularity can be captured recursively.¹⁶ However, there is a theorem, analogous

to Gödel's theorem, which applies to this case. Such a theorem says that there exists no algorithm to decide whether a given sequence is random or not.

DE FINETTI: I wonder if this problem makes sense, for I can call any sequence random or non-random depending on . . .

ALPHA: (interrupting him) But there is a definition here. A sequence is said to be random if there is no recursive rule (a law, a mathematical formula) that generates that sequence, except for the trivial one that consists in actually enumerating its elements. Clearly, a rule such as "the first digit is 0, the second is 0, the third is 1, and so forth" will always be available to us. If, besides this rule, there exists no less complex one that generates the sequence, then the sequence is random.

DE FINETTI: Perhaps this is its meaning: a sequence is not random if it can be defined with a finite number of symbols.

ALPHA: This criterion would suffice in the case of *infinite* sequences but if we consider finite sequences, then we could say that a given sequence is not random if it can be generated by means of a rule that, translated into binary code, turns out to be shorter than the sequence itself.

DE FINETTI: I cannot see the connection between this concept and probability.

ALPHA: Herein lies the connection: that through random sequences of this sort, it is possible to reconstruct von Mises' theory.

DE FINETTI: von Mises's theory has nothing to do with probability. Indeed, it concerns the limit of a sequence. And the concept of limit is not a probabilistic concept. It seems to me that theories like this one are empty of any sense. In fact, given a collective, that is a sequence of "0" and "1" digits, we could obtain from it the sub-sequence constituted only by "1" s and then the one constituted only by "0" s. According to the former sub-sequence, I should conclude that the sequence is compounded only by "1" s, whilst, according to the latter, I should conclude that it is only compounded by "0" s. It seems to me that these theories give rise to such complications as to be nothing but, to use an expression by Vailati, "big constructions full of emptiness."¹⁷ In an attempt to avoid saying what is the probability of a single well-defined event (of which everyone will be able to say whether it is true or false as soon as it is known whether it occurred or not), a countable infinity of replicas of such an event are imagined, which nonetheless cannot always repeat itself with the same conditions. And in the case of facts concerning the present, to set up a countable infinity of replicas of those facts, we should all live infinitely long, longer than Methuselah.

ALPHA: Do you consider the idea of a limiting value of the frequency to be somewhat metaphysical?

DE FINETTI: I do, for I believe that even supposing that the universe will last a billion times longer than we currently think it will, it is nonetheless of finite duration. I therefore believe that those attempts are (as it is usually put in the Italian political jargon) an "escape forward." It is as if, not being able to solve a small problem, we said: "let us tackle a huge problem in its stead, one which nobody will ever be able to solve."

ALPHA: This position of yours comes more from your positivism than from your subjectivism. As noted by Hintikka, this kind of rejection of limits is "not just de

Finetti's *subjectivism*, but as much what might be called his *positivism*" (Hintikka, 1971, p. 334).

BETA: It seems to me that we are trying to teach subjectivists how to do their job. To say to professor de Finetti that he holds a more positivist than subjectivist position is plain nonsense.

ALPHA: I am only saying that it is possible to accept the subjectivistic point of view without committing oneself to the positivistic assumption that the concept of limiting frequency is metaphysical. Hintikka, for example, interprets limits as conditional probabilities.

DE FINETTI: It seems to me that we can easily avoid the criticisms that I raise against the concept of limit by referring to a finite number n , but arbitrarily big. I wonder what the point is of putting the infinity in, when it would certainly be a considerable achievement if we could perform a billion experiments.

Editor's Notes

1. Actually neither Cantelli nor Jeffreys were mentioned during the lecture.
2. Emanuel Czeuber (1851–1925) born in fact in Prague, now the capital of the Czech Republic (and of Czechoslovakia in 1979, when de Finetti held the lectures published in the present book) and died in Austria (in Gmünd, near Salzburg). He was professor at the University of Vienna from 1890 to 1921. Like Cantelli and de Finetti himself, he studied probability theory and related areas as well as actuarial mathematics.

Czeuber wrote several books, including one on the philosophy of probability (1923). For further biographical details see O'Connor and Robertson (2005).

3. In the volume Czeuber ([1903] 1914) there is a reference to De Morgan's paper (1845), that is to an entry for an encyclopedia edited by De Morgan and published in 1845 (rather than in 1847). In 1847, however, appeared De Morgan's most famous work: *Formal Logic* (1847).

Another important essay by De Morgan is specifically dedicated to probability, of which a new edition has been recently published (De Morgan [1838] 1981). Actually, Augustus De Morgan (1806–1871) may be considered a forerunner of a logicist view of probability (rather than of a subjectivistic one).

In the past, however, the term "subjective" was often employed with the meaning of "epistemic", so encompassing both subjectivistic and logicist views of probability. For a recent study about the overlap between De Morgan's work on probability theory and logic see Rice (2003).

4. Borel ([1924] 1964).
5. See, in particular, Popper ([1935] 2004, pp. 198–205). According to Popper it is necessary "*the methodological decision never to explain physical effects, i.e. reproducible regularities, as accumulation of accidents*" (ibid., p. 199, Popper's italics).
6. From the Bayesian point of view, what play a crucial role are the initial probabilities. For those who should not have reasons to doubt the truth of a theory T , highly unlikely events in the light of T would not refute T (in fact, if a theory T has initial probability 1, this probability cannot decrease as a consequence of the obtaining of an event that, in the light of T , has a probability greater than 0). If, on the other hand, one suspected that T was false and set up experiments to test this, then a highly unlikely event in the light of T (but not similarly unlikely under the hypothesis that T was false) could substantially decrease the probability of T . Therefore — contrary to what Popper maintained — it is not only the unlikelihood of what is being observed that matters, but also the context in which the unlikely event occurred. In an empirical situation of testing, there is no certainty that the theory is true (otherwise there would be no need to test

it) and this is the reason why frequent observation of very rare events becomes relevant from the point of view of the test.

7. In other words: probability does not carry out an *explicative* office.
8. This is an example drawn from a letter by Einstein to Popper copied in the appendix *XII of *The Logic of Scientific Discovery* (Popper [1935] 2004, p. 522). By means of the latter, however, Einstein meant to defend an epistemological interpretation of physical probability (as a result of the ignorance of initial conditions).
9. The presence of the word “also” actually shows the non-radical character of Definetian instrumentalism.
10. Allusion to the positions defended by K. R. Popper. On the theme of Popperian “closeness to truth” (verisimilitude) see: *Truth, Rationality, and the Growth of Scientific Knowledge* in Popper (2002), *Two Faces of Common Sense: An Argument for Commonsense Realism and Against the Commonsense Theory of Knowledge* in Popper ([1972] 1979), *A Note on Verisimilitude* (Popper, 1976). For the problems of verisimilitude see Miller (1974b, 2006); Tichý (1974). For recent developments see Niiniluoto (1987, 2004); Festa (1993); Kuipers (1987); Goldstick and O’Neill (1988); Zamora Bonilla (2000); Zwart (2002).
11. Kolmogorov (1965, 1968). For recent essays on this subject see: van Lambalgen (1987); Delahaye (1993); Downey and Hirschfeldt (2007).
12. G. J. Chaitin. See this volume (2001).
13. Richard von Mises (1883–1953), brother of the economist Ludwig, was an Austrian mathematician, physicist and engineer who worked on fluid mechanics, aerodynamics, aeronautics, statistics and probability theory. His most famous contributions are those to probability theory and especially his theory of *random sequences*, which was one of the dominant paradigms when, towards the end of the twenties, de Finetti proposed his subjectivistic approach.
14. See Church (1940).
15. The proceedings are in Butts and Hintikka (1977). The mentioned works are by Schnorr (1977) and Jeffrey (1977).
16. Allusion to the celebrated *Church-Turing thesis*, according to which a function of positive integers is effectively calculable only if recursive or, equivalently, computable by some Turing machine. The thesis is not a mathematical *theorem* although there are strong *a priori* arguments in support of it and no counterexample has yet been found.
17. Giovanni Vailati (1863–1909), Italian mathematician and philosopher, developed a pragmatist view inspired mainly by the work of Charles Sanders Peirce (1839–1914). Between 1892 and 1895 he was assistant to Giuseppe Peano (1858–1932) and between 1897 and 1899 to Vito Volterra (1860–1940). In 1899 he decided to give up the academic career for high school teaching. In 1905 he was appointed by the Italian Minister of Education as a member of the *Royal Committee for the Reform of the Secondary School*, for which he prepared with seriousness and competence the programmes of mathematics.

For further details about the influence of Vailati (and more generally Italian pragmatism) on de Finetti’s thought see Parrini (2004, pp. 33–55).

Chapter 6

Stochastic Independence and Random Sequences*

The notion of “stochastic independence”. Variety of interpretations according to the various conceptions. Predominance of the received and simplistic views; tendency to consider “independent” all those events for which specific causal dependencies are not given in an objective sense.

Logical and Stochastic Independence

It is perhaps appropriate to begin with a terminological distinction. For the sake of simplicity I shall call “independence” *tout court* the property that more precisely should be referred to as “stochastic independence.” The adjective “stochastic” is used here “as in the calculus of probability.” Indeed, two events A and B are said to be *stochastically independent* if learning that A is true or that A is false does not alter the probability we assign to B (and *vice versa*: learning that B is true or that B is false does not alter the probability we assign to A).

We must keep in mind, however, that there is another concept of independence, namely *logical* independence. Given certain events (or propositions) $A_1, A_2, \dots A_n$, an event (or proposition) A is logically independent of $A_1, A_2, \dots A_n$ if, knowing which of the A_i 's are true and which are false, we can neither conclude that A is true nor that it is logically false. It follows that A is logically independent of $A_1, A_2, \dots A_n$ if and only if A cannot be defined from $A_1, A_2, \dots A_n$ by means of logical operations (union or logical sum, intersection or logical product, negation and so on).

The notion of stochastic independence (to which — as I said — I shall henceforth refer simply as “independence”) is a notion that is interpreted in distinct ways according to the conception one has of probability. It is typical of both the classical and the objectivistic conceptions to take independence as a condition which holds in almost every case. As for the frequentists, they introduce it as a natural property of events, whose “trials” would typically be independent. They give an approximate meaning to the word “event”. An example of an event would be: “obtaining a double six by casting two dice.” An event in this sense would be something that could be repeated *ad libitum*. And each single repetition would be considered as a “trial”

*Lecture VIII (Wednesday 28 March, 1979).

of the same event. It follows that the word “probability” cannot refer to the single case, for, in a sense, the probability of events is considered here, in one way or another and with various shades, in relation to events as general concepts, rather than in relation to the individual trials. In this case, every trial is considered (with or without mentioning this explicitly) as equally probable. I am not sure, however, whether the objectivists all agree in denying that each trial has its own probability: indeed many speak of events “whose probability does not change from trial to trial” or “whose probability varies from trial to trial.” And — by so speaking — they seem to implicitly admit that each trial has its own particular probability value.

Propensities

ALPHA: There exists a version of frequentism known as “propensity theory” according to which probability corresponds to the propensity of a certain device to cause repetitions of one sort or another with a given frequency. In this case, propensity could also be applied to each individual event.¹

DE FINETTI: How can this propensity be measured? If we say that we attach a propensity to every single trial or to every single event, we must define it in relation to the individual event.

ALPHA: Yet in this case, unlike what happens in the original frequentist interpretation, probability is not taken to be a property of a sequence, but a property of the device. It seems to me, however, that this interpretation also characterizes probability in terms of frequency. In the opinion of Popper, who introduced it, this version enables one to speak of the probability of single happenings.²

DE FINETTI: Popper does not convince me at all — with any of his positions. The only occasion on which I happened to agree with him was during a conference held in Salzburg, which both of us attended. The first thing he said was that he could not stand the smoke of cigarettes and that he would be forced to leave should anyone want to light one up. Smoke annoys me so much that — without imitating him, as I do not want to sound as if I am full of myself — if I could, I would do just like him.

If frequency is the only thing we take into account, we must ask ourselves in the first place: frequency relative to what? Let us take n arbitrary events, and suppose that m of them are verified. What can I conclude about this? I can say that the probability of each one of them was m/n . And this conclusion might well make sense, provided that we stop talking in this approximate way and introduce the notion of *exchangeability* instead. Those who intend to define probability in terms of frequency must say that all the trials are equally probable. Only in this way can they move on and say that, for each such trial, the probability is given by frequency. Yet by saying that all the trials are equally probable, one already presupposes that the probability should be referred to every single trial. If this is the case, then probability is *not* a property of the set of all the trials of an event (in their sense). If it were, in fact, the frequentists’ argument would be analogous to saying that the height of an individual is the height of the set of all the individuals of a certain country, provided that they all have the same height.

Independence and Frequentism

Let us now go back to the concept of independence. We have said that instead of defining independence (or dependence) in terms of an objective connection linking two events, we define it in terms of conditional probabilities. If the probability of A conditional on B (or conditional on \tilde{B}) is the same as the non-conditional probability, then we say that A and B are stochastically independent:

$$\mathbf{P}(A \mid B) = \mathbf{P}(A \mid \tilde{B}) = \mathbf{P}(A).$$

We said earlier that the word “stochastic” means “probabilistic”. I now add that the expression is derived from Ancient Greek.³ As I already remarked, objectivists usually include in the concept of event, besides equiprobability, the property of independence of the individual trials. Independence is endorsed by them as something heaven-sent and indisputable. Hence, given a series of events or “trials” (in their terminology) they typically assume — sometimes even without explicit mention — not only that they are equally probable but also independent. Indeed, sometimes they are even hesitant about saying (if m are the first positive outcomes among the first n trials) that m/n is the observed frequency relative to those n trials. It is in fact well known that this frequency might differ arbitrarily from probability. But this contrasts with the idea that probability is given in terms of frequency. To get around this difficulty some say that the probability p is the *limit* of the frequency (or anyway that the frequency approximates with greater precision to p as n increases). In other words, we could equally well write (according to them):

$$p =_{df} \lim_{n \rightarrow \infty} \frac{m_n}{n}$$

where m_n is the number of positive outcomes over n shots. However, this definition is utterly airy-fairy. Indeed, we could take a sequence of heterogeneous events: the limit of the frequency would not depend on the *nature* of such events but only on their *order*. This can be illustrated by the following example. Let us take the string

$$00101011100100010001110 \dots \quad (6.1)$$

Clearly, if the frequency of “0”s and “1”s in (6.1) converges to a given limit, this depends on the way the digits are ordered. Moreover, I cannot really see the point of going through the hypothesis of an infinite sequence in order to indicate the probability that an individual assigns to an event. After all, the existence of such a sequence is a purely arbitrary assumption for — even assuming that the Universe (or the Solar System or the Earth) will last to infinity without any cataclysm ever taking place to destroy it — it would be enough to refine the research on nuclear weapons a little bit to fairly quickly reach the limit beyond which we could not continue an indefinite series of experiments. Resorting to the idea of an infinite

sequence of experiments, however, is plain nonsense. But that is how it is: it really seems that the frequentist, in order to provide an adequate account of probability, must resort to concepts of no practical sense whatsoever.

Von Mises Collectives

The definition of probability as the limit of the frequency has been put forward by many authors (already in the Nineteenth Century),⁴ yet it was especially with Richard von Mises that it became fashionable (see Geiringer [1928] 1981). Although Von Mises was not the first among the contemporary authors, he is surely the most celebrated. He introduced the notion of *Kollektiv* (*collective*). A collective is an infinite sequence of trials which we must think of as ordered for, as I said above, given a sequence of digits “0” and “1”, the limit of the frequency changes according to the order in which the digits occur. If there is a more frequent occurrence of “1”s in the final segment of the sequence, then the limit would receive a higher value, maybe very close to “1”. If, on the other hand, we concentrated a greater number of “0”s in the final segment, the value of the limit would be smaller, maybe very close to 0.

In von Mises’ approach, the properties of collectives are not related to real phenomena or actual observation procedures, but are considered as *axioms*. Resorting to axioms is very convenient. Once the axioms are set, their conclusions can be deduced in a perfectly rigorous way. Unfortunately, however, those consequences can only be true if the axioms themselves are true. If the axioms are introduced only to define certain types of abstract mathematical objects, then it must be kept in mind that those are only abstract definitions and therefore, true conclusions can be derived from them only if the axioms can be given an interpretation that makes them true. All this is perfectly clear, yet frequentists take some arbitrary axioms to define probability *in general*, without checking whether their satisfaction is guaranteed in all cases or if it rather depends on contingent circumstances. And that this is a case of fortuitous satisfaction is shown by the fact that given a denumerable infinity of data, there is not, in general, a fixed order in which to arrange it. On the contrary: any ordering is allowed and the limit of the frequency varies with the ordering itself.

The worst of it all is that this resort to a convergent series is taken as a necessary axiom and not as a possibility which — under certain circumstances — might be reasonably true. If, for instance, we attach the same probability p to every event in the sequence E_1, E_2, \dots , and furthermore we take them to be mutually independent, then — as can be shown — a coherent person would judge it very probable that, on a large number of trials, the frequency will turn out to be close to p . This is an undisputable theorem. But there is no *objective* guarantee that the frequency, after any arbitrarily large number of trials, will be close to p . The fact that this is considered to be very probable cannot be used to define probability, for this would drive us straight into a vicious circle. Moreover, even assuming that the limit of the frequency is indeed p , it could well be that this happens to be the case because in

all the trials that will take place from the year 3000 onwards there will be exactly one positive outcome every p^{-1} shots. In such a case, everything that might happen from today until the year 3000, would be completely irrelevant.⁵ Of course, we are not interested at all in guessing whether the world is going to exist after the year 3000, not to mention, in that case, whether at that time anyone will be ready to fiddle around with abstractions of this sort.

Besides von Mises' theory, there are many other analogous approaches in which the probability of the single event is not taken into account or its existence is plainly denied. The reason for such a denial is that from an objective point of view, an event is a single case, which is either verified or not: a third possibility is not given. If we were to express this probabilistically, we should say that an event has probability 0 if it is not verified and probability 1 if it is. In other words, the probability coincides with the truth-values ("0" meaning "false" and "1" meaning "true"). There is no space left for any other objective probability value.

ALPHA: There are results in the calculus of probability to the effect that, in some cases, the probability of the limit of the frequency being a certain p is either 1 or 0. What sort of probability is this one from the frequentist point of view? Since it is the probability of a frequency, it should be the frequency of a frequency.

DE FINETTI: From a frequentist point of view, the frequency converges to the probability by definition. However, there is no guarantee at all that this will happen. At the bottom of frequentism lies the mystification of making probabilistic laws certain.

Another fallacy usually happens when we have to decide whether a certain sequence of digits is random or not. Typically, the reasoning goes like this. Suppose we are given a sequence as follows:

$$1110001111100001111100000. \quad (6.2)$$

Well, a frequentist would say: "this sequence is not random, for it is clear that it was contrived to make this example". However, if "random" means that it was picked by chance among the 2^{24} possibilities, then this is a sequence just like all the others and therefore, as random as any other sequence. Of course, the frequentist would reply by saying: "it is not random because it has been constructed by firstly taking three "1"s, then four "1"s, and then five "1"s, followed by as many "0"s, and so on." If I put those in some other order, then they would say: "OK, this looks rather random." Yet to be certain about it, they would run many *tests* (the so-called "randomness tests"), in order to make sure say, that "0"s do not occur more frequently in odd places than in even ones, or in order to compare distinct segments of the sequence (for instance the first hundred with the second hundred digits) and so on.

ALPHA: It seems to me that they confuse the notion of randomness with that of disorder.

DE FINETTI: Exactly. Each term in the theory of probability can be used in an objectivistic sense, more objectivistic than the objectivity of theory of probability itself. I believe that up to a certain point, this does not really matter: we could even say that

we take probability to be objective or that we consider it to be subjective and yet, at the end of the day, reason in the same way. It is a bit like the distinction between those norms which hold *de jure* and those which hold *de facto*. As soon as the *de facto* is established, the distinction becomes irrelevant. Here however, they would like to say that the term “random” has a logical meaning. What would this mean? Suppose we are looking for a definition of “sequence picked at random among the 2^n sequences of n digits” (or equivalently, of “number picked at random among the 2^n numbers of n binary digits” including, of course, the initial “0”s). For $n = 10$ there are 2^{10} sequences, that is there are around a thousand sequences, whilst for $n = 20$ there are 2^{20} sequences (that is, around a million). Many among those sequences would lead the objectivist to say: “this sequence has some regularity, there is something suspicious about it”, and hence, he would delete it. Yet he could delete just a few of them, or even all of them: indeed every such sequence is suitable for consideration. The objectivists then, contradict themselves when they firstly say that each sequence has equal probability, but then, in the face of a special sequence that was picked by chance, they claim that it should be discarded as a consequence of its regularities. At that point they say: “no way, this sequence has been written down by some trickery, for otherwise it would have not turned out in this way”. I would not really know what mental process one should adopt to test the hypothesis that a certain sequence has been chosen randomly among all the possible sequences (whether we take a hundred digits or a thousand or even a million), so as to be able to decide whether a certain sequence is random or not.

Moreover, by the word “random”, people often mean distinct things. It often happens that they swing between the following two positions (which are mutually incompatible):

- A sequence is considered to be random if it is obtained without resort to *any* selection criteria;
- A sequence is considered to be random if is obtained through a procedure which *explicitly* excludes those sequences that are easier to memorize.

In the former case, one should reasonably expect that the sequence is such that it looks somewhat irregular. But if reasoning in this way leads one to pick by chance a particular sequence, however regular it might look, it would still be random, being obtained by chance. The fact that it might look regular or irregular would then be completely irrelevant. In the latter case, the exclusion of certain sequences is simply arbitrary and unjustified.

ALPHA: Those sequences exhibiting a clear regularity are easier to memorize.

DE FINETTI: Indeed, but the notion of regularity fails to have a precise meaning. If we were to distinguish those sequences which contain an even or an odd number of “1”s, that would amount to a precise distinction. However, if we were to define the concept of the absence of any regularities, then there would be many different definitions we could put forward, yet all of them would be arbitrary and all of them would be inadequate in some respects. We might say that a sequence is regular if it has more than ten consecutive “0”s or more than ten consecutive “1”s.

Alternatively, we might say that any long sequence in which ten “0”s or “1”s never occur consecutively is regular. The truth is that the concept of an irregular sequence is a *pseudo-concept*. And pseudo-concepts always represent a serious hindrance to clear speaking about any topic. We should reject useless ideas of this sort, which ultimately originate from superstitions like the one according to which the numbers that have failed to come up in past lotto draws have a higher probability of coming out in the next draw, or similar cabbalistic beliefs. In my opinion, the first thing to do is to clear the field of all the leftovers of superstitious, fatalistic, etc., conceptions. I have debated many times with people who would claim that “cold numbers” have a higher probability of being drawn in the lotto, as it has never happened that a number did not come out in a sufficiently large series of consecutive draws. Maybe it never happened but this does not prevent it from happening at all.

ALPHA: Some people have ruined themselves playing “cold numbers.”

DE FINETTI: And I am afraid that that served them right!

BETA: You once said: “Lotto is a tax on people’s stupidity.”

DE FINETTI: Yes indeed. Lotto is like a trap to catch mice or birds, or like a trapdoor for tigers. A tiger might well be so naive as to think that men are not evil and do not mean to harm it. Yet those are persons who could reason and understand that to play the lotto is to pay a tax, without the merit of doing so intentionally. If they said: “I want to play the lotto so I can contribute towards reducing the national budget deficit,” their action would be patriotic. But doing so in order to make a profit from the state and ending up out of pocket, looks doubly contemptible to me.

I said at the beginning that I wanted to talk about stochastic dependence and independence. What I have said so far can serve as preliminary remarks to this effect. Still, as to dependence and independence, we shall have to come back to it tomorrow and, as we move on rigorously, we should consider those preliminary remarks as a kind of appetizer or aperitif, so that you will realize that things are, at the same time, a lot easier and a lot more complicated than they might seem: there are in fact bad habits into which it is very easy to slip, but which are then very hard to break. Almost all the superstitions I have mentioned do not depend as much on a misunderstanding of the concept of independence (as the majority of people have never really thought about the concept of independence) as they depend on the misbelief that there might be things which look reasonable on the basis of what one has or has not seen, and which lead one to believe that events, instead of being random in the sense that anything could happen, are rather random in an *affirmative* sense, that is to say that they should happen in a chaotic way, or that a few unsuccessful shots should be followed by a successful one. I remember that even a colleague once claimed that after three tosses resulting in Heads, the fourth toss should result in Tails.

ALPHA: It is the same fallacy as the one of “cold numbers”.

DE FINETTI: Sure. Although it sounds a bit odd that a university professor could run into such a fallacy, it is true all the same. I guess that if one had enough fantasy and wit, a whole university course (or a booklet) on the calculus of probability could be made on the basis of those examples alone, to be used as counterexamples. I even had some correspondence with people who are fiercely against that

“absurdity,” which is the calculus of probability, which does not take into account the fact that a given number cannot fail to be drawn more than a certain number of times. Sometimes, those people go so far as to make some barmy calculations (or even exact calculations, which they then interpret in a barmy way). This is a dreadful thing.

ALPHA: I guess those people reason in this way: probability corresponds to a frequency, in the sense that the phenomenon must repeat itself approximatively regularly. For example, in the case of a coin, they might think that Heads should come out more or less every second toss and that there is a roughly normal distribution. Hence, the maximal probability would occur exactly two tosses after the last time Heads came out, and would decrease as we move away from that point. In short: they interpret the process “Heads and Tails” as a periodic stochastic process.

DE FINETTI: There are many probabilistic laws that could be devised. For instance, in the case of draws, one might think that it is easy that the same ball as the previous draw will come out, because one might take into account the fact that if the ball has been put back in the urn without sufficient shuffling, this latter could remain on the surface and hence, have higher possibilities of being drawn again. This line of thought might be meaningful (even though I hardly believe that things could be this easy).

Anyway, as I said, these reflections are only a kind of appetizer and are useful to reject certain beliefs that are a lot more widespread than the ones people should have. Hence, what I said today will serve to warn you.

Before finishing, I wanted to say that von Mises went so far as to impose, by means of a specific axiom, that the sequences should be irregular. More precisely, given an event in the generic sense (I would say a “phenomenon”), von Mises not only assumes that whenever we make an infinite number of trials the frequency will necessarily converge to the limit, where the limit of the frequency will equal the probability, but also assumes that the sequence of trials must be irregular. And he called (with a nice word) *Regellosigkeitsaxiom*, the axiom which sets this condition. In German “Regellosigkeitsaxiom” means “lack-of-regularity axiom”: “Regel”, in fact, means “rule”, “los” means “lack of”, and “keit” is needed to form the abstract noun. What really makes me wonder is the fact that there are very many books on probability that follow these ideas (even though some try to fix them here and there), defining probability in terms of infinite sequences of “trials” of an event (to use their terminology).

Editor’s Notes

1. Although ubiquitous in the history of probability from the very beginning (though rendered with different words like “proclivity”, “facility” and so forth), the concept of propensity has been revived by K. R. Popper (1959) to support an objectivistic interpretation of the calculus of probability. For a recent criticism of this approach to Probability, see Eagle (2004).
2. In fact, as observed by D. W. Miller (1991), Popper’s view swung between two distinct positions: the one according to which probability as propensity is the disposition to cause certain

frequencies and the more radical one, according to which the concept of propensity is not definable in terms of frequency.

3. In fact, in Ancient Greek, "στοχαστικὸς" means "conjectural."
4. Although the relationships between frequencies and probability have been object of study since the times of J. Bernoulli, the first full exposition of a frequentist view of probability goes back to John Venn (1834–1923) (see Venn [1888] 2006).
5. In other words, the limit of the frequency is logically independent of the frequency observed in an arbitrarily long initial segment. If all "0"s and "1"s were uniformly replaced everywhere in such an arbitrarily long initial segment, the limit would not change. This all follows immediately from the definition of the limit of a sequence.

Chapter 7

Superstition and Frequentism*

Distorted (or of a “superstitious” kind) interpretations of properties or regularities and detachments from the usual “regularity.” Famous examples, expecting “cold numbers” in lotto draws. The “necessity” (instead of probability) of a small deviation of the frequency from “the probability.” The necessity of the fact that in a sequence of “independent trials” (e.g., heads/tails) there should not be long sequences with the same outcome, and so on. Variations which make such expectations reasonable.

The Frequentist Fallacy

Today I would like to put forward the reasons why I consider unacceptable those definitions according to which probability is a “relative frequency in a class of independent events.” This sentence is evoked as if it was made up of magic words and no attempt is made to clarify its meaning.

There are two key concepts involved in it: the notion of *relative frequency* and the notion of *independence*. The former is not problematic in itself: given a sequence of Heads/Tails trials, if Heads occurred m times over n trials, the relative frequency is m/n . Frequency is a matter of fact, which in itself has nothing to do with the calculus of probability. If, for instance, we toss a coin eight times, the following results are possible (where T stands for tails and H for heads):

T	T	T	T	T	T	T	T
T	T	T	T	T	T	T	H
T	T	T	T	T	T	H	T
T	T	T	T	T	T	H	H
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
H	H	H	H	H	H	H	H

The 2^8 rows of this table correspond to as many possible sequences, and to each one of them we give (provided we make the usual assumptions) probability 2^{-8} .

* Lecture IX (Thursday 29 March, 1979).

Were we to accept the definition of probability as frequency, what should we say about this situation? The only adequate answer would go as follows: every number $i/8$ included in the closed interval $[0, 1]$ corresponds to a possible value of the frequency of Heads (in the first of the above sequences the frequency is 0, while in the last one it is 1). On the other hand, the frequentist would say: “the last sequence does not count because ‘all Heads’ would be an exceptional case.” And, of course, he would also say: “the first sequence does not count because ‘all Tails’ would be an exceptional case.” By this rationale, however, we should cross out every sequence, as they are all equally probable! There is no worse conceptual distortion than, while assuming that any sequence can obtain, defining probability in terms of a property (that is the manifestation of a certain frequency), which holds only for some of the sequences.

I shall try to illustrate where this distortion comes from.

Let us consider our example. There are 2^8 possible sequences and for each one of them, the relative frequency of Heads is a number in the interval $[0, 1]$. However, we know that the 2^8 sequences can be partitioned according to the number occurrences of Heads by solving the binomial coefficients. In fact the following holds:

$$2^8 = (1 + 1)^8 = \sum_{i=0}^{i=8} \binom{8}{i} 1^i 1^{8-i} = \sum_{i=0}^{i=8} \binom{8}{i}$$

For all $i (1 \leq i \leq 8)$ the number of sequences with i occurrences of Heads is m/n , therefore, we have only one sequence with no occurrences of Heads, eight sequences with one occurrence of Heads, twenty eight sequences with two occurrences of Heads, and so on (Fig. 7.1).

What conclusion should be derived from all this? The conclusion that although the sequences with a frequency of Heads close to $1/2$ are more numerous than those with a frequency far from $1/2$, it is by no means true that probability coincides with frequency. And we cannot cross out, as a consequence of their

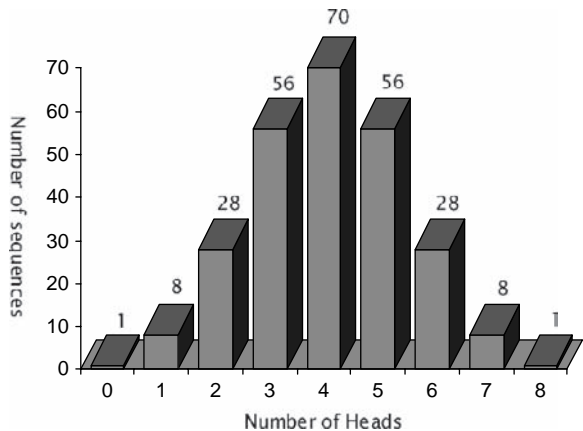


Fig. 7.1 Distribution of the possible sequences of 8 trials in the Heads-Tails model

scarcity, those sequences with a frequency far from $1/2$: in fact, each one of those — taken individually — has the same probability as any other sequence. Hence, as I already remarked, if any of the sequences were to be crossed out, then we should cross out every sequence. This shows that when we define probability as a frequency, we just give a sloppy definition. The only objective fact is the collection of all possible sequences, yet this does not tell us anything at all about their probability. The probability of the sequences is just that sensation we had before, in our way of expecting them.

Therein lies a linguistic, a logical and a commonsensical perversion. A logical mistake of this sort is not acceptable, for we cannot mistake the set of all possible sequences (which is logically determined) for the probability (which is subjective instead). If we said: “I believe that the coin is biased towards ‘heads’ ” then the first sequence, in which only Heads occurs, would have a probability higher (to a greater or lesser extent) than 2^{-8} , and the other sequences would also have greater or lesser probability, depending on whether they have a predominance of occurrences of Heads or a predominance of occurrences of Tails. Everyone can have the opinion they like: it might well look like a barmy opinion, but if one says that it is a barmy opinion, then one expresses a subjective judgment about another subjective judgment. Hence, one expresses a judgment which is, so to speak, subjective to the power of two.

An objective criticism is not possible, unless someone runs into mathematical mistakes. If when computing the probability by expanding $(1 + 1)^8$, we wrote 9 instead of 8 as the second term, we would be facing a mathematical mistake, a mistake in the computation of probability. In the other cases, it would be just a matter of deviating from the conventional opinion, the opinion that is usually examined for the purpose of illustration. This is the opinion that leads one to attach the probability $1/2$ to each shot, independently of the outcome of the previous shots.

Idealized Frameworks

These remarks apply even if we assume, to use the objectivists’ jargon, that the coin is “fair” or that it is always tossed in the same way. Indeed, when saying that the coin is always tossed in the same way, we cannot mean that it is always tossed *exactly* in the same way: indeed, if that were the case, the outcome would be always the same. The frequentist would then say that although the coin is not tossed exactly in the same way every time, it is tossed in such a way that, for every shot, either Heads or Tails can obtain *with the same probability*. By defining probability in this way we run into an obvious circularity. Indeed, when we cannot determine the value of the physical parameters on which the outcome of the single trials depend (initial velocity and impetus, angular momentum, effects of the collision between the coin and the air molecules and so on) we give the appraisal that looks more appropriate. The simplest evaluation will certainly consist in attaching, for each toss, the same probability to the occurrence of Heads and Tails (unless there are somewhat compelling reasons to reject any of the assumptions that are implicit in these examples). In this

way, all the 2^8 sequences would be considered equally probable. For the subjectivist this turns out to be a very natural evaluation, as in cases of this sort there is usually no reason to do otherwise. Yet there are situations in which this is not the case. For instance, one could suspect that the coin is biased and attach to the occurrence of Heads a probability slightly higher or slightly lower than $1/2$. Someone else might suspect that the way the coin is picked up after the toss might favour or thwart the occurrence of the same outcome on the next toss.

I do not mean to claim that it is reasonable to nourish all those doubts. What I claim is rather that probabilistic reasoning depends on subjective evaluations, which could vary for arbitrary reasons. Personally, in cases of this sort, I do not typically have reason to deviate from the usual evaluation of probability. Yet it could well happen that after observing the magnitude of the difference between the observed frequency of Heads and the value $1/2$, I might start thinking seriously that Heads is favoured. Or, after observing many consecutive tosses with the same outcome, I might form the suspicion that the person who is tossing the coin is doing it more or less in the same way all the time: in this case, I would be inclined to consider the recurrence of the same side of the coin more probable. Any consideration whatsoever of this sort is acceptable, provided it is given its due value.

When saying that Heads has probability $1/2$ for each toss, we must not give to the value $1/2$ an absolute meaning: the appraisal of this value, in fact, depends on assumptions that require the judgment of an individual. Those assumptions are usually either neglected or stated by means of clauses like “provided that the coin is fair,” and so forth. The Heads/Tails process with constant probability of $1/2$ is therefore nothing but an idealized framework. There is no reason to consider it necessary.

The Fallacy of Hypothesis Testing

I am certainly aware of the fact that in practice, everyone reasons in this way, and there is no reason to reject this way of arguing. Yet we can only admit it provided that we keep firmly in mind that “hypotheses” like the one of a “fair coin” are not, in reality, assumptions bearing on facts. From the factual point of view, every sequence of 8 shots can result in any of the 2^8 possible outcomes: that is all. Every opinion is therefore legitimate. For instance, it is legitimate to believe that the tosses are independent, yet believe that Tails has a slightly higher probability than Heads. Or that the probability is $1/2$ on average, yet owing to the way the coin is tossed, the repetition of the same side has a higher probability. Let us now question the content of these hypotheses. Suppose that the coin is not fair and the outcome is biased towards Heads. What does this assumption reduce to? It reduces to attaching to the sequences of shots with m Heads and n Tails a higher probability than the one given to those sequences with n Heads and m Tails (where m is greater than n).

There are many distortions that need to be addressed. For instance, let us consider the following sequence:

H H H T H T T H

Suppose that I predicted that this sequence would have come out and after tossing the coin, this actually is the case. What would be the typical reaction in the face of such a coincidence? It would be to say that there was a trick in it somewhere. Now, from a certain point of view, this reaction is understandable. If, for example, I did bet on that outcome and if I was the one tossing the coin, someone else might suspect that I — assuming that I knew how to do it — might have used some trickery to secure my success. If, on the other hand, one said, without having those reasons to be suspicious, that “there was some trickery because it was too improbable that exactly that sequence would obtain” this would be an inadmissible reaction. It is clear that we are facing a sequence which, exactly as any other sequence, has probability 2^{-8} . Hence, whichever sequence turns out, one could say that there was some trick in it, given that its probability was so small that it could not obtain.

In this example the sequence of trials is of no particular interest to anyone. Yet, if there was some significant interest attached to it (say a big prize), someone might say: “if exactly that result came out, it means that there was some trickery in it” I recall that Corrado Gini (the founder of the Italian National Statistical Institute — ISTAT), a statistician who would very often make acute observations, mostly directed against the tendency to attach to the results of the theory of probability a degree of certainty that they do not have, used to put forward the following example.¹ There is a very small probability — he would say — when buying a lottery ticket, of actually buying the winning ticket: if, for instance, there are a million tickets, such a probability is a millionth. Therefore, whoever turns out to be the winner, with the draw of the winning ticket, an event with probability of a millionth would be verified. If we agree that a millionth is such a small probability that an event with that probability should be considered to be practically impossible, then we should conclude from this, with practical certainty, that whoever won the lottery did so with some trickery and should therefore be arrested! Trains of thought of this sort not only show how certain views, which regrettably happen to be very widespread, are odd, but also show how (despite the fact that my conclusion does not coincide with that of statisticians, including Gini) probability cannot be identified with frequency.²

Instead, what can one simply say? Let us consider again our table with 2^8 possible sequences. 70 of them contain the same number of repetitions of T and H. The probability that the frequency of ‘heads’ will be $1/2$ (under the assumption that the probability of each toss is $1/2$) equals the ratio of the number of combinations of 8 shots with 4 Heads and 4 Tails, and the total number of sequences ($2^8 = 256$). It is therefore $\frac{70}{256} \approx 27\%$. There is nothing more to say: everyone can agree on this point.

Analogously, suppose we are to draw with replacement from an urn containing white and black balls in proportions that are not known to us. If one said: “from the frequency of the white balls drawn we can deduce that the proportion of black balls is p ” one would argue in a way that is logically groundless. Indeed, one could formally deduce nothing from the outcome of the draws except, if at least one ball of either colour was drawn, that it cannot be the case that the balls in the urn all have the same colour. And if one accepted *a priori* the belief that the probability of drawing a white ball were $1/2$ and considered the draws to be independent, then one could

not modify such an initial appraisal on the grounds of the outcome of the previous draws, for in this case, the hypothesis that the proportion of white balls in the urn was $1/2$ could neither be confirmed nor denied. If something odd happened (if, for instance, in the case of eight shots at Heads and Tails all the shots resulted in Heads) this could look somewhat odd only if there was already a suspicion — if slight — to the effect that things might not have gone according to the initial assumptions. However, if one started by considering all the sequences to be perfectly equally probable, then one would have no reason to attach to the next toss a probability value other than $1/2$, independently of the previously observed outcomes.

Editor's Notes

1. Corrado Gini (1884–1965) was an Italian statistician and economist who (among other things) made important contributions to the classical Bayesian tradition and promoted a revisitation, from a Bayesian viewpoint, of frequentist statistical techniques (cf. Scardovi, 2001, pp. xxi–xxii), in a vein similar to Harold Jeffreys' work (1939). In particular, Gini rediscovered in 1911 the class of Beta distributions, already proposed (cf. Zabell, 1989, p. 253) as priors by G. H. Hardy ([1889] 1920) and by W. A. Whitworth ([1897] 1965, pp. 224–5) — which includes as a special case the Bayes-Laplace uniform distribution — in the study of the inductive Bernoullian process (Gini [1911] 2001), so anticipating — as it may be shown via de Finetti's representation theorem — Carnap's λ -continuum (see Costantini, 1979).

The young de Finetti was highly influenced by Gini, especially during the years 1927–1931, when he worked for the ISTAT (*National Bureau of Statistics*) in Rome, founded in 1926 and of which Gini was the first president up to 1932.

A discussion of Gini's view of probability with respect to the subjectivistic position was made by de Finetti in a paper written in memory of Gini after his death (1966). A recent book (Gini, 2001), which can be browsed on the Internet, collects both in Italian and English many important contributions by Gini to statistics and inductive probability.

2. In fact, if probability were an objective property of facts, then it should be improbable — unless there is some trick — that improbable events would obtain. However, not only is it not improbable that an improbable event obtains but in certain cases, it is even certain that this is going to happen.

Chapter 8

Exchangeability*

Exchangeability: correct version of the condition which is usually referred to as “independence” with “unknown but constant probability.” Mistaken terminology because probability varies from trial to trial as a consequence of the experience acquired through recording the outcomes of previous “trials” and it is known trial-wise. What is unknown, yet has a distribution which varies according to the outcomes of the observed trials is the limiting frequency, which is often referred to (more or less appropriately) as “objective probability.”

Urn Drawings with Replacement but Without Independence

Suppose \mathcal{U}' and \mathcal{U}'' are two urns containing white balls in the proportions p' and p'' respectively (the other balls being black in both urns). Suppose further that draws are to be made from one urn only, the actual urn to be used being unknown.¹ In a situation like this, one would assign a subjective probability (e.g., 30%) to the hypothesis that the draws will be made from the urn \mathcal{U}' and the residual probability to the hypothesis that the draws will be made from the urn \mathcal{U}'' . Suppose now that some draws with replacement are made from the selected urn. Do you think these are going to be independent draws or not?

STUDENTS: No.

DE FINETTI: Why? Is there anyone who says “Yes”? No one? Then I do not know what to say: maybe you guessed the correct answer thinking that the one that looked more natural was more likely to be the wrong answer, or maybe you have already had an occasion to think of a situation of this sort.

BETA: Intuitively, it looks as if they would not.

DE FINETTI: It is good to have an intuitive grasp on things. It is even better, however, if one convinces oneself after due consideration.

BETA: There is always a reason when things appeal to one's intuition.

DE FINETTI: There is a reason also in this case. And this is what I would like to hear from you. Take your time for reflection before venturing the answer.

* Lecture X (Tuesday 3 April, 1979).

DE FINETTI: OK. To begin with we know that you gave the right answer: the thing is now to check whether you guessed it by chance, as if your brain tossed a coin to say either “Yes” or “No.”

BETA: The probability is influenced by the choice of the urn.

DE FINETTI: Sure, but the question was: are the draws independent or not?

ALPHA: No, because the probability that the draws are made from one urn or the other varies. Let us take the hypothesis h that the draws are going to be made from the first urn: the probability of h varies according to the outcome of such draws. Hence, clearly, the draws are not independent. In fact, if the opposite were true, the probability of h should not vary either.

DE FINETTI: How does the probability vary every time that a white ball or a black ball is drawn? If we draw a white ball (or, more generally, if on numerous draws the white balls prevail) what shall we conclude?

BETA: Are we trying to find out which is the urn that is being used for the draws, or is this unimportant to us?

DE FINETTI: In a way it is the same thing. The probability of drawing a white ball at each shot is going to vary according to . . .

BETA: The composition of the urn.

DE FINETTI: Yes, but although we know the composition of both urns, we do not know which one is being used for the draws. Therefore the probability will vary because . . .

BETA: It will vary with respect to the previous draws.

DE FINETTI: This has already been said, yet it has not been specified in which way it will change. Suppose we are going to make many draws: what should we expect?

BETA: That the proportion of the white balls drawn will be quite close to the proportion of white balls present in the urn.

DE FINETTI: If we were uncertain about whether the white balls were in a proportion of 30% or 70%, the outcome of the draws could confirm either hypothesis. We must keep in mind, however, that any frequency can obtain. And even if we knew the composition of the urn, we should nonetheless refrain from being superstitious and believing that the calculus of probability provides us with non-probabilistic recipes; that it can offer us certainties. And it can well happen that although the proportion of white balls is actually 30%, the outcome of each draw is always a white ball. Conversely, it might happen that although the proportion of white balls is 80%, those are never drawn. It is just a matter of probability. However, by and large (and provided that one is careful not to turn practical certainty into absolute certainty) one can say that *almost certainly*, if the urn contains 30% white balls, the frequency of the draws of white balls will be around 30%, and if one knows that it is either 30% or 70%, the hypothesis that it will be either one or the other will be supported according to whether the observed frequency is close to 30% or 70%. It must be appreciated, however, that the degree of such a confirmation depends on the initial probabilities.

If one knew which urn was selected, the draws would be independent. Then one could be led to argue like this: “there are two possible cases: either draws are made

from the first urn or they are made from the second. Given that in both cases the draws are independent, it follows that they are certainly independent.”² Yet this argument is fallacious. Where does the “trick” lie? Whether the draws are going to be made from the first urn or from the second, the draws are independent, nonetheless the draws made from a randomly selected urn are not independent. Why?

ALPHA: Because the property of independence is conditional on knowing the actual urn from which the draws are made.

DE FINETTI: Yes, fine, but . . .

BETA: At this point the outcome of the last draw changes the probability of the possible outcome of the next one.

DE FINETTI: Suppose — to reason in a more general way — that we have N urns. In this case independence is conditional on each of the hypotheses of a partition (the possible compositions of the given urns) $H_1 \dots H_N$. Suppose that the initial probabilities of drawing a white ball were $p_1 \dots p_N$. Suppose further that the initial probabilities of the hypotheses $H_1 \dots H_N$ are $c_1 \dots c_N$, respectively. What is the probability that a white ball will be drawn in the first shot? Furthermore: what is the initial probability of obtaining h positive outcomes over n shots? And what is the inductive reasoning mechanism, that is, how does the probability $\mathbf{P}_h^{(n)}$ of drawing h white balls in n draws with replacement on the basis of the previous draws?

For each hypothesis H_i , subject to such an hypothesis, the probability $\mathbf{P}_h^{(n)}$ of drawing h white balls in n draws is given by the binomial formula:

$$\mathbf{P}_h^{(n)} \mid H_i = \binom{n}{h} p_i^h (1 - p_i)^{n-h}. \quad (8.1)$$

Therefore, the probability $\mathbf{P}_h^{(n)}$ amounts to the average of the distinct values $\mathbf{P}_h^{(n)} \mid H_i$ weighted by the initial probabilities c_i of the hypotheses H_i . Hence we will have:

$$\mathbf{P}_h^{(n)} = \binom{n}{h} \sum_{i=1}^N c_i p_i^h (1 - p_i)^{n-h}. \quad (8.2)$$

In this case there is no independence among the draws. One could say: there would be independence if it were known which of the hypotheses H_i is true. But we do not possess this information and — as it can be verified — there is no independence. There is no independence because

- a) the same outcomes may obtain under a variety of hypotheses rather than one unique hypothesis, which is initially taken to be certain;
- b) the probability of the outcomes of the draws depends on the probability of those hypotheses;
- c) the probability of each hypothesis depends on the observed frequencies.

Induction and “Unknown” Probabilities

However, there exists a property which characterizes this case and it is in fact a weaker condition than independence. It is called *exchangeability*.

What does “exchangeability” mean? Let us consider the events of the form “ m white balls in n draws.” Suppose that $n = 8$ and $m = 3$. In this case we could have 3 white balls and 5 black ones in every possible order. Therefore the number of possible outcomes is

$$\binom{n}{m} = \binom{8}{3} = 56.$$

Exchangeability consists in assigning the same probability to each of those possible outcomes. In this case the order with which white and black balls are drawn is irrelevant. We begin with equation (8.1) above, that is the formula of the probability of drawing h white balls in n draws with replacement from an urn of known composition. Equation (8.2) is nothing but a *mixture* — that is to say, a weighted average of probability according to equation (8.1) — where the weights used are the probabilities of the distinct urns. The key result is as follows: *by means of exchangeability it is possible to give an intrinsic characterization of inductive inference, because the probability of the various outcomes of future draws given the outcomes of the past ones is directly derived from the condition of exchangeability according to which all the permutations of the possible outcomes are given the same probability*. Every time that probability is given by a mixture of hypotheses, with independence holding for each of them respectively, it is possible to characterize the inductive relevance that results on the basis of equation (8.2) in terms of exchangeability.³

This example shows how the concept of exchangeability is necessary in order to express in a meaningful way what is usually referred to as “independence with constant but unknown probability.” This expression is not correct because in fact

- a) there is no independence;
- b) the probability is not constant because the composition of the urn being unknown, after learning the outcome of the draws, the probability of the various compositions is subject to variation.

BETA: Let us suppose that the urn were to be chosen really randomly, that is, that the probability of the various c_i ’s were to be the same.

DE FINETTI: This would not make any difference.

ALPHA: Well, the formula would be simpler in this case.

DE FINETTI: Yes, in place of the various c_i we would have the constant $1/N$, but the mechanism would stay the same.

ALPHA: If we assigned the probability value 1 to a given c_1 and 0 to all the others, then we would be back to independence.

DE FINETTI: Yes, because that case would be covered by the model in which the selected urn is known to us. Yet we can say this: asymptotically we get close very

often to that case. Indeed, as the number of shots grows, the probability that the frequency resulting from those shots will be close to the proportion of the white balls present in the selected urn increases. If the difference in the composition of the urns is small, then it will take a really large number of shots to distinguish with sufficient certainty which of the urns is the selected one. I have to say, however, that I am not very keen on expressions like "a large number," or "a very large number," because they seem to imply the idea that a given threshold exists beyond which, as if by a miracle, a conclusion is suddenly reached.

ALPHA: Would this convergence still hold if we gave probability 0 to the urn that is actually used for the draws?

BETA: What does "probability 0" mean? Maybe that the presence of one of the possible urns has been kept hidden.

ALPHA: I did not mean to make specific hypotheses on the reasons for that 0 probability. I was simply making the assumption that one of the urns is given (without asking why) probability 0. This urn *could* however be the one used for the draws.

DE FINETTI: This is a bit of a sophistic case, yet it is worth thinking about. It is possible that such an urn existed yet, for some reason or another (because it has been kept hidden, or for any other reason) a 0 probability is attached to the fact that it is indeed the one used for the draws. As the final probability is proportional to the initial probability, if this latter is 0, then so will the final probability. Yet suppose that the frequency observed after a large number of shots turned out to deviate massively from the percentage of white balls of every urn, except for the one with probability 0. I believe that after a very large number of draws, one could say: "these results suggest that the initial scheme was not correct, because there is also the urn to which probability 0 was attached."⁴

In my opinion, my only interesting contribution was the invention of the term "exchangeability" to remove the contradiction contained in the sentence "independent event with constant but unknown probability." Exchangeability is in fact *equivalent* to the concept of independence conditional on a partition of hypotheses. Moreover, it turns out to be practically certain (asymptotically: the limit tends to 1) that the final probability of one of the urns will converge to 1, whereas it is almost certain that the probability of the others will converge to 0. Therefore, one could say (with that abuse of language I condemn because I believe that we should distinguish between logical certainty and probability 1, which is always just a degree of practical certainty) that it is almost certain that, after many experiments, in the long run we will end up singling out with practical certainty the urn from which the drawings are made. After all, this pattern can be considered as a mathematical reconstruction of the method followed by those policemen and judges who must assess the clues possessed by them. They look for new information to support the hypothesis of culpability until a degree of certainty sufficient to make a decision is reached.

ALPHA: This means that the probability distributions converge.

DE FINETTI: Yes, that is it.

ALPHA: And the probability that a white ball is drawn in the $n+1$ drawing converges, as n grows, to the frequency r/n of white balls observed in the previous draws.⁵

DE FINETTI: It is a matter of multiplying, at each draw, $r + 1/n + 1$ by the factor

$$\frac{\mathbf{P}_{r+1}^{n+1}}{\mathbf{P}_r^n},$$

where \mathbf{P}_r^n is the probability of obtaining exactly r white balls out of n draws. Your statement holds because this factor converges in probability to 1 when we let n tend to infinity, whilst $r + 1/n + 1$ converges to r/n .

Furthermore, there is a distinction that I would like to mention. It is the distinction between *bounded exchangeability* and *unbounded exchangeability*.⁶ This latter holds when the experiment is indefinitely repeatable. We have seen an example of this case above. Bounded exchangeability, on the other hand, holds for those processes that cannot be continued beyond a certain limit. To make the typical example, the case of drawings without replacement is one of bounded exchangeability.⁷

ALPHA: With a finite number of balls.

DE FINETTI: I guess I have never seen in any book a discussion of the case of urns with an infinite number of balls.⁸ To consider urns with infinitely many balls would just be a useless complication, which would just make things look more complicated than they are.

Suppose that there are N balls, of which N' are red and N'' are white. We can make at most N draws after which we will have drawn all the N' red and all the N'' white balls. Even if we did not know how many balls are of one colour or the other, we would still be in the case of bounded exchangeability.

Exchangeable Random Quantities

The concept of exchangeability can be applied not only to events (that is to say random quantities which can have only two values, usually denoted by the digits '0' and '1'), but also to arbitrary random quantities. Suppose X_1, \dots, X_n are random quantities. Then there will be a distribution function for each X_i which, if differentiable, will be representable as a density function. For example, the possible values of every X_i can be the points of the interval $[0, 1]$ and each X_i can have a probability distribution whose density is represented by the diagram in Fig. 8.1. Figure 8.2, on the other hand, shows the cumulative distribution function.⁹

In this case too we can distinguish between bounded and unbounded exchangeability. In the case of unbounded exchangeability, in place of various hypotheses on "unknown probabilities," we would have various hypotheses on "unknown distribution functions." For example, the density function represented in Fig. 8.1 could be the function f_1 . In such a case F_1 would be the corresponding cumulative distribution function. Analogously, we could have functions f_2 and F_2, \dots, f_N and F_N , each of these having a probability c_1, c_2, \dots, c_N , respectively. Since we are dealing with functions here, rather than constants (and with cumulative distribution

Fig. 8.1 Density function

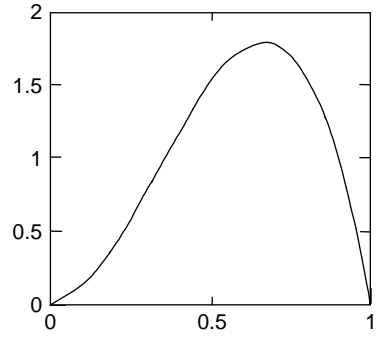
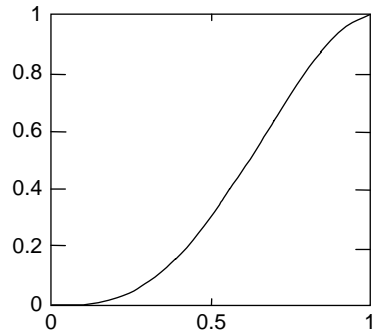


Fig. 8.2 Cumulative distribution function



rather than probability functions), the developments became notationally a bit more complicated.

From the conceptual point of view, however, it is more or less the same thing. There, the point was (to use an expression which — I do want to emphasize it every time — is wrong, yet helpful in order to single out those cases in which exchangeability holds) “independent events of equal but unknown probability.” Here is its analogue: we shall consider various hypotheses on the cumulative distribution function (or in the probability distribution relative to one of these random quantities X_1, \dots, X_N). Under the hypothesis of independence conditional on each of the hypotheses, we would have exchangeable random quantities.

We could also introduce partial exchangeability and other things, but this would take us beyond the scope of this course.

Alleged Objectivity and Convergence of Subjective Probabilities

Rather, I would like to repeat that through the concept of exchangeability I have attempted to recover the meaning of the expression “unknown probability.” And I have shown that this can be done, because making a mixture of “unknown probabilities” with respect to which all the trials are independent and equally probable amounts to satisfying the property of exchangeability. However, in this case it is not

given that either independence or equal probability hold. That sentence is not only incorrect, at least from a lexical point of view, as a way of expressing oneself, it also creates confusion.

ALPHA: One could say that the unknown probability is that probability which we would have if we knew the composition of the urn.

DE FINETTI: Yes, and so?

ALPHA: In this case the unknown probability would be subjective.

DE FINETTI: If we are uncertain about the composition of the urn we cannot have a constant probability for all the trials.

ALPHA: However, one could talk of unknown probabilities in the following sense: those would be the subjective probabilities that I would have, were I to know which is the urn from which the balls are drawn. Indeed, if I had this information, the draws would actually be independent events with constant probability.

DE FINETTI: I am not sure I fully understand what you said. What you say is correct, but it does not seem to me as if this can be used to legitimate the idea of an unknown probability.

ALPHA: Maybe because the probabilities conditional on knowing the composition of the urn are not measurable: given that I do not know the composition, I therefore cannot place any bet on it.

DE FINETTI: Why not? Everything that is not known to us is a random quantity about which we can express a probability assessment! The degree of uncertainty might be greater or smaller, but it is always possible to express it through a probability distribution.

ALPHA: In this case one would assign each urn its own probability.

DE FINETTI: It would be the same thing if there were just one urn, but we only knew that the percentage of white balls is among the values p_1, p_2, \dots, p_n . And were there more cogent reasons to believe that the composition of the urn was, say, p_2 , then we would be in a situation analogous to the one in which we knew from the beginning which was the selected urn. It should be taken into account that it is well possible that, by chance, white or black balls are drawn in a percentage that deviates significantly from p_2 .

ALPHA: Let us take an unfair die. Since all the opinions converge and since there is an "objective" convergence to the observed frequency¹⁰ can we not consider the limit to which all the opinions converge as an objective probability?

DE FINETTI: I cannot see what this would buy us. This consideration wrecks the meaning of probability. If one says that one approaches the objective probability, what would then be the difference between saying that the probabilities far from 1 are more subjective than those whose value is so close to 1 as to actually become confused with 1? The only objective probability values are 0 and 1, where by "0" I do not mean a quantity that represents the probability of a possible event yet one of probability smaller than any other arbitrary real number, but I mean *logical impossibility* and by "1" I mean *logical certainty*. One can say that the logically impossible event is the only event of objective probability 0 (taking 0 in an absolute sense) and that the certain event is the only event of objective probability 1.

Editor's Notes

1. This schema goes back to Laplace: "I suppose that I am presented two urns, A and B , of which the first contains p white tickets and q black tickets, and the second contains p' white tickets and q' black tickets. I take from one of these urns (I do not know which) $f + h$ tickets of which f are white and h black" (Laplace [1774] 1986, p. 365). This schema provides a paradigmatic model of a general inductive situation that "left an indelible imprint on statistics" (Stigler, 1986, p. 359). Suppose that an experiment, allowing exactly two outcomes R and S is repeated under identical circumstances. Suppose that there are two unknown "possible causes" C and C' such that in the presence of C the probability of R is p and the probability of S is q , while in the presence of C' the probability of R is p' and the probability of S is q' . We may apply the calculations made with reference to the urn model to this situation.

This schema may be generalized in two directions: (a) increasing the number of urns and (b) increasing the number of the colours the tickets may have, so that it provides a more general framework for the probabilistic treatment of inductive inferences. Laplace took the first line, which led him to rediscover the celebrated "rule of succession" (found by Bayes many years before (Bayes [1764] 1970)), according to which if an experiment that allows two outcomes R and S has been performed n times in which R occurred f times, then the probability that in the next experiment R will occur again is $f + 1/f + 2$.

For an excellent account of the history of the rule of succession see Zabell (2005, pp. 38–73), which traces, in the words of its author "the evolution of the rule, from its original formulation at the hands of Bayes, Price and Laplace, to its generalizations by the English philosopher W. E. Johnson and its perfection at the hands of Bruno de Finetti" (ibid., p. 38).

2. The fallacious nature of the argument shows that independence is not a relation between events but a subjective property of the probability function. This contrasts with the objectivistic point of view according to which independence is, in fact, a relation between events.
3. It should be added — and this is the content of the celebrated *de Finetti's representation theorem* — that the converse is also true: given an indefinitely extensible sequence of exchangeable events, the probability function defined on the Boolean algebra generated by this sequence can be represented as a mixture (or weighted average), with subjective weights, of Bernoullian functions (that is to say functions with respect to which the events are independent and with constant probability).

Much more generally, the representation theorem establishes the coextensivity (with respect to sequences of random quantities) of the notions of conditional independence and of exchangeability, (see Loève, 1960). The result was first published in de Finetti (1931a).

For a rigorous discussion, yet one accessible to the philosophically trained scholar, I would suggest looking at Jeffrey (1971). There is a vast literature on the representation theorem, both philosophically and technically oriented (there are many generalizations of the theorem). Among the works of the first kind I suggest: Howson and Urbach ([1989] 2005); Jeffrey ([1965] 1983); Braithwaite (1957); Good (1965); Hintikka (1971); Spielman (1976, 1977); Skyrms (1980); Suppes and Zanotti (1980); von Plato (1981); Suppes (1981); Mura (1989, 1992); Garbolino (1997); Guttman (1999); McCall (2004); Dawid (1985, 2004); Zabell (2005).

As to the mathematically-oriented works, the following are only a sample from an enormous literature:

Khinchin (1932); Hewett and Savage (1955); Ryll-Nardzewski (1957) are early classic contributions. For a more recent literature, see: Humburg (1971); Diaconis (1977); Link (1980); Jaynes (1986); Diaconis and Freedman (1980a, 1987, 2004a,b); Caves, Fuchs and Schack (2002); Fuchs, Schack and Scudo (2004).

4. This observation shows that de Finetti interpreted Bayesian procedures as revisable models, especially when some strong assumption (like exchangeability or independence) is made. Such assumptions were considered by de Finetti as approximate idealizations that may be given up at any moment if the actual degrees of belief — perhaps in the presence of unexpected pieces

of evidence — considerably differ from the probabilities computed via Bayes' theorem from the assumed priors. In such a case, according to de Finetti, it is legitimate to depart from the values inferred through Bayes' theorem, provided that one "repents" of the previously assigned initial probabilities. What is essential, according to him, is to be the bearers, at any time, of a set of coherent probability values. On this crucial point see also Chapter 4, pages 39–40 and Chapter 7, pages 71–72.

5. The statement is correct provided that convergence is understood as "convergence in probability." For convergence in the usual sense exchangeability is not sufficient. The latter holds when the probability distribution over the possible proportions of white balls gives probability 0 to the hypothesis that the proportion is included in the open interval $(0, 1)$, or, for every open interval (a, b) ($0 \leq a \leq b \leq 1$) the probability that the proportion of white balls is included in the interval (a, b) is greater than 0. This result is a special case of a more general theorem due to Gaifman (1971, p. 245).
6. Another relevant notion in this connection, introduced by de Finetti in 1938, is *partial exchangeability*, which is a generalization of exchangeability: "Imagine a game of heads and tails, played with two irregular-looking coins. If the two coins look exactly alike, one may be led to judge all tosses as exchangeable, no matter which coin is used. If, on the contrary, the coins are completely different, we may be led, at the opposite extreme, to consider as separately exchangeable the tosses made with the one and with the other of the two coins, these two types being completely independently of each other. But there is also an intermediate case, and it is precisely that case which leads us to the generalization we envisage. Suppose that the coins look only *almost* alike, even that they seem alike, but without this leading us to regard all trials as exchangeable — perhaps because we suspect hidden differences, or perhaps for some other reason; then observations of the tosses of one coin will still be capable of influencing — but in a *less direct* manner — our probability judgment regarding tosses of the other coin. . . . More generally, one can have any number of types of trials corresponding to different coins, people, temperatures, pressures, etc. . . . All the conclusions and formulas which hold for the case of exchangeability are easily extended to the present case of *partially exchangeable* events, which we could define by the same symmetry condition, specifying that the events divide into a certain number of types $1, 2, \dots, g$, and that it is only events of the same type that are to be treated as 'interchangeable' for all probabilistic purposes" (de Finetti [1938] 1980). In the case of partial exchangeability, where the events may be partitioned in g exhaustive and mutually exclusive types C_1, \dots, C_g and exchangeability holds within each type C_i ($1 \leq i \leq g$), the probability $\mathbf{P}(E)$ that any unobserved sequence E of trials on the basis of an observed sequence H of k trials is uniquely determined by k_1, \dots, k_g representing, respectively, the numbers of trials of H that belong to C_1, \dots, C_g ($\sum k_i = k$). De Finetti's representation theorem extends to this situation: $\mathbf{P}(E)$ is a mixture of g -dimensional multinomial distributions.

In the context of Carnap's inductive logic, partial exchangeability turns out to be equivalent to Carnap's *axiom of symmetry with respect to individual constants* (cf. Carnap, 1971a, pp. 131–140). De Finetti's representation theorem for partially exchangeable events is (beyond its philosophical significance) a powerful tool to investigate inductive analogical inferences. In the context of Statistics, partial exchangeability applies whenever stratified random samples are employed.

7. It should be noticed that the representation of exchangeable probabilities as mixtures of drawings with replacement (see, in this chapter, note 3 on page 83) does not hold exactly in the case of bounded exchangeability. In this case a slightly different form of de Finetti's representation theorem holds: every exchangeable probability evaluation on binary sequences of n trials is a unique mixture of draws without replacement from $n + 1$ urns containing black and white balls in proportion, respectively, $\frac{0}{n}, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}$. Diaconis and Freedman (1980a, b) proved that when an exchangeable probability on binary sequences of k trials may be extended to binary sequences of length $n > k$ (which is not the case of draws without replacement), then the minimum distance d_{\min} between the probability $\mathbf{P}_h^{(k)}$ of h successes in k trials and the corresponding probability of drawing h white balls in k draws in the Laplacean model

of drawings with replacement from urns with various proportions of white and black balls is $\leq \frac{2k}{n}$, so that as k and n tends to ∞ , d_{min} tends uniformly to 0.

8. This remark is not actually true. On the contrary, Laplace himself resorted to infinite urns in posing the general problem of Bernoullian inductive inference that led him to the rule of succession (see note 1 above): *"if an urn contains an infinity of white and black tickets in an unknown ratio and we draw $p+q$ tickets from it, of which p are white and q are black, then we require the probability that when we draw a new ticket from the urn, it will be white"* (Laplace [1774] 1986, p. 365).
9. See Chapter 9.
10. With the same proviso as in note 5.

Chapter 9

Distributions*

Distributions (in one dimension) with an outline of cases in two or more dimensions. Concentrated in “points” or diffused (with either constant or variable density) probabilities (or “masses”); somewhat special examples, such as Cantor’s distribution. Cumulative distribution functions, $F(x)$ and, if it exists, density $f(x) = \frac{dF(x)}{dx}$. (Analogous notion in two or more dimensions, $\frac{\partial^2 F(xy)}{\partial x \partial y}$ etc..)

Introductory Concepts

Today I will talk about probability distributions and frequency distributions. Although for the purposes of the present course probability distributions are of greater interest, frequency distributions still have a certain importance for they are related to the former.

To have a concrete grasp on the concept of a distribution in general, it is useful to think of a distribution of masses.

Distributions can be classified in various ways (Fig. 9.1). Besides the distinction between probability distributions and frequency distributions, we must distinguish distributions on the integers (for example the distribution of families according to the number of children) from the distributions on the real numbers. Another classification can be done according to the number of dimensions. A distribution, in fact, can have an arbitrary number of dimensions.

Of course, the distributions on the integers are such that their possible values range over the whole numbers (either positive or negative). And if a distribution is on the real numbers, its possible values range over the real numbers. In this case, if it is a distribution in one dimension, its values will all belong to the real line, whereas if it is a distribution in two dimensions, its values will be the points on a plane (like, for instance, the point with coordinates (3,8) or the point with coordinates (3,4) represented in Fig. 9.2). To every point we can associate a mass. Let us take, for instance, the distribution of families according to the number of children. Let us put on the abscissa the number of sons and on the ordinate the number of daughters. The point with coordinates (3,4) would then correspond to the families with three sons

*Lecture XI (Wednesday 4 April, 1979).

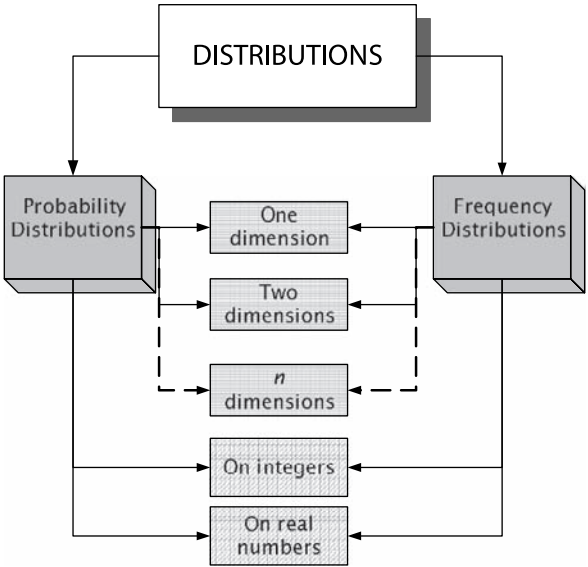


Fig. 9.1 Classification of distributions

and four daughters. And, of course, the same can be done in the continuous case, whenever for each case one is interested in a two-fold outcome. A case in point could be the statistical representation of the distribution of individuals according to weight and height. One of the two axes would represent the height and the other one the weight (of course according to the corresponding scales as those are not homogeneous quantities) and analogously there could be n -dimensional distributions for an arbitrary n . Distributions in one dimension constitute the fundamental case. Figure 9.3 represents the density of the distribution of a quantity which has a

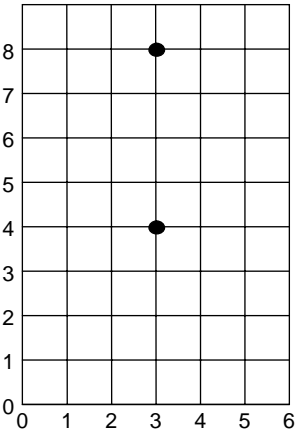
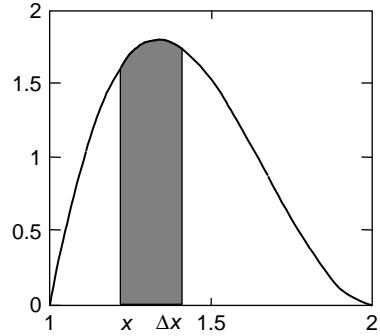


Fig. 9.2 Graphical representation of a two-dimensional distribution on integers

Fig. 9.3 Density of a distribution



minimum of 1 and a maximum of 2. As can be observed, such a distribution has a density value at every point. But what is the density of a distribution? Every value $f(x)$ of a density function can be written as a differential:

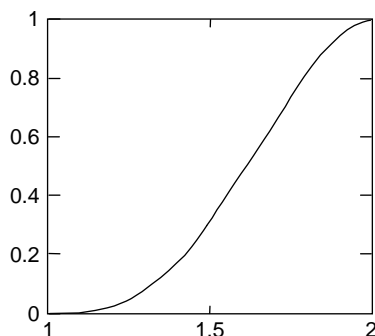
$$f(x) = \frac{dF(x)}{dx}.$$

Figure 9.3 shows how the area under a density included between the abscissas x and $x + \Delta x$ represents the mass included in the interval $[x, x + \Delta x]$. The density curve is commonly used to represent the distributions of empirical quantities. It must be said, however, that the $f(x)$ usually provides a slightly idealized representation of such a distribution. Strictly speaking, the density here should indicate the fraction of the individuals born at each instant of a given span of time. But it would be very hard to attach a well-determined meaning to the expression “instant of birth.” Indeed, even if we wanted to express the moment of birth with an approximation on the order of magnitude of microseconds or nanoseconds, this would still be only an apparently exact representation. This precision would in fact be fictitious, unless one takes a distribution to be a general mathematical entity which in practice will hold — as any other theoretical entity which is applied in practice — with those approximations that, case by case, turn out to be required.

Cumulative Distributions

There is however a way of representing distributions which requires a lesser degree of idealization. It involves plotting the cumulative distribution function $F(x)$ (Fig. 9.4). The cumulative distribution function associates to each point on the abscissa the mass included in the interval $[-\infty, x]$ (or $[0, x]$, if negative values are not possible). Whenever the density exists and is continuous, then every coordinate of the diagram of $f(x)$ represents the slope of the ordinate of $F(x)$ and therefore the maximum point of $f(x)$ corresponds to the point of $F(x)$ with the maximum slope. In short: the function $f(x)$ is the *derivative* of the function $F(x)$.

Fig. 9.4 Diagram of a cumulative distribution function $F(x)$



If density existed, yet it were discontinuous, then there would be a jump in the cumulative density function $F(x)$ and corresponding to it there would be a corner point in the diagram of $f(x)$. This is a descriptive detail which I have mentioned to show, if roughly, the relation which ties the behaviour of the function $F(x)$ to the one of the density $f(x)$. In the most general case, all cases are possible: a density could exist at each point, it might exist only in some interval, but it might also fail to exist in every interval.¹

The cumulative distribution function, on the other hand, always exists. It is always non-decreasing and — since it involves probabilities or relative frequencies — has an ordinate between 0 and 1. The curve of Fig. 9.4 is a regular, everywhere differentiable, curve and hence, has density everywhere. This is not always the case though. There could be, in fact, concentrated masses. In this case there would be a jump in the diagram of the cumulative distribution function in correspondence to those. Of course, as the sum of such jumps must be less than 1, there can only be a countable infinity of them. In this case they would give rise to a convergent series. Those distributions where all the mass is concentrated at a set of points which is at most countable (*discrete distributions*) and those where the distribution is continuous and has density, are the two most frequent pure cases.

Continuous Distributions Without Density

Yet there is also a third pure case: *the case in which all the mass is continuously distributed over a set of null measure.*² This is not a case of masses that are concentrated at certain points but of a continuous distribution, which, however does not have a density. I shall illustrate this third pure case through the construction of a specific example: *Cantor's distribution*.

The idea here is to construct a set of null measure such that all the mass is included in this set. Let us take, on the abscissa, the segment $[0, 1]$ (which contains the domain of the cumulative distribution function, since probability ranges in the interval $[0, 1]$). Let us divide this segment into three parts (they need not be equal, yet, for the sake of simplicity, I shall make this assumption). The first step (Fig. 9.5)

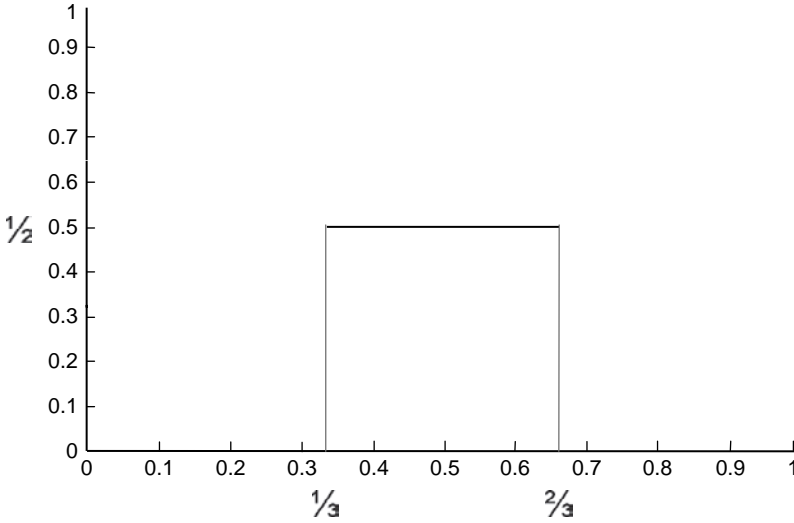


Fig. 9.5 The first step in the construction of Cantor's distribution

consists in establishing that there is no mass associated with the central segment, that is to say the interval $[\frac{1}{3}, \frac{2}{3}]$, so that $F(x)$ is a fixed constant on that segment. For symmetry reasons I shall assume (yet this one too is an irrelevant condition) that exactly half of the mass is contained in the interval $[0, \frac{1}{3}]$ whilst the remaining half is contained in the interval $[\frac{2}{3}, 1]$.

The second step (Fig. 9.6) consists in repeating the same operation on both intervals $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$. Hence, we shall divide the segment of the abscissa that spans from 0 to $\frac{1}{3}$ in three equal parts: from 0 to $\frac{1}{9}$, from $\frac{1}{9}$ to $\frac{2}{9}$, and from $\frac{2}{9}$ to $\frac{1}{3}$. Analogously, we shall divide the segment of the abscissa which spans from $\frac{2}{3}$ to 1 into segments which span respectively from $\frac{2}{3}$ to $\frac{7}{9}$, from $\frac{7}{9}$ to $\frac{8}{9}$, and from $\frac{8}{9}$ to 1. There will be no mass in the central segments (on either side). This translates into the fact that the diagram of the cumulative distribution function runs parallel to the abscissa in those segments that span from $\frac{1}{9}$ to $\frac{2}{9}$ and from $\frac{7}{9}$ to $\frac{8}{9}$. This in turn implies that in the segment spanning from 0 to $\frac{1}{9}$ we have $\frac{1}{4}$ of the total mass and analogously, for the segments from $\frac{2}{9}$ to $\frac{1}{3}$, from $\frac{2}{3}$ to $\frac{7}{9}$, and from $\frac{8}{9}$ to 1.

The third step consists in repeating the same operation for every segment with non-null mass of length $\frac{1}{9}$. Then the process will be repeated for the segments of length $\frac{1}{27}$, and so on. Cantor's distribution is nothing but that distribution in which this process converges as the number of steps tends to infinity. The result is represented in Fig. 9.7, where the flickering lines indicate that the curve is made of tiny "ladders" with increasingly smaller "steps." In Cantor's distribution the whole mass is therefore contained (with infinite density) in a set of null measure. Indeed, if in the first step the mass is contained in two intervals of length $\frac{1}{3}$, in the second

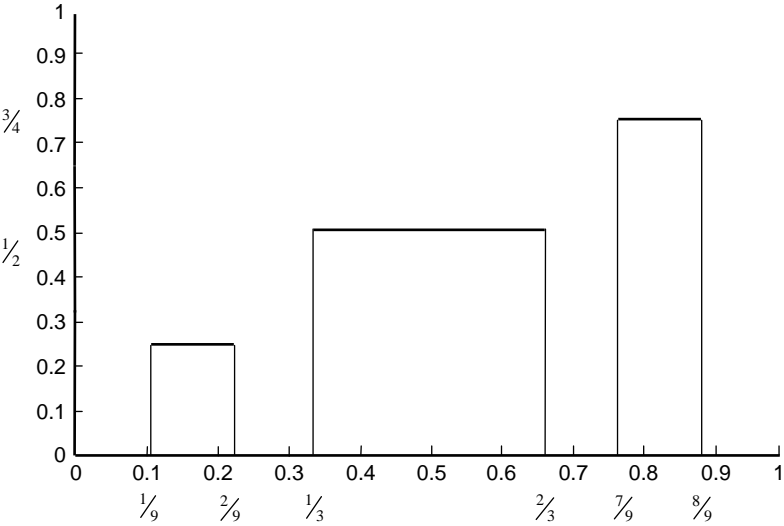


Fig. 9.6 The second step in the construction of Cantor's distribution

one it will be contained in four intervals of length $1/9$ and in general, in the n -th step, it will be contained in 2^n intervals of length $(\frac{1}{3})^n$. If we let n tend to infinity, the whole mass tends to be contained in intervals of a total length tending to 0. Hence, jumping to the limit, we obtain a distribution in which the whole mass is contained in a set of total measure 0.³

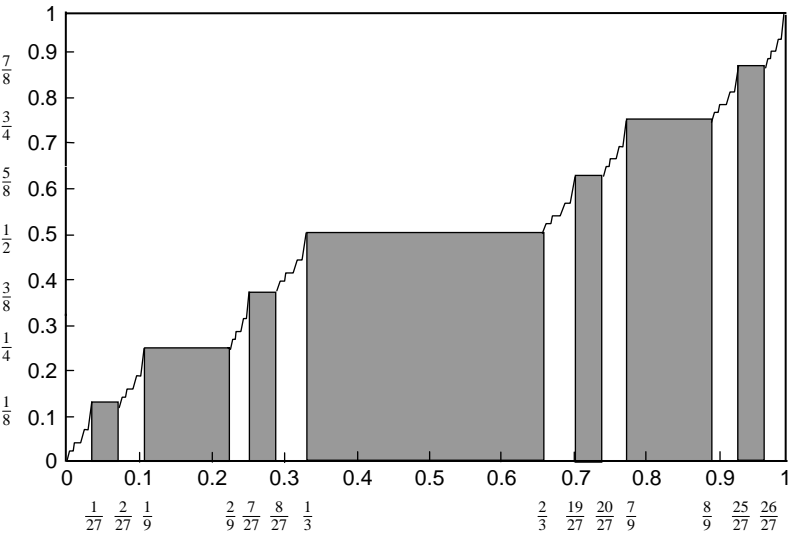


Fig. 9.7 Cantor's distribution

The General Case

What has been said so far holds for distributions of the pure kind, that is where the whole mass is distributed discretely or continuously with density, or else in a continuous way, yet on a set of null measure. In general, however, a distribution contains masses of each of those kinds. One therefore asks: given a distribution, how can we obtain, in general, its measure?

Let us observe Fig. 9.8. Corresponding to the jumps (vertical strokes) there are masses concentrated at single points. And, as has already been said, there can be 0, a finite number or even a countable infinity of those. It has also been said that if there is a countable infinity of concentrated masses, their sum $\sum c_h$ must be less than, or at most equal to 1. If we eliminate the concentrated masses, we obtain that part of the mass which is not concentrated, that is to say the continuous component. Moreover, there might be some bits where there are no jumps, yet the masses are distributed as in Cantor's distribution. In this case some intervals would exist where there would be no density. After eliminating both these latter and the jumps, we would finally obtain the segments of the third kind.

Thus, in a distribution, those three components can be distinguished. And whatever remains after we have eliminated them — so to speak — “infinitely many, infinitely small jumps” constitutes that part of the mass which has density. Clearly, however, it is not possible to distinguish *pictorially* the continuous component which has a density from the one which does not. Indeed we cannot represent pictorially “infinitely small” intervals. Not even resolution of a billionth of a micron would be enough, because it would still be infinitely inferior to the required resolution. Of course, analogous (if much more complex) remarks could be made with respect to distributions in the plane, in the space and so on.

Characteristic Functions

There is a way of working with distributions that is useful in many contexts. This amounts to considering, instead of the cumulative distribution function and the density function (provided that this latter exists), the *characteristic function*. This

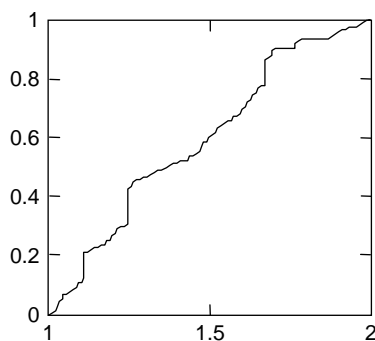


Fig. 9.8 Cumulative distribution with concentrated masses

involves a representation of the distribution in terms of an anticlockwise rotation from $-\infty$ to $+\infty$ along the circumference of the unit circle. From an analytic point of view, the characteristic function of a random quantity X is the function $\phi(u)$, which gives the prevision of e^{iuX} :

$$\phi(u) = \mathbf{P}(e^{iuX}) = \mathbf{P}(\cos uX + i \sin uX) = \mathbf{P}(\cos uX) + i\mathbf{P}(\sin uX). \quad (9.1)$$

It is useful to resort to the characteristic function when facing certain operations that require integration with respect to the density or the cumulative distribution function. If one works with the characteristic function, those integrations translate into elementary operations. Of course, first of all, one will have to convert the cumulative distribution function in terms of the characteristic function, then perform the intended operations and finally convert back the obtained result in terms of the cumulative distribution function. The conversion is not an elementary operation but when one works with characteristic functions, everything becomes really simple. Suppose, for example, that $Z = X + Y$, with X and Y independent random quantities. In this case, the following property holds:

$$\mathbf{P}(e^{iuZ}) = \mathbf{P}(e^{iu(X+Y)}) = \mathbf{P}(e^{iuX})\mathbf{P}(e^{iuY}). \quad (9.2)$$

In other words: if Z is the sum of X and Y , then the characteristic function of Z is — if X and Y are independent — the product of the characteristic functions of X and Y . The variable u can be either real or complex: the choice between one or the other is purely a matter of convenience. To each distribution corresponds its own characteristic function. Whenever a certain operation is performed between two distributions, a corresponding operation always exists between the respective characteristic functions.

Composition (Faltung) constitutes the most important case. If X and Y have cumulative distribution functions F_X and F_Y , and their respective characteristic functions are ϕ_X and ϕ_Y , then the product $\phi_Z = \phi_X \phi_Y$ provides the characteristic function of the composition $F_Z = F_X \star F_Y$. The operation of composition requires integration. Indeed, F_Z , assuming for simplicity that densities f_X and f_Y exist, is given by the formula:

$$F_Z(Z) = \int_{-\infty}^{+\infty} F_Y(Z - Y)f_X(X)dX = \int_{-\infty}^{+\infty} F_X(Z - X)f_Y(Y)dY. \quad (9.3)$$

Thanks to the method of the characteristic function, the computation of those integrals can be avoided. In fact, if X and Y are stochastically independent, then it is always possible — obtaining equivalent results — to switch to the characteristic functions ϕ_X and ϕ_Y , to make the product $\phi_Z = \phi_X \phi_Y$ and then perform the inverse operation so as to go back to F_Z .

About Means

I think I have said everything I intended to say. However, before concluding, I would like to hear what topics you would prefer me to cover during the next lectures.

DELTA: I would like to hear about means (associative means, etc.).

DE FINETTI: OK, then, I will address that topic tomorrow.

DELTA: In my opinion, means have a unifying character, yet it is really hard to see them covered adequately. Although in Italian statistics textbooks (and in particular those designed for the degree in Economics) one can indeed find a coverage of means that reflects a little bit of critical spirit, the Anglo-Saxon ones almost omit the topic. This is a bad thing because, in my opinion, means are the very essence of statistics.

DE FINETTI: Every mean provides a summary of the situation. Even when such a summary is adequate, it can never be so in an absolute sense, but only with respect to a certain goal. Consider some geometrical figures (for example cubes): ask yourselves which is the mean size. The answer is, the mean cube. By "mean cube," however, we can mean distinct things according to whether we are referring to the volume or the mass. We shall have to choose whether to integrate with respect to volume or with respect weight, depending on whether we are interested in the cube of mean weight or the one of mean volume. And this, in turn, depends on practical circumstances. If we have to carry weights, then we should be interested in the mean weight. Yet suppose that one has to fit empty and non-deformable cubical carton boxes into the boot of a car. It would obviously be inadequate, in a situation of this sort, to consider the mean with respect to weight: in such a situation, in fact, what is of interest is the volume.

Editor's Notes

1. In the standard measure theory, when a cumulative distribution function $F(x)$ is such that for every set A it holds:

$$F\{A\} = \int_A \varphi(x) m\{dx\}$$

φ is called the *density* of F with respect to the measure m . In the simplest case m is a volume in the n -dimensional cube $[0, 1]^n$. F is said to be *absolutely continuous* with respect to m if, and only if, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for every set of disjoint intervals A_1, \dots, A_n it holds:

$$\sum_{i=1}^n m(A_i) < \delta \quad \text{only if} \quad \sum_{i=1}^n F(A_i) < \varepsilon.$$

Absolute continuity should be contrasted with *tout court* continuity. F is said to be *continuous* with respect to m if, and only if, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for every single interval A it holds:

$$m(A) < \delta \quad \text{only if} \quad F(A) < \varepsilon.$$

Clearly, absolute continuity is a stronger condition than continuity. By Radon-Nikodym theorem, with respect to a certain measure m , there exists a density φ (uniquely determined up to values on a m -null set) of a cumulative distribution F if and only if F is absolutely continuous. When a cumulative distribution F is not absolutely continuous, it is called *singular*. This happens when F is a *discrete* (also called *atomic*) distribution, but it is perfectly possible that F is continuous without being absolutely continuous. In such a case, all probability is concentrated in a m -null set.

It should be added that while the Radon-Nikodym theorem and related results are typically proved, in standard textbooks, assuming countable additivity — which was rejected by de Finetti (see Chapter 12 pages 114–117) — the same results may be proved, albeit in a rather more elaborated way, using finite additivity (see Fefferman, 1967).

2. By insisting on this possibility, de Finetti is actually attacking the measure-theoretic view of probability which, in his opinion, turns out to be untenable in those cases in which probability 1 can be coherently assigned to a set of null measures.

Moreover, de Finetti rejected the dominant view according to which such distributions, being (to quote Feller's words) "not tractable by the methods of calculus" (Feller [1966] 1971, p. 139), should be avoided, so that "[f]or analytic purposes one is therefore forced to choose a framework which leads to absolutely continuous or atomic distributions." (ibid.) This widespread attitude would be justified if singular continuous distributions were irrelevant pathological exceptions. But this is not the case. On the contrary, singular continuous distributions arise naturally in many theoretical as well as practical contexts and are in general conceptually clean cut. Consider, for instance, the uniform probability distribution defined on the set V of unit vectors in \mathbb{R}^2 with random direction. Such a distribution is continuous and concentrated on the unit circle. Since the area of the unit circle is null, V is a set of null Lebesgue measures, so that F is singular with respect to the latter (Feller [1966] 1971, p. 140). Another example, discussed extensively later in the text, is Cantor's distribution. Brushing under the carpet such distributions only because they are recalcitrant to fit the framework of the measure-theoretic view of probability, is exactly what de Finetti considered as an unacceptable 'ad hoc' for mathematical convenience' (cf. Chapter 4 page 36 and note 5 page 42).

3. According to de Finetti Cantor's distribution — like several others continuous singular distributions (see note 2 above in this chapter) — is not "a pathological example with no practical meaning. On the contrary, we can give a simple practical example of a problem in probability where such a distribution arises" (de Finetti [1970] 1990a, p. 225). Suppose that a real number r in the closed interval $[0, 1]$, represented in base 3, is determined by a numerable sequence of random drawing from an urn containing only balls with the digits '0' and '2' in the same proportion. The outcomes of drawing determine, in succession, the digits to the right of the decimal point. Clearly, the fixed-point ternary representation of r cannot contain the digit '1'. Therefore the probability that r belongs to the set C of real numbers in $[0, 1]$ whose fixed-point ternary representation does not contain the digit '1' is 1. Now, there is a one-to-one correspondence between C and the points of Cantor's set resulting from the transfinite procedure described in the text. Therefore, C has measure 0 and probability mass 1. It should be stressed that de Finetti, by considering every set (including singletons) as a possible event, did not maintain that a probability value should be assigned to all subsets of a given space at the same time. On the contrary, de Finetti argued that a probability value should be assigned only to those events that are actually involved in the practical situations in which they are used. In the light of the *Fundamental Theorem of Probability* (see Chapter 12 page 117 and note 9 page 124) "we can always proceed in an 'open-ended' way, adding in new events and random entities from outside any prefabricated scheme; ... in evaluating probabilities (or a probability distribution), one should also proceed step by step, making them, little by little, more and more precise, for as long as it seems worth continuing" (de Finetti [1970] 1990a, p. 236), possibly resorting, where necessary, to transfinite induction.

Chapter 10

The Concept of Mean*

On the concept of mean (according to the general definition given by Chisini); historical outline of Chisini's idea. Some kinds of means, and in particular "associative means" and their property of being transforms of the arithmetic mean. Other means: e.g., the anti-harmonic mean, which is not associative.

Chisini's Serendipity

As a basis for this lecture I shall use the entry *Statistical Distribution*, which I have written recently for the *Enciclopedia Einaudi* (1978a, §2, pp. 1176–225). More precisely, I shall read — with appropriate comments and expansions — that part of the paper dedicated to the concept of mean. This text contains some important and instructive points. I shall begin by reading the section named *The notion of mean (Oscar Chisini: a fertile doubt)* (Section 2.2, pp. 1196–7). Chisini was in fact the first to give the general definition of the concept of mean, in place of the usual approximative definitions. I shall also read the section *The "definitions" and "the" definition* (Section 2.3, pp. 1197–8), which covers the difference that intervenes between the usual formulas with which each type of mean is defined and *the* general definition of the concept of mean under which all the single means are subsumed.

The first section that I am going to read to you concerns the history of Chisini's discovery. It illustrates how Chisini obtained his result. Oscar Chisini, for those who do not know it, was a geometry scholar. He wrote, with Federigo Enriques, the *Lezioni sulla teoria geometrica delle equazioni e delle funzioni algebriche*[†] (Chisini [1915–1934] 1985).¹ He was my professor in Milan. As we shall see, Chisini obtained his definition by chance. It must be said, however, that he did not fully understand the importance of his discovery, which he limited himself to mentioning in a short paper in the journal "Periodico di Matematiche" (1929). This work, in the author's intention, had no special scientific ambition.²

However, his idea is very important from the point of view of the philosophical foundations and constitutes a genuine conceptual advance. Hence, I begin to read:

* Lecture XII (Thursday 5 April, 1979).

[†] (Translator's note) *Lectures on the geometric theory of algebraic equations and functions*.

The notion of mean (Oscar Chisini: a fertile doubt).

Since time immemorial, men (scientists as well as laymen) have been using means without fully realizing the uniform concept and the effective requirements underlying this notion. Those were *ad hoc*eries *ante litteram*. (de Finetti, 1978a, p. 1196)

The word “ad hocery” was coined by Irving John Good, a probabilist very close to the subjectivistic approach (indeed, surely a member of it). Good criticizes as ad hoc those cut and dried rules and stereotyped notions that are often accepted without reflection.³ And indeed, to have as many definitions as there are means and to use them randomly and without reference to the goal for which they are introduced, is precisely to make use of cut and dried *ad hoc* rules. It should be checked, for each one of them, what is the goal for which they were introduced. This was, in essence, Chisini’s proposal. But let us resume the reading:

And such would they still be today if... if one year Oscar Chisini weren’t appointed as an exam commissioner in a secondary school. Owing to the reflections that those exams suggested to him, we must now refer to Chisini as the “discoverer of the concept of mean.” It might sound like a joke (like Campanile’s story on the “inventor of the Horse”),⁴ but it is not an invention of something new, rather a clarification of a concept which until then was confused and lacking of a uniform and deep characterization.

The anecdote, or the story, of those exams deserves to be told, because it is instructive from many points of view. Chisini, a university professor, pays attention to the stereotyped questions on the various kinds of mean (arithmetic, geometric, harmonic, etc.); he finds them very dull, yet instead of getting distracted, he becomes very keen on finding what, if any, is the concept underlying so many disconnected notions.

He thinks of it, not as a “pure mathematician,” yet (even greater merit!) as an intelligent person who *also* happens to be a mathematician. And he asks himself why. Why means? Why have they been introduced? Why are they USEFUL? (How many would not be appalled on posing themselves or even on hearing a similar question: “*useful* mathematics?”... then it is not science, or at least, it is not Science!).

Yet Chisini thought carefully about it and found an answer which he considered satisfactory and which accounted for any mathematical, practical and philosophical demand; however, he could not fully appreciate its importance. By chance, he came across a problem that was not familiar to him; he responded to his intimate need for clarity and he limited himself to publishing on this topic a short popular paper for the journal “*Periodico di Matematiche*” Chisini (1929), which stressed the relative and functional meaning — “corresponding to a given goal” — of the general notion of mean. We shall discuss it soon, yet we still need some supplement to the story. The small paper by Chisini, which appeared in a general interest mathematical publication, would have gone unnoticed if, among his students, there were not one who became interested in probability and statistics (it was me),⁵ and for whom it was easy to grasp the importance of Chisini’s idea and to spread and apply it in the field for which it was most appropriately conceived and in which it should have been fertile. This resulted in a rather wide and systematic discussion which, on an important point, gained new impetus thanks to a result obtained independently just around that time by the Russian Andrei Kolmogorov⁶ and the Japanese Mitio Nagumo (the “Nagumo-Kolmogorov Theorem,” see Section 2.5).⁷

Before delving into the technical aspects of the topic, it is worth insisting a little bit more on its intuitive or, at least in a certain sense, “philosophical” presentation.

The general idea of the definition given by Chisini can be expressed very well in plain words. And similarly, many other things usually considered to be “only for experts” or “only for the initiates(!),” could be explained to everyone, or at least to many, in their own exact essence yet without obfuscating its central meaning with those technicalities required to understand a specialist work. This would be a major achievement (in spite of the risk

of seeing it corrupted or even ridiculed, by lightweight “vulgarizers”). But why shouldn't there be any one else who can express in a plain, precise, passionate and poetic way the facts and theories of science, as for instance, Michael Faraday did in that small masterpiece that is the collection of conferences for younger audiences named *The Chemical History of a Candle*.⁸ (de Finetti, 1978a, pp. 1196–7)

Faraday, besides the Royal Society conferences for scientists, began to give talks for younger audiences. And he made the booklet *The Chemical History of a Candle* out of them. It is always a bit exaggerated to say that something is “the most beautiful”: yet it is not exaggerated to say that this book is one of the most beautiful things I have ever read. Faraday (1791–1867) was a poor boy: he was a bookbinder. It once happened that he had to bind a physics book. He read it from cover to cover and became almost more knowledgeable than its author, Humphry Davy.⁹ It is incredible what Faraday was able to make out of those lectures (so scientific, so valuable, so instructive) from the simple phenomenon of a lit candle! But let us go back to reading the paper:

Maybe the problem (the reason why no one writes any more masterpieces for young audiences like this one)¹⁰ is that we are too specialized, isolated, unilateral, all enclosed in our own watertight compartments from which it is hard to have a bird's eye view of the many simple and essential things that could allow us to recover, if we established contact with them, the natural intelligence and spontaneity of a child!

Section 2.3 The “definitions” and “the” definition.

The “definitions” that Chisini heard the examiners asking and the students reciting parrot-fashion — and that rightly horrified him — were nothing but the single “recipes” to compute the various means. (“Given n numbers, their arithmetic mean is their sum divided by n ; the geometric one is the n -th root of their product; the harmonic one is n divided by the sum of the reciprocals; the quadratic one is the square root of the arithmetic mean of the squares,” and so on.)

Those are correct definitions (they would be appropriate for anyone who should learn them), however they are purely formal, they fail to explain *the* “why,” which is the essential thing. (Owing to this kind of void “indoctrination” the conclusion was reached — in the comparative assessment of mathematics learning in various countries promoted by the IEA — that Italians “know it all but not its purpose!” There is some exaggeration on both sides, yet basically the problem of vacuous sciolism is our handicap, something which annihilates every possibility of obtaining a concretely educational result.)

We should “think *à la* Chisini”: as to our case (but this applies to everything), what we must do is ask, with a critical mind, what is *the meaning* of the concept of mean, which means analyzing the deep and essential reasons which constituted, if unconsciously, the goal for which that concept was introduced and which explains the intimate reasons for its usefulness.

We certainly do not perform such an analysis — as Chisini observes — when we intend to define the “mean of various quantities given a new quantity included between the smallest one and the largest one among the quantities considered”¹¹ and we avoid or neglect it whenever we prefer to define directly, case by case, the individual kinds of mean that we usually come across, consequently doing “something that is correct, yet purely formal and anti-philosophical and which can only be used, if badly, empirically.”¹²

On the contrary, we must begin, indeed as Chisini says and does, by stressing that the pursuit of a mean has as a goal that of simplifying a given question by “substituting in it two or more given quantities with one single quantity synthesizing them, without altering the general view of the phenomenon under consideration,” and we shall first of all notice

that “it makes no sense to speak of the mean of two (or more) quantities, but it makes sense to speak of their mean with respect to the synthetic evaluation of another quantity that depends on it.”

Chisini translates this concept in the form of a definition as follows: if we are interested in the function (symmetric: that is whose value is invariant under any reordering of the variables) $f(x_1, x_2, \dots, x_n)$ of n homogeneous quantities x_1, x_2, \dots, x_n and if x is the value for which $f(x, x, \dots, x)$ (x repeated n times) gives the same value (that is if with respect to the computation of the function everything is *as if* the n variables all had that same value, x), we shall express this fact by saying that x is the mean of x_1, x_2, \dots, x_n with respect to the computation of f . Besides the conceptual and concrete value, as opposed to formalistic and abstract, of the definition, it appeared to be fascinating the relative, practical, pragmatic character of the underlying concept, as well as the presence of the “as if,” reminiscent of the “*als ob* philosophy” of Vaihinger.¹³ (de Finetti, 1978a, pp. 1197–8)

Vaihinger is an acute philosopher who lived between the end of the 19th century and the beginning of the 20th and who eventually committed suicide.¹⁴ An analogous philosophical orientation, of a pragmatistic kind, can be found in Vailati (of whom a large collection of works has been published, which puts together many more or less extensive works, all of which are very interesting).¹⁵ Vailati, in turn, was very distant from “scholastic” philosophers. But let us go back to reading:

The multiplicity and specificity of means (each one in its own way, according to the purpose) is something that needs to be born in mind in order to avoid simplistic and erroneous conclusions. The example Chisini gives in his note is very practical, simple and easy to remember: it concerns a journey by car at 60 Km/h speed for two hours, and 105 Km/h for one hour; average speed of 75 Km/h. However (according to an empirical formula quoted by Chisini), the higher fuel consumption in the fastest leg is not compensated by the lower consumption in the slowest leg and therefore, the mean consumption corresponds to an average speed of 80 Km/h.¹⁶ In other words (it is worth stressing and repeating it, though it has been said and repeated in various ways), the “as if” does not have a universal value, nor one that can be extended by way of seeming analogies, yet it is intrinsically tied with the specific hypothesis corresponding to that given mean and only that one. (de Finetti, 1978a, p. 1198)

Γ -Means and the Nagumo-Kolmogorov Theorem

I will omit reading the part dedicated to means in relation to distributions. Therefore, I resume reading at the point where the deepest meaning of the concept of mean is discussed:

Section 2.5 Associative means (The Nagumo-Kolmogorov Theorem)

The most common means . . . (geometric, harmonic, quadratic, besides, of course, the arithmetic one) are all examples of associative means; means, that is, for which the following convenient property holds: the mean does not vary if the data-sets are replaced by their mean (always, of course, according to the same meaning) . . . (de Finetti, 1978a, p. 1202)

In other words: if we partition the values into many classes and then take the arithmetic mean within each class, the mean of those partial means (taking as weights the sum of the weights of the elements of each class) coincides with the mean of all the values.

Associative means have the property of being *transforms of the arithmetic mean*. In order to transform an arithmetic mean of the quantities x_1, x_2, \dots, x_n (with weights p_1, p_2, \dots, p_n) the following steps must be followed:

1. Alter the scale. For example, instead of expressing the x_i s in the natural scale, they could be expressed in the logarithmic scale. In this case, instead of the values x_1, x_2, \dots, x_n one should take the values $\log(x_1), \log(x_2), \dots, \log(x_n)$. In place of the logarithm, of course, one could use the square, or the cube and in general each increasing and invertible function $f(x)$.
2. Compute the weighted arithmetic mean

$$m_f = \frac{\sum p_i f(x_i)}{\sum p_i}$$

of the values $f(x_1), f(x_2), \dots, f(x_n)$.

3. Revert to the original scale by computing the inverse image (in the functional sense) of m_f , that is to say $f^{-1}(m_f)$.

For example, if one wanted to obtain the simple geometric mean, one should first of all compute the sum of the logarithms divided by n and then raise e to the power of the result of such an operation. With these anticipations, we can resume reading:

This property is so obvious that no one runs the risk of not being able to understand it, the risk rather being that of thinking that it always holds. A very concrete example is enough to show that this is not the case, namely that of the anti-harmonic mean which, to put it in plain words, is the “arithmetic mean of n positive quantities x_i where we take as weights the x_i ’s themselves”; of course this wording is not fully correct, yet its sense is. (Arithmetically: it is the sum of the squares divided by the sum of the values, that is, equivalently, the mean of the squares divided by the arithmetic mean.) For a physical interpretation we can say that it gives the “reduced length” of a compound pendulum, that is to say the length of an ideal simple pendulum (a punctiform mass suspended at a distance l from the fulcrum by an inextensible string of negligible mass), which oscillates at the same frequency. And it is not true that if we connect two pendulums the frequency of their joint motion will be given by the anti-harmonic mean of the corresponding reduced lengths. (Unlike its two “ingredients”: the distance between the fulcrum and the centre of mass and the moment of inertia.)

The already mentioned theorem, proved independently and almost at the same time by the Japanese Mitio Nagumo and the Russian Andrei Kolmogorov,¹⁷ tells us what is the general form of associative means: these are all and only the “transforms” of the arithmetic mean. Such are, for instance, the geometric mean, as it is the square of the product: the product of the values x_1, x_2, \dots, x_n is unchanged if those are individually replaced by the geometric mean x , that is by considering x^n ; the harmonic mean, as its reciprocal is the mean of the reciprocals; the quadratic mean (for positive values x_i), as the mean of the squares x_i is the square of the quadratic mean; and so on.

The concept of associative mean is very broad because one can consider such a mean in reference to an arbitrary increasing function $\gamma(x)$ (in place of the $\gamma(x) = x$ one could take, for instance $\gamma(x) = 1/x$, $\gamma(x) = x^2$, $\gamma(x) = \log x$ etc., all of which give rise to the above mentioned means: arithmetic, harmonic, quadratic, geometric). (de Finetti, 1978a, pp. 1202–3)

This is fairly natural. In fact, as I have already said, instead of taking

$$\frac{p_1x_1 + p_2x_2 + \dots + p_nx_n}{p_1 + p_2 + \dots + p_n},$$

we take the $\gamma(x_i)$'s, we compute their mean and then we revert to the original scale by applying the inverse function:

$$\gamma^{-1} \left(\frac{p_1\gamma(x_1) + p_2\gamma(x_2) + \dots + p_n\gamma(x_n)}{p_1 + p_2 + \dots + p_n} \right).$$

If “ γ ” meant “raise to the second power,” then we would have the quadratic mean (possibly weighted), if it meant “logarithm,” we would have the the geometric mean, and so on.

Statistical Theory of Extremes and Associative Means

DELTA: Excuse me, may I ask a question?

DE FINETTI: Yes, of course.

DELTA: By taking at most n values, we would obtain something similar to a γ -mean. Indeed, if we replaced some of those values by their maxima the maximum of those maxima would be the maximum of n values. However, the function *maximum* is not a transform of the arithmetic mean. What I wanted to know is whether this is due to the fact that the maximum of n values, once the maximum of the first $n - 1$ values is given, is compatible with many possible values for the last variable.

Everywhere in statistics one can observe a certain analogy between the sums, the observations and the maximum of the observations. Given certain quantities x_1, x_2, \dots, x_n , the maximum y of those can be seen as an associative mean (as you defined it) with respect to the *maximum* function. Indeed, y equals the maximum of a vector constituted of coordinates that are all equal to y , for it holds that $y = \max(y, \dots, y)$. Moreover, the associative property holds too, as the value is invariant under the substitution of sets of components with their maximum. However, the function *maximum* is not a transform of the arithmetic mean. Hence, what I am asking is: can this depend on the fact that for the sums of the form

$$x_1 + x_2 + \dots + x_n = r \tag{10.1}$$

there is a unique possible value of x_i for which, all the other summands being fixed, the equation (10.1) is satisfiable, whereas in the analogous relation for the maximum:

$$\max(x_1, x_2, \dots, x_n) = r \tag{10.2}$$

there are many values of x_i for which, the x_j 's being fixed for $i \neq j$, the equation (10.2) is satisfiable?

DE FINETTI: Although the operation *maximum* satisfies some of the typical properties of means, it does not look advantageous to me to characterize it as a mean.

DELTA: I doubt it would as well.

DE FINETTI: However, it is not correct to say that the function *maximum* does not belong to the γ -means. The function *maximum*, in fact, is a limiting case of γ -means. Indeed, the value of γ -mean increases (and therefore gets closer to the maximum) with the rate of growth of the function γ . If we put:

$$\gamma(x) = e^{kx}$$

we would use a function which grows very rapidly: the result would therefore be very close to the arithmetic mean. If we took:

$$\gamma(x) = e^{k^x}$$

with $k = 10^{100\,000}$ we would get a mean even closer to the maximum. If we then took:

$$\gamma(x) = e^{e^{\dots e^{kx}}}$$

we would get as close as we like to the limiting case which — trivially — is the maximum itself. Therefore, the function *maximum* can be characterized as a γ -mean, though as a limiting case. In fact it can be written as:

$$\max(x_1, x_2, \dots, x_n) = \lim_{i \rightarrow \infty} \gamma_i^{-1} \left(\frac{\gamma(x_1) + \gamma(x_2) + \dots + \gamma(x_n)}{n} \right) \quad (10.3)$$

for a suitable sequence of functions γ_i such that, for all x , the following holds:

$$\lim_{i \rightarrow \infty} \gamma_i'(x) = \infty. \quad (10.4)$$

DELTA: OK, OK, I understand.

DE FINETTI: I do not rule out the possibility that this characterization of the maximum as the limit of associative means might turn out to be useful in some cases. Yet if one raised the question as to how the function *maximum* should be framed in the context of associative means, then there is nothing to say but this: for a suitable choice of the function γ , any value included between the minimum and the maximum can be obtained. And to the limit, even the minimum and maximum themselves.

DELTA: I did not pose this problem to raise a merely formal question, but in relation to some problems in applied statistics, and in particular in the so-called *Statistical Theory of Extremes*. In this theory one makes inferences about the behaviour of the

minimum and the maximum of future experiments. I wanted to know whether in that context the property of associative means to be transforms of the arithmetic mean might turn out to be useful.

DE FINETTI: It depends on the way the problem is posed.

Inequalities Among Associative Means

I would now like to discuss the way in which a γ -mean can be represented pictorially (Fig. 10.1). Suppose that on the abscissa we put weights (masses) p_1, p_2, \dots, p_{10} at points x_1, x_2, \dots, x_{10} , respectively. The centre of mass of such points will sit itself on the abscissa, at point $(m_x, 0)$. The value m_x is the arithmetic mean of the x_i 's weighted by the p_i 's. Let us note that if we translate the masses perpendicularly to one of the axes, the coordinate of each mass is unchanged with respect to that axis. As a consequence of this, the coordinate of the centre of mass is also unchanged with respect to that axis. Let us suppose then that the masses p_i are displaced vertically from the point $(x, 0)$ to the point $(x_i, \gamma(x_i))$ on the curve $\gamma(x)$. The first coordinate of the centre of mass stays fixed at m_x . Let us translate now the same masses horizontally (parallel to the abscissa) until the ordinate axis is intercepted at point $(0, \gamma(x_i))$. The centre of mass is now at point m_y , the arithmetic, weighted mean of the $\gamma(x_i)$. Before making this last displacement (when the masses lie on the curve) the centre of mass was at (m_x, m_y) . The γ -mean m_γ , on the other hand, equals the value of the abscissa of the point of the curve with ordinate m_y .

I would like to add that if the function γ is always concave or convex upwards, then the centre of mass (whenever the masses lie on the curve) lies inside the concavity and therefore the two γ -means have a *greater* or a *smaller* value, respectively, than the arithmetic mean (all weights being equal). Moreover, the same criterion applies also to the comparison of two associative means with respect to two distinct functions γ . Indeed, if instead of considering the case of a single function γ , we con-

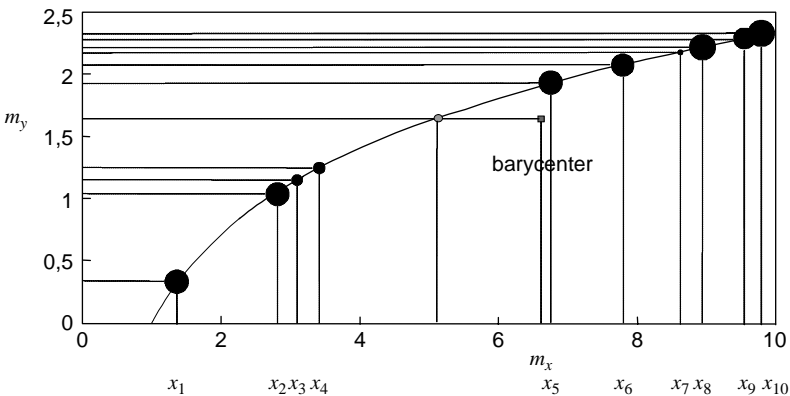


Fig. 10.1 Weighted γ -mean and weighted arithmetical mean of a function $y = \gamma(x)$

sidered two functions $\gamma_1(x)$ and $\gamma_2(x)$ and if we assumed that $\gamma_1(x)$ lay on the y-axis and that $\gamma_2(\gamma_1^{-1}(x))$ lay on the x-axis, then the frame would change but not the result. Moreover, even without the graphical comparison, it is enough to note that “greater relative concavity” corresponds (locally) to a greater value of the ratio between the second derivative of $\gamma(x)$ and the first one, $\gamma_1''(x)/\gamma_1'(x)$ (if, in the interval of interest, it is not invertible). Therefore, if we had a curve whose concavity is always on the same side, then it would be sufficient to consider this fact. In particular, it can be recalled that among the “means of powers” — those obtained from $\gamma(x) = x^n$ — the mean grows with the exponent n , thus in particular, the following inequalities hold:

$$\text{harmonic} < \text{geometric} < \text{arithmetic} < \text{quadratic} < \text{cubic}. \quad (10.5)$$

The harmonic mean is

$$\frac{N}{\sum_{i=1}^N x_i^{-1}} = \left(\frac{1}{N} \sum_{i=1}^N x_i^{-1} \right)^{-1}$$

and therefore, in this case $\gamma(x) = x^{-1}$. In general, the mean of power n is defined by the formula:

$$\mu_n = \left(\frac{1}{N} \sum_{i=1}^N x_i^{-1} \right)^{\frac{1}{n}}.$$

The case of the arithmetic mean is $\gamma(x) = x^1$, the case of the quadratic mean is $\gamma(x) = x^2$ and the case of the cubic mean is $\gamma(x) = x^3$, and so on.

I believe that this might be enough. I would like to insist on the fact that the most interesting thing about this theory is the Nagumo-Kolmogorov theorem. Does anyone have any questions?

Concluding Remarks

ALPHA: I would like to know whether the Nagumo-Kolmogorov theorem was already implicit in what Chisini was saying or in what you were saying.

DE FINETTI: Chisini had the merit of paying attention to the trivial questions of the exam he witnessed, reflecting on them and then arriving at a conclusion. But he did not work on means thereafter. He wrote those few pages for *Periodico di Matematiche*, of which I believe he was the managing editor (unless it was Enriques).¹⁸ For him it was worth showing that the cut and dried questions that were asked of high-school students specializing in accountancy (or maybe building survey) were of conceptual, scientific interest. Moreover, I learnt from Chisini that even the most complicated formal notions become intuitive if one looks at them from an appropriate point of view. Chisini always tried to explain things intuitively. For example, in algebraic geometry, Chisini used to show how to turn around or, how

to go from one sheet of paper to the other. He even managed to “let people see” the objects of a hyperspace. I find this effort to make things intuitive useful.¹⁹ Many colleagues — on the other hand — believe that any attempt to state the results in an accessible language rather than with pompous formulas, is childish and degrading. For instance, it once happened that a colleague asked me for a suggestion about how to explain a given topic. After illustrating to him my teaching proposal, he exclaimed in the Triestine dialect: “it is like giving daisies to pigs: in this way everyone understands it!” I believe, on the other hand, that it is always worth making an effort to let people understand things. This is one of the reasons why I took part in the organization of mathematical competitions among students of secondary schools.²⁰ In my opinion mathematical competitions are much more instructive than studying textbooks. The latter are inevitably boring, for they are books that have to be studied compulsorily. And nothing is taken more seriously than what is done for fun.

BETA: This applies to me as well. In fact, I registered for the degree in mathematics because I could not solve a problem. At a certain point I asked myself: “can it be possible that I am not able to do it?”

DE FINETTI: It often happens, when facing a difficult mathematical problem, that one does not know where to start.

ALPHA: Do you believe that mathematics is valuable only because it is susceptible to being applied? I refer to pure mathematics. In fact one can attempt to render the theorems of pure mathematics concrete in terms of some application.

DE FINETTI: Certainly. Otherwise mathematics would become some sort of machine with no purpose. Even an electronic machine performs calculations by applying certain rules and so arrives at certain results. But by so doing it does not learn anything. The machine has performed nothing but purely mechanical transformations. I do not want to claim that the only relevant aspect of mathematics is the intuitive one. However, this should never be completely neglected. For instance, it might look as if intuition has nothing to do with solving a system of linear equations in many variables. Yet, if one represented such a system in vectorial form, one could *see* what it amounts to.²¹ If, instead of considering the n -tuples of numbers a_1, a_2, \dots, a_n , one looked at the vector \mathbf{a} of components a_1, a_2, \dots, a_n in an n -dimensional space, then it would be possible to “see” (if only in an approximate representation, maybe after reducing the n -dimensional vectors to 3-dimensional vectors, as no one has ever managed to take a trip in a 4-dimensional space). And if we thought, with Einstein, of time as a fourth dimension, then space-time would be in four dimensions and, as a consequence of this picture . . .

ALPHA: [interrupting] We would not go beyond the fourth dimension anyway . . .

DE FINETTI: 4 is already 1 more than 3!

BETA: I suppose that there are applications where spaces in many dimensions are used.

DE FINETTI: Of course, especially in the area of operational research, where matrices of dimensions as big as one hundred are often used. The same holds for game theory. Indeed, to solve such complicated problems intuition is not enough: a computer is needed!

Editor's Notes

1. Oscar Chisini (1889–1967) studied with Federigo Enriques at the University of Bologna, from where he graduated in 1912. In 1923 he obtained the chair of Geometry at the University of Cagliari. From 1925 until his retirement (1959) he taught at the Technical University of Milan. Apart from the editing (with many additions) Enriques's lectures, he must be remembered as the author of important original contributions, especially in the field of algebraic geometry (see Caprino, 1981).
2. Chisini writes in the cited work:

[H]owever the concept of mean is so simple and so perspicuous that in order to recover its very nature, and its corresponding mathematical definition, it is enough to pay a little bit of attention to it, as the reader will have occasion to see from the following short reflections. Those Reflections I had an occasion to make accidentally, during a conference in the Milanese section of "Mathesis"; and they are so obvious that I would not have published them unless the many persons who were consulted (among whom many capable lecturers and distinguished scientists) did not appear to consider these observations actually interesting. (1929, p. 106)

3. Cf. Chapter 4 at page 36 and note 5 at page 42.
4. Achille Campanile (1899–1977) was an Italian writer and journalist. He wrote a number of plays and novels — among them, in 1925, the short drama *The Inventor of the Horse* (1995) — whose humour, resting on bizarre events and paradox, forerun the so-called "theatre of the absurd."
5. Orally added remark.
6. I correct here the original, which reads "Antonij Kolmogorov."
7. Reference to the paper "Sul concetto di media" de Finetti (1931b).
8. See Faraday ([1885–1889] 2002).
9. The English chemist Sir Humphry Davy (1778–1829) is celebrated for his pioneering studies on the electrolytic phenomena. It should also be added that after reading Davy's work in 1812, Faraday sent the reflections that he came up with following the reading to the author. Those were so highly appreciated that Davy appointed Faraday as his laboratory assistant at the Royal Institution in London see (see Williams, 1987).
10. Orally added parenthesis, not included in the original text.
11. Chisini (1929, p. 106).
12. Chisini (1929, *ibid.*). I rectify here the use of the inverted commas in various points of the paragraph.
13. I rectify the original text, which reads "Veihinger".
14. See Vaihinger ([1911] 1965). According to Vaihinger, German philosopher (1852–1933), all the concepts and theoretical principles of both science and philosophy are *fictions* whose value lies in their pragmatic use. The expression "as if" (*als ob*) was previously used by Kant (of whom Vaihinger was a great scholar, being the author of an important commentary on the *Critique of Pure Reason* and the founder of the "Kantstudien" journal) when he theorized the so-called *regulative use of the ideas of pure reason*. Kant, however, did not go so far as to consider those ideas as mere fictions. See the *Appendix to the Transcendental Dialectic* of the *Critique of Pure Reason* (Kant [1781–1787] 2004, A 642–704 and B 671–732).
15. See note 17 at page 58.
16. The empirical equation proposed by Chisini Chisini (1929, p. 107) for the average speed with respect to the consumption is the following:

$$v_m = 60 + \sqrt{\frac{s_1(v_1 - 60)^2 + s_2(v_2 - 60)^2}{s_1 + s_2}}.$$

17. Again (see note 6 above) I correct here “Antonij” with “Andrei.” See Kolmogorov (1930); Nagumo (1930). On this topic I refer to the already mentioned paper by de Finetti (1931b) and the following, later, works: Hardy, Littlewood and Polya (1934); Daboni (1974); Norris (1976); Ben-Tal (1977); Weerahandi and Zidek (1979); Chew (1983); Fishburn (1986); Muliere and Parmigiani (1993); Kotz et al. (2006).
18. Indeed at that time, the managing editors of the journal were F. Enriques and G. Lazzari. Oscar Chisini was the editorial secretary. “Periodico di Matematica” (a journal mainly addressed at secondary school teachers) began to be published in 1886 under the direction of Davide Besso. In 1921 Enriques became its managing editor and changed the journal name into “Periodico di Matematiche.” Enriques held the managing editorship until 1946. Chisini took it over and kept it until 1967, the year of his death. De Finetti succeeded Chisini and kept the managing editorship until 1978.
19. See the paper *Oscar Chisini and his teaching* published by de Finetti in memory of Chisini on “Periodico di Matematiche” de Finetti (1968a). The title of the lecture in which Chisini illustrated for the first time his analysis of the concept of mean: *Quello che vorremmo insegnare: la veduta matematica delle questioni* (Translator’s note: *What we would like to teach: the mathematical way of seeing questions*) (see Chisini (1929), p. 106 n. 2) is illuminating in order to appreciate Chisini’s influence on de Finetti. De Finetti, in fact, drawing his inspiration from this approach, wrote two demanding educational works: *Matematica logico-intuitiva* (Translator’s note: *Logico-intuitive mathematics*) (de Finetti [1944] 2005) and *Il saper vedere in matematica* (Translator’s note: *Knowing how to see in Mathematics*) (1967). The title of this latter work echoes the one of Chisini’s lecture.
20. This mostly happened during the seventies, when de Finetti, director of the *Mathesis* association, could organize such competitions.
21. See note 19 above.

Chapter 11

Induction and Sample Randomization*

Recollection and discussion of some of the topics covered before the Easter break. Initial sketches of controversial topics concerning some “definitions” of probability, and in particular the so-called “axiom of non-regularity” (Regellosigkeitsaxiom).

Exchangeability and Convergence to the Observed Frequency

I would like to discuss the relation between the concepts of *random experiment* and *exchangeable experiment*. After all, there is only a lexical difference between the two notions, which can be summarized as follows: *the expression “equally probable events with unknown but constant probability,” used by the objectivists does not make any sense from the subjectivistic point of view*, simply because there is no such a thing as an unknown probability (the probability being that which a certain person assigns at a certain time). However, what is typical of these cases is exchangeability: those cases in which one speaks of independent events with unknown but constant probability are, in fact, all cases of exchangeability.

However, behind this terminological difference lies a conceptual difference concerning the problem of inductive inference. The objectivists do not answer this question satisfactorily and in fact, they almost completely neglect it. Their argument goes as follows: since, in the long run, frequency coincides with probability, in order to determine the probability it is sufficient to observe a somewhat large number of experiments. From the subjectivistic point of view, this argument is unacceptable. Indeed, for us subjectivists, probability cannot be determined empirically but it is evaluated by everyone, at any instant, on the basis of one's own experience. Probability, in fact, changes with every new experience.

Suppose we are drawing from an urn containing white and black balls in unknown proportions. Suppose, however, that we know that the percentage of white balls is one of the following: 30%, 50%, 70%, 80%. I shall call the four possible hypotheses about the percentage of white balls H_1, H_2, H_3 and H_4 , respectively. Suppose that an initial probability is assigned to each of the hypotheses H_i respectively. As we continue the draws, those probabilities change according to Bayes' theorem. In fact, the probability of the hypothesis that is closest to the observed

* Lecture XIII (Friday 27 April, 1979).

frequency undergoes an increase. And it is probable that certain sequences obtain such that, in the long run, the probability of one of the hypotheses H_i will get really close to 1. And the probability relative to a single shot would be very probably very close to the observed frequency. However, we must always bear in mind the influence of the initial probabilities assigned to the hypotheses H_i .¹

ALPHA: However, the subjective differences are always tempered by this convergence. Therefore, the Bayesian method, provided that the condition of exchangeability is satisfied, is in some sense a self-corrective method (to use Reichenbach's term).²

DE FINETTI: Yes, it is. Who uses this term?

ALPHA: Reichenbach who, however, referred to the estimation of frequencies rather than subjective probabilities. According to him an estimation rule is self-corrective when the limit of the difference between the estimate obtained with that rule and the observed frequency is 0.

BETA: Hence, the subjective probability of one of the hypotheses converges to the value 1 as the number of experiments grows.

DE FINETTI: Yes, provided that it is borne in mind that all this does not hold necessarily but depends on the premises (exchangeability).³

BETA: Let us suppose that there are three urns: the first one containing only black balls, the second one only white balls and the third one half white and half black.

DE FINETTI: This is a very simple case. In fact, as soon as two balls of distinct colours were drawn, it would be known with certainty which urn is being used for the draws. If, on the other hand, only white or only black balls were drawn, then — as the number of shots grows — the probability that the draws are being made using the first of the second urn would rapidly increase.

BETA: At the beginning, the probability reflects the personal state of mind of whoever makes the evaluation. But then, as new draws are carried out, differences among people's opinions tend to disappear. Therefore, the growth of knowledge leads the opinions to converge.

DE FINETTI: Yes, the differences in the initial opinions have no other consequence than delaying the preponderance of the observed frequency over the initial opinion itself.

Bayesian Statistics and Sample Randomization

ALPHA: Let us now tackle the problem of the methods of random selection of statistical samples. Savage, in this booklet, which you might be familiar with . . .

DE FINETTI: What is the title?

ALPHA: *The Foundations of Statistical Inference*⁴ (Barnard and Cox, 1962). It is a short summary of the course that Savage taught for the *International Mathematical Summer Centre* in Italy (Savage, 1959). Immediately after that course, as explained in the book, Savage went to London.

DE FINETTI: OK, I understand: it is the report that Savage presented at the conference in London.

ALPHA: As Savage writes: “the problem of analysing the idea of randomization is more acute and at present more baffling, for subjectivists than for objectivists, more baffling because an ideal subjectivist would not need randomization at all” (Savage, 1962, p. 34). Perhaps Savage intended to say that the subjectivist, since he should not neglect any piece of information, would have no reason to resort to randomization by means of which some of the information available is actually excluded. But, Savage continues, “[t]he need for randomization presumably lies in the imperfection of actual people and, perhaps, in the fact that more than one person is ordinarily concerned with an investigation.” (ibid.) This sentence suggests a new argument supporting the rationality of the randomization of statistical samples: thanks to the randomization, the likelihood can be computed more intersubjectively. In fact, the Bayesian method produces the convergence when the likelihood is the same for everyone.⁵ But if the draws are not randomized, then the likelihood varies, in general, from person to person and this might preclude convergence. What is your opinion about this justification of the use of randomization in the formation of statistical samples?

DE FINETTI: I seem to agree with this. But I should think more carefully about it.

ALPHA: Savage adds: “the imperfections of real people with respect to subjective probability are vagueness and temptation to self-deception . . . and randomization properly employed may perhaps alleviate both these defects.” (ibid.) Do you believe that Savage’s analysis is correct or do you believe that there could be other reasons that make rational the use of the randomization of samples? It seems to me that the practice of randomization could be justified by means of the need for the intersubjectivity of science. A scientific community, in fact, accepts a result when the majority of its members recognize its value. Is it possible to use the method of randomization in order to facilitate the agreement of many peoples’ judgments?

DE FINETTI: The problem of the randomization of the samples has a mixed character, as it does not have a probabilistic nature only. Randomization is a measure that guards us from the instinctive tendency — which is often followed *bona fide* — to fiddle the results. This can be done in many ways. For instance, it can happen that a researcher excludes some abnormal piece of data thinking that it might be the consequence of a typo or it might be due to a faulty measurement. This would be legitimate if it turned out, for instance, that a certain individual’s height is 170 metres: it would be reasonable to assume that in reality the value of the height is 170 centimetres. But in other cases there could be a tendency to alter the real data because it is considered unreliable. Or there could be a tendency to round off. If many people in a sample turn out to have a height of exactly 170 centimetres and very few people a height of 169 or 171 centimetres, then it would be natural to suspect that a rounding off of the data has taken place. Randomization is a procedure that guards the data from some forms of manipulation and in particular, a biased selection of the sample.

ALPHA: An observation that occurred to me at this moment is as follows. The randomization of the sample makes it easier to determine the state of information. Taking into account all the information that one possesses would be a lot more complicated if the choice was not random. When the sampling is random, the influence

of many relevant pieces of information present on the state of information of the single individuals is eliminated.

DE FINETTI: Also those considerations need to be made cautiously. Suppose, for instance, that despite the fact that the selection has been done correctly from the point of view of precautions (reshuffling, etc.), the sample turns out to be decidedly skewed towards heights that are clearly too big. The suspicion could then arise that this might be due to a systematic tendency to choose tall people. In any case, the problem of the random selection of statistical samples is a very complicated problem and I have never managed to find a completely satisfactory solution to it.

ALPHA: The problem consists in this: strictly speaking one should try to maximize the quantity of empirical information, whereas with the random selection, one intentionally deprives oneself of some information that could turn out to be relevant. If it were known that one individual satisfies some relevant property, this information should also be taken into account rather than neglected because that individual does not belong to the randomly selected sample.⁶

Editor's Notes

1. For precise details see Chapter 8.
2. "The inductive procedure, therefore, has the character of a method of *trial and error* so devised that, for sequences having a limit of the frequency, it will automatically lead to success in a finite number of steps. It may be called a *self-corrective method*, or an *asymptotic method*" (Reichenbach, 1949, p. 446). Reichenbach points out (ibid., note 1) that C. S. Peirce had already stressed in 1878, without however explaining the reason for it, the "constant tendency of the inductive process to correct itself" (Hartshorne and Weiss, 1960, vol. 2, p. 456).
3. Important observation, often neglected: Bayes's theorem alone is not sufficient to guarantee the convergence.
4. The book contains a contribution by Savage (1962).
5. To put the matters in more Definetian terms, all random samples are exchangeable and all stratified random samples are partially exchangeable.
6. The problem of random samples has been addressed by many Bayesian authors. See, for example, Stone (1969); Rubin (1978); Swijtink (1982); Kadane and Seidenfeld (1990); Spiegelhalter, Freedman, and Parmar (1994); Papineau (1994); Berry and Kadane (1997); Frangakis, Rubin, and Zhou (2002); Kyburg and Teng (2002); Berry (2004); Localio, Berlin, and Hane (2005); Worral (2007).

Chapter 12

Complete Additivity and Zero Probabilities*

Discussion of topics suggested by the course attendees concerning various controversial questions:

- “zero” probabilities (and their comparability)
- circularity of some ostensible definitions;
- “complete” additivity (the case of countably many events).

The Betting Framework and Its Limits

I would like to begin by explaining briefly why probability appraisals must satisfy the axioms of the calculus of probability. The reason is very simple: that a set of probability appraisals satisfies the axioms of the calculus of probability is a necessary and sufficient condition in order to avoid manifestly harmful consequences. For instance, the calculus forces a person X who gives probability p to a certain event E , to give probability $100\% - p$ to the event \tilde{E} . Well, if X gave probability q distinct from $100\% - p$ to \tilde{E} , she would offer an opponent the chance to gain without fail. In fact, if $q < 100\% - p$ (if for instance p was 40% and q was 50%) the opponent could argue as follows: “I will place two bets at the same time, one on E and one on \tilde{E} ; I will pay 40 cents for the former and 50 cents for the latter (for a total cost of 90 cents), in return for the promise to receive (for each of the two bets) 100 cents should I win. As it is certain that I shall win one (and only one) of the two bets, I will certainly get 100 cents against the 90 cents paid, with a sure, net gain of 10 cents.” If, on the other hand $q > 100\% - p$ (if for instance p was 40% and q was 70%), then the opponent could reason as follows: “I will accept two bets, one on E and the other one on \tilde{E} and I will charge 40 cents for the former and 20 cents for the latter (and hence, 110 cents in total). I will promise that in return I shall pay 100 cents (for each of the bets) in case of my opponent’s win. As it is certain that I will win one and only one of those two bets, it is certain that I will pay 100 cents against the 110 collected, with a sure, net gain of 10 cents.” Within the betting framework, *coherence* reduces to this.¹

ALPHA: Coherence is equivalent to the so-called *sure-thing principle*.

* Lecture XIV (Friday 2 May, 1979).

DE FINETTI: Yes, I suppose one could say so. It seems to me that the principle you mention says the same thing but in a more sophisticated way.²

ALPHA: You use “coherence” also to refer to an analogous requirement that applies to the appraisal of probabilities in the framework of proper scoring rules (e.g., Brier’s rule). That is the condition according to which probabilities should be assigned in such a way as to ensure that, should a different assignment take place, this would not result in a certainly smaller penalization.

DE FINETTI: The goal of proper scoring rules is to make advantageous for an individual to state sincerely the degree of probability that she attaches to a given event. Among them, Brier’s rule is the simplest. But in order to explain what probability is, the betting framework is even simpler: it is enough, in fact, to say that the probability of an event E is the price that a person believes it is fair to pay in order to get 1 in case E obtains. However, if one uses the betting framework to measure probability in an operational way, it is not enough to ask the individual X to state the amount she would indifferently exchange with the offer to receive 1, should E obtain. In fact, it is not given that it would be advantageous for X to give a sincere answer. Had X , for instance, the option of choosing the direction of the exchange, X would have an interest in giving an answer which might favour one of the sides, and choose this latter for herself. And if the direction were chosen by the opponent, X would find herself in a position of disadvantage, had the opponent more information than X . In this case, it would not always be convenient for X to state the fair price of the offer.

ALPHA: Furthermore, it must be observed that X could be influenced by her opinion concerning the opponent’s probability evaluations. For instance, should X believe that the opponent gives a high probability to a certain event which, on the other hand, she considers improbable, X would take into account the opponent’s possible choices when appraising the probabilities.

DE FINETTI: Absolutely! In order for coherence to force the betting quotients to coincide with the subjective probabilities, a neutral situation must be set up, one in which considerations about the others’ choices play no role whatsoever.³

ALPHA: In such a case the figure of the opponent would be in a way eliminated, as his opinions and preferences should not be taken into account by the bettor.

DE FINETTI: Yes indeed.

Finite and Countable Additivity

BETA: At some point in your book you criticize the so-called *Kolmogorov’s fourth axiom*, which contains the requirement of countable additivity.⁴ You also say that by assuming simple additivity, one can take as a Boolean algebra the algebra of all possible events (that is to say, set theoretically, the σ -algebra of the subsets)⁵ provided that coherence is satisfied. Maybe I am not understanding correctly, but I wonder how can we preserve additivity here, even if I agree with the requirement that any sum of probabilities defined on the elements of a finite partition must equal

1. I doubt that given a simply additive probability function defined on a Boolean algebra it is always going to be possible to extend it to the σ -algebra generated by it.⁶

DE FINETTI: In which way would you extend the probability function?

BETA: If I understand correctly, you say in your book that whenever the condition of coherence is satisfied, it is always possible, given an infinite set of points S and a Boolean algebra \mathcal{A} generated by a subset of S , to extend any simply additive function defined on \mathcal{A} to the σ -algebra constituted by power set of S .⁷ I doubt that any extension of this kind will preserve at once coherence and additivity.

DE FINETTI: Simple additivity is a consequence of the rule of coherence. Coherence reduces to the requirement that there should be no combinations of bets whose ultimate outcome will turn out to be certainly negative. I do not think it is particularly useful, however, to consider a probability function defined on all the subsets of an infinite number of events.

BETA: There are many studies on the extensions in the literature and there are some very complicated theorems. I did not know that things were this simple. I did not know, that is, that it was enough to consider a finitely additive function defined on the whole algebra of subsets in order to obtain the required extension. I would like to know how it is possible to achieve this result.

ALPHA: It seems to me that from the subjectivistic point of view, the extensibility of one own's probability evaluations to all the possible events is not required, as one can play only a finite number of bets at a time. It is sufficient to be able to extend coherently the probability assignments to the events, which time after time are involved in those situations in which one has to take a decision or place a bet.

DE FINETTI: The extensibility requirement derives from the fact that events are thought of as "regular" sets (like "potatoes," just to make things simple). This representation suggests always taking probability as an area, on some suitable scale. This is possible under restrictive conditions but not in general. Let us suppose that instead of a regular (that is to say, convex) set, we are to consider sets with a jagged outline or the set of rational numbers or some similar set. Those sets might have a null measure. In those circumstances, if we took probability to be an area, it would be impossible to give them a non-zero probability.⁸ However, there would be nothing incoherent in this. Luckily, this limitation does not apply if we introduce the events one by one, in a countable sequence of probability evaluations. At each step of the sequence the interval of the possible values compatible with the previous assignments should be determined. Carrying on in this way would never lead one into any contradictions.

The conception of probability as a measure derives from the mistake of taking too seriously the geometrical representation provided through Venn's diagrams. This representation, as any other representations (e.g., mechanical), might turn out to be a useful aid to intuition whenever an event is known as such. But we must not think that what we are representing is a compact set. This representation is fine to picture, with a figure that helps intuition, a generic set of points. But we must not assume that the event that we are representing with that figure also satisfies the topological properties of the figure. If we take real events, independently of whether

they are representable or not in any way, it is enough to attach, time after time, without contradictions, a probability value to the new events, which we take into consideration. If, for example, we considered a set with n points, there would be at most 2^n subsets. It is always possible to assign coherently a probability to each one of those. By doing so, even if the events satisfied none of those topological properties that permit the representation of probability as an area, no contradiction could ever take place. On the contrary, a greatest and a smallest bound would exist. It is therefore preferable to introduce one event at the time and then find, for the event E_n the probability that one assigns to every intersection of E_n with the 2^{n-1} “bits” (or *parts*), provided that they are all compatible, which we obtain with the first $n - 1$ events.

ALPHA: Perhaps the problem discussed in the literature is another one, namely the — exquisitely mathematical — problem of *defining* probability functions whose domain is a field of subsets.

BETA: Indeed. This has to do with the problem of defining an additive function. More precisely, the problem is, given the values assumed on a specific σ -algebra, to define an additive measure on the whole set of subsets.

ALPHA: Once again, it seems to me that the need for a *definition* of a function on the entire σ algebra falls outside the scope of the subjectivistic perspective. After all, the problem of definability must not be confused with that of σ -additivity: one might well oppose the requirement of σ -additivity and nonetheless require that probability should always be assigned by means of a function defined on the set of all the events.

DE FINETTI: If we think of the case of a σ -algebra with countably many, mutually disjoint events, the axiom of complete additivity forces us to produce for the sequence E_1, E_2, \dots a corresponding sequence $\mathbf{P}(E_1), \mathbf{P}(E_2), \dots$ of probability values whose sum converges to 1. In this case, every subset \mathcal{A} would receive a well-defined probability because this would amount to adding up the terms $\mathbf{P}(E_i)$ relative to the E_i s contained in \mathcal{A} . Moreover, in such a situation, if we partitioned the elements of \mathcal{A} into two subsequences S_i and S'_i , the sum of the corresponding series would be 1. In other words, the following would hold:

$$\sum_{i=1}^{\infty} \mathbf{P}(S_i) + \sum_{i=1}^{\infty} \mathbf{P}(S'_i) = 1.$$

Hence, within that context, complete additivity cannot cause any problems. Yet this is not the only possibility: if, for instance, one accepted the definition of probability as the limit of the frequency in a sequence, then every single sequence would have probability 0 and the sum of all the probabilities of all the successions would have probability 1. However, simple additivity would be satisfied here too.

BETA: One could also look at the continuous case.

DE FINETTI: There are many things one might look at. . . .

BETA: It is not possible to measure everything in a σ -additive way when the algebra is continuous.

DE FINETTI: One might say: let us consider only the rational values as possible results and let us assign to each interval a probability corresponding to the length of the interval. So long as we limit ourselves to considering closed intervals, or sets, then everything proceeds as in the finite case. The point is that not all sets are of this type.

ALPHA: Have you ever tried studying infinite sets of bets? Indeed, simple additivity is justified by the fact that it is equivalent to coherence, which is always defined on finite sets of bets. What would happen if we took infinite sets of bets?

DE FINETTI: The idea of an infinite set of bets is rather artificial. One could perhaps characterize it asymptotically, by taking firstly a finite sequence of n bets and then moving on to the limit by letting n tend to infinity. Yet this is *not* equivalent to taking an actual infinity of bets.

ALPHA: Would you therefore say that it is not realistic to think of infinitely many bets because in practice, these cannot be made?

DE FINETTI: It seems to me that the infinite case is dealt with more reasonably if one thinks of it as an impossible limiting case of all the possible cases, with n finite but increasingly greater. I believe, in fact, that actual infinity is — in this context — too indeterminate.

BETA: Taking a function that is defined on an infinite structure implies the guarantee of having a probability evaluation for any set of events (be it finite or not), while preserving coherence. In other words, we could rely on the certainty that a measure (and hence, a probability value) exists for any event we might want to consider. Although, in practice, we always consider a finite number of events.

DE FINETTI: As I already said, it seems to me that this would always be guaranteed to happen even if we restricted ourselves to accepting simple additivity. Indeed, if under simple additivity we take one of the parts to which a probability has been assigned, it is always possible to fix a minimum and a maximum (or, in the case of an infinity of events, a lower bound and an upper bound) for the probabilities that can be given to a new event. Given certain events E_1, E_2, \dots, E_n we call the *constituents generated by the E_i s* those events C_1, C_2, \dots, C_n of the form $X_1 \wedge X_2 \wedge \dots \wedge X_n$, where, for all i ($1 \leq i \leq n$), X_i stands either for E_i or for \tilde{E}_i .⁹ Every constituent C_i can be classified into one of the following three categories, according to whether

- i) $C_i \subset E$;
- ii) $C_i \subset \tilde{E}$;
- iii) $C_i \not\subset E$ and $C_i \not\subset \tilde{E}$.

If we considered all the constituents of type iii) as if they were contained in E , then our probability evaluation would be incoherent “by excess,” whereas if they were considered to be all external with respect to E , we would make an incoherent evaluation “by defect.” On the basis of considerations of this sort it is possible to determine, once the probabilities of some other events E_1, E_2, \dots, E_n have been fixed, the interval of the possible probability values of E .¹⁰

‘Strict’ Coherence

ALPHA: I would like to hear something about the concept of *strict coherence*, which you mention in your book. The notion was introduced in 1955 by Abner Shimony, a Bayesian philosopher, subjectivist in his own way.

DE FINETTI: Of course, Shimony.

ALPHA: I recall that “strict coherence” is a strengthening of the condition of coherence. It amounts to requiring that the evaluation of the betting quotients not only rule out the possibility that the opponent might ensure a win without fail, but also that he could not guarantee not to lose if the possibility of a net win for her exists. Of course, Shimony, referred to observable events, which always involve a finite number of elementary alternatives — for example the outcomes of horse races or those of football matches.

DE FINETTI: Could you please repeat the formulation of this condition?

ALPHA: The bettor should not only avoid exposing herself to sure loss (coherence), but should also avoid accepting bets involving a risk of loss for her in certain cases, and for the remaining cases the certainty of a draw. The fact that such bets are not advantageous derives from the fact that they yield the possibility of losing without this being counterbalanced by the possibility of winning.

DE FINETTI: And what are the rules of this situation?

ALPHA: It is the usual betting framework.¹¹ The gain function could be such that for certain constituents or atoms (suppose we are referring to a given Boolean algebra) the bettor would lose, while for all the other atoms she would get a null gain. According to Shimony it would be incoherent (or more precisely, weakly incoherent) to appraise probability in such a way as to admit systems of bets of this sort. Strict coherence is a more restrictive condition than simple coherence, because it implies — unlike simple coherence — the equivalence between probability 0 and logical impossibility.

DE FINETTI: I see, this is the case in which it is possible to lose, but with probability 0.

ALPHA: Correct. Indeed it is clear that given a combination of bets that permits one to lose but not to win, whenever simple coherence is satisfied, every case leading to a loss must get probability 0. Thus, if one placed a bet with a quotient 0 on an event, the following situation would hold: one would pay a certain amount of money in case of loss, to get nothing in case of win. According to Shimony it is irrational to accept such a bet whenever it is possible to avoid it. It is not rational, in fact, to run the objective risk of losing without having in return the *logical possibility* of winning. This restrictive condition leads to the addition of the (so-called) *axiom of regularity* to the basic axioms of the calculus of probability, which, in those situations characterized by a finite number of elementary cases, is equivalent to considering probability 1 as logical certainty and probability 0 as logical impossibility. It must be stressed that the range of application of this axiom is restricted to the partitions with a finite number of elementary cases. This does not impose non-zero probabilities in the case of infinite sets, randomly chosen numbers, etc. Let us take the case of a football match with three possible outcomes. Here — according to

Shimony — probability 0 should mean logical impossibility and as a consequence it would not be rational to give probability 0 to any of the three possible outcomes 1, X , 2, whereas coherence allows one to do this.

DE FINETTI: Let us suppose that the bettor gives to that event probability 0. Should she state another value? I do not understand.¹²

ALPHA: She would be incoherent because she would accept a bet where the possibility of losing is not compensated by the possibility of winning. Is the requirement of strict coherence not a sound one in your opinion?

DE FINETTI: One can split hairs over this as much as one likes and there is no doubt that no-one would accept — being in a position to avoid it — a bet in which the only possible cases are not losing anything and losing something. And this independently of considerations of probability. However, if the point is defining the notion of a fair bet, then it must be admitted that if we added to a fair bet another bet yielding either a draw or a loss with probability 0, the (prevision of the) balance would be unchanged and the set of the two bets would therefore be fair, despite the presence of zero probabilities assigned to logically possible events.

ALPHA: The point is that probability 0 means this: that the offer of a certain amount of money x conditional on the occurrence of an event E has value 0 (independently of the value of x , which in fact could be multiplied by an arbitrary coefficient).

DE FINETTI: It has a null value in the scale of natural currency. If we were to place a bet on the toss of a coin, where the bet is considered valid only if a zero probability event occurs, there would be no reason to reject it, independently of the betting quotient. And this is because it is practically certain that a bet conditional on the occurrence of an event of probability 0 will be called off.

ALPHA: Therefore it is not possible to measure a person's X subjective probability with a bet of this sort. As X is subjectively sure that the bet will be called off, she could state whatever she likes. X would have no incentive to be sincere.

DE FINETTI: Exactly! It seems to me that this is the most important point. To raise objections against this point would amount to splitting hairs to no purpose. Indeed, placing a bet conditional on an event of probability 0 is — from the point of view of the bettor who assigns probability 0 to that event — almost like not betting at all.¹³

ALPHA: Suppose that a person believes with probability 1 that the case yielding a draw will occur. It is hard to see why one should consider irrational a person who accepted a bet of this sort. After all, she would have the practical certainty that the final outcome will be a draw (the probability of the contrary being 0).

DE FINETTI: In this case she should be indifferent between placing or not a bet under bizarre conditions.

ALPHA: "Under bizarre conditions" means "conditional on an event of probability 0"?

DE FINETTI: Yes. Even the right to receive a million, on condition that an event of probability 0 obtains, should be a lot less valuable than the right to receive with certainty one cent (no matter how worthless a cent is). My impression is that one could twist around this argument in many ways without being able to find anything particularly . . .

ALPHA: [interrupting] Deep.

DE FINETTI: No, on the contrary: it is in fact too deep to have some really concrete meaning!

ALPHA: As you will remember, the fact that whenever the event on which a bet is conditional has probability 0 the bettor loses her incentive to give a unique probability evaluation, was discussed in connection with the way in which coherence can be represented in terms of systems of bets. More precisely, it was as follows. Suppose that the bettor makes certain probability evaluations. Then certain combinations of bets are made, amounting to indirect bets on that event.¹⁴ It can be shown that if the betting quotients of the direct bet violate the axioms of the calculus of probability, then a serious contradiction arises. In this case, in fact, there would be an event E whose quotient q , stated to be the maximal quotient at which one would be willing to bet on E , would in practice fail to be such because, indirectly (that is through a series of bets), it would be possible to bet on E with a betting quotient greater than q . Moreover, the unconditional requirement of the identity between direct and indirect quotients implies the axiom of regularity. Indeed, when the event on which the bet is conditional has probability 0, it is always possible to bet indirectly on the impossible event at any quotient whatsoever. Therefore, in order for uniqueness to be always satisfied, the axiom of regularity must be satisfied as well. But, given that whenever the conditioning event has probability 0, uniqueness cannot be required, this cannot be a way to justify the axiom of regularity.

I also notice that there is a *petitio principii* in the argument for strict coherence. Those who resort to such an argument claim that it is irrational to violate strict coherence because otherwise one would be indifferent between accepting offers which yield the possibility of losing but not that of winning on the one hand and not betting at all on the other. If a person X gives probability 0 to an event E , this means that X is the bearer of a *subjective* degree of certainty with respect to the obtaining of E identical to the degree of certainty that she bears with respect to the impossible event. Supposing that this were allowed, there would be no reason to consider the violation of strict coherence as irrational. In such a case, in fact, X would be subjectively certain that her bet will have a null outcome. And, coherently, X would therefore be indifferent between accepting that bet and refusing to bet. From this I conclude that in order to say that the violation of strict coherence is irrational, one must assume that it is irrational to give to a possible event the same degree of subjective certainty that one gives to the impossible event. That is, one must assume precisely what is forced by the axiom of regularity, which is meant to be justified by the argument of strict coherence. Hence, the circularity of the argument.

Conditioning on Events of Zero Probability

DE FINETTI: What could be done is to consider a ratio between two zero probabilities. Although this is a slightly fictitious thing to do, I presented some ideas to this effect in a talk by the title *Les probabilités nulles* (1936).

ALPHA: There is a publication of yours in 1936. Yet I did not know about the conference.

DE FINETTI: The publication contains that series of talks.

ALPHA: Is it the same series including *Foresight*? (de Finetti [1937] 1980).

DE FINETTI: No, this is a separate thing, which I presented at a conference.

ALPHA: You were speaking of zero probabilities.

DE FINETTI: Yes. Let us take two logically incompatible events E_1 and E_2 having both probability 0. One would be tempted to write like this:

$$\mathbf{P}(E_1) = 0 \qquad \mathbf{P}(E_2) = 0.$$

Yet this notation would be misleading because it suggests that the values $\mathbf{P}(E_1)$ and $\mathbf{P}(E_2)$ are identical, being both equal to 0. Actually, it might be that it is the *same* 0. Let us take the sum event $E_1 \vee E_2$. In my opinion, it makes sense to evaluate the probability of E_1 conditionally on $E_1 \vee E_2$, as it makes sense to evaluate the probability of E_2 conditionally on $E_1 \vee E_2$. I believe that a comparison of those probabilities is possible. And of course it should hold that

$$\mathbf{P}(E_1 \mid E_1 \vee E_2) + \mathbf{P}(E_2 \mid E_1 \vee E_2) = 1.$$

Although I never took these considerations seriously, it seemed to me that it was appropriate to show that it is possible to establish a hierarchy among zero probabilities. A zero probability, in fact, can be infinitely greater or infinitely smaller than another zero probability: suffice it to think of a single point in comparison with two or three points or to a small segment of a curve in relation with a long curve. Or even a curve in comparison with an area or a volume.

BETA: Those would be, so to say, various degrees of impossibility.

DE FINETTI: It is something similar to a theory of measure initially set up in a three-dimensional space where the sets of points had a volume for measure. So long as solid bodies only were considered, no problem would arise, but surfaces would have null volume. However, the surfaces are comparable on the grounds of their area. If we considered very tiny layers, the volume would essentially depend on their area. Analogously, a line has usually a length: but even a single point, if one thought of it materially in the physical space, would be something that could always be increasingly “infinitesimized,” so to speak. If a certain number X had a certain value, one could represent it as a point on a line (Fig. 12.1).

But one could take as well another quantity Y (Fig. 12.2). In this case, the choice of a point X on the line would amount to the choice of the whole vertical, with a certain abscissa. In order to identify the point, two conditions would now be necessary. If we added a further dimension, the conditions would rise to three. In this case we would have three 0s, the second one being the square of the first one, and the third one being the cube of the first one. Although from the strictly arithmetical point of view 0^3 equals 0, from the point of view of zero probabilities this equality no longer holds. Therefore, if in the first case the probability is 0, in the second

Fig. 12.1 Representation of a number as a point on a line

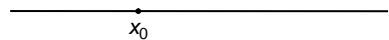
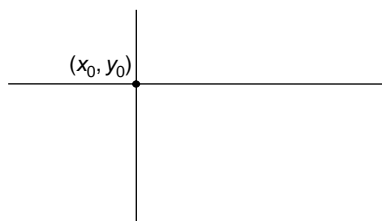


Fig. 12.2 Representation of a pair of numbers as a point on a plane



case the probability will be 0^2 and in the third 0^3 . I am saying this without any intention to make a theory of the dimensions of zero. Yet this is the basic intuitive idea. Given two independent events of probability 0, the probability that both will obtain is $0^2 \neq 0$.

ALPHA: You maintain that in order for probability to make sense, it must be operationally measurable (for instance by means of bets). However, those differences between zero probabilities have no counterpart in the betting framework.

DE FINETTI: Probabilities conditional on events of probability 0 could be defined as limiting cases. If one learns that a certain event H , which previously had probability 0 is verified, then, for any event E , the conditional probability $\mathbf{P}(E \mid H)$ would equal the ratio between the 0 corresponding to $\mathbf{P}(EH)$ and the 0 which corresponds to $\mathbf{P}(H)$. However, let me emphasise — to cross the t s — that although I wrote that paper, I do not want to give the impression that I am “the advocate of zero probabilities” or that I pride myself on having elaborated theories by means of which one can work with zero probabilities. However, what I really cannot stand is the identification of probability 0 with impossibility. My speculations were primarily motivated by my rejection of that identification.

ALPHA: Recently, the so-called *non-standard analysis*¹⁵ has been developed, which legitimates — to put it intuitively — the formulation of the theorems of analysis in terms of the notion of an actual infinitesimal, rather than that of a limit. My question is: do you consider it plausible that this hierarchy of zero probabilities could be replaced by a hierarchy of actual infinitesimals in the sense of non-standard analysis?

DE FINETTI: I only attended a few talks on non-standard analysis and I have to say that I am not sure about its usefulness. On the face of it, it does not persuade me, but I think I have not delved enough into this topic in order to be able give an well thought-out judgment.

ALPHA: It seems to me that this institute¹⁶ hosted a conference on this topic in 1960. Robinson too, the inventor of non-standard analysis, took part in it.

DE FINETTI: I made those speculations on infinitely small probabilities to see the extent to which the idea of a comparison between zero probabilities is plausible. However, I did not attach much importance to it and I am not sure whether one needs sophisticated theories, such as non-standard analysis, for that goal. A motivation for my considerations was to reply to some observations by Borel that looked absurd to me. He said that a very small probability, for instance 10^{-1000} is so small that is exactly identical to 0. This view looked a bit too simplistic to me.¹⁷

Allais' Paradox

ALPHA: Objections of this sort are grounded on the analysis of actual behaviour.¹⁸ The same holds for Allais' paradox. It is also discussed in Savage's book.

DE FINETTI: Yes.

ALPHA: Allais' paradox is a typical paradox based on actual behaviour: there are choice situations under risk in which the majority of people is intuitively led to make choices which, after careful analysis, turn out to be incoherent.¹⁹

DE FINETTI: I cannot recall either the formulation of Allais' paradox, or the objections that I raised against Allais' thesis in the full details. Yet roughly my objection was that Allais counted twice the decreasing behaviour of the curve of the marginal utility of monetary values.²⁰

ALPHA: Allais' paradox shows that there exist choices, motivated by prudence, that cannot be preserved by any utility function — not even by "straightening" the curve of monetary values. In other words: there exists a form of risk aversion that cannot be explained in terms of the curve of the marginal utility of monetary values.

DE FINETTI: It seems to me Allais is wrong in claiming this. I reckon that he applies twice the idea that marginal utility decreases with the increase of the monetary amount. I admit that taken literally, this sentence does not make much sense, but I believe that it conveys the intuitive point of my objection.

ALPHA: It must be so, because if we assumed the perfect linearity of the values, there would be no space for risk aversion. On the other hand, people's actual behaviour can also be irrational, without this posing any difficulties to a normative theory.

Editor's Notes

1. Allusion to the Dutch Book argument. See note 12 on page 43.
2. The *sure-thing principle*, introduced by L. J. Savage (Savage [1954] 1972, p. 23), can be formulated as follows: "given two uncertain acts *A* and *B* yielding the same utility *g* if an event *E* obtains, the preference relations between *A* and *B* would be unchanged should *A* and *B* be modified so as to yield, should *E* obtain, a utility *g'* in place of *g*." Although the satisfaction of this principle is a necessary condition for the existence of a class of *linear* utility functions (equivalent up to positive linear transformations) preserving the preference relation (Savage [1954] 1972, p. 75), it is neither necessary nor sufficient to the characterization of subjective probability as a fair betting quotient (see Mura, 1994).
3. As remarked by de Finetti (1981) the betting framework presents game-theoretical aspects that may destroy the convenience of the bettor to declare "sincerely" those odds that she believes to be fair.

Although — as far as I know — there is no detailed study of de Finetti's betting framework for measuring probabilities from a game-theoretical viewpoint, this kind of game-theoretic interaction (called "strategic manipulability") has been extensively studied with respect to other contexts in the theory of group choices — especially in the study of voting procedures. On this topic see Gibbard (1973); Satterthwaite (1975); Pattanaik (1978); Tanaka (2001); Bochet and Sakai (2004).

I am indebted to Nicholas H. Bingham for the suggestion of an additional note on strategy-proofness with respect to de Finetti's betting framework.

4. The condition of *countable additivity* (also known as *complete additivity*) is satisfied whenever, for each countable sequence of pairwise disjoint events E_1, E_2, \dots , the probability of their union $\bigcup E_i$ equals the value of the series $\sum \mathbf{P}(E_i)$. The condition of *finite additivity* (also referred to as *additivity tout court*) on the other hand, is satisfied whenever given a finite sequence of pairwise disjoint events E_1, E_2, \dots, E_n , the probability of their union $E_1 \cup E_2 \cup \dots \cup E_n$ equals the sum $\mathbf{P}(E_1) + \mathbf{P}(E_2) + \dots + \mathbf{P}(E_n)$.
5. A Boolean σ -algebra is a Boolean structure closed under countable sum and product.
6. Many studies on the theme of the reciprocal relation between countable additivity and coherence appeared in the wake of the publication of the English edition, in 1974, of *Theory of Probability* (de Finetti [1970] 1990a). Within this particular context it is particularly interesting to note that (de Finetti [1970] 1990a, p. 177) if we admit simple additivity but not complete additivity, then the so-called *conglomerative property* of probability (see Chapter 17) is not always satisfied on infinite partitions. Among the contributions in this area, I would like to point to Dubins (1975); Heath and Sudderth (1978, 1989); Hill (1980); Seidenfeld and Schervish (1983); Sudderth (1980); Kadane, Schervish, and Seidenfeld (1986); Hill and Lane (1986); Goldstein (1983, 2001); Seidenfeld (2001); Bartha (2004).
7. "It has been proved... that if one is not content with finite additivity but insists on countable additivity, then it is no longer possible to extend the "measure" to *all the sets* (whereas there is nothing to prevent the extension to all sets of a finitely additive function" (de Finetti [1970] 1990a, p. 229)). This possibility (which in general is far from being unique) is grounded on de Finetti's *fundamental theorem of probability*, discussed below in this chapter (see 117 and note 9 on page 124). De Finetti reasoning appealed to transfinite induction as well as to the axiom of choice (de Finetti [1970] 1990b, p. 337). It should be added that these results may be generalized to conditional probabilities as well as to previsions (absolute and conditional) of random quantities.
8. De Finetti's standpoint against the measure-theoretic approach is so summarized in de Finetti ([1970] 1990b, pp. 258–9):

The current practice of reducing the calculus of probability to modern measure theory... has resulted in the following:

probability is obliged to be not merely additive (as is necessary) but, in fact, σ -additive (without any good reason);

events are restricted to being merely a subclass (technically, a σ -ring with some further conditions) of the class of all subsets of the space (in order to make σ -additivity possible but without any real reason that could justify saying to one set 'you are an event', and to another 'you are not');

people are led to extend the set of events in a fictitious manner (i.e., not corresponding to any meaningful interpretation) in order to preserve the appearance of σ -additivity even when it does not hold (in the meaningful field), rather than abandoning it.

In this vein, any adopted algebraic structure \mathcal{E} and any ultimate partition \mathcal{Q} of "points" must be considered as provisional. "This implies that we can only consider as meaningful those notions and properties which are, in a certain sense, invariant with respect to 'refinements' of \mathcal{Q} " (de Finetti [1970] 1990b, p. 268) as well as invariant with respect to embedding \mathcal{E} in a larger structure whatsoever and with respect to the chosen set of generators of \mathcal{E} . Events, ultimately, belong to a minimally structured "unrestricted field" (de Finetti [1970] 1990b, p. 267). According to de Finetti, "[t]o approach the formulation of a theory by starting off with a preassigned, rigid and 'closed' scheme" is like shutting "oneself inside a tank in order to journey through a peaceful and friendly country" (de Finetti [1970] 1990b, p. 269 note).

9. Using a standard terminology, the constituents generated by E_1, E_2, \dots, E_n are the atoms of the finite Boolean algebra generated by E_1, E_2, \dots, E_n . It is clearly possible that E does *not* belong to this algebra.
10. De Finetti found out already in *Foresight* that "[i]f E' and E'' are respectively the sum of all the constituents [see note 9 above] contained in E or compatible with E , p' will be the smallest

value admissible for the probability of E' and p'' the greatest for E''' (de Finetti [1937] 1980, p. 108, notation style adapted).

Now, if in determining p' and p'' we assumed either that some constituent not contained in E was contained in E or that some constituent not contained in \tilde{E} was contained in \tilde{E} , then we would be allowed to assign to E a probability value which respectively is, in the presence of coherently assigned values $\mathbf{P}(E_1), \mathbf{P}(E_2), \dots, \mathbf{P}(E_n)$, either greater than the maximum coherent value (*incoherence by excess*) or less than the minimum coherent value (*incoherence by defect*). The general problem consists in determining the actual numerical minimum and maximum coherent value for the probability of E in the presence of the values $\mathbf{P}(E_1), \mathbf{P}(E_2), \dots, \mathbf{P}(E_n)$. The answer to this problem is provided by an important result, which de Finetti called the *Fundamental theorem of probability* (de Finetti [1970] 1990a, p. 111). It shows that the problem of determining p' and p'' amounts to solving a linear programming problem, such that thanks to the linear programming algorithms the computation of p' and p'' may be carried out. See also Hailperin (1965, 1996) and Lad (1996).

11. For a detailed exposition of such a schema see Mura (1995b, pp. 19–23).
12. This reaction suggests that de Finetti maintained that the assignments of probability that violated the basic axioms, being contradictory, would not represent a genuine state of belief, whereas this cannot be said of those assignments that would violate the axiom of regularity.
13. The relevance of this observation with respect to the axiom of regularity will become clear later on.
14. Allusion to a personal communication. In a nutshell, an “indirect bet” on $E \mid H$ is a linear combination of bets whose balance amounts to a constant gain g if EH obtains, a constant gain g' ($0 \leq g > g'$) if $\bar{E}H$ obtains and a null gain otherwise. In 1975, while working on my doctoral dissertation, I proved that given a finite Boolean algebra \mathcal{A} , every function \mathbf{P} assigning a conditional betting quotient to all the conditional events of the form $E \mid H$ (where both E and H belong to \mathcal{A}) satisfies the axioms of finite probability if and only if the betting quotient of every indirect bet on $E \mid H$ (such that $\mathbf{P}(H) > 0$) is $\mathbf{P}(E \mid H)$.

Now, if degrees of belief do exist, they should be assumed to be uniquely determined in the bettor's mind and therefore to be represented by indirect betting quotients. In such a case, they necessarily satisfy the axioms of finite probability. This shows that betting quotients violating the laws of finite probability cannot consistently represent uniquely determined degrees of belief.

For detailed expositions of this argument and discussions of its philosophical import, see Mura (1986, 1995b). Similar results were later published by Colin Howson (2000, pp. 127–34).

15. Non-standard analysis was introduced by A. Robinson (1918–1974). His major contributions on this topic are (1961, 1996). There are several good textbooks on this subject, for example, Davis ([1977] 2005); Robert (2003); Pinto (2004). In connection with the application of non-standard analysis to the problem of zero probabilities see Carnap (1980, Chapter 21); Hoover (1980); McGee (1994); Halpern (2001); .
16. That is the *Institute for Advanced Mathematics* in Rome.
17. See Chapter 5 on page 49.
18. Probably allusion to the fact that when probabilities are very small it is impossible to discern them from the zero probability and to measure them with bets.
19. Allais (1953) considers two situations S and T in which one must choose between two options (which I shall refer to as S_1, S_2 and T_1, T_2 , respectively). Those are defined as follows:

$S_1 = \$500000$ with probability 1

$$S_2 = \begin{cases} \$2500000 & \text{with probability } 0,1, \\ \$500000 & \text{with probability } 0,89, \\ \$0 & \text{with probability } 0,01. \end{cases}$$

$$T_1 = \begin{cases} \$500000 & \text{with probability } 0,1, \\ \$0 & \text{with probability } 0,89. \end{cases}$$

$$T_2 = \begin{cases} \$2500000 & \text{with probability } 0,1, \\ \$0 & \text{with probability } 0,9. \end{cases}$$

The paradox consists in the fact that, according to the standard theory of decision, one should — as a consequence of the sure thing principle — prefer S_2 over S_1 whenever T_2 is preferred over T_1 , whereas the majority of “cautious” people prefer S_1 over S_2 despite preferring T_2 over T_1 .

Savage and de Finetti solved the paradox by appealing to the normative character of the standard theory: preferring S_1 over S_2 in case T_2 is preferred over T_1 is irrational and must be explained as a slip that the theory helps to correct.

Many authors, on the other hand, have tried to solve the paradox by weakening the axioms of the theory. Among many, I would suggest: Allais (1953, 1979); Levi (1986); Loomes and Sugden (1982, 1998); Machina (1982); Quiggin (1982); Quiggin and Wakker (1994); Yaari (1987); Chew (1983); Fishburn (1981); Fishburn and LaValle (1988); Gilboa (1987); Birnbaum (1999); Luce and Marley (2000); Nielsen and Jaffray (2006).

In some relatively recent works (1979, 1984), Allais presents the results of experimental research based on his theory. De Finetti himself took part in that research. Allais commented on de Finetti's contribution in this way: “[o]f all the answers I received, de Finetti was without doubt *the most remarkable in terms of the scientific conscientiousness and great intellectual honesty* with which he responded to the various questions.” (Allais [1979], 673 note 19, Allais' italics).

The list of similar experimental studies on this subject is enormous. I shall limit myself to mentioning two recent interesting reports, which seem to support opposite conclusions about the accordance of the preferences of trained people with standard decision theory in Allais-like situations: Wakker (2001b); List and Haigh (2005).

20. De Finetti's argument is clearly expressed in (1979, p. 161): “[u]tility is . . . a convex function of the monetary value, since aversion to risk usually exists (and is admitted as the ‘normal’ assumption in economic theory), so that a sure amount is preferred to an uncertain one with the same expectation The objection by Allais . . . consists in asserting that the same correction should be repeated about the utility. This seems tantamount to asserting that, when the deflection of a bridge owing to a given load is computed, the deflection from the deflected position should be computed again because the load also acts on the deflected line (missing to note that, by definition, the deflected line is just that one for which elasticity exactly reacts so that the weight of the load is balanced).”

The key point of Allais' response to de Finetti's objection was that “[t]he neo-Bernoullian formulation based on consideration of mathematical expectation [of the index of psychological value \bar{s}] $\sum p_i \bar{s}(x)$ neglects an essential element, the probability distribution of $\bar{s}(x)$ about its average In actual fact, the analysis of [the] 1952 survey showed that for every subject (and for de Finetti in particular) there is no way to determine a [neo-Bernoullian] indicator taking account simultaneously of the curvature of the index of cardinal preference (cardinal utility) and the greater or lesser degree of preference for security, *for different sets of questions*.” (Allais, 1979, p. 565).

To properly understand de Finetti's tenet it should be reminded that the mean-variance approach to risk aversion, advocated by Allais for neo-Bernoullian utility, was developed by de Finetti (1940) with respect to the utility value of money. De Finetti's ideas were independently rediscovered a dozen years later by Harry M. Markowitz (1952), who in 1990 shared the Nobel prize for Economics with Merton H. Miller and William F. Sharpe precisely for this contribution applied to portfolio selection (see Pressacco, 2006). Allais too was awarded (in 1988) the Nobel prize for Economics.

Chapter 13

The Definitions of Probability*

Presentation, illustration and discussion of the various definitions of probability and of the distinctions that arise from them in the interpretation of the approach and the conclusions of the corresponding theory.

Axioms: formal properties in abstract form, “equally probable cases” and frequentist visions.

Axiomatic, Classical, and Frequentistic Approaches

There is no other notion in Mathematics that is defined in so many distinct and diverging ways as the notion of probability. Therefore, I think it is well worth going through a short critical comparison among the various definitions of probability.

Many authors provide a purely formal characterization of probability and maintain that one should limit oneself to stating its formal properties. These are well-known principles, like the principle of *additivity* (“the probability of the logical sum of two incompatible events is the sum of their probability”), the principle of *compound probabilities* (“the probability of the logical product of two independent events is the product of the respective probabilities”), the principle of *total probability* (“the certain event has probability 1”) and so on. Those principles establish some purely formal properties, which actually do not say anything about what we mean by “probability.” To limit ourselves to a formal characterization is like developing a theory of mechanics just by introducing certain symbols and certain formulas, which however are treated as arbitrary mathematical formulas, that is without thinking that they actually apply to something (solid, liquid and gaseous bodies; motion, the space in which motion takes place; etc.). The result is an abstract theory, empty of a genuinely probabilistic meaning, until an interpretation is given to the symbols and the formulas occurring in it. Unfortunately, according to the axiomatic point of view, everyone should have the liberty to give the interpretation she likes, with the consequence that the analysis of the meaning of what we are talking about would end up being of no importance whatsoever.

According to another, no less abstract, orientation at first, concepts like *field of events* are introduced. This is sometimes intended as a space in which every figure can be taken into consideration (each figure being an event). Some go so far as to

* Lecture XV (Thursday 3 May, 1979).

interpret the points of such an abstract space as if they corresponded to as many histories of the world (or of that portion of the world that it is being considered). In this way every point would provide a precise way of describing the whole set of facts (what is their status at a given instant or how they develop over a given time interval, or else in the entire eternity). According to such an approach, one should decompose all the imaginable events that take place in the world into points and then assign to each set of points its own probability. Here too, however, those abstract structures are built up without saying what it is meant by probability. The outcome is again a completely abstract and formal theorization.

In the elementary textbooks on the calculus of probability, the formal properties are derived from the so-called *classical definition*, according to which probability is the “ratio between the favourable cases and the number of equally probable cases.” Suppose that probability is distributed uniformly across the members of a certain class of mutually exclusive and jointly exhaustive alternatives and that we are interested in the logical sums which we can construct with them. In order to have a measure satisfying the usual conditions of linearity — which, however, are not only satisfied by probability but also by other quantities, including mass (the union of two bodies, in fact, is a body which has the mass of the sum of the masses) — it will be necessary to assign to each event a probability identical to the ratio between the number of the alternatives that verify it and the total number of possible alternatives. It is illusory however, to think that by defining probability in this way one is actually clarifying its meaning. Indeed, this does not go beyond the usual axiomatic characterization, to which only the further constraint of the equiprobability of the elementary alternatives is added. In order to turn this ostensible definition into a proper definition one should clarify *in primis* what it is meant by ‘equal probability.’ And since the definition of equal probability constitutes an essential part of the definition of probability itself, the result is a circular definition. To see this, it is enough to note that a person who did not know what ‘probability’ means, could not learn it from the classical definition, for an understanding of the term ‘equal probability’ would be needed to that effect.

One could try to get rid of the circularity by substituting the condition that the elementary alternatives should be equiprobable with a condition that would justify the equal probability given to them. But when is it that two events have equal probability? If you asked someone who adopted the classical view: “why are those elementary alternatives equiprobable?”, such a person could answer in very different ways, none of them being sufficient to *define* what it is meant by the expression ‘equally probable.’

Therefore, anyone adopting the classical definition runs into circularity, independently of the way in which the statement of equiprobability required by its application is justified. Both those who ground the judgment of equiprobability on considerations of physical symmetry and those who ground it on considerations of frequency run into such a circularity. The latter should be also charged with the aggravating circumstance of mixing up frequency and probability. The frequency with which certain events obtained or will obtain cannot be identified with probability. Frequency is a mere fact, independent of both the meaning of probability and the probability values assigned to the events.

There are orientations in which the role of frequency becomes even more fundamental, for it is included in the definition of probability itself. According to such orientations, the probability of an event increases with its frequency. To make this definition plausible one must understand the word event in a broad sense, as a general term of which every single event would be a “trial.” However, this is nothing but a verbal trick in order to speak of “trials of the same event” instead of events which, being analogous (in certain respects) are considered to be equally probable a priori. Talk of “trials of an event” instead of, simply, “events” is a mere terminological convention which cannot alone be sufficient to eliminate the singular and one-off character of each “trial.”

In order to express the same idea, yet without any commitment and any risk of giving rise to the confusion generated by the expression ‘trials of the same event,’ I introduced the expression ‘trials of the same phenomenon.’ To be very honest I do not consider ‘phenomenon’ entirely satisfactory either, but as yet I have failed to find something more adequate. The point in favour of ‘phenomenon,’ however, is that it does not suggest, unlike ‘trial,’ that the repetitions are objectively identical entities. On the contrary: it is meant to suggest that the trials, despite the fact that they might share some common descriptive feature, could well differ in respects that are not directly related to their general description, yet can nonetheless be of some influence.

ALPHA: One can always assume that the person who is making her own evaluation is not aware of those differences.

DE FINETTI: Yes, OK. But it would be a subjective evaluation. Indeed, when the initial hypotheses correspond to an individual’s opinion, all the so-called definitions become, from the subjectivistic point of view, true propositions. In this case, all the consequences drawn from them would conform to the opinions of that individual, of course provided that she was coherent.

ALPHA: This notwithstanding, the concept of *frequency of successes* is a rigorously definable concept.

DE FINETTI: The frequency of successes must be understood in reference to a certain number of events, which must be judged similarly *from some point of view* in order for one to be able to say that they are equal. Hence, if one wants to make sense of the statement that certain events are equally probable, one must first define the probability of a single case. Of course it might happen that a person thinks of giving a priori equal probability to certain events. In this case, it would be perfectly legitimate to run experiments in order to measure the frequency and consequently assign to the events a probability close or even identical to the observed frequency.

ALPHA: However, the frequencies of successes, defined with respect to a generic sequence of bets, always satisfy the principles of the calculus of probability.

DE FINETTI: Fine, but those principles are satisfied by almost every physical quantity, including mass. What is peculiar about probability is not the fact that it satisfies certain properties typical of linear quantities, but the fact that it always requires the judgment of a single person. Of course, if facing cases which no one is able to distinguish according to some feature commonly regarded as relevant, everyone would agree in assigning an equal probability to them. It seems to me, however, that

this has nothing to do with the *definition* of probability. It is rather a clarification: when a person believes that in her own view, none of the properties for which the events differ either favours or opposes the occurrence of the events that satisfy it, it is natural for her to assign equal probability to all the events. This is not an objective condition but a condition that depends on the point of view of each single person.

Indistinguishable Events and Equal Probability

ALPHA: How do you consider the fact that a person assigns equal probability to indistinguishable events? Is it a mere empirical fact without any normative value? Do you think that one *should* give equal probabilities to a pair of indistinguishable events?

DE FINETTI: If an individual states that one event is more probable than another, one must assume that that individual sees something distinctive in it. If a person said that she considered an event more probable than another one where, on the face of it, it would seem natural to give them the same probability, I would ask her: "do you have a reason to assign dissimilar probabilities to these two events?" If the person replied: "Yes, I do have a reason" and then explained it to me, I might reply that according to me that is not the case, or on the other hand, I could accept her arguments.

ALPHA: Suppose that person said: "I cannot distinguish between these two events, yet I believe that E_1 is more probable than E_2 ." Would you consider such a person irrational?

DE FINETTI: Irrational? Well, it is not a rational thing, yet she might . . .

BETA: [interrupting] If she said: "I think that the probability of E_1 is greater than the probability of E_2 " this would mean that she could distinguish between E_1 and E_2 .

ALPHA: Let us suppose that this person said: "I know that E_1 is distinct from E_2 but I do not know in which respect they are different." Or that she said: "All the properties that I know are satisfied by E_1 , are also satisfied by E_2 and conversely. I also know, however that there exist properties with respect to which E_1 and E_2 differ, but I do not know what they are."

BETA: In my opinion, a person should know at least, if from her own point of view, that the probability of E_1 is greater than that of E_2 . This is a difference between E_1 and E_2 .

ALPHA: That is not a property of the object but a property of the individual who is making a judgment. Otherwise, a person could change the object by changing her probability evaluation. This is the same reason why a noun is not a property of the object to which it refers.

BETA: Evidently a property must exist with respect to which E_1 and E_2 can be distinguished. Suppose a coin is tossed and a person X says that Heads on the fifth toss is more probable than Heads on the fourth toss. Clearly, X must distinguish those two events, at the very least on the basis of the temporal coordinates of their occurrence. X must surely use a property in order to distinguish between the two events. Otherwise, one could not even say that they are *two* trials of the same phenomenon.

ALPHA: One could object here that a bettor who failed to distinguish the events on which she is betting could be easily deceived, as she herself would not be able to decide which one of the two events actually obtained. If she bet on one of them, it could be said to her that she had bet on the other, so that she could not realize the deception.

DE FINETTI: These are digressions. If one wanted to split hairs over the way two events could differ or not, one would surely arrive at no conclusion whatsoever.

ALPHA: If one took (as Carnap used to do) the ideal limiting case of a person with a null initial state of information, the axiom of symmetry would be equivalent — in the presence of infinitely many events — to the statement that any two absolutely indistinguishable events should receive the same probability.¹

DE FINETTI: If a person did not know anything, for her, every event E would be indistinguishable from every other and in particular, from its negation \tilde{E} . Then she should conclude that every event E has the same probability of its negation \tilde{E} , and therefore probability $1/2$. But such a conclusion would be incoherent in the presence of more than two mutually exclusive events.

Frequentism and Exchangeability

These are simple things. Much more dreadful difficulties arise when frequency is brought into the question. Indeed, the latter is even taken as the *definition* of probability. If one does that, there is nothing left to say. If ‘probability’ *means* frequency, then one could no longer ask what is the probability that the frequency will take this or that value: frequency would, in fact, be identical to probability *by definition*. Since an assertion of this sort would be completely senseless, the frequentist would say that in order to give a probability evaluation, one must observe the frequency through many trials or groups of trials. For example, in the familiar example of the draw of white and black balls from an urn of unknown composition, he could have observed, after many draws, that the frequency of white balls is around $2/3$ and around $1/3$ for black balls. He would conclude from this that the probability of drawing a white ball is close to $2/3$ and the probability of drawing a black ball is close to $1/3$.

Of course, one would try to make sure that the repetitions take place in such conditions as to make the statement of independence reasonable. One will therefore resort to a number of devices: for example, one will make sure that the balls are carefully reshuffled before each draw; that the person making the draw is changed after each trial, and so on. After making many sequences of experiments and after computing the mean among the events of each sequence (the so-called *relative frequency*)², one would compute the mean of all those means obtaining, in this way, the relative frequency of a bigger sample. And one would compare the differences observed among the relative frequencies of the various series in order to make sure that they are not too big.

This is an attempt to reverse the relation between probability and frequency so that one is then able to say, given the results of the experiments (given, that is, the

frequencies observed overall or in each of the subgroups of trials): “since the overall frequency turned out to be stable (alternatively: the frequency in each subgroup turned out to be stable and all the frequencies turned out to be closely grouped), the probability is close to those frequencies.”

Even a subjectivist would consider those conclusions reasonable: yet the subjectivist is conscious of the fact that in order to give a meaning to those considerations, it is necessary to initially assign the probabilities in such a way that the condition of exchangeability is satisfied. I have discussed exchangeability already:³ in what follows I will quickly recap the main idea.

To say that certain events are exchangeable is the correct way to say what, in a way which is only apparently more comprehensible (but which is, in fact, contradictory), is usually expressed as “independent events with constant but unknown probability.” In the case of an urn of unknown composition, this expression means that conditionally on every possible composition of the urn, the events are independent and with constant probability. What is unknown here however, is the composition of the urn, not the probability: this latter is always known and depends on the subjective opinion on the composition, an opinion which changes as new draws are made and the observed frequency is taken into account. The outcomes of the single draws are therefore exchangeable but not independent: the probability of every future event will not depend on the order of their occurrence but only on the observed frequency.

If one were told the composition of the urn, then the events would become independent. But until such a piece of information becomes available, they are not, for at each draw one would have a reason to change one’s probability evaluation in favour of the colour that has been drawn. After a large number of draws, one might even be practically certain that the composition of the urn is reflected with very good approximation by the observed frequencies.

Unfortunately however, the objectivist is not satisfied with this conclusion and says that for him the probability of drawing a white ball in a single draw on the basis of the observed frequencies is nothing but a subjective *estimate* of the “true” probability (which coincides with the frequency of the white balls in the urn, or in the class of infinite draws from the urn), an estimate which would become more and more reliable as the size of the sample increased. In this way she runs into the fallacy of identifying the composition of the urn with the probability. The composition of the urn, however, is a mere fact and cannot be identified with the probability.

These considerations clarify the sense in which one can say that whenever there is a reason to think of independence conditionally on a partition of hypotheses, the method of evaluating probability on the basis of the observed frequency appears to be reasonable. It must also be said that it is essential that those cases typically satisfy the condition of exchangeability. The fact that the events are considered to be independent conditionally on the composition of the urn (rather than conditionally on any other arbitrary partition) is, on the other hand, a detail of secondary importance. To conclude, we can say that among the objectivists (and in particular among statisticians) there is a tendency to mix up probability and frequency and consequently, to say: “probability is an idealized frequency.” This expression would have — if interpreted in the light of the concept of exchangeability — a kernel of

sensibleness. But — unfortunately — the interpretation they give to it is nowhere near being sensible.

There is yet another, even more theoretical, way to define probability. It is based on the recognition that the observed frequency is subject to change at each trial without there being a way to fix the number of trials needed in order for it to coincide with the probability. The argument is that in the long run, however, the frequency stabilizes. And by taking this consideration to extremes they say: “let us take as probability value the *limit* of the frequency over infinitely many trials.”

Let us consider such an infinity of trials. Let us write in a table the results of a sequence of trials as follows:

Trials	Results	Frequencies
E_1	0	0
E_2	0	0
E_3	1	1/3
E_4	1	2/4
E_5	1	3/5
\vdots	\vdots	\vdots

In the column on the right-hand side we have, corresponding to each “trial” E_i of E , the frequency f_i of successes in the first i trials and in the central column we list the results (0 or 1) relative to the single trials. The sequence $f_1, f_2, \dots, f_i, \dots$ tends to a certain value (the limiting frequency) and it therefore holds that:

$$\lim_{n \rightarrow \infty} f_n = f.$$

In this way f would be — by definition — the probability of E . All the objections that we have already raised against the identification of probability with frequency also apply to the present case. Yet to those objections we must add a new one, namely that the idea of an infinite sequence of trials is a purely fictitious idea: the number of trials actually made will in fact always be finite. And then one must base one’s judgment on the trials that have actually taken place. If a million trials have been carried out, then on the basis of the observed frequency, one will take as probability value a number close to that frequency, or even the frequency itself. But this is not *by definition*, rather it is because one realizes that it is not worth further refining the subjective evaluation with new experiments.

If one had an infinite sequence of trials at one’s disposal, a countable infinity \mathbb{N} of trials could be carried out. In such a sequence, the number of successes would again be a countable infinity: but the ratio between the two countable infinities in itself would have no meaning whatsoever, but would only make sense as the *limit* of the frequency. This latter, however, — and this is the key point — changes with the ordering. Therefore, one cannot speak of the limit of the frequency as if it were a *proportion* within the class of all trials (in a false analogy with the finite case, where the order is irrelevant instead).

ALPHA: Furthermore, it must be observed that the limit of the frequency does not always exist.

DE FINETTI: In the case of an indefinitely extensible sequence of exchangeable events, as a consequence of the fact that the deviation of the observed frequency from the probability is expected to decrease as $(\sqrt{n})^{-1}$, one can say — roughly — that it is almost certain that the frequency will tend to probability in the limit. Yet one can say that provided that one bears in mind that “almost certain” means just “with high probability.”

BETA: There is an analogy between those attempts to determine the limit of the frequency by means of finite sequences and the computer-based attempts to solve Fermat’s problem.⁴ One could compute the frequency over a billion trials, but still another countable infinity of billions of trials would be needed to reach infinity.

Von Mises’ “Regellosigkeitsaxiom”

DE FINETTI: There is an alleged refinement (actually this is an even more radical degeneracy) of this orientation. It is very much the same thing yet with the addition of a further postulate, the *Regellosigkeitsaxiom*.⁵ “Regel” in German means “rule,” “los” means “without” and “keit” is a suffix for abstract terms: therefore “Regellosigkeit” means “irregularity.” The meaning of the word, hence, is: *axiom of the lack of regularity* (that is, of disorder). The notion of disorder is very vague indeed. If a binary sequence consisted of alternated “0” s and “1” s, i.e., 0101010101 . . . , then it would be considered a regular or ordered sequence. And whoever knew that a certain sequence of digits contained exactly the first 20 digits of π , would consider it to be regular. But if a person came across the following sequence:

11001101100110110101111011000100111110011110111000100110

without knowing that it corresponds to the juxtaposition of the first 20 digits (expressed in binary form) of the decimal part of π , she would in all probability say that it is a random sequence.

BETA: The irregularity axiom does not say anything, for no one is able to determine whether a sequence is random or not, unless one is so lucky as to face a sequence with the “0” s and the “1” s all nicely arranged — so to speak.

DE FINETTI: The point here is that even that could happen by chance. Those who rely on the *Regellosigkeitsaxiom* mistake very small probability with impossibility. Underpinning this confusion is the fallacy highlighted by the *Lottery Paradox**: although everyone who buys a lottery ticket has a minimal probability of winning, there is always going to be a winner. Why exactly her if it was almost impossible that she would win? Commenting on this paradox, Corrado Gini⁶ (a statistician, former President of the *Italian National Statistical Bureau*) used to observe ironically: since

* (Translator’s note:) In English in the original text.

the occurrence of an event of very small probability is always hardly credible, the winner of the lottery should be arrested.⁷ The probability, in fact, that exactly that person would win is so small that one should always presume that she won due to some trickery.

At any rate, it is completely bizarre to ground the definition of probability on an axiom like the *Regellosigkeitsaxiom*, for this latter has no precise meaning whatsoever. According to von Mises⁸ — who introduced it (von Mises, 1919) every random sequence must necessarily satisfy such an axiom. For example, not only the frequency of the digits '0' and '1' in the entire sequence, but also every infinite subsequence should tend to $1/2$, provided it is not determined as a function of the occurrences of '0' and '1'.

BETA: Does it have to tend to $1/2$? What if there is a coin that is not fair: would the probability of Heads still be $1/2$?

ALPHA: Supposing the coin was fair.

DE FINETTI: I just made an example involving $1/2$, but one could take any other value, say $1/3$. If the probability were $1/3$, one should have, on average, m successes every $3m$ tosses. There are asymptotic laws to the effect of determining, given a sequence of independent events with probability p , the expected difference between p and the actual frequency, in terms of the number n of trials. One of those laws is the so-called *Law of the iterated logarithm*. This law singles out an upper bound to the fluctuations of the difference between the actual number of successes S_n and the prevision np of such a number, given in terms of the number n of trials.⁹

I would like to conclude by saying that the only frequentist thesis that I find acceptable is that knowledge of frequencies can reasonably suggest attributions of probabilities close to such frequencies. Yet this is not a necessary result, rather it is the conclusion of an argument that is based on *subjective* premises. From the objective point of view however, the fact that a certain frequency occurred or not in the past is completely irrelevant to the occurrence or not of future events.

Editor's Notes

1. See, Mura (1992, pp. 154–157).
2. Recall that de Finetti identifies an event (or proposition) E with the random quantity which can assume either the value 1 or 0, according to whether E turns out to be true or false, respectively. Keeping this in mind, given the events E_1, E_2, \dots, E_n , their mean value — being equal to the sum of the events whose value is 1 (that is to say to the number of true events) divided by the total number of events — coincides with the relative frequency of true events.
3. See Chapter 8.
4. The so-called *Fermat's last theorem* states that $x^n + y^n = z^n$ has no non-zero integer solutions for x, y and z when n is an integer greater than 2. Pierre de Fermat (1601–1665) raised the problem in 1630 by writing a marginal note in a copy of Diophantus's *Arithmetica*, where he claimed that he had discovered a proof, which the margin could not contain. After a long history of efforts towards solving the problem, finally Andrew John Wiles (1953–) proved Fermat's last theorem in 1995.
5. On this topic see also Chapter 5 pages 54–57 and Chapter 6 pages 62–66.
6. See note 1 page 74.

7. Cf. Chapter 7 page 73.
8. For biographical details on Richard von Mises see Chapter 5, note 13 page 58.
9. The law of the iterated logarithm is due to A. Khinchin (1924). Intuitively, it states that it is almost certain that the difference between the actual number of successes and the relative prevision lies (beyond a non-determinable n) within the interval

$$np \pm \lambda \sqrt{2np(1-p) \log \log n}$$

provided that $\lambda > 1$. Whereas if $\lambda < 1$ it would be almost certain that the number of successes would somewhat frequently (according to the value of λ) lie outside such an interval. More precisely, the theorem states that given a countable sequence of independent events with constant probability p , for each $\theta \in (0, 1)$ there exists an indeterminate n such that the number of successes S_i satisfies, for each $m > 0$ and every real coefficient $\lambda > 1$, the following relation:

$$\mathbf{P} \left(\bigcap_{i=1}^m S_{n+1} \in np \pm \lambda \sqrt{2np(1-p) \log \log n} \right) > \theta. \quad (13.1)$$

Moreover, for $\lambda < 1$ and for all $\theta \in (0, 1)$ there exists a natural number n such that:

$$\mathbf{P} \left(\bigcup_{i=1}^n S_{n+1} \notin np \pm \lambda \sqrt{2np(1-p) \log \log n} \right) > \theta. \quad (13.2)$$

Chapter 14

The Gambler's Fallacy*

Discussion on various kinds of misunderstandings and in particular those deriving from transposing the judgment of "great" or "small" probability from the single case to the collective case.

Against the Measure-Theoretic Approach

BETA: Could you explain why you are against the introduction of the axiom of countable additivity?

DE FINETTI: Countable additivity is an *ad hoc*ery, which is endorsed, as Good says, "for mathematical convenience".^{†1}

BETA: Nonetheless, countable additivity gives rise to unsurmountable pathologies in measure theory. For instance there are Vitali's null sets,² which in my opinion, are among the most pestiferous things one could possibly come across. On the other hand, by endorsing simple additivity, although one has to struggle a lot more to prove Vitali's theorems, one gets in return the advantage of being able to measure everything. From the measure-theoretic point of view — rather than from the probabilistic point of view — it is unpleasant to face non-measurable sets.

ALPHA: There are cases in which the problem does not arise. This happens for instance in inductive logic, where one takes the space \mathcal{S} whose points are the products of maximal ideals (or ultrafilters) of the algebra of first-order logic (where logically equivalent sentences are identified). \mathcal{S} is a Hausdorff space. Moreover, Gödel's completeness theorem (Gödel [1930] 1986) — which is considered to be the most important theorem in first-order logic — shows that \mathcal{S} is compact.³ Of course, every probability function defined over \mathcal{S} satisfies — as a consequence of the compactness of \mathcal{S} — the axiom of countable additivity.⁴

BETA: In this way, however, one obtains only a part of the events.

ALPHA: A probability is assigned to everything that can be expressed in the language. If something cannot be expressed with the resources of the language, then it fails to receive a probability.

* Lecture XVI (Friday 4 May, 1979).

[†] (Translator's note:) In English in the original text.

BETA: In measure theory one sometimes needs to avoid Vitali sets and measure *everything*.

ALPHA: It seems to me that it is legitimate to assign a probability only to what can be expressed in some way in the language.

BETA: Yes, but measure does not have just a probabilistic aspect: it also has a fundamental analytic aspect.

ALPHA: What cannot be said, cannot be thought either. And after all, an event is thought of in words.

DE FINETTI: If this were true, we should limit ourselves to considering (to take the geometric image) all the sets of points that form "potatoes," or bits of "potatoes" (finitely many intersections of "potatoes").⁵ In this case, one would have a finite partition and no problem would therefore arise.

ALPHA: By means of those languages we can reach infinity, by using certain devices such as quantifiers.

BETA: One could subjectively evaluate the probability of hitting the Vitali set with an arrow in the interval $[0, 1]$ (assuming, as I think is reasonable to do so, that the probability of hitting an interval is proportional to its length). If we assume that probability is proportional to the Lebesgue measure, we cannot define the probability of a Vitali set. Therefore, countable additivity gives rise to a decidedly serious shortcoming.

ALPHA: Has anyone ever tried to pose the question of probability in relation to *definability*? There is an extremely rich logical theory which addresses the limits of definability in a language.

DE FINETTI: In my opinion, all these discussions pertain to the structure of certain spaces considered as sets of points, rather than to probability. As a consequence, those problems are totally foreign to the theory of probability. Consider a "potato" (Fig. 14.1).

Each point corresponds to a possible "elementary case." But it would be equally legitimate to take as elementary cases all the straight lines passing through a certain point of the abscissa. This could be extended to n dimensions, so that we could take into account all those new quantities in which we might become interested. We can introduce as many of those quantities as we like, in theory even a countable infinity or a continuous infinity. Of course, those would be functional spaces, which in no way could be represented through figures. One could nonetheless resort to colour. For instance, one could associate to each point the number corresponding to the wavelength of the colour used to represent the point.

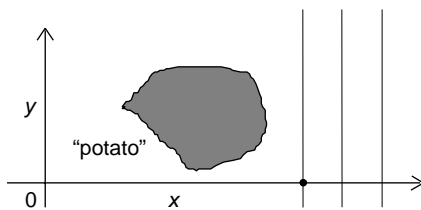
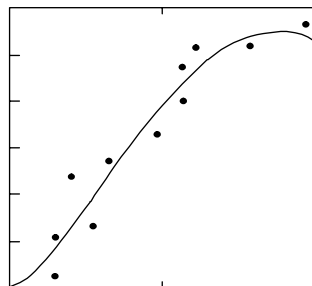


Fig. 14.1 Graphical representation of events by convex sets ("potatoes")

Fig. 14.2 Plot of empirical data along with a function fitting them



If one took the behaviour of an empirical function (Fig. 14.2) — corresponding, that is, to the behaviour of temperature or of some other quantity — one could look at the probability that y will always be included in the interval $f(x) \pm \varepsilon$. There is not a unique way of confining y within an interval in such a way as to satisfy an equation of the form $y = f(x) \times (1 \pm \varepsilon)$. Whatever choice of ε would be ad hoc, though there is nothing necessarily bad about this, for it is not clear what definition would not take into account a certain intrinsic inexactness of things. The same would hold for an arbitrary random quantity X . Of course, if it were known that X can only take a finite number of possible values, it would always be possible to distinguish them. But suppose that X could take all the rational values $1/n$ (which constitute a countable infinity). In this case, almost all the values of X would be concentrated around the origin and still, a probability could be assigned to all of them.

ALPHA: It is interesting to consider the *families of properties*, as Carnap called them (Carnap [1950] 1962, pp. 76–78). In particular colours, flavours, and so forth.

DE FINETTI: Who is it that you mentioned?

ALPHA: Carnap. The families are always decomposable in mutually exclusive sub-cases (it is impossible that the same point is of two colours . . .)

DE FINETTI: In *Der logische Aufbau der Welt* Carnap needs half a page of symbols to say that the space of colours is a three-dimensional space, as all the colours can be obtained by trichotomy, starting with just three fundamental colours (Carnap [1928] 2003, §118). Now, such a space could just as well correspond to the space in this room, which — like the space of colours — has three dimensions. What I mean is this: it is one thing to propose a merely formal definition characterizing colours up to a bijective and continuous transformation, and quite another thing to *define* colours as we can see them with our own eyes. Not every animal, nor every camera in fact, represents colours in the same way: we all noted that in photography colours are sometimes altered (usually through a bijective transformation). Carnap's approach fails to capture the distinction between colours as we see them and the way they are represented by the camera. Every colour should be characterized as a mixture of various wavelengths, rather than in the formalistic way of Carnap.

ALPHA: What you say is correct, but what I wanted to bring to your attention is Carnap's idea that certain properties by means of which the whole of experience

could be described, belong to families of properties which form a partition (after all, everyone sees colours, tastes flavours, and so on.).

BETA: To represent tastes with points is a rather bizarre idea.

DE FINETTI: I wonder whether we would swallow the points too while eating!

Gambler's Fallacy and Frequentist Fallacy

But let us go back to probability. One thing that should be borne in mind is that the probability of a single case and the probability of many cases belonging to some class of cases are often confused. For example, it is sometimes claimed that "a sequence constituted only by '1' s does not represent a possible sequence of outcomes of the toss of a coin." Surely, there are 2^n possible outcomes for n shots: as the outcome "all Heads" is unique, it will have probability 2^{-n} . But if one took a certain (somewhat random) sequence, one could say: "it cannot be this one either, as it has itself probability 2^{-n} ."

Usually, sequences are classified into random and non-random. This can only make sense in relation to the idea that one might form about why that very sequence occurred, or it has been written down exactly in that way. While I was working at the *Italian National Statistical Bureau* immediately after my graduation, professor Gini wanted to carry out an experiment. He asked all the employees to write down 200 digits at random. No one had a clear idea about what we were supposed to write. But almost everyone showed the tendency to write numbers more at random — so to speak — than they would have resulted from a random process. If one used a ten-faced die, a somewhat irregular sequence (depending on "the die's will") would have come up. But suppose that no one had an interest in using dice and concentrated instead on writing the 200 digits at random. Very probably one would have made sure that every digit occurred in the sequence with a frequency $1/10$. Then one would have been tempted to touch up the digits in an *ad hoc* way, rather than keep them as they randomly came to mind. In fact, it turned out that the sequences constituted, for instance, by eight consecutive occurrences of the same digit, were a lot less frequent than one should have expected. I wrote down repeated digits (but I did so because I knew that one should have expected them to be there, rather than because the same digits came to my mind eight times in a row). This shows that it is contradictory to say that in order for a sequence to be considered random, it should satisfy certain general properties.

However, the worst mistake is (as I said in a previous lecture)⁶ the introduction of the *Regellosegekeitsaxiom*, which is allegedly meant to provide some rules to exclude all those cases that can occur, but would not be sufficiently random. Thus, for example, should the set of five 1 2 3 4 5 came out in a lotto draw, one would be tempted to say: "this is an exceptional combination, it cannot be the case that it came out randomly." The truth of the matter is that this is a set of five as any other: there are no special reasons why it is harder for that set of five to be drawn than another one. Notwithstanding, I firmly believe, for the little I know about lotto players, that none of them have ever played the set of five 1 2 3 4 5!

ALPHA: Another fallacy is the one into which players of the *totocalcio* run when making the so-called *development of a system*.^{*} It is as follows. To construct a system means to determine the class of columns to be played. In order to determine such a class one starts with all the 3^{13} possible columns. At first, those matches for which one of the three results 1 X 2 is considered significantly more probable than the other two are identified. Then all the columns that do not contain that result are excluded. Secondly, the matches whose outcome is more uncertain, in the sense that only one of the three possible results can be excluded as improbable, are chosen. All the columns listing the excluded result are eliminated from the system. After all these exclusions, the resulting set of remaining columns constitutes the "system." So far nothing wrong: the columns are excluded on the basis of the probability assigned to the various outcomes. But the "development" of the system consists in "reducing" it further by excluding also the system columns which are considered to be "too regular," or to have "too infrequent" features. On this basis, the column with thirteen "2s" would be excluded even if, by chance, it belonged to the system.

DE FINETTI: This is an excellent example. And this practice is all the more absurd given that the occurrence of consecutive identical results depends only on the order in which the matches are listed on the card.

BETA: Yet those who play the system do not do things randomly: they think about it.

ALPHA: The fallacy does not consist in making the system, but in *developing it*, that is in eliminating those columns that are considered to be too regular or to have infrequent properties.

DE FINETTI: The fallacy derives from the fact that the properties of a single case are mistaken for the properties of a whole set of cases that are considered to be analogous and in which that single case is included.⁷

ALPHA: As to the "cold numbers" in the Lotto, one could object as follows: "were the cold numbers really favourite, it would be enough to buy a Lotto urn, to make many trials at home and, as soon as a certain number became cold according to those trials, run to the lottery office, play that number and therefore win with practical certainty." Of course, players would reply that the home-made trials should not be taken into account, for only the official draws matter.

DE FINETTI: Sure. But this is only one part of the superstition. If the fallacy consisted only in this, it would actually be a different superstition: the superstition according to which it is not possible to predict the winning number by playing alone. Perhaps such a superstition would be too subjectivistic to be considered acceptable!

ALPHA: The fact that such a piece of information is taken to be relevant suggests that, after all, the origin of the fallacy is not in the same mistakes that give rise to the objectivistic interpretation of probability.

DE FINETTI: It is hard to analyse crack-brained ideas but my guess is that those players think that the regularity exists urn by urn, and then in the whole system of urns, (according to whether a given number is cold, e.g. in Rome's urn, or cold in general). Thus, it looks as if many think that if a certain number fails to be drawn *in*

^{*} (Translator's note:) See the Chapter 1, footnote ^{*} page 6.

Rome for a long time, then it will have to come up soon *in Rome*. If a number failed to be drawn for a long time from any urn, then it should be played for every urn, though it is not known from which urn it will be drawn. Then, maybe, they would even consider it probable that it will be drawn from all the urns at the same time!⁸

ALPHA: But if one thought that the phenomenon of lotto is governed by some probabilistic law, as believed by the majority of . . .

DE FINETTI: [Interrupting] Those who play cold numbers? They say (and this is the most obvious example of the gamblers' fallacy): "since it never happened that a number failed to be drawn (alternatively: it was extremely improbable that it would have been drawn) for a hundred weeks consecutively, one should expect that this time is going to be the right time."

ALPHA: Many bet just for fun. There are many who without believing these superstitions, pretend to do so: superstition is for them just an excuse to play.

DE FINETTI: Betting is disadvantageous in the long term, whereas it could be lucrative in the case of a single shot. Suppose, for example, that one had to pay, threatened with death, a debt amounting to a billion (which one does not possess). In such a situation, it would be advantageous to risk one's all in Montecarlo, in the hope of winning. If the only way to avoid death was to pay a billion and the gambler, by risking all had a probability — if a very slim one — of succeeding in avoiding death, then the most profitable choice for her would be to try her own luck. Savage has worked on this problem.

ALPHA: Profitability should be shown in terms of the curvature of the utility function.

DE FINETTI: In a case like this, there is no *curvature* of the utility any more, as the utility is 0 for amounts below a billion and ∞ for amounts above it (and even if it were not exactly ∞ , it would still be the upper bound of the value of everything in the world).

ALPHA: On other occasions I brought with me Savage's book, which I am reading.⁹

DE FINETTI: I seem to recall that this example is discussed in the book. I know that he talked about it at a conference in Bressanone.

ALPHA: The CIME Conference?¹⁰

DE FINETTI: Perhaps it was organized by the CIME but it is anyway an event that took place some years ago.

ALPHA: Was it in 1959? That year Savage came to Italy.¹¹

DE FINETTI: He came several times: the 1959 visit was his penultimate one here. Afterwards he came another time, even if he was only passing through Italy on his way back from the conference in Bucharest. He died shortly afterwards.¹² The problem discussed by Savage concerned the optimization of the sum available to determine whether it is profitable to bet immediately the whole amount, or to place various successive bets. Savage carried out a simple and interesting investigation.¹³ I cannot recall the precise details now, though I remember that the solution was shown by a curve which illustrated how, in order to maximize the probability of winning the amount required to pay the debt, one should make several attempts.

ALPHA: Similar situations also occur "in the real world," during economic crises. During those periods, in fact, the number of players and the amount of money bet,

especially in those games which — like the *totocalcio* — promise considerable wins, increase dramatically. And it also happens that the game becomes more uniformly distributed than usual on the range of possible results (in fact even the weirdest combinations are played in the hope of a big win).

DE FINETTI: There are two contrary tendencies that compete in the psychology of the gambler. On the one hand, there is the tendency to bet in such a way that one can hope for a considerable win (at the expense of the probability of winning), whilst on the other hand, there is the tendency to maximize the probability of winning (at the expense of the amount of the win).

ALPHA: It all depends on the gambler's attitude towards risk.

DE FINETTI: Those are psychological tendencies that are subject to change at any moment. For instance, should the news spread about big wins that have occurred recently, one could think that it is a favourable time for gamblers. On the other hand, one could also believe that it would not be worth risking: given that — so to speak — Fortune has been so generous for a few days, she will presumably allow herself a period of rest.

BETA: My sister won a 12 and a 13 at the *totocalcio* in the span of three consecutive weeks by playing systems. Of course, the win was not big, being those probable results. After the wins, she decided that she would stop playing, because she believed that Fortune would have not smiled on her anymore.

ALPHA: Let us take the toss of a coin. Here the advantage depends on the cumulative gain. In other words: the gain depends, at each shot, on the sum of the gains obtained in the previous shots. In this way periods of seeming "fortune" can easily occur during which one of the players leads for a long time. This phenomenon is studied in Feller's book.¹⁴

DE FINETTI: To say that there is a period of fortune is to change the terms of the problem. It simply happens that by chance, certain sequences of various lengths of consecutive wins occur. Every time the sign changes, it does not make any difference if it is the first or the millionth shot. But the probability that the better who played the last shot will win just once before losing a shot is $1/2$. And the probability that she will win twice in a row before losing a shot is $1/4$; $1/8$ three times; $1/16$ four times; $1/32$ five times; $1/64$ six times; $1/128$ seven times; and so on. Probabilities like $1/32$, $1/64$, or $1/128$ are not that small: therefore not-so-long repetitions should not come as a surprise.

Many times one wonders: how is it possible that it was exactly her who won? Or, why a person with such an odd name? Actually, whoever is the person to whom strange things happen, she is always going to be "her" for those who know her. Whatever misfortune might happen, it must happen to someone and she is going to be a well-determined individual. As to the name, if we knew that the name of the winner is Asdrubale,* one could be tempted to say: "how is it possible that a person by the name of Asdrubale won the lottery? It is such an uncommon name!"

* (Translator's note:) Name of Phoenician origin. It is, by antonomasia, synonymous with "uncommon first name."

BETA: What if there were two winners, both by the name of Asdrubale? In such a case many would surely say: "there must be a trick in it!"

Events and Propositions

DE FINETTI: If one gave a sufficiently precise and detailed description of what happens, one could always say that an event of probability 10^{-n} , for an arbitrarily large n , obtained. Analogously, if we included with sufficient precision in the description of the event E the temporal coordinates at which E occurred, the probability of E could be made arbitrarily small. And the probability that something happens at a punctiform instant is typically 0.

ALPHA: This example shows that probability, instead of being given to the events themselves, is given to the propositions. Were an event to be described in its full details, its probability would be very small. On the other hand, were it to be described very generically, then it would become very probable. Therefore, the probability of an event depends on the words used to describe it.

DE FINETTI: It is precisely to avoid this ambiguity that I call *event* the single case.

ALPHA: Yes, the single case, but described in a certain way.

DE FINETTI: I do not speak of "trials of an event," because this way of putting it presupposes that the single trials are not themselves events, and therefore not objects to which a judgment of probability can be attached. Although this is a terminological question, it is quite important. Indeed, many believe that the evaluation of probability can only be applied to *repeatable* events. According to them, if that were not the case, the evaluations of probability could not be controlled through the observation of the frequencies. Moreover, by saying "trials of an event" one takes it for granted that they are equally probable events. Usually, the independence of the trials is also taken for granted (perhaps even before independence itself has been defined). Unless one speaks, as if it were a special case, of "events whose probability changes from trial to trial." Then everything collapses because in this case probability would not be assigned to events any more, but to the single "trials."

ALPHA: Perhaps the most appropriate term is "proposition," as probability depends on the way facts are described. And if one described a fact in a progressively more detailed way, such a fact would become a *distinct* event every time, despite being the same fact all along. This shows that probability refers to the propositions of a language, rather than to the events in the world which, in themselves, are so complicated that once described with sufficient precision, would all get probability 0.

DE FINETTI: I agree, but it seems advisable to me to keep the term "event" for the object of the evaluations of probability. The reason for this being for the sake of uniformity with the usual terminology of the calculus of probability. Even though the interpretation I propose of the word "event" is very different from the one endorsed by the frequentists, given that for me an event is intended as something well-determined rather than something generic and repeatable.

ALPHA: There are several advantages deriving from such a choice: on the one hand, it allows one to keep using the familiar term “event”; on the other hand it allows everyone to interpret it according to their own preferences.

DE FINETTI: There is a risk of being misunderstood, yet every term can be misunderstood. I would not know which other term I could use, unless one did not invent a new word and said, for instance, “*abracadabra*.”

ALPHA: I repeat what I have already said: the word “proposition” seems appropriate to me.

DE FINETTI: The proposition does not always show all the conditions that need to be checked in order to decide the outcome of a bet placed on it. Moreover, I do not like “proposition” because it suggests that probability applies to grammatical objects. It is of course true that an event can be expressed by a proposition of the language. Yet an event should be something which can be thought of without any need for expressing it, not even mentally, through a proposition of the language.

ALPHA: To touch upon the nature of propositions is like stirring up a hornets’ nest, for it is a really controversial problem. Philosophers distinguish, in general, between *sentences* and *propositions*. Sentences correspond to the sequence of letters by means of which a proposition is expressed. Propositions, on the other hand, correspond to *what the sentences express*. It can be also put in this way: a proposition corresponds to the common content of all those sentences (possibly belonging to distinct languages) which — if by means of distinct words — say the same thing. And although the proposition is not tied to a specific phrase or language, it is nonetheless a linguistic object. One could maintain that propositions belong to the world. This, for instance, was precisely Wittgenstein’s idea.¹⁵ Indeed, according to him, the structure of the world mirrors the language: as there are elementary sentences there would be, in the world, *elementary events* corresponding to the former, as well as compound events (disjunctions, conjunctions, negations, etc. of other events). Then, unless one rejects Wittgenstein’s theses, where could such an intersection of events take place if not in the language? It does not seem to me that the events perform those operations unknown to us. They intersect, they negate. After all, — I guess — they do so in the language.

DE FINETTI: These considerations remind me of a famous film on World War I, in which two French prisoners, after escaping from prison with a trick, walk until they reach Switzerland. Crossed the Swiss border, one of them, realizing the fact, says to the other: “we are safe now: we just entered Switzerland.” And the other replies: “where did you see the border? I hadn’t realized at all that we had entered Switzerland.” And the first soldier replies: “No way! The border is not something you can see: Nature doesn’t give a fuck!”

ALPHA: He wanted to say that the border is just a matter of convention.

DE FINETTI: A human, political convention.

ALPHA: The comparison is quite appropriate. One could think that some elementary events existed in the world, which Wittgenstein used to call *Sachverhalten*, that is to say the events or atomic facts that are not decomposable.

DE FINETTI: “Halten” means behaving. What would that mean?

ALPHA: The atomic fact. According to the theory of logical atomism, by progressively decomposing facts, one would eventually reach some elementary facts, constituting relations between simple objects, which cannot be decomposed any further. And all the facts in the world would be particular combinations of those elementary facts.

DE FINETTI: Would not this be the theory Dante refers¹⁶ to by saying of Democritus that he “ascribes the world to chance”?*

ALPHA: In a certain sense yes, because in fact, Wittgenstein claimed that nothing is necessary in the world apart from what is *logically* necessary.¹⁷ Wittgenstein discovered the Boolean tautologies and introduced the so-called *truth-tables* (I believe he was the first one to discover them).¹⁸ He was very impressed by the existence of propositions which are “always” true or “always” false, i.e., propositions such that their truth-value is independent of the truth-value of their constituents. For instance, the truth-value of the proposition $p \vee \bar{p}$ does not depend on the truth value of p , because $p \vee \bar{p}$ is true both in case p is true and in case p is false. Wittgenstein used to say that propositions like this are *sinnlos*.¹⁹

DE FINETTI: It means: empty of meaning.

ALPHA: And he added that they are limiting cases of propositions, which do not contain any factual information.²⁰ He believed that every proposition is a truth-function of elementary propositions, i.e., that it can be obtained from these latter through connectives such as conjunction, disjunction and so on.²¹

Editor's Notes

1. On “ad hoceries” see Chapter 4 at page 36 and note 5 at page 42.
2. Vitali's null sets are examples of sets that are non Lebesgue measurable. A Vitali's null set is always defined by means of the axiom of choice. For example, in the real line, it may be obtained by picking one (and only one) point in each coset of \mathbb{Q} of the additive subgroup of rational numbers.
3. The completeness theorem states that if a sentence p is a logical consequence of a set of sentences M (in the sense that whenever all the elements of M are true, then p is true), then there exists a derivation (according to the rules of inference of the logical calculus) of p from M . As in such a deduction the set of actual premises is constituted by a finite subset M' of M , and every derivation within the calculus is carried out *salva veritate*, p is a logical consequence of M' . Therefore, whenever a sentence p is a logical consequence of a set of sentences M , p is also a logical consequence of a *finite* subset M' of M . This corollary is known as the *finiteness theorem*.

Moreover, for every inconsistent set M there exists — by definition — an inconsistent logical consequence p of M . By the finiteness theorem, there exists a finite subset M' of M from which p follows logically and which is therefore inconsistent. Therefore, every inconsistent set M contains a finite subset, which is itself inconsistent.

This result (known as the *special finiteness theorem*) enables the derivation of the compactness of the Hausdorff space mentioned in the main text. Indeed, in such a space, a *closed*

* (Translator's note:) The translation of Dante's line “che il mondo a caso pone” adopted in the main text follows that of Charles Singleton's, Princeton University Press, 1970.

set is associated with every sentence and in particular, the empty set is associated with every inconsistent sentence. The special finiteness theorem therefore implies that in that space every family of disjoint closed sets contains a finite family whose intersection is, in turn, empty. This condition is equivalent to the condition of compactness.

The topological representation of the logical metatheorems is due to E. W. Beth (1951).

4. In a compact metric space \mathcal{S} generated by a countable set, every sequence has a limiting value. It follows that every simply additive measure defined on \mathcal{S} is also completely additive.
5. De Finetti referred to Venn diagrams as “potatoes.”
6. See Chapter 6 page 66 and Chapter 13 page 134.
7. According to de Finetti, at the root of the gambler's fallacy lie some of the fallacies that, in his opinion, can also be found at the root of frequentism. In particular, the assumption that probability depends necessarily on the way the events are ordered in a sequence as well as on the choice of a “reference class” is a mistake made both by those gamblers who believe in cold numbers and by those probability scholars who embrace frequentism.
8. The Italian Lotto, also called “Lotto di Genova” derives from the practice that emerged in the Republic of Genoa during the XVI century of betting on the candidates to public charges. In that Republic, twice in a year, 90 names of worthy people were numbered and 5 of them were elected to public charges by drawing without replacement 5 numbers from an urn (called “seminajo”) containing 90 ballot-boxes each with a different number ranging from 1 to 90. The game of betting on various events resulting from the draws became very popular.

Later, the game was repeated even without connection with elections as a pure game of chance and, as such, replicated in several other States of Italy. Two years after the unification of Italy into a single State (1863), the Lotto was reorganized by the government (who managed it directly) and its revenues considered as fiscal revenues. Moreover, all the separate lotto games were unified in a single national game, although the draws continued to be made in the original cities and new cities were added with an urn for drawing. The system of bets changed because fixed odds were established. The game has remained substantially unchanged ever since.

The fixed odds are very unfair (for the draw of a single number in a specified urn they are about 1:11 while the corresponding fair odds are 1:17). For this reason (considering the fiscal character of Lotto revenues), de Finetti often spoke about the Lotto as a “tax on stupidity” (see Chapter 6 page 65).

9. The reference is to *The Foundations of Statistics* (Savage [1954] 1972). Actually, this work does not include the example discussed in the present lecture.
10. CIME (*International Mathematical Summer Centre*) is a still active Italian organization founded in 1954 by the UMI (*Italian Mathematical Union*) to promote the internationalization of the Italian mathematical community after its long isolation due to Fascism and the Second World War. The Summer CIME courses are directed by leading Italian academic scholars. The lectures are held by (typically foreign) mathematicians of world-wide renown. De Finetti himself organized CIME courses in 1959, 1965, and 1966.
11. This is an allusion to the CIME course directed by de Finetti in Varenna (Italy) in June 1959, entitled “Induzione e statistica” (“Induction and Statistics”) with lectures by de Finetti and Savage. These lectures (in Italian) were recorded and then put in written form. A transcription is available from the library of the Department of Mathematics “Guido Castelnuovo” in Rome. An English translation of a revised version of de Finetti's lectures is included in de Finetti (1972). Savage's lectures, entitled “La probabilità soggettiva nella pratica della statistica” (“Subjective Probability in the Statistical Practice”) actually do not contain any reference to the problem de Finetti refers to in the text.
12. The allusion is to the *Logic, Methodology and Philosophy of Science* conference, which took place in 1971. The proceedings are in Suppes, Henkin, Joja and Moisil (1973). Savage died in New Haven (US) on 1 November 1971, aged 54.
13. These investigations led to the book Dubins and Savage (1965), where, through gambling concepts, a general theory of discrete-time stochastic control was developed.
14. See Feller ([1950] 1967), pp. 77–88.

15. Here and in the following paragraphs, the reference is to the *Tractatus Logico Philosophicus* (Wittgenstein [1921] 2001). As is well-known that Wittgenstein disowned, in a subsequent phase of his philosophy, the theses that are briefly pointed out here.
16. Cf. *The Divine Comedy, Inferno*, Canto IV, 136.
17. Reference to proposition 6.37 of the *Tractatus* (Wittgenstein [1921] 2001): "A necessity for one thing to happen because another has happened does not exist. There is only *logical* necessity."
18. Truth tables, partly anticipated by C. Peirce (1880), were introduced in 1921 by Wittgenstein ([1921] 2001) and independently by E. Post (1921).
19. See *Tractatus* (Wittgenstein [1921] 2001), proposition 4.461: "Tautology and contradiction are without sense. (Like the point from which two arrows go out in opposite directions.)"
20. See *Tractatus* (Wittgenstein [1921] 2001), proposition 4.466: "Tautology and contradiction are the limiting cases of the combination of symbols, namely their dissolution." and *Tractatus* (Wittgenstein [1921] 2001), proposition 4.461: "I know, e.g., nothing about the weather, when I know that it rains or does not rain."
21. See *Tractatus* (Wittgenstein [1921] 2001), proposition 5: "Propositions are truth-functions of elementary propositions."

Chapter 15

“Facts” and “Events”*

Distinction between:

“Fact” which can “happen”;

“Event” which can “be verified”;

“Phenomenon” which can “repeat itself”;

Fact: in a generic sense, something which happens, or which can happen (objective);

Event: a precise, detailed statement which is ultimately either verified or disproved;

Phenomenon: generic term to speak about the “trials of the same phenomenon of similar cases” (instead of “trials of the same event”).

A Pragmatic View of Events

DE FINETTI: Today we will address a terminological question. The goal is to clarify the meaning of the word “event” and other related terms.¹

ALPHA: The word “happening” could be used. This is the term used by Popper ([1935] 2004). It stands for the single case, distinct from ‘event’ in the frequentist sense.

DE FINETTI: I prefer to carry on using the term ‘event’ to denote the single case. However, it could also be useful to the subjectivist to have a term designating what frequentists mean by ‘event,’ provided one always avoids saying that two cases that fall within the scope of that term are of the same sort. I propose the word ‘phenomenon.’ In the presence of a class of events exhibiting some reciprocal analogy, I say that every event is a trial of the same phenomenon without making any assumptions on the nature of such an analogy and without presuming that it constrains the probability value of each trial.

ALPHA: In the logical lexicon, instead of saying that certain events are *trials of the same phenomenon*, one would say that they are *instances of a particular property*. The property is what is common to all the examples. The rolls of a die would then all be instances of the property “to be the roll of a die.”

* Lecture XVII (Thursday 8 May, 1979).

DE FINETTI: But would it sound right to say, for instance, "I perform many trials of a property"? Properties would rather seem to be things like, for instance, colour and weight.

At any rate, I shall now explain the solution that appeared reasonable to me yesterday. It looked appropriate to me to say of what happens that it is simply a fact. For instance, that Heads came up on the seventh shot is a fact in its own right, whether one judges its probability or not.

ALPHA: I have some difficulties here.

DE FINETTI: To say 'fact'? What would be your alternative proposal?

ALPHA: Probability seems to refer to some entity connected with language, not with a particular language but with language in a broad sense. In fact, we speak of the probability of the negation, of the probability of the union, of the intersection and so on. To think of facts that can be negated seems to be a bit of an artificial thing. Similar remarks apply to unions and intersections.

DE FINETTI: No, no, wait a second. The distinction I propose goes as follows. A fact is such, whether one is thinking of betting on it or not. It is a *fact* that at this moment there are three people in this room. It is a state of affairs.

ALPHA: Yes, fine, it is a state of affairs, but . . .

DE FINETTI: [Interrupting] 'State of affairs' would however be less appropriate because it suggests a more specific concept than the one I have in mind when speaking of *facts*.

ALPHA: I insist: one could use the term 'proposition', which is used by logicians. It seems appropriate to me because it does not have anything to do with the grammatical form of the sentence.

DE FINETTI: Yes, but 'proposition' corresponds, in any case, to what I would like to call 'event.' A fact is a fact in its own right. One can say: "this fact happened." And, mark well, it does not have to be an elementary fact. Let us consider, for example, the possible fact that in a given election one of the parties obtained a percentage of votes included in a certain interval. This is a state of affairs that depends on millions of elementary cases (the single votes).

I shall start by saying what I mean by 'event.' Well, the characteristic feature of what I refer to as an *event* is that the circumstances under which the event will turn out to be "verified" or "disproved" have been fixed in advance. I call instead *facts* those circumstances that verify or disprove an event. If I said, before knowing the event: "I believe that the percentage of votes obtained by party A will lie between 22% and 28%," then this statement could be verified by many distinct facts.

ALPHA: I mentioned, during the last lecture,² Wittgenstein's *Sachverhalten*: your *facts* seem to me to correspond to those, to the so-called *atomic facts* of Wittgenstein's.

DE FINETTI: As I said already, facts need not necessarily be *elementary*, because the fact, say, that the percentage of voters who go to the polling stations lies between 80% and 85% though not elementary, is indeed a fact.

ALPHA: In other words, an event is a *possible* fact, which can either exist or not.

DE FINETTI: What distinguishes 'fact' from 'event' is that by 'fact' I mean *something which exists independently of me*. Even if I did not think of that fact, it would still exist.³

ALPHA: I have the impression that there are metaphysical assumptions underpinning this distinction.

DE FINETTI: At any rate, a *fact* can either happen or not. An *event*, on the other hand, can correspond to a fact but differs from the fact it corresponds to because it depends on the way in which, before knowing that it happened, a person formulates a possible doubt about it. For example, a person could wonder: “what value is a given percentage going to take?” Yet there are many ways of formulating the question. One could ask, for instance, whether such a percentage will lie between certain bounds, say between 30% and 50%. In this case, the event could be represented by many distinct facts (because these could be all the values lying between 30% and 50%). That which actually happens is a fact; the event, on the other hand, is the answer to a question raised at a previous time. This distinction is also relevant to the concept of conditional probability. Indeed, if one stipulates a conditional bet, one must specify exactly the conditions under which the bet will be called off. When writing $P(E \mid H)$, what is actually meant by H is an *event*, not a *fact*.⁴ There can be very many facts corresponding to an event.

ALPHA: You say that there are very many facts in an event. My question is: are those facts decomposable in turn?

DE FINETTI: I do not think that facts bear a direct relation with events. I have been looking for different verbs to distinguish the two concepts. My proposal consists in saying that a *fact* happens (or takes place).

ALPHA: Savage uses the term *obtains**

DE FINETTI: Yes. Rather than being Savage’s own, this is a typical English expression.

ALPHA: I do not know. However, Savage says at some point (I quote from memory): “I failed to find a better verb than this one.”⁵

DE FINETTI: Then he probably wants to say that the English language offers various, more or less appropriate, words without there being one which fully satisfies him. It is the same situation in which I find myself now. The essential point is this: in order for something to be a *fact* it is irrelevant whether it has been taken into consideration or not. If a train accident happened, it would be a fact that has happened, but it would not be something which has been previously taken into consideration in order to determine whether it happened or not. On the other hand — as I said — I call *event* whatever is the object of an explicit question or curiosity. In other words, an event is something which has been previously figured out and subsequently checked in order to see whether it took place or not. Both facts and events should be distinguished from *phenomena* (a term which I adopted a long time ago, and to which I have never found a better alternative). ‘Phenomenon’ is a generic name for events of a specific kind. For example, the draws from an urn are *phenomena*. The meaning of the term ‘phenomenon’ does not have, in itself, any relation with probability. I introduced the term ‘phenomenon’ in order to replace the expression ‘trial of an event’ which I do not like, as it presupposes that events are repeatable, rather than well-determined single cases. Therefore, instead of speaking of *trials of a certain event*, I prefer to

* (Translator’s note:) In English in the original text.

speak of *trials of the same phenomenon*. Yet the way I use the term 'phenomenon' does not coincide with the way frequentists use the term 'event.' Indeed, to say that a class of events constitutes a class of *trials of a certain phenomenon*, I do not require (unless otherwise stated) the satisfaction of any specific condition. In particular, I do not presuppose that such trials are equally probable.

ALPHA: If I understand correctly, the single trials of a phenomenon are themselves *facts* and possibly also *events*.

DE FINETTI: Yes, either facts or events, which nonetheless have something in common.

ALPHA: Some properties.

DE FINETTI: Yes, but expressible in plain language and without presupposing anything concerning the evaluations of probability. By symmetry we could say that a fact *happens*, that an event *is verified* and that a phenomenon . . .

ALPHA: An appropriate expression would be 'presents itself.' Another one is 'appears.' Yet another one is 'manifests itself.' But perhaps the most appropriate is 'repeats itself.' It seems to me that in fact, the characteristic feature of a phenomenon is that it can repeat itself.

BETA: In my opinion, by saying *repeats itself* we impose a severe restriction on the meaning of the term 'phenomenon.' In fact the trials of a phenomenon are not exact repetitions of the same phenomenon: there is always some difference between any two repetitions.

DE FINETTI: I propose to say that a phenomenon *realizes*.⁶ I would like to know your opinion concerning this proposal. There are pros and cons, without the balance being sensitive enough to tilt decidedly in favour of either side. However, I have set up a very simple example to clarify these concepts. It involves a sequence of ten tosses of a coin. I wrote down something, which I shall now read to you:

It is a *fact* that Heads came up on the seventh shot. This could be transformed into an event by making it the object of a bet which is going to be won or lost according to whether that fact is verified or not. To clarify the concept within the scope of this simple example, let us think of some other distinct events that could be realized by the same fact. We can take as an event the fact of obtaining 'heads' on the seventh shot.

Therefore the event Heads on the seventh shot *is verified* when the *fact* that heads comes up on the seventh shot happens.

ALPHA: A further difference between 'event' and 'fact' is as follows: whilst events admit of being negated, facts do not. The negation of a fact as such does not exist, if we distinguish between fact and event.

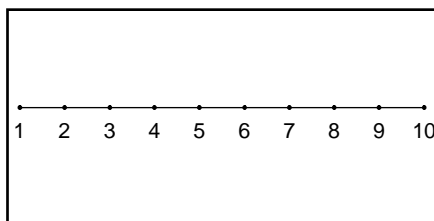
DE FINETTI: Yes, I agree.

BETA: There is also a further difference: an event might correspond to many facts. And the same fact might be involved in a multiplicity of events.

On Elementary Facts

DE FINETTI: This is exactly what I wanted to show with the examples of the sequence of ten tosses of a coin that I set up (Fig. 15.1). I have already mentioned the first example: (a) obtaining Heads on the seventh shot. The remaining examples are:

Fig. 15.1 Graphical representation of a sequence of ten tosses of a coin



(b) obtaining Heads on at least four out of the ten shots; (c) obtaining Heads in at most four shots; (d) obtaining Heads in one shot. More precisely: (a) is verified only on the seventh shot. After any of the first six shots, we do not know if the event is verified or not. Once we get to the seventh shot, it is possible to decide whether the event is verified or not according to whether the outcome is Heads or Tails, the successive shots being irrelevant. Thus, as far as this event is concerned, the only relevant shot is the seventh. On the contrary, in the example (b) the realization of four specific results Heads is one of the *facts* that verify this *event*.

ALPHA: Why is (b) is not a fact whereas (a) can also be seen as a fact? Maybe because (a) is not decomposable? If so, then (b) should be characterized as the disjunction of many facts (better: the disjunction of the intersection of facts). In other words, as “the first *and* the second *and* the third *and* the fourth *or* the first *and* the second *and* the third *and* the fifth, and so on.” Therefore (a) could be called an *atomic fact*. In other words, non decomposable and therefore in a certain sense, elementary.

DE FINETTI: Fine, it might well be elementary but it can be combined as much as one likes.

ALPHA: Of course: otherwise it would not be elementary.

DE FINETTI: Not only could one add further shots to the sequence of ten shots, but one could also combine it, for instance, with the fact *it is going to rain tomorrow* or *it is not going to rain tomorrow*. In the context of this problem, the possible cases are the 2^{10} sequences of Heads and Tails. Then one could compare this situation with the case of a single draw from an urn containing exactly $2^{10} = 1024$ balls. And just as in this situation, the elementary cases would consist in the individual draws of the 1024 balls, in the situation of the ten tosses of a coin, one could take as elementary the 1024 possible sequences that specify the possible outcomes of the ten shots.

ALPHA: Elementary cases must have the property of logical independence, whereas the facts you mention are not logically independent. Indeed, the 1024 sequences are logically mutually exclusive. On the contrary, two events of the kind *Heads on the seventh shot* and *Heads on the fourth shot* are not logically exclusive, nor do they imply one another. It seems to me that this suggests that facts, the way you intend them, have the characteristic feature of logical independence.

DE FINETTI: I cannot see why one should take logically independent facts as elementary facts. In the case of drawings from an urn, for example, it seems natural to take as elementary facts the draws of the single balls instead which, as you say, are not logically independent. Yet from the point of view of the logical representation,

there is no difference between tossing a coin ten times and drawing a ticket from an urn containing 1024 tickets, each one of them carrying a distinct sequence of results among the 1024 possible ones.⁷

ALPHA: I am afraid I do not understand. Do you take an individual sequence to be a *fact* or an *event*?

DE FINETTI: The difference between fact and event is that whatever happens in "the external world" is a fact. There is no need to evaluate the probability of *all* facts: usually, it is sufficient to restrict oneself to the probability evaluation of some of them. And had one bet on the event *Head on the seventh shot*, then the results of the first six shots would all be irrelevant to the end of deciding whether the bet has been won or not.

ALPHA: But is (b) an event, or is it also a fact?

DE FINETTI: It depends. That Heads came out on the seventh shot *independently of fixing the event in which I am interested* is a fact as any other. Yet it becomes an event if . . .

ALPHA: [Interrupting] If one places a bet on it.

DE FINETTI: Yes, but it is not necessary to actually place a bet: the point is rather that one takes that fact into consideration in order to give it a probability value. In this case, it will be necessary to define precisely the conditions that enable one to decide whether the event is verified or not.

BETA: In essence: that a fact happens is a matter apart, although the occurrence of that fact can coincide, at the same time, with an event being verified. The fact is objective.

DE FINETTI: It seems to me that you expressed what I have in mind quite clearly. In short: an event is something whose truth conditions have been carefully specified before carrying out the experiment. An event is — unlike a fact — open to be either in agreement or disagreement with the state of affairs.

BETA: Thus, even *four Heads out of ten shots* could be considered both as a fact and as an event. It is an *event* whenever one asks oneself what is the probability that it will be verified (taking into account the features of the coin, and so on). It is instead a *fact* that after ten tosses the coin landed Heads four times.

DE FINETTI: Let us take the event (b).⁸ It is a fact, for instance, that the first three results are all Heads. This fact can be used to verify (b). After those initial shots (b) is verified as soon as the coin lands on Heads. Conversely, if out of the first nine shots only two resulted in Heads, the event would be disproved.

BETA: The fact, when speaking about the trials of an event, is the outcome of the trial.

DE FINETTI: Yes, but the event can take place on the basis of the results of ten shots and therefore, can have 1024 possible results. It would be a *fact* if 2 Heads and 8 Tails came out. If we considered in advance the conditions of its verification, this would be an event. I acknowledge that on the face of it, the distinction between fact and event might appear rather abstract. But this is not the case. By *event* I mean something with the property that all the circumstances that enable one to decide without ambiguity whether it is verified or not have been fully specified. But in

order to check whether a certain event is verified or not, it is necessary to know that certain facts have happened.

ALPHA: Let us say, for example, that the proposition describing the fact *Heads on the seventh shot* together with the propositions describing the corresponding facts *Heads on the first shot*, *Heads on the second shot* and *Heads on the sixth shot* logically implies the truth of the proposition describing (b).

DE FINETTI: Yes, sure, it is correct. Indeed I just wanted to say that the fact of obtaining Heads on the seventh shot can settle the question.

ALPHA: But then one could always put it in terms of logical implication. One could in fact say that the description of that event logically implies the proposition describing a fact relative to the event (b).

DE FINETTI: This is a correct formulation of what I wanted to say. In my opinion, what seems to be conceptually and psychologically more relevant, and the thing about which I would like to hear your opinion, is that what happens is not an event unless one has previously posed the question of its happening or not.

ALPHA: But then by 'event' we mean a *question* that has been posed. Then I would like to ask: is the interrogative aspect a fundamental feature?

DE FINETTI: I cannot see why this should be the case. By putting it in affirmative terms, one obtains a proposition that could turn out to be either true or false. By putting it in interrogative terms, one gets the same thing, with the only difference that in place of truth and falsity, one would have the two horns of a dilemma.

BETA: And what is, in this situation, the phenomenon?

Events and "Phenomena"

DE FINETTI: In our example the phenomenon consists in the series of trials involving the toss of a coin or, if we want to be more precise, in the sequence of trials involving ten tosses of a coin. The point here is to recover just what could be useful from the frequentist idea of *repeated trials* (though in my own view, the expression 'trial of a phenomenon' also stresses too much the circumstance that the ten trials are of the *same* phenomenon). They are ten experiments. Even if one changed the coin every time, or added to the series the event *even at roulette* and so on, one could still speak of trials of the same phenomenon.

ALPHA: The phenomenon is a property that is common to certain facts.

DE FINETTI: Yes. 'Phenomenon' is a generic term to say that its own "trials" are in some sense analogous. Should that be appropriate, one could explain what the analogy consist in.

I set up other examples. Unfortunately, it is rather hard to find a fully satisfactory one because they tend to be either trivial, or too complicated. It is probable, after some reflection, that two or three more interesting ones will come up.

ALPHA: Going back to the notion of a fact, I would like to ask: is the outcome of two successive shots also a fact?

DE FINETTI: Yes, of course.

BETA: Whichever manifestly visible thing is a fact.

ALPHA: Is it fact to say: “the sun shines *and* the second shot landed Heads?”

DE FINETTI: This is a fact as well.

ALPHA: Let us now turn to events: do they have to possess the characteristic feature of verifiability?

DE FINETTI: Sure. Then let us say that a phenomenon realizes, an event is verified and a fact happens, or takes place. Do you believe that there could be more appropriate words? I wrote down half a page to clarify the distinction. I have rewritten it many times because I could not manage to find a satisfactory solution.

BETA: I have some doubts about the adequacy of the verb ‘to realize’ in reference to phenomena.

ALPHA: Maybe it would be more appropriate to say that a phenomenon *manifests itself*.

BETA: I have not grasped the concept of phenomenon very well. In the previous example the phenomenon was: tossing a coin ten times. How can the phenomenon realize or not? I cannot quite see what it means to say that *the phenomenon realized*.

ALPHA: The phenomenon manifests itself. Or appears.

BETA: How do they say in English?

DE FINETTI: I guess they say *obtains**.

ALPHA: Also for phenomena?

DE FINETTI: I am not sure. At any rate it is not appropriate to base ourselves on the English terminology, as the vast majority of the works written in English belong to the frequentist approach. It is better not to use the same terms used by the frequentists because otherwise, we would run the risk that those terms, though used in a different way, could be interpreted in a frequentist sense, thus generating ambiguities and misunderstandings.

Editor’s Notes

1. This discussion, stimulated by the one carried out in the immediately preceding lecture, induced de Finetti to modify his traditional distinction between *events* and *phenomena* by introducing a new class of objects: *facts*. At the bottom of the discussion lies a genuinely philosophical problem: is the object of probabilistic judgment *linguistic* or is it *extra-linguistic*?

During some private conversations that I had with him, de Finetti spoke in favour of the latter alternative, the argument being that also animals are bearers of expectations which, admitting of degrees, could be expressed as probability values. I replied that the laws of the calculus of probability refer directly to logical constants or their set-theoretic counterparts (negations, unions, intersections, and so forth) whose linguistic nature appeared to me beyond dispute.

It is noteworthy that de Finetti devoted a specific Section 2.2 of the entry *Probability* of the *Enciclopedia Einaudi*. The end of the section reads as follows:

The opportunity to make these terminological clarifications arose from the critical discussions on those issues which took place during a series of lectures at the Italian Institute for Advanced Mathematics (Rome, March–May 1979). (de Finetti, 1980, p. 1162)

* (Translator’s note:) In English in the original text.

De Finetti's distinction between "facts" and "events" (as introduced in the first Italian edition of the present book) attracted the attention of Glenn Shafer (2001) who commented on it in a lecture held in Italy.

2. See Chapter 14, p. 145.

3. It now seems clear to me that de Finetti is here defending a *pragmatic* view of events rather than a purely semantic one. So, according to de Finetti, only in a pragmatic context of decisions does the notion of event, as used in probability theory, make sense.

This standpoint is, in turn, in accordance with de Finetti's aversion to the measure-theoretic approach which "starts off with a preassigned, rigid and 'closed' scheme" (see de Finetti [1970] 1990b, p. 269 note) where "points" have an absolute meaning, while they should be pragmatically seen "as indicating the limit of subdivision beyond which it is not necessary to proceed (at a given 'moment'; i.e., with respect to the problems under consideration)" (ibid., p. 269).

4. I clearly remember that during the nineteenth lecture (which accidentally failed to be recorded) de Finetti claimed that the state of information H_0 is not an event, in the sense specified by him but a set of facts. I then objected that if H and H_0 are heterogeneous, it is hard to see how one could make their logical product, as it appears in the notation $\mathbf{P}(E \mid HH_0)$. His reply was that the distinction between facts and events is not to be taken as heterogeneity: after all — he said — events are also facts, though described more specifically.

5. This is in fact a distortion of Savage's thought. The original text reads: "It is important to be able to express the idea that a given event contains the true state among its elements. English usage seems to offer no alternative to the rather stuffy expression, 'the event obtains'" (Savage [1954] 1972, p. 10).

6. In contrast to this proposal, de Finetti says in the summary of the lecture that a phenomenon can repeat itself. This is not an oversight: indeed de Finetti eventually adopted the expression "the phenomenon repeats itself," as proved by Section 2.2 of the entry *Probability* of the *Enciclopedia Einaudi*, where de Finetti writes: "the 'flash of lightning' (in a broad sense, whenever and wherever that could be) is a *phenomenon* (that can repeat itself, always and in every place)" (de Finetti, 1980, p. 1162).

7. These observations confirm how de Finetti was fully aware of the conventional nature of the distinction between atomic (or elementary) propositions on the one hand, and molecular (or compound) propositions on the other. It was in fact very clear to him that equivalent languages exist such that the same proposition can appear to be atomic in one and molecular in the other.

Unfortunately, nowadays, despite the fact that no one is willing to defend logical atomism, theories of logical independence, of content, and so on, which depend essentially on the distinction between atomic and molecular propositions keep being proposed, giving rise in this way to theorizations which depend essentially on a linguistic convention.

This point was stressed, in the Anglo-American philosophical literature, by David Miller (1974a, 1977), but it was already underlined by de Finetti: "[i]f we do not choose to ignore the way in which [the field] \mathcal{S} has been derived from the basis \mathcal{B} , the possibility arises that we could single out certain events as being somewhat *special*: for example, belonging to the basis or logically expressible in terms of a finite or countable number of basis elements" (de Finetti [1970] 1990b, p. 271).

8. Recall that the event (b) is *obtaining Heads on at least four out of the ten shots*.

Chapter 16

“Facts” and “Events”: An Example*

Elaboration of a concrete example to distinguish among the three concepts discussed during the previous lecture. Analysis of which “fact” determines the truth or falsity of an event (if it is verified or disproved). Illustrations on the results of 12 successive tosses of a coin.

A Sequence of Coin Tosses

Today we will try to work out together an example to illustrate the distinction upon which we insisted during the last lecture, between *facts* and *events*. I thought of taking into consideration a sequence of 12 tosses of a coin and some events that are definable in this situation, for instance, the event *four consecutive Heads*.

BETA: I would immediately observe that, in that situation, it is only after the ninth shot that one can be certain that the event is not verified.

DE FINETTI: That was just one of the possible examples: there are many others that could be considered. For instance *at least h successes*, or *an even number of successes*, or else *all the outcomes Heads separated by at least one Tails*. Again, these are very simple examples, yet I would like to find some characteristic examples to distinguish the various cases. It is interesting to study those examples in which until a certain shot takes place, it is undecided whether the event is verified or not, though at each shot one can say, on the basis of the previous outcomes, which are the facts that would verify or, on the contrary, disprove the event under consideration.

BETA: Hence, is the point not here to show how the truth-value of many events can be decided as a consequence of one and the same fact happening?

DE FINETTI: The point is, given an arbitrary sequence, to determine first of all whether the first *h* shots are sufficient to decide the event under consideration. Thus, were the event *E* in question *at least four Heads in any order*, as soon as the fourth ‘heads’ comes up, one would know that *E* is verified. More generally, the idea here is to determine, given the result of *h* tosses of the coin, which are the results that can either verify or disprove the event.

BETA: We must distinguish between necessary conditions and sufficient conditions. For instance, if no sequence of four consecutive Heads appears among the first eight tosses, it will be *necessary*, in order for the event *four consecutive Heads* to be

* Lecture XVIII (Wednesday 9 May, 1979).

verified, that Heads comes out on the ninth shot. After three consecutive Heads on the other hand, it will be *sufficient* that the following shot comes up Heads as well.

DE FINETTI: For some events it is necessary to wait until the very last shot in order to decide their truth-value. One example of this is *the number of shots resulting in Heads is even*. Here, the last shot is a priori decisive, independently of the results of the previous shots. I would like to develop an interesting example in full detail. I tried it out a little bit at home but after struggling for half an hour, I became irritated. With your help, I would like to resume and carry on the work I have already done.

In this context, an event E can be identified with the class S of the k sequences that verify it, whereas its negation \tilde{E} , can be identified with the class S' of the remaining $2^{12} - k$ sequences. Thus, for each event taken into consideration, we have a partition into two classes of the set of the 2^{12} sequences of twelve tosses of a coin. To know whether the truth value of E can be decided after a given number h of tosses, one must determine whether it is possible, after those h shots, to establish with certainty to which of the two classes S and S' the sequence of twelve shots will belong. If that were the case, one could say that after those shots, the event is *decided*, no matter what happens afterwards.

ALPHA: Is the idea here to establish, for every event, the number of shots after which its truth-value is determined?

DE FINETTI: Among other things. Keep in mind that the number of shots required to decide the truth-value of a given event E varies according to the outcome of the first shots. Such a number is therefore a random quantity.

BETA: If I understand correctly, the problem is to identify, for every event E : (a) the distinct ways in which E can be verified; (b) the k sequences that realize it and (c) the final step deciding its truth-value in different cases.

DE FINETTI: The problem here is just to work out a framework to illustrate the point. To make the example more effective, it is better to assume that events are temporarily ordered. This latter assumption, however, is not necessary to describe a situation involving the tosses of a coin. One could imagine, for instance, that the tosses were performed by two distinct persons, instead of being labeled by an ordinal number (the first, the second, and so on), so that every shot could be identified on the basis of the person who made it. In this case one would say, for instance, “Anna’s shot,” “Alberto’s shot,” and so on.¹ Yet, as regards the examples that I have taken into consideration like *at least four consecutive Heads*, the order is necessary: indeed it makes sense to speak of consecutive shots only in the presence of an ordering. Let us now take this other event: *a sequence of at least six alternating shots*.

BETA: There are seven possible sequences starting with Heads that verify this event. And there are just as many starting with Tails that verify it. Therefore, the number of sequences that realize it are $7 \times 2 = 14$.

DE FINETTI: That is not the case. Indeed, there are only 6 out of 12 alternating shots. And if there are 6 alternating shots, there are 2^6 possible combinations for the others.²

BETA: Yes, but the possible *subsequences* of the six shots verifying the event are fourteen in total. In fact, the six alternating shots can begin with the first shot but

they can also begin with the second, the third, the fourth, the fifth, the sixth and the seventh. There are therefore seven initial points for the subsequence. And since the first shot can either be Heads or Tails, there are — as I said — fourteen subsequences satisfying the event.

EPSILON: If a characterizing condition exists, which is a condition which allows one to say whether a certain combination is realized or not, then I guess it would be useful to represent it graphically. For instance, one could associate a distance or a thickness, to every event. In this way, whenever final the distance (or the thickness) reached a certain value, the truth-value of the event under consideration would be determined.

A Graphical Representation

DE FINETTI: Yes, I have already in mind the following graphical representation (Fig. 16.1). Every step upward in the diagram corresponds to the coin landing on Heads, whereas every step downward corresponds to Tails. The ordinate is therefore identical with the difference between the number of Heads and that of Tails. Thus, it is going to be positive, null or negative, according to whether the number of Heads is greater, equal or smaller than that of Tails, respectively. This representation allows us to reduce the study of a problem to the study of the properties of the segments representing the relevant sequences. Our goal is not that of discovering new examples, rather that of isolating an example, permitting us to highlight the nature of the process.

BETA: Whenever six alternating shots occur, then there must be at least one line of the diagram made up of three consecutive triangles.

DE FINETTI: In order to be alternating they should be exactly as the first six shots shown in Fig. 16.1. Another interesting example consists in imposing the condition that two and only two equilibria (that is to say equal number of Heads and Tails) should occur: one by crossing and the other as a corner (Fig. 16.2). This investigation should be taken further, but I do not have enough patience to do it. Therefore I address you. At your age, I would have done it, for fun, in five minutes. I suggest you try out many examples so that we can decide later which ones might be worth including, in terms of idea and in terms of picture, in this paper.³

In order to facilitate your task, I would like to give you, as a hint, a list of examples I have been thinking of:

1. *at least 5 Heads;*
2. *no more than 2 consecutive shots, both either Heads or Tails;*
3. *exactly 5 consecutive Heads;*
4. *at least 5 consecutive Heads;*
5. *at least 4 Heads followed immediately by Tails.*

Fig. 16.1 Graphical representation of the game Heads-Tails

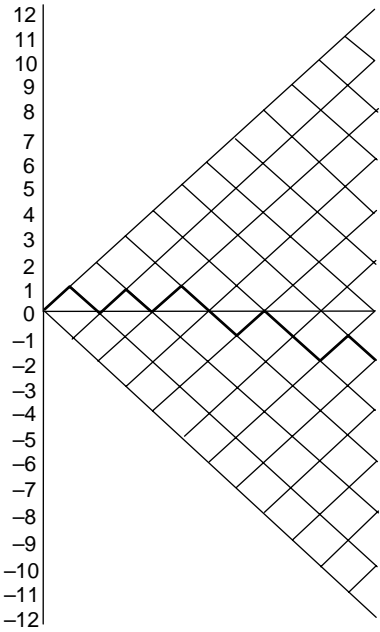
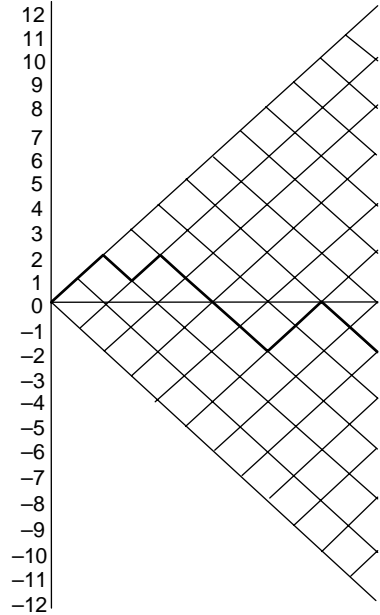


Fig. 16.2 An example of the game Heads-Tails with the occurrence of exactly two equilibria



The idea here is to find out, given any of the above conditions, when it is that the event becomes determined, so that the information about what happened after that point becomes redundant.

Editor's Notes

1. This remark also shows that exchangeability does not presuppose an underlying ordered set (or sequence). In fact, exchangeability may be defined in a general way as probability invariance with respect to any permutation of events, instead of order irrelevance. The last definition is obviously equivalent to the former whenever the events are identified by their ordinal position in a sequence. However, it does not make sense with respect to an unordered set of events, while the former is always meaningful.
2. In the discussion of this example de Finetti assumes throughout that the first event of the alternating sequence is Heads.
3. The paper in question is the entry "Probabilità" for the *Enciclopedia Einaudi* (de Finetti, 1980). However, no such example has actually been included in that work.

Chapter 17

Prevision, Random Quantities, and Trievents*

The theory of probability as based on the concept of “coherence,” in the sense of forbidding combinations of bets that lead to sure loss.¹

Probability as a Special Case of Prevision

Today we shall go back to the interpretation of probability and prevision as *price*. Let me start by recalling that the simplest way — though some might not consider it to be noble enough — to characterize both probability and prevision consists in thinking of them as the *price* of a bet. The concept of prevision is more general than that of probability. In fact, if we think of an event E as that random quantity whose only values are 1 and 0 (according to whether E is true or false), the probability of E coincides with the prevision of E .

In general, the prevision $\mathbf{P}(X)$ of a random quantity X , is the fair price of the random amount X , according to the evaluation of a person who is uncertain about the value that X will actually take. X can either be a discrete or a continuous quantity. If X is an event E , the prevision (or probability) of E is the price of the offer of E , that is to say of the offer:

$$\begin{cases} 1 & \text{if } E \text{ is verified,} \\ 0 & \text{if } \tilde{E} \text{ is verified.} \end{cases} \quad (17.1)$$

Whenever prices are expressed through monetary amounts (rather than utility values), this definition must be taken cautiously because the *linearity* of the prices is always assumed (which, if on the one hand always holds for utility values, on the other hand need not be true for monetary amounts).

The above definition refers to *absolute* random quantities and events, that is to say non *conditional*. Were X conditional on H , $\mathbf{P}(X \mid H)$ would be the fair price p of the offer of:

* Lecture XX (Tuesday 15 May, 1979).

$$\begin{cases} X & \text{if } EH \text{ is verified,} \\ 0 & \text{if } \tilde{E}H \text{ is verified} \\ p & \text{if } \tilde{H} \text{ is verified.} \end{cases} \quad (17.2)$$

In particular if X is an event E , then the prevision (or probability) of $\mathbf{P}(E \mid H)$ is given by the price p of the offer:

$$\begin{cases} 1 & \text{if } EH \text{ is verified,} \\ 0 & \text{if } \tilde{E}H \text{ is verified} \\ p & \text{if } \tilde{H} \text{ is verified.} \end{cases} \quad (17.3)$$

Since probability can be seen as a special case of prevision, it is possible to take the concept of prevision as primitive and that of probability as derived. Usually, the opposite happens. More precisely, the concept of probability is taken as primitive and the concept of prevision $\mathbf{P}(X)$ of a random quantity X is defined as the mean of the possible values x_1, x_2, \dots, x_n of X weighted by the corresponding probabilities p_1, p_2, \dots, p_n :

$$\mathbf{P}(X) = \sum_{i=1}^n x_i p_i. \quad (17.4)$$

It is however preferable to derive the relation (17.4) from the definition of $\mathbf{P}(X)$ as fair price. This is obtained as follows. For all i ($1 \leq i \leq n$), let E_i be the event $X = x_i$. Then the following identity holds:

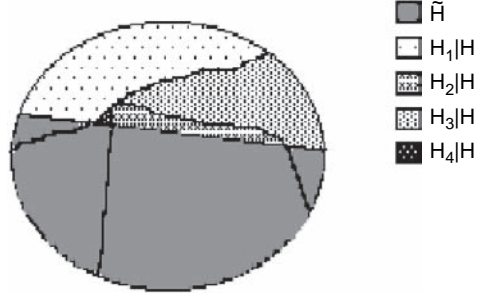
$$X = \sum_{i=1}^n x_i E_i, \quad (17.5)$$

from which, by the linearity of prices, we can deduce:

$$\mathbf{P}(X) = \mathbf{P}\left(\sum_{i=1}^n x_i E_i\right) = \sum_{i=1}^n \mathbf{P}(x_i E_i) = \sum_{i=1}^n x_i \mathbf{P}(E_i) = \sum_{i=1}^n x_i p_i. \quad (17.6)$$

Let us now take an arbitrary finite partition of events H_1, H_2, \dots, H_n and suppose that we are to assign to each element of the partition H_i , a probability $\mathbf{P}(H_i)$. To illustrate the point we can represent each H_i as a portion of a unit area. Were we to consider a condition H , how would the values of $\mathbf{P}(H_i)$ be linked with the values of $\mathbf{P}(H_i \mid H)$? The answer is very easy: it will be sufficient to equate the area of H to 1 and then consider, for each H , the portion $H_i \mid H$ of such an area. This will correspond to the probability $\mathbf{P}(H_i \mid H)$ (Fig. 17.1).²

Fig. 17.1 Diagram illustrating probabilistic conditioning as a normalization of probabilities



The Conglomerative Property

Let us now have a look at some general properties of the notion of *prevision*. First of all, let us observe that again by the linearity of the prices, the prevision of the sum of random quantities equals the sum of the respective previsions:

$$\mathbf{P}(X + Y) = \mathbf{P}(X) + \mathbf{P}(Y). \quad (17.7)$$

Note that the relation (17.7) holds for arbitrary random quantities X and Y , whether X and Y are independent or not.

Let us now take X to be conditional on a finite partition of hypotheses \mathcal{H} (we shall write “ $X \mid \mathcal{H}$ ”). A partition is — as is well-known — a class \mathcal{H} of events H_1, H_2, \dots, H_n , which satisfies the following properties:

$$\sum_{i=1}^n H_i = 1 \quad (17.8)$$

and

$$H_i H_j = 0 \quad (1 \leq i \leq j \leq n). \quad (17.9)$$

By equations (17.8) and (17.9), the following holds:

$$X = H_1(X \mid H_1) + \dots + H_n(X \mid H_n), \quad (17.10)$$

whence

$$\mathbf{P}(X) = \mathbf{P}(H_1)\mathbf{P}(X \mid H_1) + \dots + \mathbf{P}(H_n)\mathbf{P}(X \mid H_n). \quad (17.11)$$

Let us observe that the relations (17.5) and (17.6) are special cases of the relations (17.10) and (17.11), respectively. Since the probability of the H_i ’s is 1, equation (17.11) shows that $\mathbf{P}(X)$ is the weighted mean of the $\mathbf{P}(X \mid H_i)$ ’s, with relative weights $\mathbf{P}(H_i)$. As to events, it is enough to replace X with E . In this way we obtain, as a special case, the important identity:³

$$\boxed{\mathbf{P}(E) = \mathbf{P}(H_1)\mathbf{P}(E \mid H_1) + \dots + \mathbf{P}(H_n)\mathbf{P}(E \mid H_n)} \quad (17.12)$$

Infinite partitions could also be considered: in this case the weighted mean will have to be replaced by an integral. In some cases this will result in a probability distribution endowed with density:

$$\int_{-\infty}^{+\infty} x f(x) dx.$$

However, it must be recalled that in order for the density to exist, the cumulative distribution function $F(x)$ must be differentiable. We have seen when we discussed distributions, that in the general case we can distinguish, given an arbitrary probability density function, between a component which is endowed with density (if it exists) and a component constituted of n ($1 \leq n \leq \infty$) concentrated masses, whose total mass is:

$$\sum_{i=1}^n m_i x_i.$$

If n is ∞ then the series will converge. It is worth recalling that there is also an intermediate case, namely the case illustrated by Cantor's distribution.⁴ However, if one adopts Stieltjes' integral, integration is always possible. Therefore, one single integral is sufficient to measure all the components of a distribution:

$$\int_{-\infty}^{+\infty} x dF(x).$$

The above considerations show how the interpretation of the probability and prevision of random quantities as price makes all the key properties of probability and prevision immediately clear.

Let us turn to the relation holding between the probability of an intersection of events and the probability of the individual events. Let us take, as our example, a sequence of tosses of a coin. The probability that a single toss will land on 'heads' is $1/2$, whereas the probability that n Heads will obtain equals 2^{-n} . This follows from the definition of mutual independence. Indeed, if n events are mutually independent, the probability that all of them will be simultaneously verified is the product of the single probabilities:

$$\mathbf{P}(E_1, E_2, \dots, E_n) = \mathbf{P}(E_1)\mathbf{P}(E_2) \dots \mathbf{P}(E_n),$$

and if $\mathbf{P}(E_i) = \frac{1}{2}$ for all i ($1 \leq i \leq n$), the probability of $\mathbf{P}(E_1, E_2, \dots, E_n)$ is going to be 2^{-n} .

We have characterized the probability conditional on a hypothesis H as the price of a bet which would be called off should H fail to be satisfied. Drawing on this

concept, it is possible to characterize, in a general way, the probability of an intersection of events in terms of the probabilities of the single events (without requiring mutual independence among the events). For instance, given four events A, B, C, D the following relation holds:

$$\mathbf{P}(ABCD) = \mathbf{P}(A)\mathbf{P}(B | A)\mathbf{P}(C | AB)\mathbf{P}(D | ABC). \quad (17.13)$$

Of course, in the general case of n events $E_1, E_2, \dots E_n$, in place of equation (17.13) we would have the formula:

$$\mathbf{P}(E_1 E_2 \dots E_n) = \mathbf{P}(E_1)\mathbf{P}(E_2 | E_1) \dots \mathbf{P}(E_n | E_1 E_2 \dots E_{n-1}). \quad (17.14)$$

Trievents

To conclude I shall discuss a critical problem concerning the concept of *conditional event*. What does the expression “ $B | A$ ” exactly denote? On some occasions I have used the term “trievent” to suggest that the usual two-valued logic is not sufficient to characterize it adequately. The following are the possible cases:

- a) $A = 0$ and $B = 0 \quad B | A = \emptyset$
- b) $A = 0$ and $B = 1 \quad B | A = \emptyset$
- c) $A = 1$ and $B = 0 \quad B | A = 0$
- d) $A = 1$ and $B = 1 \quad B | A = 1$

On the basis of this table, if the hypothesis A is not verified, then $B | A$ is not informative and its value is *null*, which I denote by the symbol \emptyset . Under the hypothesis that A is verified, $B | A$ becomes equivalent to B .⁵

BETA: $1 | 0$ and $0 | 0$ therefore have no meaning.

DE FINETTI: They mean “null.”

ALPHA: Is \emptyset understood as a truth-value?

DE FINETTI: It indicates that the hypothesis under which the event A is being considered ceases to hold. The event B is taken into account only if A is verified.

BETA: Does it not simply mean, for instance, that if the horses fail to start, one cannot say that the horse A loses the race, for none turned out to be the winner?

DE FINETTI: Indeed: the bet would be called off. There are, therefore, *three* cases.

ALPHA: If I understand correctly, \emptyset is a truth-value. In your opinion there are three truth-values: 0, 1 and \emptyset . One could have written ‘ $1/2$ ’ instead of \emptyset .

DE FINETTI: ‘ $1/2$ ’ is not appropriate because it somewhat suggests that it is an intermediate value between true and false.⁶

ALPHA: It seems to me that the truth-value \emptyset corresponds to the one Aristotle attributes to the so-called “contingent futures.”⁷

DE FINETTI: I have to say I am not particularly strong in this subject.

ALPHA: Aristotle says that an event is true if it is verified and false if it is not. Yet, for Aristotle, a sentence about the future is neither true nor false: rather it has another truth-value.⁸

DE FINETTI: If it is something that will either take place or not, then I would say that the truth-value is either true or false, nothing else. It exists, but it is unknown.

ALPHA: Aristotle believed that everything that happens, happens because it is necessary for it to happen.⁹

DE FINETTI: I would rather avoid the analysis of the term 'necessary' because it seems too philosophical a question to me.

ALPHA: 'Necessary' means this: that it is not possible that it might not happen. Yet there are several kinds of possibility and hence, of necessity. There is logical necessity but there is also physical necessity. If one believes that everything is predetermined, then everything that happens is necessary.

DE FINETTI: The fact that events might be predetermined or not is irrelevant unless we *know* in which way they are predetermined.

ALPHA: Of course, from the point of view of our knowledge.

DE FINETTI: It is like when one reads a book. One knows the story up to the point that has been reached by reading, yet one has to wait until the end of the book to know its conclusion.

ALPHA: There are, however, many-valued logics.

DE FINETTI: The introduction of the truth-value \emptyset is needed in order to distinguish between reference to an event as a *statement* and reference to an event as a *condition*. Whenever the *condition* B is satisfied, then $A \mid B$ is either true or false (1 or 0). But unless the condition B is satisfied, one can neither say that the event $A \mid B$ is true, nor that the event $A \mid B$ is false. It is *void* or *null* in the sense that the premise under which it is considered either true or false no longer holds. In my opinion, these three cases should be treated as distinct.

ALPHA: It seems to me that \emptyset means *indeterminate*.

DE FINETTI: It does not even mean *indeterminate* because an indeterminate event is an event whose truth conditions are unknown.

ALPHA: Yes, you are right.

DE FINETTI: Before learning its truth-value, every event E is indeterminate, unless E is either a tautology or a contradiction (in which cases one would know *a priori* if E is true or false). Yet if one makes a statement conditional on a hypothesis H , the truth of H is a necessary condition for such a statement to make any sense.

ALPHA: What you are saying is that the meaning of a conditional statement depends on the truth of some other proposition. More precisely: it does not have a well-defined truth-value unless the proposition on which it is conditional is true. In my opinion, this point of view is analogous to the one you maintained about definite descriptions.¹⁰ Your point of view can be rendered as follows: the proposition "the present king of France is bald," can be true or false only conditionally on the hypothesis that there exists, at present, a king of France.

DE FINETTI: You should explain this to the others, as they were not present when we discussed it.

ALPHA: At the beginning of the Twentieth century, Bertrand Russell raised the question concerning the meaning of propositions like “the present king of France is bald.” The problem here is related to the meaning of the so-called *definite descriptions*.¹¹ Not all definite descriptions are of this form. Yet this is the paradigmatic case. By the excluded middle principle, one could argue as follows: “there are two possible cases: either the present king of France is bald or he is not.” However, none of the two possibilities can be true if there exists no present king of France. How to solve this paradox? Let us suppose that a Martian who has just landed on Earth were told: “the present king of France is bald.” Suppose that after acquiring this piece of information the Martian went to France and discovered that there is no such person as the king of France. The following problem would then arise: has he been told a proposition that has no truth-value or has he been told lie? In the former one admits that the truth-value (true or false) of the proposition exists only provided that there exists a present king of France and is null if the present king of France does not exist. Russell found another solution. According to Russell, when saying: “the present king of France is bald,” one actually makes three statements. In particular, one also states the *existence* of the present king of France. More precisely the sentence “the present king of France is bald” is equivalent, according to Russell, to the conjunction of the following three propositions: (a) “there exists at least one king of France”; (b) “there exists at most one king of France”; and (c) “there exists no present king of France who is not bald.” Therefore according to Russell the meaning of the description is not conditional: the proposition is always either true or false, as any other proposition. There are some authors, however, like Strawson,¹² who maintain — just like you — that should the king of France fail to exist, the proposition would not have a well-defined truth-value despite having a well-defined *meaning*.

BETA: Hence, according to Russell’s analysis, if the proposition “the king of France exists” were false, the proposition “the present king of France is bald” would also be false.

ALPHA: Exactly. This analysis was used by Russell as a reply to the point of view that considers definite descriptions as *nouns*. Should that be the case, they would always refer to something. But the referent of “the present king of France” cannot be a king in the flesh, with the consequence that purely ideal referents should be admitted.¹³ Russell considered this unacceptable.

BETA: On the other hand, there are propositions that do not make sense even in a two-valued logic.

ALPHA: For instance?

BETA: “Pythagoras’ theorem is green.” This is a statement that makes no sense. The one you mentioned earlier does make sense.

ALPHA: “Pythagoras’ theorem is green” makes no sense because we know *a priori* that the property “green” cannot be applied to a mathematical object.

Editor's Notes

1. As a matter of fact, the present lecture does not deal with coherence at all. It rather develops the idea that the properties of probability are deducible from the linear relationships between prices and events. Moreover, this chapter contains some interesting remarks about the use of a three-valued logic to formalize conditional events.
2. Cf. Chapter 4 on page 35.
3. De Finetti named the property captured by the relation (17.12) the *conglomerative property*. The importance of this property derives from the fact that it does not always hold for *infinite* partitions, unless the condition of finite additivity is strengthened. This in turn entails that finite additivity implies exposure to countable Dutch books (Seidenfeld and Schervish, 1983). A more general analysis shows that “Dutch book arguments have no force in infinite cases” (Arntzenius, Elga and Hawthorne, 2004). Were the axiom of countable additivity accepted, the conglomerative property would be satisfied for all countable partitions. Yet de Finetti rejected this axiom (see Chapters 12) and never endorsed the idea of infinite systems of bets.
4. See Chapter 9 on pages 90–92.
5. At the *Congress of Scientific Philosophy* held in Paris in 1935, de Finetti presented a paper (de Finetti [1936] 1995) — overlooked until it was translated into English in 1995 — in which he outlined a three-valued logic containing (beyond Łukasiewicz’s connectives for the negation ‘ \neg ,’ Sum ‘ $+$ ’ and product ‘ \cdot ’) two conditional connectives: ‘implication’ (whose truth-table coincides with the connective proposed by Kleene (1938) for material implication), and the conditioning ‘ $|$ ’ whose truth-table is presented in the text. Moreover, de Finetti introduced two two-valued unary connectives ‘ T ’ and ‘ H ’ whose truth tables are (using the same notation as above):

A	$T(A)$	$H(A)$
1	1	1
\emptyset	0	0
0	0	1

The connective ‘ T ’ coincides with Bochvar’s so-called *external assertion* “ A_* ” (Bochvar [1938] 1981). A set \mathcal{A} containing the ordered set of constants $< 0, \emptyset, 1 >$ with the connectives introduced by de Finetti is a *De Morgan algebra*, namely a bounded distributive lattice with an order inverting operation (the negation). Moreover, if the so-called “Kleene’s condition” $\neg A \mid A \leq A \mid A$ is added, one obtains a special case of three-valued Łukasiewicz algebras, characterized by the presence of the self-dual element ‘ \natural ,’ for which it holds that $\neg \natural = \natural$. Milne (2004), who proved this result, calls such a structure a *de Finetti algebra*.

Clearly, the subset \mathcal{B} of \mathcal{A} containing all two-valued “ordinary events” with negation, sum and product is a Boolean algebra. It is easy to show that $\mathcal{B} = \{X \in \mathcal{A} \mid X = T(X)\}$ or, equivalently, $\mathcal{B} = \{X \in \mathcal{A} \mid H(X) = 1\}$. De Finetti observed that every element X of \mathcal{A} may be “decomposed” as $B_1 \mid B_2$ where both B_1 and B_2 belong to \mathcal{B} . It suffices to note that for every $X \in \mathcal{A}$ it holds that $X = T(X) \mid H(X)$. As a consequence, we can always think of de Finetti’s trievents as *conditional events* without any need for considering the application of the truth-function ‘ $|$ ’ to pairs of ordinary events as a special case. Moreover, a probability function \mathbf{P} may be defined on $\mathcal{A} - \{X \mid \mathbf{P}(H(X)) = 0\}$ by extending a probability function defined on \mathcal{B} by the following rule:

$$\frac{\mathbf{P}(T(X) \wedge H(X))}{\mathbf{P}(H(X))} = \frac{\mathbf{P}(T(X))}{\mathbf{P}(H(X))}$$

(see Mura, 2008).

De finetti's ideas on trievents were rediscovered, as I shall explain briefly, after a long history of debates on conditional events and conditional assertions, motivated by work in computer science (especially in AI) on the one hand and philosophical research on the other. Philosophical research on conditionals is a classical topic. Moreover, the idea that a sentence may lack a truth value without being meaningless goes back at least to Frege ([1892] 1980) and, as mentioned in the text, the thesis that a sentence ' q presupposes a condition p ' may be logically explained by saying that q is neither true nor false if p is false was argued by P. F. Strawson (1952, p. 175). On the pragmatic (rather than the semantical) side Quine claimed in 1959 that "an affirmation of the form 'if p then q ' is commonly felt less as an affirmation of a conditional than as a conditional affirmation of the consequent. . . If . . . the antecedent turns out to have been false, our conditional affirmation is as if it had never being made." (Quine [1950] 2004, p. 12).

Fifteen years later, Ernest W. Adams, developing an idea that goes back to F. P. Ramsey ([1929] 1990, p. 247), outlined a theory of conditional assertions based on the idea that the conditionals are nothing but bearers of conditional probability (Adams, 1965, 1975). The degree of pragmatic assertability of a conditional "if p then q " is given by its probability value $\mathbf{P}(q \mid p)$. An argument involving conditional assertions as premises is valid if it preserves degrees of probability in the sense that for every probability function the improbability $1 - \mathbf{P}(q \mid p)$ of the conclusion is not lower than the the sum of the improbabilities of the premises.

If Probability is meant (unlike the original Adams' proposal) as *subjective* (as suggested by B. Skyrms (1984), this approach is not at odds with the spirit of de Finetti's philosophical views. However, it should be remarked that for de Finetti *all* factual propositions (and not only conditional propositions) are nothing but bearers of probability values, and that from his pragmatic viewpoint truth-conditions are indispensable in the context of decision-making (cf. Chapter 15 note 3 on page 157). In the betting situation, for example, truth-conditions are necessary to specify when a bet is won, lost or called off. Unfortunately, a theorem proved by McGee (1981) shows that no relation of logical consequence based on a finite many-valued logic, as it is currently characterized, including de Finetti's logic of conditional events, is compatible with the account provided by Adams.

Adams' seminal work had a large impact on the American philosophical community. An inquiry to provide truth-conditions for Adams' probabilistic conditionals was made by Robert Stalnaker (1975), by means of the so-called "semantics of possible worlds." Adams (1977) proved that his probabilistic validity is equivalent to Stalnaker's model based validity. Another attempt, based on a sharp distinction between *truth-conditions* and *assertability conditions* was made by Frank Jackson (1979).

Commenting on Adam's account, Lewis ([1976] 1986) noticed that several philosophers, including Richard C. Jeffrey (1964) and Brian Ellis (1969), maintained the view according to which degrees of assertability are always represented by degrees of probability. This view may be spelled out by saying that "probabilities of conditionals are conditional probabilities" (ibid., p. 134). If so, the problem arises of finding a connective ' \rightarrow ' (by which a new sentence $p \rightarrow q$ may be formed combining any two sentences p and q) such that $\mathbf{P}(p \rightarrow q) = \mathbf{P}(q \mid p)$. Lewis proved that except in trivial cases, there is no connective (let alone a truth-functional one) satisfying such a condition (see Chapter 4, note 2 on page 41).

Although many philosophers accepted the view that probability conditionals are not connectives (in the sense explained above), other scholars tried to face Lewis' challenge by relaxing the the assumption (which was an essential premise in Lewis' arguments) that the underlying algebraic structure to which conditional assertions belong is a Boolean algebra.

The need for a truth-functional theory of conditional assertions derived mainly from researches in the AI field, where the interest to overcome some limitations of the non truth-functional approach (especially its inability to address compound assertions involving conditionals) found a practical motivation, in particular the need to provide a formalization of the connections between conditional probabilities and the underlying production rules in expert

systems. This research programme, forerun by the previous work of Geza Schay (1968), was pursued by several scholars, including Philip G. Calabrese (1987, 1994), Didier Dubois and Henri Prade (1994), Irwing R. Goodman, Hung T. Nguyen and Elbert Walker (1991, 1995). These researches had some impact on philosophical discussions. Michael McDermott (1996) defended the truth-functional approach and Peter Milne (1997) pointed out that “Bruno de Finetti had laid the foundations for a unified treatment *in* 1935” (ibid., p. 212), that ultimately coincides with the Goodman-Nguyen-Walker approach and has been fully developed by Milne himself (2004).

However, although de Finetti’s approach to conditional events showed *ante litteram* how Lewis’ challenge may be solved by resorting to a larger and well defined structure, it is not free of difficulties. For example, in the standard calculus of probability defined over a Boolean algebra, if an assertion q is such that for every finitely additive probability function \mathbf{P} , it holds that $\mathbf{P}(q) = 1$, then q equals \top , so that it holds also, for any r , $\mathbf{P}(q \wedge r) = \mathbf{P}(\top \wedge r) = \mathbf{P}(r)$. This property is lost in de Finetti’s lattice of trievevents, with unexpected consequences. For example, for every probability function, it holds $\mathbf{P}(q | q) = 1$, while $\mathbf{P}((q | q) \wedge (r | \neg q)) = \mathbf{P}(\perp | \neg q \wedge \neg r) = 0$ even if $\mathbf{P}(r | \neg q) = 1$ (cf. Edgington, Spring 2006, pp. 18–9). This result is, of course, consistent with de Finetti’s interpretation of truth-conditions as conditions under which conditional bets are valid and of truth-functional connectives as providing the truth-conditions (in the above sense) of compound sentences. As a matter of fact, a bet on $(q | q) \wedge (r | \neg q)$ is, according to de Finetti’s truth tables, called off in all cases except when both q and r turn out to be false, in which case the bet is lost. Its “fair betting quotient” is thus 0.

6. What de Finetti means is clearly that ‘ \emptyset ’ cannot be interpreted as a “partial truth” or a “degree of truth.” From a purely formal viewpoint, however, ‘ \emptyset ’ actually *is* intermediate between ‘0’ and ‘1’ in de Finetti’s lattice of conditional events (see note 5 above). Indeed, in such a lattice it holds that $0 < \emptyset < 1$. This may be pragmatically justified in terms of bets. A bet on a trievevent $E|H$ is won if $E|H$ is true, called off if $E|H$ is void and is lost if $E|H$ is false. Now, in terms of payoffs, calling off a bet is clearly intermediate between winning and losing.
7. See Aristotle, *De Interpretatione* (Barnes, 1984), 9, 18^a28 – 19^b3.
8. Aristotle wanted to reply to the argument that if the truth-value of a future event is already decided in the present, then the future is determined and every rational deliberation would be useless. Łukasiewicz ([1957] 1987) was the first to introduce a three-valued logic to explain the Aristotelian argument.
9. Indeed, Aristotle does not make such a statement.
10. Allusion to a private conversation.
11. The main reference here is to the paper *On Denoting* (Russell, 1905).
12. “For a sentence of the statement-making type to have a meaning, it is not necessary that every use of it, at any time, at any place, should result in a true or false statement. It is enough that it should be possible to describe or imagine circumstances in which its use would result in a true or false statement. For a referring phrase to have a meaning, it is not necessary that on every occasion of its use there should be something to which it refers” (Strawson, 1952, p. 185).

In the private conversation mentioned in note 10, de Finetti maintained that a proposition (say p) like “the author of *Divine Comedy* was born in Florence” may be translated in a logically rigorous language by the conditional assertion “every author of *Divine Comedy* was born in Florence provided there exists one and only author of *Divine Comedy*.” In symbols, if ‘ A ’ stands for “author of *Divine Comedy*” and ‘ F ’ stands for “born in Florence,” p can be written:

$$\forall x(Ax \rightarrow Fx) \mid (\exists xAx \wedge \forall x\forall y(Ax \wedge Ay \rightarrow x = y)).$$

Clearly, if there is no author (or more than a single author) of *Divine Comedy*, the truth-value of such a sentence would be, according to de Finetti, null. According to Russell, it would be, by contrast, false, since the Russellian correct logical translation of p would be

$$\forall x(Ax \rightarrow Fx) \wedge (\exists x Ax) \wedge \forall x \forall y((Ax \wedge Ay) \rightarrow x = y).$$

Thus, according to de Finetti's view, a definite description is a conditional assertion in which the antecedent is a *presupposition* for the whole assertion being true or false. He maintained that his logic of conditional assertions captures the idea of presupposition in Strawson's sense as far as pure *logic* is concerned. The differences between presuppositions and conditional events are pragmatical (rather than logical) in character. They are related to the epistemic state of who makes the assertion and of her audience as well as to the context of use. Supposing that a conditional event is used in the paradigmatic context of betting (which is not, however, the only context in which conditional assertions are used) some differences appear to be obvious. For example, in the case of presuppositions the antecedent of a conditional assertion is taken for granted both by the speaker and her audience, while the antecedent of a conditional event (as characterized in terms of conditional bets) is typically considered as uncertain, at least by the speaker. Moreover, presuppositions are typically not explicitly proffered, while the truth-conditions under which a bet is considered valid are normally established explicitly and in detail to avoid ambiguity about the outcome of the bet.

13. Allusion to the theses of the Austrian philosopher Meinong (1853–1920), which motivated Russell's reflections on the topic of definite descriptions.

Chapter 18

Désiré André's Argument*

Problems related to the "gamblers' ruin," duration of a process, and so on. Désiré André's argument and applications.¹ Ballot problems. Further examples of the same type.

During this penultimate lecture, I would like to outline some very elementary, yet at the same time very interesting, examples of probabilistic reasoning. They are so simple that even children could understand them.²

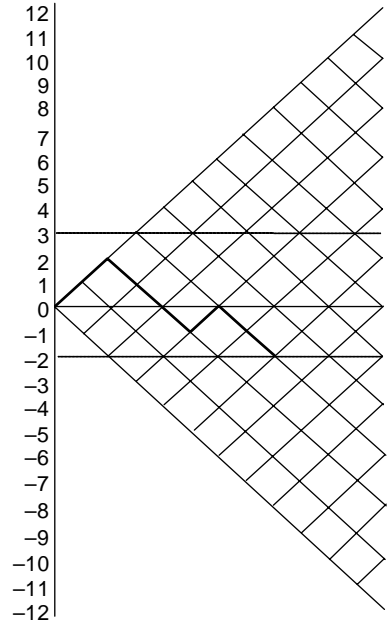
Heads and Tails: The Gambler's Ruin

Figure 18.1 represents the behaviour of a game of Heads and Tails between two players A and B who possess, at the beginning of the game, an initial fortune of 2 and 3, respectively. A game of Head and Tails is nothing but a sequence of fair bets on Heads, which terminates when one of the two players exhausts his fortune. At each shot, A gains 1 (at the expense of B) if the shot results in Heads and loses 1 (that is to say, his gain is -1) if the outcome of the toss is Tails. The order of the shots is displayed on the horizontal (discrete) axis, whereas on the ordinate is A 's net gain after each shot. The possible values on the ordinate are those included between the minimum and maximum gain for A . The minimum (-2) is identical to A 's initial fortune with the opposite sign and the maximum is identical to the initial fortune of B . The game terminates when the ordinate hits either of those two values for the first time. In fact, if this happens, one of the two players is left with no money to carry on playing. In this case, one says that he is *ruined*. Various interesting problems arise in a context of this sort. One of them is the problem of determining the prevision of the duration of the game in terms of the number of shots. Our goal now is to see what kind of argument can lead us to a solution for this problem.

Clearly, if after k shots ($0 \leq k$) the value of the ordinate is i , after the $k + 1$ -th shot, it will go to either $i + 1$ or $i - 1$. Thus, there will be a certain probability p_k that the next value will be $i + 1$ and a probability $1 - p_k$ that the next value will

* Lecture XXI (Wednesday 16 May, 1979).

Fig. 18.1 Graphical representation of Heads and Tails process



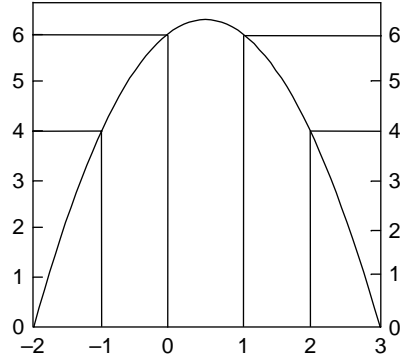
be $i - 1$. For the sake of simplicity, let us assume that $p = 1/2$, for all k . Let us observe that after k shots the prevision of the number of shots needed to reach one of the player's ruin is not dependent on k , but only on the ordinate i_k . For example, if $i_k = 3$ or $i_k = -2$, then it would be certain that the number of shots needed to reach the end of the game is 0, so the prevision of its duration is itself 0. Let us denote by n_i , for each value i ($-2 \leq i \leq 3$) of the ordinate, the prevision of the number of shots needed to reach the end of the game starting from the point of ordinate i . We know that n_3 and n_{-2} both equal 0. Let us observe now that all the remaining values of i_k, i_{k+1} is going to be either $i_k + 1$ or $i_k - 1$, according to whether the outcome of the $k + 1$ -th shot is Heads or Tails. Therefore, for all i ($-2 \leq i \leq 3$), it holds that $n_i = 1 + \frac{1}{2}(n_{i+1} + n_{i-1})$. It can be shown analytically (though I shall omit the proof here) that the result is a parabola (Fig. 18.2) whose equation is:

$$n_i = -(i + 2)(i - 3) = -i^2 + i + 6. \quad (18.1)$$

And more generally, if c' and c'' are the two ordinates of ruin, for $c' \leq i \leq c''$, it holds that:

$$n_i = (c' - i)(c'' - i) = -i^2 + i(c' + c'') - c'c''. \quad (18.2)$$

Fig. 18.2 Diagram representing the prevision of the duration of a Heads and Tails game at a given step. The diagram shows that, at each step, the prevision of duration equals the product of the fortunes that the two players possess at that moment



Therefore by equation (18.1), the values of n_i are:

$$n_{-2} = 0 \times 5 = 0$$

$$n_{-1} = 1 \times 4 = 4$$

$$n_0 = 2 \times 3 = 6$$

$$n_{+1} = 3 \times 2 = 6$$

$$n_{+2} = 4 \times 1 = 4$$

$$n_{+3} = 0 \times 5 = 0$$

If follows from equation (18.2) that

$$n_0 = -c'c''.$$

Therefore the initial prevision of the duration of the game equals the product of the initial fortune of A and the initial fortune of B . And in general, at each step, the duration of the game (expressed in terms of the number of shots) equals the product of the fortunes that the two players possess at that moment. Therefore, if A and B started off with a fortune of 100 each, the prevision of duration of the game would equal $100 \times 100 = 10\,000$ shots. However, it is not certain at all that the actual duration will be of 10 000 shots: it could well happen that the game will in fact have an infinite duration (though the probability that this will be the case is 0), and even that it will terminate after just 100 shots (namely if the first 100 shots all turned out to be either 'heads' or 'tails').

In spite of its simplicity, this situation lends itself to very enlightening remarks. For example, we have already observed that if both A and B possess an initial fortune of a 100, the prevision of duration of the game is 10 000 shots. This is the *maximum* value for the prevision when the sum of the fortunes of the two players equals 200: indeed, if after a certain number of shots, A 's and B 's fortunes were for instance, 80 and 120 respectively, the prevision of duration would be 9 600 shots, that is, less than 10 000. On the other hand, were the fortunes equal to 10 and 190, respectively, the prevision would be $10 \times 190 = 1\,900$, which is considerably less

than 10 000. And of course, were one of the fortunes zero, the prevision of the duration would likewise be zero.

Let us now observe that 10 000 is not only the result of the product 100×100 , but also of the product $10 \times 1\,000$. Therefore, the number of expected shots is the same regardless of whether one of the players has 10 cents whilst the other has 1 000, or they have 100 cents each (under the hypothesis that the stake for each shot is one cent). This might sound counterintuitive. For instance one could say: "10 is a small number; it doesn't take long to get ruined if one has only 10 cents!" And it is indeed true that it is very probable that the game will terminate soon. However, this is counterbalanced by the fact that in those cases in which the game terminates on the opposite side, it is highly probable that the play will last for quite some time.

The Wiener-Lévy Process

The game of Heads and Tails can be used to illustrate the important process known as the *Wiener-Lévy process* (which is applied to the physical theory of Brownian motion). Let us suppose that the game of Heads and Tails is performed with regular tosses of the coin, say one toss per minute: then it is possible to interpret the values displayed on the abscissa of Fig. 18.1 as temporal values. Suppose further that the stake is lowered by dividing it by a certain number $N > 0$: for the purposes of the computation of the duration of the game, this amounts to multiplying A 's and B 's fortune by N . It follows — by what has been said so far — that the initial prevision of the duration of the game will grow at the rate of N^2 . By increasing the frequency of the shots, one can ensure that the duration of the game is unchanged in temporal terms. Of course, given that the temporal duration of the game decreases linearly with the frequency of the shots, in order to keep the duration constant, one will have to increase the frequency of the shots at a rate of N^2 (and not at the rate of N , as one might be *prima facie* inclined to do). When the initial prevision of duration is constant, the standard deviation of the gains stays constant too. Indeed, the prevision of the duration of the game decreases linearly with the size of the variations of gains per temporal unit. More precisely, if the stake for each shot is $1/N$, the standard deviation of each temporal interval is $(1/N)^2$. This quantity remains unchanged if the number of shots in that interval is multiplied by N^2 . Therefore, in order to keep the standard deviation constant, the coefficient by which the frequency of the shots is multiplied must equal the square of the coefficient by which the size of the increment of the gains is divided at each shot. Let us suppose now that N tends to ∞ in such a way as to keep the standard deviation constant: the limiting case obtained in this way is precisely the *Wiener-Lévy process*.

Again on Gambler's Ruin

Let us go back to the problem of the gamblers' ruin. What is the probability that the game will end before the k -th shot?

EPSILON: Does the game go on after the ruin?

DE FINETTI: For the purposes of this problem, this is an irrelevant point. The *ruin* of one of the players means that either the prearranged maximal or minimal gain is reached for the first time. Now, the idea here is to compute the probability that such an event will take place in less than k shots. Whatever happens after this event is verified does not change the nature of the problem. We could assume that the game goes on by introducing a rule according to which the “ruined” player can be put back in the game. For instance it could be arranged that whenever one of the two players drains his fortune the opponent is forced to give him 1. In this way the game would go on forever. Yet this would not change the terms of our problem.

In order to determine the probability that the game will terminate in less than k shots we can make use of the celebrated *Désiré André's argument*. It is a proof technique based on the following principle. Let P and Q be two points with positive ordinate in the diagram of a Heads and Tails process. Let $p > 0$ be the ordinate of P . Let P' be the point with abscissa identical to P and with ordinate $-p$. The number of paths (that is to say, segments) connecting P and Q whose ordinate takes the value 0 in least one point equals the number of all the paths connecting P' with Q . The justification for this principle is illustrated in Fig. 18.3.

To each path S connecting P and Q corresponds one and only one path S' connecting P' and Q such that: (a) it coincides with S on the segment $O - Q$ and (b) on the segment $P' - O$ it is obtained by vertically reflecting the segment $P - O$ of S . On the other hand, every path S' connecting P' and Q *must* pass through the x -axis and therefore through it, there will always be a first shot at which its ordinate will take

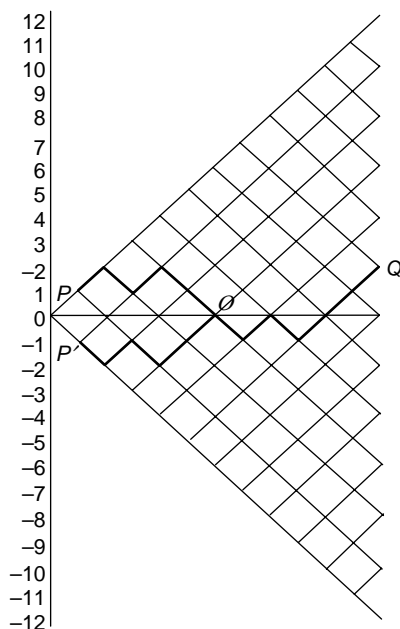


Fig. 18.3 Graphical illustration of Désiré André's argument

the value 0 at a point of abscissa O . After vertically reflecting the segment $P' - O$, to \mathcal{S}' will correspond one and only one path connecting P and Q , which also hits the x -axis at O . Therefore, as said above, there exists a bijection between the set of the paths connecting P and Q , which also reach or intersect the x -axis and the set of all paths connecting P' and Q .

By making use of this principle, we can dramatically simplify the solution to our problem. By applying the above argument with respect to the point of ruin rather than the x -axis, one can immediately see that the number of paths that would terminate (ideally even continuing the game after the ruin) in k shots beyond the point of ruin equals the number of paths which terminate before the point of ruin, but reach it at least once. Therefore, the probability q_k that the ruin will take place before the k -th shot equals the sum of twice the probability that the gain G_k at the k -th shot will be beyond the point of ruin c and the probability that at the k -th shot the gain G_k will be exactly at the point of ruin c :

$$q_k = 2\mathbf{P}(G_k > c) + \mathbf{P}(G_k = c).$$

The Ballot Problem

Another problem that can be easily solved by using the principle of Désiré André is the *Ballot problem* (Fig. 18.4). Suppose that the votes in an election between two candidates have to be counted. Suppose that it is known that the two candidates obtained the same number of votes. Let us ask ourselves what is the probability that — during the counting — one of the candidates will be constantly in the lead until the very last examined vote (a vote which will then be for her opponent).³ Let us take as positive steps (+1) those votes which are in favour of the candidate who was voted on the first examined vote and as negative steps (−1) those votes in favour of the other candidate. One immediately observes that all the paths terminating in a draw after N steps being otherwise always positive have ordinate $y = 1$ at the $N - 1$ -th step. Therefore, the number of paths we are interested in is the number of

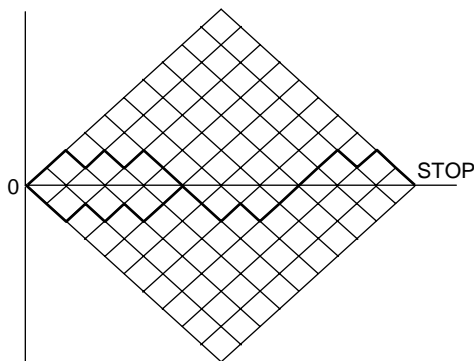


Fig. 18.4 Graphical illustration of the Ballot problem

always positive paths from $(0, 0)$ to $(N - 1, 1)$. In order to compute these latter, one observes, in the first place, that their number equals the total number of paths minus the number of paths which either hit or cross the axis of draw at least once. By the Désiré André principle there exists, on the set of all paths going from $(0, 0)$ to $+1$ in $N - 1$ shots, a bijection between the paths that goes to the point -1 on the first step, and the set of all paths which on the first step goes to the point $+1$ and either hits or crosses the draw at least once. Let us observe that the probability p which, starting at the point $(0, 0)$, the point $(N - 1, 1)$ will be reached while hitting the draw axis, equals the sum $q + r$ of the probability q that this will happen the first shot going to -1 and the probability r that this will happen the first shot going to $+1$. Of course it holds that $q = p$. Therefore, the probability that the path going from the point $(0, 0)$ to the point $(N - 1, 1)$ will hit the axis of draw at least once is $2r$. Now the value of r (considered that the last examined vote is negative and therefore cannot be examined first), is provided by the formula

$$r = \frac{N - 1 - \frac{N}{2}}{N - 1} = \frac{N - 2}{2(N - 1)}. \quad (18.3)$$

As p is — as already noted — twice r by equation (18.3) it holds that

$$p = 2 \frac{N - 2}{2(N - 1)} = \frac{N - 2}{N - 1}, \quad (18.4)$$

where the probability p' that the segment connecting the origin with the point $(N - 1, 1)$ will never hit the axis of draw is $1 - p$. Therefore, it follows from equation (18.4) that

$$p' = 1 - \frac{N - 2}{N - 1},$$

and finally:

$$\boxed{p' = \frac{1}{N - 1}}.$$

The Power of Désiré André's Argumentative Strategy

Many other problems can be simplified by means of the repeated use of Désiré André's principle. For instance, if there are strips of equal width that are delimited by straight lines parallel to the x -axis with respective ordinates of $1, 2, 3, 4, -3$, etc., it is possible fold up the curve representing the behaviour of the diagram on itself in a way as to keep it always inside the strip, by means of reflections performed at the height of each of the straight lines which mark the bounds of the stripes. Just think of a folding door.

A sequence that is often useful for all these cases is the following:

$$u = (1 - \frac{1}{2})(1 - \frac{1}{4})(1 - \frac{1}{6})(1 - \frac{1}{8}) \dots (1 - \frac{1}{2n}).$$

This sequence satisfies the following recursive relation:

$$u_{n+2} = u_n \left(1 - \frac{1}{n+2}\right) = \left(u_n \frac{n+1}{n+2}\right).$$

In expanded form, the value of u_n is given by the formula:

$$u_n = \sum_{k \leq n/2} \frac{u_{2k}}{2k-1} u_{n-2k}.$$

Furthermore, the value u_n equals the probability that the Heads and Tails process will never reach the equilibrium point until the n -th step inclusive.

Editor's Notes

1. The so-called "Désiré André's argument" (also known as the "reflection principle") goes back to André's solution (1887) to the ballot problem (discussed below) which was posed (and otherwise solved) by Joseph Bertrand (1887). It should be stressed that André's contribution is confined to the purely combinatorial aspect of the reflection principle, without resorting to the geometric representation of trajectories explained by de Finetti in this lecture. The geometric representation is actually due to J. Aebly (1923) and D. Mirimanoff (1923) (see Renault, 2006, p. 6).
Désiré André (1840–1917) was a student of Bertrand (1822–1900) and a mathematics teacher in a Parisian *lycée*. In spite of his excellent work in combinatorics, published between 1870 and 1880, he never obtained an academic position (see Taqqu, 2001, p. 11). André's solution to the ballot problem was included by Bertrand in his classic textbook on probability calculus (Bertrand [1888] 2005).
2. It may appear surprising that topics like the gambler's ruin or the ballot problem are dealt with in this series of lectures, which were devoted to the general and philosophical aspects of probability rather than to the development of the probability calculus. However, it should be kept in mind that de Finetti's goal here is *not* to teach a fragment of probability theory as such. He rather aims at providing paradigmatic examples of *good* probabilistic thinking and teaching. In his opinion, these examples show in a straightforward way what is wrong with the axiomatic-formalistic view, criticized in the first lecture (Chapter 1) and provides concrete examples of his "logico-intuitive" approach to doing and teaching mathematics, which was inspired by his teacher Oscar Chisini (cf. Chapter 10 on page 105 and note 19 on page 108).
3. This is actually a special case of the Ballot problem as originally posed by Bertrand. In the original formulation, the two candidates being A and B , a votes are supposed to be cast for candidate A and b votes for B , where $a > b$. In such a case, the probability that A stays ahead of B throughout the counting of the votes is $(a-b)/N$, where $N = a+b$ (Bertrand, 1887). De Finetti's formulation is equivalent to a special case in which $a = b + 1$.

Chapter 19

Characteristic Functions*

*The method of the characteristic function.
Characteristic function of the sum-distribution and
independent distributions.*

Special cases:

*Heads and Tails;
Normal distribution;
Poisson's distribution;
Uniform distribution.*

Prevision and Linearity

This final lecture is dedicated to illustrating the method of characteristic functions. Before going into this topic however, I would like to go back to some notational proposals that in my opinion could be useful to highlight the linear character of the logical operations on events, as well as to show how those operations are related to the concept of probability and more generally, to the concept of prevision.

First of all, I would like to insist on the fact that probability is nothing but a special case of prevision (usually called “mathematical expectation.”) One of the reasons I prefer the term “prevision” to the term “mathematical expectation” is that it begins with the same initial “P” as “probability,” so that by writing $\mathbf{P}(X)$ one can mean, in those cases in which X is an event, both (in general) the prevision of X and (in particular) the probability of X . No ambiguity arises from this notation for we can think of an event E as that random quantity which takes value 0 if it is false and 1 if it is true. Moreover, the logical operations on events can be seen as linear operations on random quantities. Many probabilistic relations can be immediately derived by applying the linear operator of prevision to the events.

The main logical operations on events can be defined as arithmetical operations as follows:¹

Logical sum: $E_1 \vee E_2 = E_1 + E_2 - E_1 E_2 = \max(E_1, E_2)$.

Logical product: $E_1 \wedge E_2 = E_1 E_2 = \min(E_1, E_2)$.

Negation: $\tilde{E} = 1 - E$.

* Lecture XXII (Thursday 17 May, 1979).

Also the concepts of *number of successes* and *relative frequency of successes* can be characterized in terms of arithmetical operations on events. The number of successes S_n can be defined as follows:

$$S_n = E_1 + E_2 + \cdots + E_n ,$$

whereas the relative frequency (or percentage) of successes can be thought of as the arithmetic mean of the events. Therefore, it can be written as:

$$\frac{S_n}{n} = \frac{1}{n}(E_1 + E_2 + \cdots + E_n) .$$

The operations of logical sum and logical product not only apply to events, but also to random quantities. In general in fact, one can write :

$$X \vee Y = \max(X, Y);$$

$$X \wedge Y = \min(X, Y);$$

$$\tilde{X} = 1 - X.$$

Rotations in the Complex Plane

Let us turn now to characteristic functions. There are many variations. The most practical one consists in taking the function $\phi(u)$ in place of the distribution of the random quantity X , which gives the prevision of the following random quantity:

$$e^{iuX} = \cos uX + i \sin uX ,$$

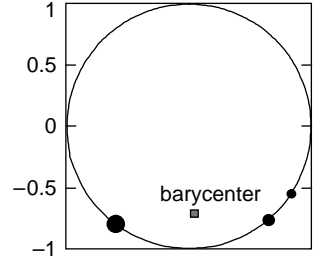
so that $\phi(u)$ is defined as follows:²

$$\phi(u) = \mathbf{P}(e^{iuX}) = \mathbf{P}(\cos uX) + i\mathbf{P}(\sin uX) .$$

In order to make the concept of characteristic function intuitive we can draw on the theory of rotation in the complex plane. Suppose that a certain distribution of masses is lying on a circle. For the sake of illustrating the point (though the same conclusions hold in general) let us suppose it is a discrete distribution of masses concentrated at three points of the circumference of the unit circle (Fig. 19.1).

In order to obtain a geometrical representation of the concept of characteristic function it is enough to jointly rotate the three masses along the circumference anticlockwise. The rotations of an angle u — as is well known — can be represented in the complex plane as exponentials e^{iu} . By rotating the system of masses, its barycenter will move accordingly. The problem here is to determine which position of the barycenter corresponds to each possible value of u (which ranges from

Fig. 19.1 Geometrical representation of the characteristic function of a discrete random number as a rotation in the complex plane



$-\infty$ to $+\infty$). The prevision of e^{iu} corresponds to such a position. Were X (contrary to Fig. 19.1) a continuous quantity endowed with density f , the prevision would be given by the following equation:

$$\mathbf{P}(e^{iuX}) = \int_{-\infty}^{+\infty} f(x)e^{iux} dx .$$

In the discrete case (as in Fig. 19.1) we would instead obtain a sum or a converging series (provided that X could only take countably many possible values). Yet in general, as we well know, all those cases can be grouped together thanks to the Stieltjes integral:

$$\mathbf{P}(e^{iuX}) = \int_{-\infty}^{+\infty} e^{iux} dF(x) . \quad (19.1)$$

Formula (19.1) — where $F(x)$ denotes the cumulative distribution function of X — allows us to convert, in all cases, the distribution of X into the characteristic function of X . Of course this is not sufficient to ensure that we can operate indifferently either on the characteristic function of X or on the distribution of X . To this effect, the inverse transformation is required as well. However this latter always exists. In the simplest case, where the cumulative distribution function $F(x)$ admits a continuous and bounded density $f(x)$, the following relation holds:

$$f(x) = \frac{1}{2\pi} \left(\lim_{a \rightarrow \infty} \int_{-a}^{+a} e^{-iux} \phi(u) du \right) .$$

Of course, if one also wanted to consider those distributions containing discrete masses, then one should resort again to the Stieltjes integral.

The key advantage offered by characteristic functions consists in the fact that given two independent random quantities X and Y , with characteristic functions $\phi_1(u)$ and $\phi_2(u)$, respectively, the characteristic function $\phi(u)$ of $X + Y$ is given by the product of $\phi_1(u)$ and $\phi_2(u)$:

$$\boxed{\phi(u) = \phi_1(u)\phi_2(u)} .$$

This is indeed rather natural given that:

$$e^{i(X+Y)} = e^{iX} e^{iY} .$$

Some Important Characteristic Functions

Let us now turn to some important examples of characteristic functions. Let us begin with Heads and Tails. In this case the gain after n shots — G_n — equals $2S_n - n$ (where S_n is the number of successes $E_1 + E_2 + \dots E_n$). The characteristic function of the gain of a single shot is:

$$\phi_1(u) = \frac{1}{2}(e^{iu} + e^{iu(-1)}) = \frac{1}{2}(e^{iu} + e^{-iu}) = \cos u .$$

so that it holds:

$$\boxed{\phi_1(u) = \cos u} .$$

whereas that of G_n is the sum of n independent shots and therefore equals (by what has been said above) the *product* of n characteristic functions:

$$\boxed{\phi_n(u) = \cos^n u} .$$

The case of the normal distribution is interesting as well. This distribution, when normalized (with mean = 0 and variance = 1) has the following density function:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

and its characteristic function is:

$$\boxed{\phi(u) = e^{-\frac{1}{2}u^2}} .$$

Another important example is the characteristic function of the family of Poisson's distributions (characterized by the λ parameter). For a given λ , Poisson's distribution is given by the formula

$$p_k = e^{-\lambda} \frac{\lambda^k}{k!} ,$$

which provides the probability of k successes in a unit interval. The corresponding characteristic function is the following:

$$\boxed{\phi(u) = e^{\lambda(e^{iu}-1)}} .$$

Another interesting case is that of the uniform distribution on the interval $[-1, +1]$, which has constant density $1/2$ and characteristic function

$$\phi(u) = \frac{\sin u}{u}.$$

I would like to know if you are already familiar with these things.

BETA: Yes. Although I never came across an explanation of characteristic functions in terms of rotations.

DE FINETTI: There is nothing new in it; it is just a way of presenting the topic aimed at making intuitive the fundamental idea.³

BETA: Geometric representations often help in understanding things better.

Concluding Remarks

DE FINETTI: I would be very happy if that happened to be the case: it is very important to me that one is able to internalize the concepts. Once concepts have been grasped at the root, it is indeed easy to work out the formulae. This also happens — and this is my case — when one is not able to remember them by heart. Someone once gave the following definition of culture: “culture is what remains once we have forgotten everything we have learnt.” Maybe things are not this drastic in reality. However, it seems to me that there is a kernel of truth in this sentence: it might well not be possible to remember everything correctly, parrot-fashion. Indeed, it is enough to keep ready, in various compartments of our minds, all that we need in the various circumstances. After all, not even this much is necessary: it would be enough to recall the reference to a source containing what we are interested in. In this way, even without knowing things by heart, we would be able to find what we need. It must be said, however, that one must remember at least the conceptual structure.

Editor's Notes

1. See Chapter 4 on pages 31–33.
2. See Chapter 9 on page 94.
3. Here we have to repeat the considerations made in Chapter 18, note 2 on page 184.

Cited Literature

- Adams, Ernest W. [1961]. "On Rational Betting Systems." *Archive für Mathematische Logik und Grundlagenforschung* **6**, pp. 112–28.
- [1965]. "The Logic of Conditionals." *Inquiry* **8**, pp. 166–97.
- [1975]. *The Logic of Conditionals*. Amsterdam: D. Reidel.
- [1977]. "A Note on Comparing Probabilistic and Modal Logics of Conditionals." *Theoria* **43**, pp. 186–94.
- Aebly, J. [1923]. "Démonstration du problème du scrutin par des considérations géométriques." *L'enseignement mathématique* **23**, pp. 185–6.
- Allais, Maurice [1953]. "Le comportement de l'homme rationel devant le risque: critique des postulats et axiomes de l'école américaine." *Econometrica* **21**, pp. 503–46.
- [1979]. "The so-called Allais Paradox and Rational Decision under Uncertainty." In Allais and Hagen, eds., pp. 437–699.
- [1984]. "The Foundations of the Theory of Utility and Risk." In Ole Hagen and Fred Wenstøp, eds., *Progress in Utility and Risk Theory, Theory and Decision Library*, vol. 42. Dordrecht: Reidel, pp. 3–131.
- Allais, Maurice and Ole Hagen, eds. [1979]. *Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions and Rational Decisions under Uncertainty with Allais' Rejoinder, Theory and Decision Library*, vol. 21. Dordrecht: Reidel.
- Armstrong, Stefan [1977a]. "Caso/probabilità." In *Enciclopedia Einaudi*, vol. 2. Torino: Einaudi, pp. 668–87.
- [1977b]. "Causa/Effetto." In *Enciclopedia Einaudi*, vol. 2. Torino: Einaudi, pp. 823–45.
- André, Désiré [1887]. "Solution directe du problème résolu par M. Bertrand." *Comptes Rendus de l'Académie des Sciences* **105**, pp. 436–7.
- Arden, Brigitta [2004]. "A Guide to the Papers of Bruno de Finetti, 1924–2000." Tech. rep., Special Collections Department, University of Pittsburgh Library System, Pittsburgh. Archives of Scientific Philosophy. URL: <http://proxima.library.pitt.edu/libraries/special/asp/finetti.pdf>
- Armendt, Brad [1980]. "Is There a Dutch Book Argument for Probability Kinematics?" *Philosophy of Science* **47** (4), pp. 583–88.
- Arntzenius, Frank, Adam Elga, and John Hawthorne [2004]. "Bayesianism, Infinite Decisions and Binding." *Mind* **113**, pp. 251–83.
- Ayer, Alfred, Michael Dummett, Michel Foucault, et al. [1980]. "Letter: The Flying University." *The New York Review of Books* **26** (21–22). URL: <http://www.nybooks.com/articles/7545>
- Baillie, Patricia [1973]. "Confirmation and the Dutch Book Argument." *The British Journal for the Philosophy of Science* **24**, pp. 393–97.
- Barnard, George A. and David R. Cox, eds. [1962]. *The Foundations of Statistical Inference*. New York: John Wiley & Sons.
- Barnes, Jonathan, ed. [1984]. *The Complete Works of Aristotle: The Revised Oxford Translation*. Princeton, NJ: Princeton University Press.

- Bartha, Paul [2004]. "Countable Additivity and the de Finetti Lottery." *The British Journal for the Philosophy of Science* **55**, pp. 301–21.
- Bayes, Thomas [(1764) 1970]. "An Essay Towards Solving a Problem in the Doctrine of Chances." In Egon S. Pearson and Maurice G. Kendall, eds., *Studies in the History of Statistics and Probability*, vol. 1. London: Charles Griffin, pp. 134–53. Original edition *Philosophical Transactions of the Royal Society of London* **53**, pp. 370–418.
- Ben-Tal, Aharon [1977]. "On Generalized Means and Generalized Convex Functions." *Journal of Optimization Theory and Applications* **21**, pp. 1–13.
- Bernardo, José M., Morris H. DeGroot, Dennis V. Lindley, and Adrian F. M. Smith, eds. [1980]. *Bayesian Statistics*. Valencia: Valencia University Press. Proceedings of the First International Meeting Held in Valencia (Spain), May 28 to June 2, 1979.
- [1985]. *Bayesian Statistics*, vol. 2. Amsterdam: North Holland. Proceedings of the Second International Meeting Held in Valencia (Spain), September 6th to 10th, 1983.
- Berry, Donald A. [2004]. "Bayesian Statistics and the Efficiency and Ethics of Clinical Trials." *Statistical Science* **19** (1), pp. 175–87.
- Berry, Scott M. and Joseph B. Kadane [1997]. "Optimal Bayesian Randomization." *Journal of the Royal Statistical Society B* **59** (4), pp. 813–19.
- Berti, Patrizia, Eugenio Regazzini, and Pietro Rigo [1991]. "Coherent Statistical Inference and Bayes Theorem." *Annals of Statistics* **19** (1), pp. 366–81.
- Bertrand, Joseph [1887]. "Solution d'un problème." *Comptes Rendus de l'Académie des Sciences* **105**, p. 369.
- [(1888) 2005]. *Calcul des probabilités*. Providence, RI: AMS Chelsea Pub., 3th ed. Reprint. Original edition Paris: Gauthier-Villars, 1888.
- Beth, Evert W. [1951]. "A Topological Proof of the Theorem of Löwenheim-Skolem-Gödel." *Indagationes Mathematicae* **13**, pp. 436–44.
- Birnbaum, Michael H. [1999]. "Paradoxes of Allais, Stochastic Dominance, and Decision Weights." In J. Shanteau, B. A. Mellers, and D. A. Schum, eds., *Decision Science and Technology: Reflections on the Contributions of Ward Edwards*. Norwell, MA: Kluwer Academic, pp. 27–52.
- Blattenberger, Gail and Frank Lad [1985]. "Separating the Brier Score into Calibration and Refinement Components: A Graphical Exposition." *American Statistician* **39** (1), pp. 26–32.
- Bochet, Olivier and Toyotaka Sakai [2004]. "Strategic Manipulations of Multi-valued Solutions in Economies with Indivisibilities." To appear in *Mathematical Social Sciences*.
- Bochvar, D. A. [(1938) 1981]. "On a Three Valued Logical Calculus and its Application to the Analysis of the Paradoxes of the Classical Extended Functional Calculus." *History of Philosophical Logic* **2**, pp. 87–112.
- Border, Kim C. and Uzi Segal [2002]. "Coherent Odds and Subjective Probability." *Journal of Mathematical Psychology* **46** (3), pp. 253–68.
- Borel, Émile [(1924) 1964]. "Apropos of a Treatise on Probability." In Kyburg Jr. and Smokler (1964), pp. 45–60. Translated by Howard E. Smokler. Original edition "À Propos d'un Traité de Probabilités," *Revue philosophique*, **98**, pp. 321–36.
- Borkar, Vivek S., Vijay R. Konda, Goldman Sachs and Sanjoy K. Mitter [2004]. "On De Finetti Coherence and Kolmogorov Probability." *Statistics and Probability Letters* **66** (4), pp. 417–21.
- Bradley, Darren and Hannes Leitgeb [2006]. "When Betting Odds and Credences Come Apart: More Worries for Dutch Book Arguments." *Analysis* **66**, pp. 119–27.
- Braithwaite, Richard B. [1957]. "On Unknown Probabilities." In Stephan Körner, ed., *Observation and Interpretation: A Symposium of Philosophers and Physicists*. London: Butterworth, pp. 3–11.
- Brier, Glenn W. [1950]. "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review* **78**, pp. 1–3.
- Bröcker, Jochen and Leonard A. Smith [2006]. "Scoring Probabilistic Forecasts: The Importance of Being Proper." Tech. rep., London School of Economics, London (UK). URL: <http://www.lse.ac.uk/collections/cats/>

- Butts, Robert E. and Jaakko Hintikka, eds. [1977]. *Basic Problems in Methodology and Linguistics*. Dordrecht: North-Holland.
- Calabrese, Philip G. [1987]. "An Algebraic Synthesis of the Foundations of Logic and Probability." *Information Sciences* **42**, pp. 187–237.
- [1994]. "A Theory of Conditional Information with Application." *IEEE Transactions on Systems, Man, and Cybernetics* **24** (12), pp. 1676–84.
- Campanile, Achille [1995]. *The Inventor of the Horse and Two Other Short Plays, Drama series*, vol. 7. Toronto: Guernica. Translated and with an introduction by Francesco Loriggio.
- Caprino, Silvia [1981]. "Chisini, Oscar." In *Dizionario biografico degli italiani*, vol. 25. Roma: Istituto della enciclopedia italiana, pp. 48–51.
- Carnap, Rudolf [(1928) 2003]. *The Logical Structure of the World and Pseudoproblems in Philosophy*. Open Court classics, Chicago: Open Court. Translated by Rolf A. George. Original editions *Der logische Aufbau der Welt*, Berlin-Schlachtensee: Weltkreis-Verlag, 1928 and *Scheinprobleme in der Philosophie; das Fremdpsychische und der Realismusstreit*, Berlin-Schlachtensee: Weltkreis-Verlag, 1928.
- [(1950) 1962]. *Logical Foundations of Probability*. Chicago: University of Chicago Press, 2nd ed.
- [1952]. *The Continuum of Inductive Methods*. Chicago: University of Chicago Press.
- [(1962) 1971]. "Inductive Logic and Rational Decisions." In Carnap and Jeffrey, eds., pp. 5–31. Modified and expanded version of the paper "The Aim of Inductive Logic." In Nagel et al. (1962).
- [1971a]. "A Basic System of Inductive Logic, Part 1." In Carnap and Jeffrey, eds., pp. 33–165.
- [1971b]. "Inductive Logic and Rational Decisions." In Carnap and Jeffrey, eds., pp. 5–31.
- [1980]. "A Basic System of Inductive Logic, Part 2." In Jeffrey, pp. 7–155.
- Carnap, Rudolf and Richard C. Jeffrey, eds. [1971]. *Studies in Inductive Logic and Probability*, vol. 1. Berkeley: University of California Press.
- Caves, Carlton M., Christopher A. Fuchs, and Rüdiger Schack [2002]. "Unknown Quantum State: The Quantum de Finetti Representation." *Journal of Mathematical Physics* **43** (9), pp. 4537–59. http://puhep1.princeton.edu/~mcdonald/examples/QM/caves_jmp-43-4537-02.pdf
- Cervera, J. L. [1996]. "Discussion of Scoring Rules and the Evaluation of Probabilities." *Test* **5**, pp. 27–31.
- Chaitin, Gregory J. [2001]. *Exploring Randomness*. London: Springer-Verlag.
- Chew, Soo Hong [1983]. "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox." *Econometrica* **51** (4), pp. 1065–92.
- Chisini, Oscar, ed. [(1915–1934) 1985]. *Federigo Enriques: Lezioni sulla teoria geometrica delle equazioni e delle funzioni algebriche*, vol. 1–4. Bologna: Zanichelli. Reprint.
- Chisini, Oscar [1929]. "Sul concetto di media." *Periodico di matematiche* **9**, pp. 106–16.
- Christensen, David [1991]. "Clever Bookies and Coherent Beliefs." *Philosophical Review* **100** (2), pp. 229–47.
- [1996]. "Dutch-Book Arguments Depigmatized: Epistemic Consistency for Partial Believers." *The Journal of Philosophy* **93** (9), pp. 450–79.
- [2001]. "Preference-Based Arguments for Probabilism." *Philosophy of Science* **68** (3), pp. 356–76.
- Church, Alonzo [1940]. "On the Concept of a Random Sequence." *Bulletin of the American Mathematical Society* **46**, pp. 130–5.
- Clemen, Robert T. [2002]. "Incentive Contracts and Strictly Proper Scoring Rules." *Test* **11** (1), pp. 167–89.
- Coletti, Giulianella and Romano Scozzafava [2002]. *Probabilistic Logic in a Coherent Setting*, vol. 15. Trends in logic. Dordrecht: Kluwer Academic.
- Cooke, R. M. and L. H. J. Goossens [2004]. "Expert Judgement Elicitation for Risk Assessments of Critical Infrastructures." *Journal of Risk Research* **7** (6), pp. 643–56.

- Corfield, David and Jon Williamson [2001]. *Foundations of Bayesianism*. Dordrecht: Kluwer Academic.
- Cornfield, Jerome [1969]. "The Bayesian Outlook and Its Applications." *Biometrics* **25**, pp. 617–57.
- Costantini, Domenico [1979]. "Il 'metodo dei risultati' e le ipotesi profonde." *Statistica* **39** (1).
- Cubitt, Robin P. and Robert Sugden [2001]. "On Money Pumps." *Games and Economic Behavior* **37** (1), pp. 121–60.
- Czuber, Emanuel [(1903) 1914]. *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung*. Leipzig: B. G. Teubner, 3th ed.
- [1923]. *Die Philosophischen Grundlagen der Wahrscheinlichkeitsrechnung*. B. G. Teubner.
- Daboni, Luciano [1974]. *Calcolo delle probabilità ed elementi di statistica*. Torino: UTET, 2nd ed.
- Davidson, Barbara and Robert Pargetter [1985]. "In Defence of the Dutch Book Argument." *Canadian Journal of Philosophy* **15**, pp. 405–23.
- Davis, Martin [(1977) 2005]. *Applied Nonstandard Analysis*. Mineola, NY: Dover. Unabridged republication of the work originally published by John Wiley & Sons: New York.
- Dawid, A. P. [1982a]. "The Well-Calibrated Bayesian." *Journal of the American Statistical Association* **77** (379), pp. 605–10.
- [1982b]. "The Well-Calibrated Bayesian: Rejoinder." *Journal of the American Statistical Association* **77** (379), pp. 612–3.
- [1985]. "Probability, Symmetry and Frequency." *The British Journal for the Philosophy of Science* **36**, pp. 107–28.
- [2004]. "Probability, Causality and the Empirical World: A Bayes–de Finetti–Popper–Borel Synthesis." *Statistical Science* **19** (1), pp. 44–57.
- [2007]. "The Geometry of Proper Scoring Rules." *Annals of the Institute of Statistical Mathematics* **59** (1), pp. 77–93.
- de Finetti, Bruno [1931a]. "Funzione caratteristica di un numero aleatorio." *Atti della Reale accademia nazionale dei lincei. Memorie della classe di scienze fisiche, matematiche e naturali* **4**, pp. 86–133.
- [1931b]. "Sul concetto di media." *Giornale dell'istituto italiano degli attuari* **2**, pp. 369–96.
- [(1931) 1989]. "Probabilism. A Critical Essay on the Theory of Probability and on the Value of Science." *Erkenntnis* **31** (2–3), pp. 169–223. English translation of *Probabilismo. Saggio critico sulla teoria della probabilità e sul valore della scienza*. Napoli: Perrella, 1931, reprinted in Mondadori (ed.) (1989), pp. 3–70.
- [(1931) 1992]. "On the Subjective Meaning of Probability." In Monari and Cocchi (1992), pp. 291–321. English translation of "Sul significato soggettivo della probabilità," *Fundamenta mathematicæ*, **17** (1931), pp. 298–329.
- [1936]. "Les probabilités nulles." *Bulletin des sciences mathématiques* **40**, pp. 275–88.
- [(1936) 1995]. "The Logic of Probability." *Philosophical Studies* **77** (1), pp. 181–190. Translated by R. B. Angell from "La logique de la probabilité." *Actualités Scientifiques et Industrielles* **391** (1936), Paris: Hermann et Cie, 1936, pp. 31–9. Actes du Congrès International de Philosophie Scientifique, Sorbonne, Paris, 1935, IV. Induction et Probabilité.
- [(1937) 1980]. "Foresight: Its Logical Laws, Its Subjective Sources." In Kyburg Jr. and Smokler (1980), pp. 93–158. Translated by Henry E. Kyburg Jr. with new notes added by the author of "La prévision: ses lois logiques, ses sources subjectives," *Annales de l'Institut Henri Poincaré*, **7** (1937), pp. 1–68.
- [(1938) 1980]. "On the Condition of Partial Exchangeability." In Jeffrey, ed., pp. 193–205. English translation by Paul Benacerraf and Richard Jeffrey of "Sur la condition d'équivalence partielle," *Actualités scientifiques et industrielles*, No. 739 (Colloque Genève d'Octobre 1937 sur la théorie des probabilités, 6^{ème} partie), Paris: Hermann et Cie, 1938.
- [1940]. "Il problema dei « pieni »." *Giornale dell'Istituto Italiano degli Attuari* **11**, pp. 1–88.

- [(1944) 2005]. *Matematica logico intuitiva: Nozioni di matematiche complementari e di calcolo differenziale e integrale come introduzione agli studi di scienze economiche statistiche attuariali*. Biblioteca IRSA di cultura del rischio, Milano: Giuffrè Editore, 3th ed. Reprint of 2nd edition, Roma: Cremonese, 1959.
- [1951]. "Recent Suggestions for the Reconciliation of Theories of Probability." In Jerzy Neyman, ed., *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (July 31–August 12, 1950)*. Berkeley: University of California Press, pp. 217–25.
- [1962]. "Does it Makes Sense to Speak of 'Good Probability Appraisers'?" In I. J. Good, A. J. Mayne, and J. M. Smith, eds., *The Scientist Speculates: An Anthology of Partly-Baked Ideas*. London: Heinemann, pp. 357–64.
- [1966]. "La probabilità secondo Gini nei rapporti con la concezione soggettivistica." *Metron* 25 (1–4), pp. 85–8.
- [1967]. *Il saper vedere in matematica*. Torino: Loescher. German translation *Die Kunst des Sehens in der Mathematik*, Basel: Birkhauser, 1974; Polish translation *Sztuka Widzenia w Matematyce*, Warsaw: Państwowe Wydawnictwo Naukowe, 1983.
- [1968a]. "Oscar Chisini e il suo insegnamento." *Periodico di matematiche* 46, pp. 26–33.
- [1968b]. "Probability: the Subjectivistic Approach." In Raymond Klíbanksky, ed., *La philosophie contemporaine*. Florence: La Nuova Italia, pp. 45–53.
- [1969]. "Initial Probabilities: A Prerequisite for any Valid Induction." *Synthese* 20 (1), pp. 2–16.
- [1970]. "Logical Foundations and Measurement of Subjective Probability." *Acta Psychologica* 34, pp. 129–45.
- [(1970) 1990a]. *Theory of Probability. A critical Introductory Treatment*, vol. 1. Chichester: John Wiley & Sons. Translated by Antonio Machi and Adrian Smith. Original edition *Teoria delle probabilità*, vol. 1–2. Torino: Einaudi, 1970 (reprinted Milano: Giuffrè, 2005).
- [(1970) 1990b]. *Theory of Probability. A critical Introductory Treatment*, vol. 2. Chichester: John Wiley & Sons. Translated by Antonio Machi and Adrian Smith. Original edition *Teoria delle probabilità*, vol. 2. Torino: Einaudi, 1970 (reprinted Milano: Giuffrè, 2005).
- [1972]. *Probability, Induction and Statistics: The Art of Guessing*. Wiley Series in Probability and Mathematical Statistics, London: John Wiley & Sons.
- [(1976) 1980]. "Probability: Beware of Falsifications!" In Kyburg Jr. and Smokler (1980), pp. 194–224. Translated by I. McGilvray from "La probabilità: guardarsi dalle contraffazioni!" *Scientia*, 111 (1976), pp. 255–81 (reprinted in Mondadori (1989), pp. 149–88).
- [1978a]. "Distribuzione statistica." In *Enciclopedia Einaudi*, vol. 4. Einaudi, pp. 1176–225.
- [1978b]. "Rischi di una « matematica » di base « assiomatica »." *Tuttoscuola* 4 (62), p. 13.
- [1979]. "A Short Confirmation of My Standpoint." In Allais and Hagen, eds., p. 161.
- [1980]. "Probabilità." In *Enciclopedia*, vol. 10. Torino: Einaudi, pp. 1146–87.
- [1981]. "The Role of 'Dutch Books' and of 'Proper Scoring Rules'." *The British Journal for the Philosophy of Science* 32 (1), pp. 55–6.
- [1982a]. "Probability and My Life." In Joseph Gani, ed., *The Making of Statisticians*. New York: Springer, pp. 3–12.
- [1982b]. "The Proper Approach to Probability." In G. Koch and F. Spizzichino, eds., *Exchangeability in Probability and Statistics*. Amsterdam: North-Holland, pp. 1–6. Proceedings of the International Conference on Exchangeability in Probability and Statistics, Rome, 6th–9th April, 1981, in Honour of Professor Bruno de Finetti.
- de Finetti, Bruno and Leonard J. Savage [1962]. "Sul modo di scegliere le probabilità iniziali." *Biblioteca del Metron C (Note e commenti)* 1, pp. 82–154.
- de Finetti, Fulvia [2000]. "Alcune lettere giovanili di B. de Finetti alla madre." *Nuncius* 15, pp. 721–40.
- De Morgan, Augustus [(1838) 1981]. *An Essay on Probabilities*. New York: Arno Press. Reprint. Original edition *An Essay on Probabilities, and their Application to Life Contingencies and Insurance Offices*. London: Longman, Orme, Brown, Green & Longmans.

- [1845]. “Theory of Probability.” In *Encyclopædia Metropolitana*, vol. 2. London: B. Fellowes, p. 396.
- [1847]. *Formal Logic: Or the Calculus of Inference, Necessary and Probable*. London: Taylor & Walton.
- DeGroot, Morris H. and Erik A. Eriksson [1985]. “Probability Forecasting, Stochastic Dominance, and the Lorenz Curve.” In Bernardo, DeGroot, Lindley, and Smith (1985), pp. 99–118. Proceedings of the Second International Meeting Held in Valencia (Spain), September 6th to 10th, 1983.
- DeGroot, Morris H. and Stephen E. Fienberg [1982]. “Assessing Probability Assessors: Calibration and Refinement.” In S. Gupta and J. O. Berger, eds., *Statistical Decision Theory and Related Topics III*, vol. 1. New York: Academic Press, pp. 291–314.
- [1983]. “The Comparison and Evaluation of Forecasters.” *Statistician* **32** (1–2), pp. 12–22. Proceedings of the 1982 I.O.S. Annual Conference on Practical Bayesian Statistics (March–June, 1983).
- [1986]. “Comparing Probability Forecasters: Basic Binary Concepts and Multivariate Extensions.” In Goel and Zellner, pp. 247–64.
- Delahaye, Jean-Paul [1993]. “Randomness, Unpredictability, and Absence of Order: The Identification by the Theory of Recursivity of the Mathematical Notion of Random Sequence.” In Jacques-Paul Dubucs, ed., *Philosophy of Probability*. Dordrecht: Kluwer Academic, pp. 145–67.
- Dempster, A. P. [1967]. “Upper and Lower Probabilities Induced by a Multivalued Mapping.” *Annals of Mathematical Statistics* **38**, pp. 325–39.
- Diaconis, Persi [1977]. “Finite Forms of de Finetti’s Theorem on Exchangeability.” *Synthese* **36**, pp. 668–87.
- Diaconis, Persi and David Freedman [1980a]. “De Finetti’s Generalizations of Exchangeability.” In Jeffrey (1980), pp. 233–49.
- [1980b]. “Finite Exchangeable Sequences.” *The Annals of Probability* **8** (4), pp. 745–764.
- [1987]. “A Dozen de Finetti Style Results in Search of a Theory.” *Annales de l’Institut Henry Poincaré* **23** (S2), pp. 397–423. URL: http://www.numdam.org/item/char63relaxid=AIHPB_1987__23_S2_397_0
- [2004a]. “The Markov Moment Problem and de Finetti’s Theorem: Part I.” *Mathematische Zeitschrift* **247** (1), pp. 183–99.
- [2004b]. “The Markov Moment Problem and de Finetti’s Theorem: Part II.” *Mathematische Zeitschrift* **247** (1), pp. 201–12.
- Diecidue, Enrico [2006]. “Deriving Harsanyi’s Utilitarianism from De Finetti’s Book-Making Argument.” *Theory and Decision* **61** (4), pp. 363–71.
- Diecidue, Enrico and Peter P. Wakker [2002]. “Dutch Books: Avoiding Strategic and Dynamic Complications and a Comonotonic Extension.” *Mathematical Social Sciences* **43**, pp. 135–49.
- Döring, Frank [2000]. “Conditional Probability and Dutch Books.” *Philosophy of Science* **67** (3), pp. 391–409.
- Downey, Rod G. and Denis Roman Hirschfeldt [2007]. *Algorithmic Randomness and Complexity*. Berlin: Springer-Verlag.
- Dubins, Lester E. [1975]. “Finitely Additive Conditional Probabilities, Conglomerability and Disintegrations.” *Annals of Probability* **3**, pp. 89–99.
- Dubins, Lester E. and Leonard J. Savage [1965]. *How to Gamble if You Must: Inequalities for Stochastic Processes*. Mineola, NY: Dover, 2nd ed.
- Dubois, Didier and Henri Prade [1994]. “Conditional Objects as Nonmonotonic Consequence Relationships.” *IEEE Transactions on Systems, Man, and Cybernetics* **24** (12), pp. 1724–40.
- Eagle, Antony [2004]. “Twenty-one Arguments against Propensity Analyses of Probability.” *Erkenntnis* **60** (3), pp. 371–416.

- Eaton, Morris L. and David A. Freedman [2004]. "Dutch Book against Some Objective Priors." *Bernoulli* **10** (5), pp. 861–72.
- Edgington, Dorothy [Spring 2006]. "Conditionals." In Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*. Stanford University. URL: <http://plato.stanford.edu/archives/spr2006/entries/conditionals/>
- Eells, Ellery and Brian Skyrms, eds. [1994]. *Probability and Conditionals: Belief Revision and Rational Decision*. Cambridge Studies in Probability, Induction and Rational Decision, Cambridge, Eng.: Cambridge University Press.
- Ellis, Brian [1969]. "An Epistemological Concept of Truth." In Robert Brown and C. D. Rollins, eds., *Contemporary Philosophy in Australia*. Library of philosophy, London: Allen & Unwin, pp. 52–72.
- Faraday, Michael [(1885–1889) 2002]. *The Chemical History of a Candle*. Mineola, NY: Dover. Original edition *A Course of Six Lectures on the Chemical History of a Candle*. New York: Chautauqua Press, 1885–1889.
- Fefferman, Charles [1967]. "A Radon-Nikodym Theorem for Finitely Additive Set Functions." *Pacific Journal of Mathematics* **23** (1), pp. 35–45.
- Feller, William [(1950) 1967]. *An Introduction to Probability Theory and Its Applications I*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 3th ed.
- [(1966) 1971]. *An Introduction to Probability Theory and Its Applications II*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 2th ed.
- Festa, Roberto [1993]. *Optimum Inductive Methods: A study in Inductive Probability, Bayesian Statistics and Verisimilitude*. Amsterdam: Kluwer Academic.
- Fishburn, Peter C. [1981]. "An Axiomatic Characterization of Skew-Symmetric Bilinear Functionals, with Applications to Utility Theory." *Economic Letters* **8**, pp. 311–3.
- [1986]. "Implicit Mean Value and Certainty Equivalence." *Econometrica* **54** (5), pp. 1197–206.
- Fishburn, Peter C. and Irving H. LaValle [1988]. "Context-Dependent Choice with Nonlinear and Nontransitive Preferences." *Econometrica* **56** (5), pp. 1221–39.
- Frangakis, Constantine E., Donald B. Rubin, and Xiao-Hua Zhou [2002]. "Clustered Encouragement Designs with Individual Noncompliance: Bayesian Inference with Randomization and Application to Advance Directive Forms." *Biostatistics* **3** (2), pp. 147–64.
- Frege, Gottlob [(1892) 1980]. "On Sense and Reference." In Peter Geach and Max Black, eds., *Translations from the Philosophical Writings of Gottlob Frege*. Oxford: Blackwell, 3d ed., pp. 56–78. Original edition "Über Sinn und Bedeutung." *Zeitschrift für Philosophie und philosophische Kritik*, **100**, pp. 25–50.
- Fuchs, Christopher A., Rüdiger Schack, and Petra F. Scudo [2004]. "De Finetti Representation Theorem for Quantum-Process Tomography." *Physical Review A* **69** (6). URL: <http://adsabs.harvard.edu/>
- Gaifman, Haim [1971]. "Applications of de Finetti's Theorem to Inductive Logic." In Carnap and Jeffrey, eds., pp. 235–51.
- Galavotti, Maria Carla [1989]. "Anti-realism in the Philosophy of Probability: Bruno de Finetti's Subjectivism." *Erkenntnis* **31**, pp. 239–61.
- [1991]. "The Notion of Subjective Probability in the Work of Ramsey and de Finetti." *Theoria* **57**, pp. 239–59.
- [1995]. "F.P. Ramsey and the Notion of Chance." In Jaakko Hintikka and Klaus Puhl, eds., *The British Tradition in the 20th Century Philosophy*. Vienna: Holder-Pichler-Tempsky, pp. 330–40. Proceedings of the 17th International Wittgenstein Symposium.
- [1999]. "Some Remarks on Objective Chance (F. P. Ramsey, K. R. Popper and N. R. Campbell)." In Maria Luisa Dalla Chiara, Roberto Giuntini and Federico Laudisa, eds., *Language, Quantum, Music*. Dordrecht: Kluwer, pp. 73–82.
- [2003a]. "Harold Jeffreys' Probabilistic Epistemology: Between Logicism and Subjectivism." *British Journal for the Philosophy of Science* **54**, pp. 43–57.

- [2003b]. “Kinds of Probabilism.” In Paolo Parrini, Wesley Salmon, and Merrilee Salmon, eds., *Logical Empiricism. Historical and Contemporary Perspectives*. Pittsburgh: Pittsburgh University Press, pp. 281–303.
- [2005]. *Philosophical Introduction to Probability*. Stanford: CSLI.
- Garbolino, Paolo [1997]. *I fatti e le opinioni: La moderna arte della congettura*. Roma: Laterza. Presentazione di Giulio Giorello.
- Geiringer, Hilda, ed. [(1928) 1981]. *Richard von Mises: Probability, Statistics and Truth*. New York: Dover Publications. Reprint. Original edition London: Allen and Unwin, 1957. Translation of *Wahrscheinlichkeit, Statistik, und Wahrheit*. Wien: J. Springer, 1928.
- Gibbard, Alan [1973]. “Manipulation of Voting Schemes: A General Result.” *Econometrica* **41** (4), pp. 587–601.
- Gilboa, Itzhak [1987]. “Expected Utility with Purely Subjective Non-Additive Probabilities.” *Journal of Mathematical Economics* **16**, pp. 65–88.
- Gini, Corrado [(1911) 2001]. “Considerations on A Posteriori Probability.” In Gini, ed., pp. 233–55. URL: <http://diglib.cib.unibo.it/diglib.php>
- [2001]. *Statistica e induzione*. Biblioteca di STATISTICA, Bologna: CLUEB. URL: <http://diglib.cib.unibo.it/diglib.php>
- Gneiting, Tilmann, Fadoua Balabdaoui, and Adrian E. Raftery [2005]. “Probabilistic Forecasts, Calibration and Sharpness.” Tech. Rep. 483, Department of Statistics, University of Washington.
- Gneiting, Tilmann and Adrian E. Raftery [2007]. “Strictly Proper Scoring rules, Prediction and Estimation.” *Journal of the American Statistical Association* **102** (477), pp. 359–78.
- Gödel, Kurt [(1930) 1986]. “Die Vollständigkeit der Axiome des logischen Funktionenkalküls / The Completeness of the Axioms of the Functional Calculus of Logic.” In Solomon Feferman, John W. Dawson, Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jean van Heijenoort, eds., *Collected Works*, vol. 1. New York: Oxford University Press, pp. 102–23. Original edition *Monatshefte für Mathematik und Physik* **37**, 1930, pp. 173–98.
- Goel, Prem K. and Arnold Zellner, eds. [1986]. *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*. Amsterdam: North-Holland.
- Goldstein, Michael [1983]. “The Prevision of a Prevision.” *Journal of the American Statistical Association* **78** (384), pp. 817–9.
- [2001]. “Avoiding Foregone Conclusions: Geometric and Foundational Analysis of Paradoxes of Finite Additivity.” *Journal of Statistical Planning and Inference* **94**, pp. 73–87.
- Goldstick, D. and B. O’Neill [1988]. “Truer.” *Philosophy of Science* **55**, pp. 583–97.
- Good, I. J. [(1952) 1983]. “Rational Decisions.” In Good, pp. 3–14. Originally published in “Journal of the Royal Statistical Society B,” **14** (1952), pp. 107–14.
- [(1962) 1983]. “Subjective Probability as the Measure of a Non-measurable Set.” In Good, ed., pp. 73–82. Original edition in Ernest Nagel, Patrick Suppes, and Alfred Tarski, eds., *Logic, Methodology, and Philosophy of Science*. Stanford: Stanford University Press, 1962, pp. 319–29. Proceedings of the 1960 International Congress.
- [1965]. *The Estimation of Probabilities: An Essay on Modern Bayesian Methods*. Cambridge, Mass.: MIT Press.
- [1969]. “Discussion of Bruno de Finetti’ Paper ‘Initial Probabilities: A Prerequisite for any Valid Induction’.” *Synthese* **20** (1), pp. 17–24.
- [1983]. *Good Thinking: The Foundations of Probability and its Applications*. Minneapolis: University of Minnesota Press.
- Goodman, Irwing R. and Hung T. Nguyen [1995]. “Mathematical Foundations of Conditionals and Their Probabilistic Assignments.” *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems* **3**, pp. 247–339.
- Goodman, Irwing R., Hung T. Nguyen, and Elbert Walker [1991]. *Conditional Inference and Logic for Intelligent Systems: A Theory of Measure-Free Conditioning*. Amsterdam: North-Holland.

- Grayson Jr., C. J. [1960]. *Decisions under Uncertainty: Drilling Decisions by Oil and Gas Operators*. Boston: Harvard University Press.
- Greaves, Hilary and David Wallace [2006]. "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility." *Mind* **115** (459), pp. 607–32.
- Guttman, Y. M. [1999]. *The Concept of Probability in Statistical Physics*. Cambridge Studies in Probability, Induction and Decision Theory, New York: Cambridge University Press.
- Hailperin, Theodore [1965]. "Best Possible Inequalities for the Probability of a Logical Function of Events." *The American Mathematical Monthly* **72**, pp. 343–59.
- [1996]. *Sentential Probability Logic: Origins, Development, Current Status and Technical Applications*. Bethlehem: Lehigh University Press.
- Hájek, Alan [2005]. "Scotching Dutch Books?" *Philosophical Perspectives* **19** (1), pp. 139–51.
- Halpern, Joseph Y. [2001]. "Lexicographic Probability, Conditional Probability, and Nonstandard Probability." In *Theoretical Aspects of Rationality and Knowledge*. San Francisco: Morgan Kaufmann, pp. 17–30. Proceedings of the 8th Conference on Theoretical Aspects of Rationality and Knowledge, Session: Probability and Uncertainty. URL: <http://portal.acm.org/citation.cfm?char63relaxid=1028131>
- Hardy, G. H., J. E. Littlewood, and G. Polya [1934]. *Inequalities*. Cambridge (Eng.): Cambridge University Press.
- Hardy, Godfrey Harold [(1889) 1920]. "Letter." *Transactions of the Faculty of Actuaries* **8**, pp. 180–181. Reprinted. Originally published on *Insurance Record* 457.
- Hartshorne, Charles and Paul Weiss, eds. [1960]. *Charles Sanders Peirce: Collected papers*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Heath, David and William Sudderth [1978]. "On Finitely Additive Priors, Coherence and Extended Admissibility." *Annals of Statistics* **6** (2), pp. 333–45.
- [1989]. "Coherent Inference from Improper Priors and from Finitely Additive Priors." *Annals of Statistics* **17** (2), pp. 907–19.
- Hewett, E. and L. J. Savage [1955]. "Symmetric Measures on Cartesian Products." *Transactions of the American Mathematical Society* **80**, pp. 470–501.
- Hild, Matthias [1998]. "The Coherence Argument against Conditionalization." *Synthese* **115** (2), pp. 229–58.
- Hill, Bruce M. [1980]. "On Finite Additivity, Non-Conglomerability, and Statistical Paradoxes." In Bernardo, DeGroot, Lindley, and Smith (1980), pp. 39–49. Proceedings of the First International Meeting Held in Valencia (Spain), May 28 to June 2, 1979.
- Hill, Bruce M. and David Lane [1986]. "Conglomerability and Countable Additivity." In Goel and Zellner (1986), pp. 45–57.
- Hintikka, Jaakko [1971]. "Unknown Probabilities, Bayesianism, and de Finetti's Representation Theorem." In R. C. Buck and R. S. Cohen, eds., *In Memory of Rudolf Carnap, Boston Studies in the Philosophy of Science*, vol. 8. Philosophy of Science Association, pp. 325–341. Proceedings of the 1970 Second Biennial Meeting of the Philosophy of Science Association, Boston, Fall 1970.
- Hoover, Douglas N. [1980]. "A Note on Regularity." In Jeffrey (1980), pp. 295–7.
- Hosni, Hykel and Jeff B. Paris [2005]. "Rationality as Conformity." *Synthese* **144** (2), pp. 249–85.
- Howson, Colin [1993]. "Dutch Book Arguments and Consistency." In *PSA 1992: Proceedings of the Biennial Meetings of the Philosophy of Science Association*, vol. 2, pp. 161–8.
- [2000]. *Hume's Problem: Induction and the Justification of Belief*. Oxford: Clarendon Press.
- Howson, Colin and Peter Urbach [(1989) 2005]. *Scientific Reasoning: The Bayesian Approach*. Chicago: Open Court, 3th ed.
- Humburg, Jürgen [1971]. "The Principle of Instantial Relevance." In Carnap and Jeffrey, (1971), pp. 225–33.
- Jackson, Frank [1979]. "On Assertion and Indicative Conditionals." *Philosophical Review* **88**, pp. 565–89.

- Jaynes, Edwin T. [1986]. "Some Applications of the de Finetti Representation Theorem." In Goel and Zellner (1986), pp. 31–42.
- Jeffrey, Richard C. [1964]. "If." *Journal of Philosophy* **61**, pp. 702–3.
- [(1965) 1983]. *The Logic of Decision*. Chicago: University of Chicago Press, 2nd ed. First edition New York: McGraw-Hill.
- [1971]. "Probability Measures and Integrals." In Carnap and Jeffrey, eds., pp. 167–223.
- [1977]. "Mises Redux." In Butts and Hintikka (1977), pp. 213–22.
- ed. [1980]. *Studies in Inductive Logic and Probability*, vol. 2. Berkeley: University of California Press.
- [1992a]. "De Finetti's Radical Probabilism." In Monari and Cocchi (1992), pp. 263–75. A collection of de Finetti's papers in Italian and English.
- [1992b]. *Probability and the Art of Judgment*. Cambridge, UK: Cambridge University Press.
- [1992c]. "Radical Probabilism (Prospectus for a User's Manual)." In Enrique Villanueva, ed., *Rationality in Epistemology*. Ridgeview, pp. 193–204.
- [2004]. *Subjective Probability: The Real Thing*. Cambridge, UK: Cambridge University Press.
- Jeffreys, Harold [1939]. *Theory of Probability*. Oxford: Clarendon Press.
- Kadane, Joseph B. [1982]. "The Well Calibrated Bayesian: Comment." *Journal of the American Statistical Association* **77** (379), pp. 610–1.
- Kadane, Joseph B., Mark J. Schervish, and Teddy Seidenfeld [1986]. "Statistical Implications of Finitely Additive Probability." In Goel and Zellner (1986), pp. 59–76.
- Kadane, Joseph B. and Teddy Seidenfeld [1990]. "Randomization in a Bayesian Perspective." *Journal of Statistical Planning and Inference* **25** (3), pp. 329–45.
- Kant, Immanuel [(1781–1787) 2004]. *Critique of Pure Reason*. Mineola, NY: Barnes & Noble. Translated by J. M. D. Meiklejohn. Introduction to the new edition by Andrew Fiala.
- Kemeny, John G. [1955]. "Fair Bets and Inductive Probabilities." *Journal of Symbolic Logic* **20**, pp. 263–73.
- Kennedy, Ralf and Charles Chihara [1979]. "The Dutch Book Argument: its Logical Flaws, its Subjective Sources." *Philosophical Studies* **36** (1), pp. 19–33.
- Keynes, John Maynard [(1921) 2004]. *A Treatise on Probability*. Mineola, NY: Dover Publications. Original edition London: Macmillan and Co., 1921.
- Khinchin, A. Y. [1924]. "Über einen Satz der Wahrscheinlichkeitsrechnung." *Fundamenta Mathematicae* **6**, pp. 9–20.
- [1932]. "Sur le classes d'événements équivalents." *Matematičeskii Sbornik* **39**, pp. 40–3.
- Kleene, Stephen C. [1938]. "On a Notation for Ordinal Numbers." *Journal of Symbolic Logic* **3**, pp. 150–55.
- Kolmogorov, Andrei Nikolaevich [1930]. "Sur la notion de la moyenne." *Atti della Reale accademia nazionale dei lincei. Rendiconti della classe di scienze fisiche, matematiche e naturali* **12**, pp. 388–91.
- [1965]. "Three Approaches for Defining the Concept of Information Quantity." *Information Transmission* **1**, pp. 3–11.
- [1968]. "Logical Basis for Information Theory and Probability Theory." *IEEE Transactions on Information Theory* **14**, pp. 662–64.
- Koopman, B. O. [1940]. "The Axioms and Algebra of Intuitive Probability." *Annals of Mathematics* **41**, pp. 269–92.
- Kotz, Samuel et al. [2006]. "Weighted Quasilinear Mean." In Samuel Kotz, Campbell B. Read, N. Balakrishnan, and Brani V. Radaković, eds., *Encyclopedia of Statistical Sciences*. John Wiley & Sons, 2nd ed.
- Krämmer, Walter [2006]. "Evaluating Probability Forecasts in Terms of Refinement and Strictly Proper Scoring Rules." *Journal of Forecasting* **25** (3), pp. 223–26.
- Kuipers, Theo A. F. [1987]. *What is Closer-to-the-Truth? A Parade of Approaches to Truthlikeness*. Poznań Studies in the Philosophy of the Sciences and the Humanities, Amsterdam: Rodopi.

- Kyburg, Henry E. and Choh Man Teng [2002]. "Randomization and Uncertain Inference." In J. G. Carbonell and J. Siekmann, eds., *PRICAI 2002: Trends in Artificial Intelligence*. Berlin: Springer, p. 598. Proceedings of the 7th Pacific Rim International Conference on Artificial Intelligence, Tokyo, Japan, August 18-22, 2002.
- Kyburg Jr., Henry E. and Howard E. Smokler, eds. [1964]. *Studies in Subjective Probability*. New York: John Wiley & Sons.
- [1980]. *Studies in Subjective Probability*. Huntington, NY: Krieger, 2nd ed. Substantially revised edition of Kyburg and Smokler (1964).
- Lad, Frank [1996]. *Operational Subjective Statistical Methods: A Mathematical, Philosophical and Historical Introduction*. Wiley Series in Probability and Statistics. Applied Probability and Statistics, New York: John Wiley & Sons.
- Laplace, Pierre Simon [(1774) 1986]. "Memoir on the Probability of the Causes of Events." *Statistical Science* **1** (3), pp. 364–78. Translated by Stephen M. Stigler. Original edition "Mémoire sur la probabilité des causes par les évènements," *Mémoires de Mathématique et de Physique*, **6**, pp. 621–56. Reprinted in *Œuvres complètes de Laplace*, Paris: Gauthier-Villars, 1878–1912, Vol. 8, pp. 27–65.
- Lehman, Sherman R. [1955]. "On Confirmation and Rational Betting." *Journal of Symbolic Logic* **20**, pp. 251–62.
- Levi, Isaac [1974]. "On Indeterminate Probabilities." *Journal of Philosophy* **71**, pp. 391–418.
- [1986]. *Hard Choices: Decision Making under Unresolved Conflict*. Cambridge, Eng.: Cambridge University Press.
- Lewis, David [(1976) 1986]. "Probabilities of Conditionals and Conditional Probability." In *Philosophical Papers*, vol. 2, chap. 20. Oxford: Oxford University Press, pp. 133–56. First publication *Philosophical Review* **85** (1976), pp. 297–315.
- Lindley, D. V. [1982]. "The Well Calibrated Bayesian: Comment." *Journal of the American Statistical Association* **77** (379), pp. 611–2.
- Lindley, Dennis V. [2000]. "The Philosophy of Statistics." *The Statistician* **49** (3), pp. 293–337.
- Link, Godehard [1980]. "Representation Theorem of the de Finetti Type for (Partially) Symmetric Probability Measures." In Jeffrey (1980), pp. 207–31.
- List, John A. and Michael S. Haigh [2005]. "A Simple Test of Expected Utility Theory Using Professional Traders." *Proceedings of the National Academy of Sciences of the United States of America* **102** (3), pp. 945–8.
- Localio, A. Russell, Jesse A. Berlin, and Thomas R. Ten Have [2005]. "Longitudinal and Repeated Cross-Sectional Cluster-Randomization Designs Using Mixed Effects Regression for Binary Outcomes: Bias and Coverage of Frequentist and Bayesian Methods." *Statistics in Medicine* **25** (16), pp. 2720–36.
- Loève, Michel [1960]. *Probability Theory*. The University series in higher mathematics, Princeton, NJ: Van Nostrand, 2nd ed.
- Loomes, Graham and Robert Sugden [1982]. "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty." *Economic Journal* **92** (368), pp. 805–24.
- [1998]. "Testing Alternative Stochastic Specifications for Risky Choice." *Economica* **65**, pp. 581–98.
- Luce, R. Duncan and A. A. J. Marley [2000]. "On Elements of Chance." *Theory and Decision* **49** (2), pp. 97–126.
- Łukasiewicz, Jan [(1957) 1987]. *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*. Greek & Roman Philosophy, New York: Garland. Reprint. Original edition Oxford: Clarendon Press, 1957.
- [1970]. *Selected Works*. Amsterdam: North-Holland.
- Machina, Mark J. [1982]. "'Expected Utility' Analysis without the Independence Axiom." *Econometrica* **50** (2), pp. 277–324.
- Maher, Patrick [1997]. "Depragmatized Dutch Book Arguments." *Philosophy of Science* **64** (2), pp. 291–305.
- Manski, Charles F. [2004]. "Measuring Expectations." *Econometrica* **72** (5), p. 1329.

- Markowitz, Harry Max [1952]. "Portfolio Selection." *Journal of Finance* **6**, pp. 77–91.
- McCall, John J. [2004]. "Induction: From Kolmogorov and Solomonoff to De Finetti and Back to Kolmogorov." *Metroeconomica* **55** (2–3), pp. 195–218.
- McCarthy, John [1956]. "Measures of the Value of Information." *Proceedings of the National Academy of Sciences of the United States of America* **42** (9), pp. 654–5.
- McDermott, Michael [1996]. "On the Truth Conditions of Certain 'If'-Sentences." *Philosophical Review* **105** (1), pp. 1–37.
- McGee, Vann [1981]. "Finite Matrices and the Logic of Conditionals." *Journal of Philosophical Logic* **10** (3), pp. 349–51.
- [1994]. "Learning the Impossible." In Ellery Eells and Brian Skyrms (1994), pp. 179–99.
- [1999]. "An Airtight Dutch Book?" *Analysis* **59** (264), pp. 257–65.
- Mellor, D. H., ed. [1990]. *Frank Plumpton Ramsey: Philosophical Papers*. Cambridge: Cambridge University Press.
- Miller, David W. [1974a]. "On the Comparison of False Theories by their Bases." *The British Journal for the Philosophy of Science* **25**, pp. 178–88.
- [1974b]. "Popper's Qualitative Theory of Verisimilitude." *The British Journal for the Philosophy of Science* **25**, pp. 166–77.
- [1977]. "The Uniqueness of Atomic Facts in Wittgenstein's *Tractatus*." *Theoria* **43** (3), pp. 174–85.
- [1991]. "Single Case Probabilities." *Foundations of Physics* **21**, pp. 1501–16.
- [2006]. *Out of Error: Further Essays on Critical Rationalism*. Aldershot, Eng.: Ashgate.
- Milne, Peter [1990]. "Scotching the Dutch Book Argument." *Erkenntnis* **32** (1), pp. 105–26.
- [1997]. "Bruno de Finetti and the Logic of Conditional Events." *The British Journal for the Philosophy of Science* **48** (2), pp. 195–232.
- [2004]. "Algebras of Intervals and a Logic of Conditional Assertions." *Journal of Philosophical Logic* **33**, pp. 497–548.
- Mirimanoff, D. [1923]. "À propos de l'interprétation géométrique du problème du scrutin." *L'enseignement mathématique*, pp. 187–9.
- Monari, Paola and Daniela Cocchi, eds. [1992]. *Bruno de Finetti: Probabilità e induzione (Induction and Probability)*. Bologna: CLUEB. A collection of de Finetti's papers in Italian and English.
- Muliere, Pietro and Giovanni Parmigiani [1993]. "Utility and Means in the 1930s." *Statistical Science*, **8** (4), pp. 421–32.
- Mura, Alberto [1986]. "Il carattere regolativo della teoria delle decisioni." *Economia delle scelte pubbliche. Journal of Public Finance and Public Choice* **4**, pp. 177–93.
- [1989]. "Concezione logico-trascendentale della probabilità e induzione." *Annali dell'università di Ferrara*, New series, section 3, Discussion paper no. 6. Summary in English.
- [1992]. *La sfida scettica: saggio sul problema logico dell'induzione*. Pisa, Italy: Edizioni ETS.
- [1994]. "Fairness and Expectation." In Carlo Cellucci, Maria Concetta Di Maio, and Gino Roncaglia, eds., *Logica e filosofia della scienza: problemi e prospettive*. Pisa: Edizioni ETS, pp. 455–74. Atti del congresso triennale della Società Italiana di Logica e Filosofia delle Scienze, Lucca 7–10 gennaio 1993.
- Mura, Alberto, ed. [1995a]. *Bruno de Finetti: Filosofia della probabilità*. Theoria, Milano: Il Saggiatore.
- Mura, Alberto [1995b]. "Probabilità soggettiva e non contraddittorietà." In Mura (1995a), pp. 35–58.
- [2008]. "Probability and the Logic of de Finetti's Trievents." In *De Finetti Radical Probabilist*. Proceedings of the Workshop held in Bologna, 26–28 October 2006 (forthcoming).
- Nagumo, Mitio [1930]. "Über eine Klasse der Mittelwerte." *Japan Journal of Mathematics* **7**, pp. 71–9.
- Nielsen, Thomas D. and Jean-Yves Jaffray [2006]. "Dynamic Decision Making without Expected Utility: An Operational Approach." *European Journal of Operational Research* **169** (1), pp. 226–46.

- Niiniluoto, Ilkka [2004]. "Verisimilitude: The Third Period." *The British Journal for the Philosophy of Science*.
- [1987]. *Truthlikeness*. Dordrecht: Reidel.
- Norris, Nilan [1976]. "General Means and Statistical Theory." *American Statistician* **30** (1), pp. 8–12.
- O'Connor, John J. and Edmund F. Robertson [2005]. "Emanuel Czuber." MacTutor History of Mathematics. URL: <http://www-history.mcs.st-andrews.ac.uk/Biographies/Czuber.html>
- Papineau, David [1994]. "The Virtues of Randomization." *The British Journal for the Philosophy of Science* **45** (2), pp. 437–50.
- Paris, Jeff B. [2001]. "A Note on the Dutch Book Method." *Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*. Ithaca, NY, USA, June.
- Parrini, Paolo [2004]. *Filosofia e scienza nell'Italia del Novecento: Figure, correnti, battaglie*. Epistemologica, Milano: Guerini e Associati.
- Pattanaik, Prasanta K. [1978]. *Strategy and Group Choice, Contributions to Economic Analysis*, vol. 113. Amsterdam: North-Holland.
- Peirce, Charles Sanders [1880]. "On the Algebra of Logic." *American Journal of Mathematics* **3**, pp. 15–57.
- Pinto, J. Sousa [2004]. *Infinitesimal Methods of Mathematical Analysis*. Chichester, UK: Horwood. Translated by R. F. Hoskins.
- Popper, Karl R. [(1935) 2004]. *The Logic of Scientific Discovery*. Routledge Classics, London: Routledge. Reprint. First English edition London: Hutchinson & Co, 1959. Original edition *Logik der Forschung*, Wien: J. Springer, 1935 (11th German edition Herbert Keuth, ed. *Karl R. Popper: Logik der Forschung in Karl R. Popper: Gesammelte Werke*, vol. 3, Tübingen: Mohr Siebeck, 2005).
- [1959]. "The Propensity Interpretation of Probability." *The British Journal for the Philosophy of Science* **10**, pp. 25–42.
- [(1972) 1979]. *Objective knowledge: An Evolutionary Approach*. Oxford: Clarendon Press, 2nd ed. First edition Oxford: Clarendon Press, 1972.
- [1976]. "A Note on Verisimilitude." *The British Journal for the Philosophy of Science* **27**, pp. 147–59.
- [2002]. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge Classics, London: Routledge.
- Post, Emil L. [1921]. "Introduction to a General Theory of Elementary Propositions." *American Journal of Mathematics* **43**, pp. 163–85.
- Pressacco, Flavio [2006]. "Bruno De Finetti, the Actuarial Sciences and the Theory of Finance in the XX Century." Tech. rep., INA and Accademia Nazionale dei Lincei, Rome. URL: <http://www.brunodefinetti.it/>
- Quaranta, Mario, ed. [1987]. *Giovanni Vailati: Scritti*. Bologna: Forni.
- Quiggin, John C. [1982]. "A Theory of Anticipated Utility." *Journal of Economic Behaviour and Organization* **3**, pp. 323–43.
- Quiggin, John C. and Peter P. Wakker [1994]. "The Axiomatic Basis of Anticipated Utility: A Clarification." *Journal of Economic Theory* **64** (2), pp. 486–99.
- Quine, Willard Van Orman [(1950) 2004]. *Methods of Logic*. Cambridge, USA: Harvard University Press, 4th ed. Reprint. First edition New York: Holt, 1950.
- Ramsey, Frank Plumpton [(1926) 1990]. "Truth and Probability." In Mellor (1990), pp. 52–94.
- [(1929) 1990]. "General Propositions and Causality." In Mellor (1990), pp. 145–63.
- Reichenbach, Hans [1949]. *The theory of Probability: An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability*. Berkeley: University of California Press, 2nd ed. English translation by Ernest H. Hutten and Maria Reichenbach.
- Renault, Marc [2006]. "A Reflection on André's Method and Its Application to the Generalized Ballot Problem." URL: <http://www.ship.edu/~msrena/ballotproblem/ReflectionAndre.pdf>

- Rice, Adrian [2003]. “‘Everybody Makes Errors’: The Intersection of De Morgan’s Logic and Probability, 1837–1847.” *History and Philosophy of Logic* **24** (4), pp. 289–305.
- Robert, Alain M. [2003]. *Nonstandard Analysis*. Mineola, NY: Dover. Translated and adapted by the author.
- Robinson, Abraham [1961]. “Non-Standard Analysis.” *Proceeding of the Royal Academy of Sciences* **64**, pp. 432–40.
- [1996]. *Non-standard Analysis*. Princeton Landmarks in Mathematics and Physics, Princeton, NJ: Princeton University Press.
- Rossi, Carla [2001]. “Bruno de Finetti: the Mathematician, the Statistician, the Economist, the Forerunner.” *Statistics in Medicine* **20**, pp. 3651–66.
- Rubin, Donald B. [1978]. “Bayesian Inference for Causal Effects: The Role of Randomization.” *The Annals of Statistics* **6** (1), pp. 34–58.
- Russell, Bertrand [(1901) 1957]. “Mathematics and the Metaphysicians.” In *Mysticism and Logic*. Garden City, NY: Doubleday, pp. 70–92. Originally published “Recent Work on the Principles of Mathematics,” *International Monthly* **4**, pp. 83–101.
- [1905]. “On Denoting.” *Mind* **14**, pp. 479–93.
- [(1919) 2002]. *Introduction to Mathematical Philosophy*. London: Routledge, 2 ed. Reprint. Original publication G. Allen & Unwin, 1920. First edition 1919.
- [1944]. “My Mental Development.” In Paul Arthur Schilpp, ed., *The Philosophy of Bertrand Russell, the Library of Living Philosophers*, vol. 5. Evanston, IL (USA): Northwestern University, pp. 1–20.
- Ryll-Nardzewski, C. [1957]. “On Stationary Sequences of Random Variables and the de Finetti Equivalences.” *Colloquia Mathematica* **4**, pp. 149–56.
- Satterthwaite, Mark Allen [1975]. “Strategy-proofness and Arrows Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions.” *Journal of Economic Theory* **10** (2), pp. 187–217.
- Savage, Leonard J. [(1954) 1972]. *The Foundations of Statistics*. New York: Dover Publications, 2nd ed. First edition New York: John Wiley & Sons, 1954.
- [1959]. “La probabilità soggettiva nella pratica della statistica.” Typed text of the recorded lectures held for the *International Mathematical Summer Centre* (CIME) in Varenna in 1959, June 1–10, session “Induzione e statistica.” The typed material is available at the library of the Department of Mathematics “Guido Castelnuovo,” University of Rome “La Sapienza.”
- [1962]. “Subjective Probability and Statistical Practice.” In Barnard and Cox (1962), pp. 9–35.
- [1971]. “Elicitation of Personal Probabilities and Expectations.” *Journal of the American Statistical Association* **66** (336), pp. 783–801.
- Scardovi, Italo [2001]. Preface to *Statistica e induzione*, by Corrado Gini. Bologna: CLUEB. URL: <http://diglib.cib.unibo.it/diglib.php>
- Schay, Geza [1968]. “An Algebra of Conditional Events.” *Journal of Mathematical Analysis and Applications* **24**, pp. 334–44.
- Schervish, Mark J. [1989]. “A General Method for Comparing Probability Assessors.” *Annals of Statistics* **17** (4), pp. 1856–79.
- Schervish, Mark J., Teddy Seidenfeld, and Joseph B. Kadane [1998]. “Two Measures of Incoherence: How Not to Gamble if You Must.” Tech. Rep. 660, Department of Statistics, Carnegie Mellon University, Pittsburgh PA 15213.
- [2000]. “How Sets of Coherent Probabilities May Serve as Models for Degrees of Incoherence.” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **8** (3), pp. 347–56.
- Schnorr, Claus Peter [1977]. “A Survey of the Theory of Random Sequences.” In Butts and Hintikka (1977), pp. 193–210.
- Seidenfeld, Teddy [1985]. “Calibration, Coherence and Scoring Rules.” *Philosophy of Science* **52** (2), pp. 274–94.

- [2001]. "Remarks on the Theory of Conditional Probability: Some Issues of Finite versus Countable Additivity." *Tex/Latex 92*, PhilSci Archives (University of Pittsburgh). URL: <http://philsci-archive.pitt.edu/archive/00000092/>
- [2004]. "A Contrast between Two Decision Rules for Use with (Convex) Sets of Probabilities: Γ -Maximin versus E-Admissibility." *Synthese* **140** (1), pp. 69–88.
- Seidenfeld, Teddy and Mark J. Schervish [1983]. "A conflict between Finite Additivity and Avoiding Dutch Book." *Philosophy of Science* **50** (3), pp. 398–412.
- Seillier-Moisewitsch, F. and A. P. Dawid [1993]. "On Testing the Validity of Sequential Probability Forecasts." *Journal of the American Statistical Association* **88** (421), pp. 355–9.
- Selten, Reinhard [1998]. "Axiomatic Characterization of the Quadratic Scoring Rule." *Experimental Economics* **1**, pp. 43–62.
- Shafer, Glenn [2001]. "The Notion of Event in Probability and Causality. Situating myself relative to Bruno de Finetti." Unpublished paper presented in Bologna and Pisa in March 2001. URL: <http://www.glennshafer.com/articles.html>
- Shimony, Abner [1955]. "Coherence and the Axioms of Confirmation." *Journal of Symbolic Logic* **20**, pp. 1–28.
- Silber, Daniel [1999]. "Dutch Books and Agent Rationality." *Theory and Decision* **47** (3), pp. 247–66.
- Skyrms, Brian [1980]. *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws*. New Haven: Yale University Press.
- [1984]. *Pragmatics and Empiricism*. Yale: Yale University Press.
- Smith, C. A. B. [1961]. "Consistency in Statistical Inference and Decision." *Journal of the Royal Statistical Society* **23**, pp. 1–25.
- Spiegelhalter, David J., Laurence S. Freedman, and Mahesh K. B. Parmar [1994]. "Bayesian Approaches to Randomized Trials." *Journal of the Royal Statistical Society A* **157** (3), pp. 357–416.
- Spielman, S. [1976]. "Exchangeability and the Certainty of Objective Randomness." *Journal of Philosophical Logic* **5**, pp. 399–406.
- [1977]. "Physical Probability and Bayesian Statistics." *Synthese* **36**, pp. 235–69.
- Stalnaker, Robert [1975]. "Indicative Conditionals." *Philosophia* **5**, pp. 269–86.
- Stigler, Stephen M. [1986]. "Laplace's 1774 Memoir on Inverse Probability." *Statistical Science* **1** (3), pp. 359–63.
- Stone, Mervyn [1969]. "The Role of Experimental Randomization in Bayesian Statistics: Finite Sampling and Two Bayesians." *Biometrika* **56** (3), p. 681.
- Strawson, Peter F. [1952]. *Introduction to Logical Theory*. London: Methuen & Co.
- Sudderth, William [1980]. "Finitely Additive Priors, Coherence and the Marginalization Paradox." *Journal of the Royal Statistical Society* **42**, pp. 339–41.
- Suppes, Patrick [1981]. *Logique du probable: démarche bayésienne et rationalité*. Paris: Flammarion. Postface by Jules Vuillemin.
- Suppes, Patrick, Leon Henkin, Athanase Joja, and Grigore C. Moisil, eds. [1973]. *Logic, Methodology and Philosophy of Science IV*. North-Holland. Proceedings of the 1971 International Congress.
- Suppes, Patrick and Mario Zanotti [1980]. "A New Proof of the Impossibility of Hidden Variables Using the Principles of Exchangeability and Identity of Conditional Distributions." In *Studies in the Foundations of Quantum Mechanics*. East Lansing, Mich.: Philosophy of Science Association, pp. 173–91.
- [1989]. "Conditions on Upper and Lower Probabilities to Imply Probabilities." *Erkenntnis* **31**, pp. 163–9.
- Swijtink, Zeno G. [1982]. "A Bayesian Argument in Favor of Randomization." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1982, pp. 159–68.
- Tanaka, Yasuhito [2001]. "Generalized Monotonicity and Strategy-proofness: A Note." *Economics Bulletin* **4** (11), pp. 1–6.

- Taqqu, Murad S. [2001]. "Bachelier and His Times: A Conversation with Bernard Bru." *Finance and Stochastics* 5, pp. 3–32.
- Teller, Paul [1973]. "Conditionalization and Observation." *Synthese* 26 (2), pp. 218–258.
- Tichý, Pavel [1974]. "On Popper's Definitions of Verisimilitude." *The British Journal for the Philosophy of Science* 25, pp. 155–60.
- Titiev, Robert [1997]. "Arbitrage and the Dutch Book Theorem." *Journal of Philosophical Research* 22, pp. 477–82.
- Vaihinger, Hans [(1911) 1965]. *The Philosophy of 'As if': A System of the Theoretical, Practical and Religious Fictions of Mankind*. International Library of Psychology, Philosophy and Scientific Method, London: Routledge & K. Paul, 2nd ed. Translated from the 6th German edition by C. K. Ogden. Original edition *Die Philosophie des Als ob*, Berlin: Reuther & Reichard, 1911.
- Vailati, Giovanni [(1904) 1987]. "La più recente definizione della Matematica." vol. 1. Bologna: Forni, pp. 7–12. Original edition *Leonardo* 2, 1904.
- van Fraassen, Bas C. [1984]. "Belief and the Will." *The Journal of Philosophy* 81 (5), pp. 235–56.
- van Lambalgen, Michiel [1987]. *Random Sequences*. Ph.D. thesis, Department of Mathematics and Computer Science, University of Amsterdam. URL: <http://staff.science.uva.nl/michiell/docs/fFDiss.pdf>
- Venn, John [(1888) 2006]. *The Logic of Chance*. Mineola, NY: Dover, 3th ed. Original edition London: Macmillan, 1888.
- Vineberg, Susan [1997]. "Dutch Books, Dutch Strategies and What They Show about Rationality." *Philosophical Studies* 86 (2), pp. 185–201.
- von Mises, Richard [1919]. "Grundlagen der Wahrscheinlichkeitsrechnung." *Mathematische Zeitschrift* 5, pp. 52–99.
- von Plato, Jan [1981]. "On Partial Exchangeability as a Generalization of Symmetry Principles." *Erkenntnis* 16, pp. 53–9.
- Waidacher, C. [1997]. "Hidden Assumptions in the Dutch Book Argument." *Theory and Decision* 43 (3), pp. 293–312.
- Wakker, Peter P. [2001a]. "History of the Term Dutch Book." URL: <http://www1.fee.uva.nl/creed/wakker/miscella/Dutchbk.htm>
- [2001b]. "Testing and Characterizing Properties of Nonadditive Measures through Violations of the Sure-Thing Principle." *Econometrica* 69 (4), pp. 1039–59.
- Weerahandi, Samaradasa and James V. Zidek [1979]. "A Characterization of the General Mean." *Canadian Journal of Statistics* 7, pp. 83–90.
- Whitworth, William Allen [(1897) 1965]. *DCC Exercises, Including Hints for the Solution of All Questions in Choice and Chance*. New York: Hafner Pub. Co. Reprint of the edition of 1897.
- Williams, L. Pearce [1987]. *Michael Faraday: A Biography*. The Da Capo Series in Science, New York, NY: Da Capo Press. Reprint. Originally published: Basic Books, New York, 1965.
- Willmann, Gerald [1999]. "The History of Lotteries." Tech. rep., Department of Economics, Stanford University, Stanford, CA 94305-6072, U.S.A. URL: <http://willmann.bwl.uni-kiel.de/~gerald/history.pdf>
- Winkler, Robert L. [1994]. "Evaluating Probabilities: Asymmetric Scoring Rules." *Management Science* 40 (11), pp. 1395–405.
- Wittgenstein, Ludwig [(1921) 2001]. *Tractatus Logico-Philosophicus*. Routledge Classics, London: Routledge. Translated by D. F. Pears and B. F. McGuinness. With an introduction by Bertrand Russell. Reprint. First English edition London: Kegan Paul, 1922. Original edition "Logisch-philosophische Abhandlung." *Annalen der Naturphilosophie*, 14 (1921), pp. 185–262.
- Worrall, John [2007]. "Why There's No Cause to Randomize." *The British Journal for the Philosophy of Science* 58, pp. 451–88.
- Yaari, Menahem E. [1987]. "The Dual Theory of Choice under Risk." *Econometrica* 55 (1), pp. 95–115.

- Zabell, Sandy L. [1989]. "R. A. Fisher on the History of Inverse Probability." *Statistical Science* **4** (3), pp. 247–56.
- [2005]. *Symmetry and Its Discontents: Essays on the History of Inductive Philosophy*. Cambridge Studies in Probability, Induction and Decision Theory, New York: Cambridge University Press.
- Zamora Bonilla, Jesus P. [2000]. "Truthlikeness, Rationality and Scientific Method." *Synthese* **122** (3), pp. 321–35.
- Zelen, Marvin [1982]. "The Contributions of Jerome Cornfield to the Theory of Statistics." *Biometrics* **38**, pp. 11–5. Proceedings of "Current Topics in Biostatistics and Epidemiology." A Memorial Symposium in Honor of Jerome Cornfield.
- Zwart, Sjoerd D. [2002]. *Refined Verisimilitude*. Synthese library, Dordrecht: Kluwer Academic.

Name Index

A

Adams, Ernest W., 45, 173
Aebly, J, 184
Alighieri, Dante, 25, 146, 148
Aliotta, Antonio, xvi
Allais, Maurice, 123, 125, 126
Amsterdamski, Stefan, 22, 26
André, Désiré, 184
Arden, Brigitta, xi
Aristotle, 169, 170, 174
Armendt, Brad, 45
Arntzenius, Frank, 172
Ayer, Alfred, 26

B

Baillie, Patricia, 45
Balabdaoui, Fadoua, 14
Barnard, George A., 110
Barone, Francesco, viii
Bartha, Paul, 124
Bayes, Thomas, 36, 74, 83
Ben-Tal, Aharon, 108
Berkeley, George, xvi
Berlin, Jesse A., 112
Berry, Donald A., 112
Berry, Scott M., 112
Berti, Patrizia, 45
Bertrand, Joseph, 184
Besso, Davide, 108
Beth, Evert W., 147
Bingham, Nicholas Hugh, ix, 123
Birnbaum, Michael H., 126
Blattenberger, Gail, 14
Bochet, Olivier, 123
Bochvar, D. A., 172
Border, Kim C., 45
Borel, Émile, 4, 11, 48–50, 57, 122
Borkar, Vivek S., 45
Bröcker, Jochen, 14

Bradley, Darren, 45
Braithwaite, Richard B., 83
Brier, Glenn W., 12
Butts, Robert E., 58

C

Calabrese, Philip G., 174
Calderoni, Mario, xvi
Campanile, Achille, 98, 107
Cantelli, Francesco Paolo, 47, 57
Caprino, Silvia, 107
Carnap, Rudolf, xvii, xx, 12, 40, 41, 43, 45, 47, 51, 74, 125, 131, 139
Caves, Carlton M., 83
Cervera, J. L., 14
Chaitin, Gregory J., 58
Chew, Soo Hong, 108, 126
Chihara, Charles, 45
Chisini, Oscar, 97–100, 105, 107, 108, 184
Christensen, David, 45
Church, Alonzo, 55, 58
Clemen, Robert T., 14
Coletti, Giulianella, 45
Cooke, Roger M., 14
Corfield, David, 45
Cornfield, Jerome, 35, 42
Costantini, Domenico, 74
Cox, David R., 110
Cubitt, Robin P., 45
Czuber, Emanuel, xv, 47, 57

D

Daboni, Luciano, 108
Dante, *see* Alighieri, Dante
Davidson, Barbara, 45
Davis, Martin, 125
Davy, Humphry, 99, 107
Dawid, A. Philip, 13, 14, 83
de Fermat, Pierre, 134, 135

de Finetti, Bruno, vii, viii, xi, xiii–xx, 2, 6, 7, 9–14, 22, 23, 25, 26, 29, 38, 41–44, 55, 57, 58, 74, 83, 84, 96, 98–101, 107, 108, 123–126, 135, 147, 156, 157, 163, 172–175, 184

de Finetti, Fulvia, ix, xiv

De Morgan, Augustus, 47, 57

DeGroot, Morris H., 14

Delahaye, Jean-Paul, 58

Democritus, 146

Dempster, Arthur P., 43

Diaconis, Persi, 83

Diecidue, Enrico, 45

Diophantus, 135

Döring, Frank, 45

Downey, Rod G., 58

Dubins, Lester E., 124, 147

Dubois, Didier, 174

Dummett, Michael, 26

E

Eagle, Antony, 66

Eaton, Morris L., 45

Edgington, Dorothy, 174

Einstein, Albert, 54, 58, 106

Elga, Adam, 172

Ellis, Brian, 173

Enriques, Federigo, 97, 105, 107, 108

Eriksson, Erik A., 14

F

Faraday, Michael, 99, 107

Feller, William, 96, 147

Fermat, Pierre, *see* de Fermat, Pierre

Festa, Roberto, 12, 58

Fienberg, Stephen E., 14

Finetti, Bruno, *see* de Finetti, Bruno

Finetti, Fulvia, *see* de Finetti, Fulvia

Fishburn, Peter C., 108, 126

Foà, Carlo, xiii

Foucault, Michel, 26

Frangakis, Constantine E., 112

Freedman, David A., 45, 83

Freedman, Laurence S., 112

Frege, Gottlob, 173

Fuchs, Christopher A., 83

G

Gaifman, Haim, 84

Galavotti, Maria Carla, ix, xx, 10, 45

Garbolino, Paolo, 83

Gauss, Karl Friedrich, 10

Geiringer, Hilda, 62

Gibbard, Alan, 123

Gilboa, Itzhak, 126

Gini, Corrado, xiii, xiv, 73, 74, 134, 140

Gneiting, Tilmann, 14

Gödel, Kurt, 56

Goldstein, Michael, 124

Goldstick, D., 58

Good, Irving John, 12, 31, 36, 42, 43, 98, 137

Goodman, Irwing R., 174

Goossens, L. H. J., 14

Grayson Jr., C. J., 8, 15, 19

Greaves, Hilary, 45

Guttmann, Y. M., 83

H

Haigh, Michael S., 126

Hailperin, Theodore, 125

Hájek, Alan, 45

Halpern, Joseph Y., 125

Hardy, Godfrey Harold, 74, 108

Hartshorne, Charles, 112

Hawthorne, John, 172

Heath, David, 124

Henkin, Leon, 147

Hewett, E., 83

Hild, Matthias, 45

Hill, Bruce M., 124

Hintikka, Jaakko, 56–58, 83

Hirschfeldt, Denis Roman, 58

Hoover, Douglas N., 125

Hosni, Hykel, ix, 45

Howson, Colin, 45, 83, 125

Humburg, Jürgen, 83

Hume, David, xvi

J

Jackson, Frank, 173

Jaffray, Jean-Yves, 126

Jaynes, Edwin T., 83

Jeffrey, Richard C., xx, 42, 43, 45, 58, 83, 173

Jeffreys, Harold, xvii, xix, xx, 42, 47, 57, 74

Joja, Athanase, 147

K

Kadane, Joseph B., 14, 45, 112, 124

Kant, Immanuel, 107

Kemeny, John G., 45

Kennedy, Ralf, 45

Keynes, John Maynard, 11, 40, 49

Khinchin, Aleksandr Yakovlevich, 83, 136

Kleene, Stephen Cole, 172

Kolmogorov, Andrei Nikolaevich, 14, 55, 58, 98, 100, 101, 105, 108, 114

Konda, Vijay R., 45

Koopman, Bernard Osgood, 43

Kotz, Samuel, 108
 Krämer, Walter, 14
 Kuipers, Theo A. F., 58
 Kyburg, Henry E., 112

L

Lad, Frank, 14
 Lambalgen, Michiel, *see* van Lambalgen, Michiel
 Lane, David, 124
 Laplace, Pierre Simon, 74, 83, 85
 LaValle, Irving H., 126
 Lazzari, G., 108
 Lehman, Sherman R., 45
 Leitgeb, Hannes, 45
 Leoni, Bruno, xvi
 Levi, Isaac, 43, 126
 Lewis, David, 41, 45, 173, 174
 Lindley, Dennis V., 14, 42, 45
 Link, Godehard, 83
 List, John A., 126
 Littlewood, John Edensor, 108
 Localio, A. Russell, 112
 Loève, Michel, 83
 Loomes, Graham, 126
 Luce, R. Duncan, 126
 Łukasiewicz, Jan, 172, 174

M

Mach, Ernst, xvi
 Machina, Mark J., 126
 Maher, Patrick, 45
 Manski, Charles F., 14
 Markowitz, Harry Max, 126
 Marley, Anthony Alfred John, 126
 Martellotti, Anna, xi
 McCall, John J., 83
 McCarthy, John, 6, 12–14
 McDermott, Michael, 174
 McGee, Vann, 45, 125, 173
 Meinong, Alexius von, 175
 Miller, David W., 58, 66, 157
 Miller, Merton Howard, 126
 Milne, Peter, 45, 172, 174
 Minkowski, Hermann, 54
 Mirimanoff, D., 184
 Mises, Richard, *see* van Mises, Richard
 Mitter, Sanjoy K., 45
 Moisil, Grigore C., 147
 Mondadori, Marco, xi
 Muliere, Pietro, 108
 Mura, Alberto, xiii, 13, 26, 29, 42, 45, 83, 123, 125, 135, 172

N

Nagumo, Mitio, 6, 14, 98, 100, 101, 105, 108
 Nguyen, Hung T., 174
 Nielsen, Thomas D., 126
 Niiniluoto, Ilkka, 58
 Norris, Nilan, 108
 Nozick, Robert, 26

O

O'Connor, John J., 57
 O'Neill, B., 58

P

Papineau, David, 112
 Pargetter, Robert, 45
 Paris, Jeff B., 45
 Parmar, Mahesh K. B., 112
 Parmigiani, Giovanni, 108
 Parrini, Paolo, 10, 58
 Pattanaik, Prasanta K., 123
 Peano, Giuseppe, 9, 10, 58
 Peirce, Charles Sanders, 10, 58, 112, 148
 Pinto, J. Sousa, 125
 Plato, xv, 25
 Plato, Jan, *see* von Plato, Jan
 Polya, G., 108
 Popper, Karl R., 47, 49, 57, 58, 60, 66, 149
 Post, Emil L., 148
 Prade, Henri, 174
 Pressacco, Flavio, 126
 Price, Richard, 83
 Putnam, Hilary, 26

Q

Quiggin, John C., 126
 Quine, Willard Van Orman, 26, 173

R

Radaelli, Anna, xiii
 Radaelli, Carlo Alberto, xiii
 Raftery, Adrian E., 14
 Ramsey, Frank Plumpton, xv, xix, xx, 41, 45, 173
 Regazzini, Eugenio, 45
 Reichenbach, Hans, xviii, 47, 48, 110, 112
 Renault, Marc, 184
 Rice, Adrian, 57
 Rigo, Pietro, 45
 Robert, Alain M., 125
 Robertson, Edmund F., 57
 Robinson, Abraham, 122, 125
 Rossi, Carla, xv
 Rubin, Donald B., 112
 Russell, Bertrand, 1, 9, 10, 171, 174, 175
 Ryll-Nardzewski, Czeław, 83

S

Sachs, Goldman, 45
 Satterthwaite, Mark Allen, 123
 Savage, Leonard Jimmie, xiv, xix, xx, 13, 14, 42, 43, 83, 110–112, 123, 126, 142, 147, 151, 157
 Scardovi, Italo, 74
 Schack, Rüdiger, 83
 Schay, Geza, 174
 Schervish, Mark J., 14, 45, 124, 172
 Schnorr, Claus Peter, 58
 Scozzafava, Romano, 45
 Scudo, Petra F., 83
 Segal, Uzi, 45
 Seidenfeld, Teddy, 14, 45, 112, 124, 172
 Seillier-Moiseiwitsch, F., 14
 Selten, Reinhard, 14
 Shafer, Glenn, 157
 Sharpe, William Forsyth, 126
 Shimony, Abner, 45, 118, 119
 Silber, Daniel, 45
 Skyrms, Brian, 83, 173
 Smith, Cedric A. B., 43
 Smith, Leonard A., 14
 Spiegelhalter, David J., 112
 Spielman, Stephen, 83
 Stalnaker, Robert, 173
 Stigler, Stephen M., 83
 Stone, Mervyn, 112
 Strawson, Peter F., 171, 173–175
 Sudderth, William, 124
 Sugden, Robert, 45, 126
 Suppes, Patrick, 43, 83
 Swijtink, Zeno G., 112

T

Tanaka, Yasuhito, 123
 Taqqu, Murad S., 184
 Teller, Paul, 45
 Ten Have, Thomas R., 112
 Teng, Choh Man, 112
 Thom, René, 26
 Tichý, Pavel, 58
 Titiev, Robert, 45
 Turing, Alan Mathison, 58

U

Urbach, Peter, 45, 83

V

Vaihinger, Hans, 100, 107
 Vailati, Giovanni, xvi, 9, 10, 56, 58, 100
 van Fraassen, Bas C., 45
 van Lambalgen, Michiel, 58
 Venn, John, 67, 115, 147
 Vinciguerra, Giovanni, 2, 10
 Vineberg, Susan, 45
 Vitali, Giuseppe, 137
 Volterra, Vito, 58
 von Mises, Richard, xviii, 47, 48, 55, 56, 58, 62, 63, 66, 134–136
 von Plato, Jan, 83

W

Waidacher, Christoph, 45
 Wakker, Peter P., 45, 126
 Walker, Elbert, 174
 Wallace, David, 45
 Weerahandi, Samaradasa, 108
 Weiss, Paul, 112
 Whitworth, William Allen, 74
 Wiles, Andrew John, 135
 Williams, L. Pearce, 107
 Williamson, Jon, 45
 Willmann, Gerald, 45
 Winkler, Robert L., 14
 Wittgenstein, Ludwig, 145, 146, 148, 150
 Worrall, John, 112

Y

Yaari, Menahem E., 126

Z

Zabell, Sandy L., 74, 83
 Zamora Bonilla, Jesus P., 58
 Zanotti, Mario, 43, 83
 Zelen, Marvin, 42
 Zhou, Xiao-Hua, 112
 Zidek, James V., 108
 Zwart, Sjoerd D., 58

SYNTHESE LIBRARY

1. J. M. Bochénski, *A Precís of Mathematical Logic*. Translated from French and German by O. Bird. 1959 ISBN 90-277-0073-7
2. P. Guiraud, *Problèmes et méthodes de la statistique linguistique*. 1959 ISBN 90-277-0025-7
3. H. Freudenthal (ed.), *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences*. 1961 ISBN 90-277-0017-6
4. E. W. Beth, *Formal Methods*. An Introduction to Symbolic Logic and to the Study of Effective Operations in Arithmetic and Logic. 1962 ISBN 90-277-0069-9
5. B. H. Kazemier and D. Vuysje (eds.), *Logic and Language*. Studies dedicated to Professor Rudolf Carnap on the Occasion of His 70th Birthday. 1962 ISBN 90-277-0019-2
6. M. W. Wartofsky (ed.), *Proceedings of the Boston Colloquium for the Philosophy of Science, 1961–1962*. [Boston Studies in the Philosophy of Science, Vol. I] 1963 ISBN 90-277-0021-4
7. A. A. Zinov'ev, *Philosophical Problems of Many-valued Logic*. A revised edition, edited and translated (from Russian) by G. Küng and D.D. Comey. 1963 ISBN 90-277-0091-5
8. G. Gurvitch, *The Spectrum of Social Time*. Translated from French and edited by M. Korenbaum and P. Bosserman. 1964 ISBN 90-277-0006-0
9. P. Lorenzen, *Formal Logic*. Translated from German by F.J. Crosson. 1965 ISBN 90-277-0080-X
10. R. S. Cohen and M. W. Wartofsky (eds.), *Proceedings of the Boston Colloquium for the Philosophy of Science, 1962–1964*. In Honor of Philipp Frank. [Boston Studies in the Philosophy of Science, Vol. II] 1965 ISBN 90-277-9004-0
11. E. W. Beth, *Mathematical Thought*. An Introduction to the Philosophy of Mathematics. 1965 ISBN 90-277-0070-2
12. E. W. Beth and J. Piaget, *Mathematical Epistemology and Psychology*. Translated from French by W. Mays. 1966 ISBN 90-277-0071-0
13. G. Küng, *Ontology and the Logistic Analysis of Language*. An Enquiry into the Contemporary Views on Universals. Revised ed., translated from German. 1967 ISBN 90-277-0028-1
14. R. S. Cohen and M. W. Wartofsky (eds.), *Proceedings of the Boston Colloquium for the Philosophy of Sciences, 1964–1966*. In Memory of Norwood Russell Hanson. [Boston Studies in the Philosophy of Science, Vol. III] 1967 ISBN 90-277-0013-3
15. C. D. Broad, *Induction, Probability, and Causation*. Selected Papers. 1968 ISBN 90-277-0012-5
16. G. Patzig, *Aristotle's Theory of the Syllogism*. A Logical-philosophical Study of *Book A* of the *Prior Analytics*. Translated from German by J. Barnes. 1968 ISBN 90-277-0030-3
17. N. Rescher, *Topics in Philosophical Logic*. 1968 ISBN 90-277-0084-2
18. R. S. Cohen and M. W. Wartofsky (eds.), *Proceedings of the Boston Colloquium for the Philosophy of Science, 1966–1968, Part I*. [Boston Studies in the Philosophy of Science, Vol. IV] 1969 ISBN 90-277-0014-1
19. R. S. Cohen and M. W. Wartofsky (eds.), *Proceedings of the Boston Colloquium for the Philosophy of Science, 1966–1968, Part II*. [Boston Studies in the Philosophy of Science, Vol. V] 1969 ISBN 90-277-0015-X
20. J. W. Davis, D. J. Hockney and W. K. Wilson (eds.), *Philosophical Logic*. 1969 ISBN 90-277-0075-3
21. D. Davidson and J. Hintikka (eds.), *Words and Objections*. Essays on the Work of W. V. Quine. 1969, rev. ed. 1975 ISBN 90-277-0074-5; Pb 90-277-0602-6
22. P. Suppes, *Studies in the Methodology and Foundations of Science. Selected Papers from 1951 to 1969*. 1969 ISBN 90-277-0020-6
23. J. Hintikka, *Models for Modalities*. Selected Essays. 1969 ISBN 90-277-0078-8; Pb 90-277-0598-4

SYNTHESE LIBRARY

24. N. Rescher *et al.* (eds.), *Essays in Honor of Carl G. Hempel*. A Tribute on the Occasion of His 65th Birthday. 1969 ISBN 90-277-0085-0
25. P. V. Tavanec (ed.), *Problems of the Logic of Scientific Knowledge*. Translated from Russian. 1970 ISBN 90-277-0087-7
26. M. Swain (ed.), *Induction, Acceptance, and Rational Belief*. 1970 ISBN 90-277-0086-9
27. R. S. Cohen and R. J. Seeger (eds.), *Ernst Mach: Physicist and Philosopher*. [Boston Studies in the Philosophy of Science, Vol. VI]. 1970 ISBN 90-277-0016-8
28. J. Hintikka and P. Suppes, *Information and Inference*. 1970 ISBN 90-277-0155-5
29. K. Lambert, *Philosophical Problems in Logic*. Some Recent Developments. 1970 ISBN 90-277-0079-6
30. R. A. Eberle, *Nominalistic Systems*. 1970 ISBN 90-277-0161-X
31. P. Weingartner and G. Zeche (eds.), *Induction, Physics, and Ethics*. 1970 ISBN 90-277-0158-X
32. E. W. Beth, *Aspects of Modern Logic*. Translated from Dutch. 1970 ISBN 90-277-0173-3
33. R. Hilpinen (ed.), *Deontic Logic*. Introductory and Systematic Readings. 1971
See also No. 152. ISBN Pb (1981 rev.) 90-277-1302-2
34. J.-L. Krivine, *Introduction to Axiomatic Set Theory*. Translated from French. 1971
ISBN 90-277-0169-5; Pb 90-277-0411-2
35. J. D. Sneed, *The Logical Structure of Mathematical Physics*. 2nd rev. ed., 1979
ISBN 90-277-1056-2; Pb 90-277-1059-7
36. C. R. Kordig, *The Justification of Scientific Change*. 1971
ISBN 90-277-0181-4; Pb 90-277-0475-9
37. M. Čapek, *Bergson and Modern Physics*. A Reinterpretation and Re-evaluation. [Boston Studies in the Philosophy of Science, Vol. VII] 1971 ISBN 90-277-0186-5
38. N. R. Hanson, *What I Do Not Believe, and Other Essays*. Ed. by S. Toulmin and H. Woolf. 1971 ISBN 90-277-0191-1
39. R. C. Buck and R. S. Cohen (eds.), *PSA 1970*. Proceedings of the Second Biennial Meeting of the Philosophy of Science Association, Boston, Fall 1970. In Memory of Rudolf Carnap. [Boston Studies in the Philosophy of Science, Vol. VIII] 1971
ISBN 90-277-0187-3; Pb 90-277-0309-4
40. D. Davidson and G. Harman (eds.), *Semantics of Natural Language*. 1972
ISBN 90-277-0304-3; Pb 90-277-0310-8
41. Y. Bar-Hillel (ed.), *Pragmatics of Natural Languages*. 1971
ISBN 90-277-0194-6; Pb 90-277-0599-2
42. S. Stenlund, *Combinators, γ Terms and Proof Theory*. 1972 ISBN 90-277-0305-1
43. M. Strauss, *Modern Physics and Its Philosophy*. Selected Paper in the Logic, History, and Philosophy of Science. 1972 ISBN 90-277-0230-6
44. M. Bunge, *Method, Model and Matter*. 1973 ISBN 90-277-0252-7
45. M. Bunge, *Philosophy of Physics*. 1973 ISBN 90-277-0253-5
46. A. A. Zinov'ev, *Foundations of the Logical Theory of Scientific Knowledge (Complex Logic)*. Revised and enlarged English edition with an appendix by G. A. Smirnov, E. A. Sidorenka, A. M. Fedina and L. A. Bobrova. [Boston Studies in the Philosophy of Science, Vol. IX] 1973
ISBN 90-277-0193-8; Pb 90-277-0324-8
47. L. Tondl, *Scientific Procedures*. A Contribution concerning the Methodological Problems of Scientific Concepts and Scientific Explanation. Translated from Czech by D. Short. Edited by R.S. Cohen and M.W. Wartofsky. [Boston Studies in the Philosophy of Science, Vol. X] 1973
ISBN 90-277-0147-4; Pb 90-277-0323-X
48. N. R. Hanson, *Constellations and Conjectures*. 1973 ISBN 90-277-0192-X

SYNTHESE LIBRARY

49. K. J. J. Hintikka, J. M. E. Moravcsik and P. Suppes (eds.), *Approaches to Natural Language*. 1973 ISBN 90-277-0220-9; Pb 90-277-0233-0
50. M. Bunge (ed.), *Exact Philosophy*. Problems, Tools and Goals. 1973 ISBN 90-277-0251-9
51. R. J. Bogdan and I. Niiniluoto (eds.), *Logic, Language and Probability*. 1973 ISBN 90-277-0312-4
52. G. Pearce and P. Maynard (eds.), *Conceptual Change*. 1973 ISBN 90-277-0287-X; Pb 90-277-0339-6
53. I. Niiniluoto and R. Tuomela, *Theoretical Concepts and Hypothetico-inductive Inference*. 1973 ISBN 90-277-0343-4
54. R. Fraïssé, *Course of Mathematical Logic – Volume 1: Relation and Logical Formula*. Translated from French. 1973 ISBN 90-277-0268-3; Pb 90-277-0403-1
(For Volume 2 see under No. 69).
55. A. Grünbaum, *Philosophical Problems of Space and Time*. Edited by R.S. Cohen and M.W. Wartofsky. 2nd enlarged ed. [Boston Studies in the Philosophy of Science, Vol. XII] 1973 ISBN 90-277-0357-4; Pb 90-277-0358-2
56. P. Suppes (ed.), *Space, Time and Geometry*. 1973 ISBN 90-277-0386-8; Pb 90-277-0442-2
57. H. Kelsen, *Essays in Legal and Moral Philosophy*. Selected and introduced by O. Weinberger. Translated from German by P. Heath. 1973 ISBN 90-277-0388-4
58. R. J. Seeger and R. S. Cohen (eds.), *Philosophical Foundations of Science*. [Boston Studies in the Philosophy of Science, Vol. XI] 1974 ISBN 90-277-0390-6; Pb 90-277-0376-0
59. R. S. Cohen and M. W. Wartofsky (eds.), *Logical and Epistemological Studies in Contemporary Physics*. [Boston Studies in the Philosophy of Science, Vol. XIII] 1973 ISBN 90-277-0391-4; Pb 90-277-0377-9
60. R. S. Cohen and M. W. Wartofsky (eds.), *Methodological and Historical Essays in the Natural and Social Sciences. Proceedings of the Boston Colloquium for the Philosophy of Science, 1969–1972*. [Boston Studies in the Philosophy of Science, Vol. XIV] 1974 ISBN 90-277-0392-2; Pb 90-277-0378-7
61. R. S. Cohen, J. J. Stachel and M. W. Wartofsky (eds.), *For Dirk Struik. Scientific, Historical and Political Essays*. [Boston Studies in the Philosophy of Science, Vol. XV] 1974 ISBN 90-277-0393-0; Pb 90-277-0379-5
62. K. Ajdukiewicz, *Pragmatic Logic*. Translated from Polish by O. Wojtasiewicz. 1974 ISBN 90-277-0326-4
63. S. Stenlund (ed.), *Logical Theory and Semantic Analysis*. Essays dedicated to Stig Kanger on His 50th Birthday. 1974 ISBN 90-277-0438-4
64. K. F. Schaffner and R. S. Cohen (eds.), *PSA 1972. Proceedings of the Third Biennial Meeting of the Philosophy of Science Association*. [Boston Studies in the Philosophy of Science, Vol. XX] 1974 ISBN 90-277-0408-2; Pb 90-277-0409-0
65. H. E. Kyburg, Jr., *The Logical Foundations of Statistical Inference*. 1974 ISBN 90-277-0330-2; Pb 90-277-0430-9
66. M. Grene, *The Understanding of Nature*. Essays in the Philosophy of Biology. [Boston Studies in the Philosophy of Science, Vol. XXIII] 1974 ISBN 90-277-0462-7; Pb 90-277-0463-5
67. J. M. Broekman, *Structuralism: Moscow, Prague, Paris*. Translated from German. 1974 ISBN 90-277-0478-3
68. N. Geschwind, *Selected Papers on Language and the Brain*. [Boston Studies in the Philosophy of Science, Vol. XVI] 1974 ISBN 90-277-0262-4; Pb 90-277-0263-2
69. R. Fraïssé, *Course of Mathematical Logic – Volume 2: Model Theory*. Translated from French. 1974 ISBN 90-277-0269-1; Pb 90-277-0510-0
(For Volume 1 see under No. 54)

SYNTHESE LIBRARY

70. A. Grzegorzczak, *An Outline of Mathematical Logic*. Fundamental Results and Notions explained with all Details. Translated from Polish. 1974 ISBN 90-277-0359-0; Pb 90-277-0447-3
71. F. von Kutschera, *Philosophy of Language*. 1975 ISBN 90-277-0591-7
72. J. Manninen and R. Tuomela (eds.), *Essays on Explanation and Understanding*. Studies in the Foundations of Humanities and Social Sciences. 1976 ISBN 90-277-0592-5
73. J. Hintikka (ed.), *Rudolf Carnap, Logical Empiricist*. Materials and Perspectives. 1975 ISBN 90-277-0583-6
74. M. Čapek (ed.), *The Concepts of Space and Time*. Their Structure and Their Development. [Boston Studies in the Philosophy of Science, Vol. XXII] 1976 ISBN 90-277-0355-8; Pb 90-277-0375-2
75. J. Hintikka and U. Remes, *The Method of Analysis*. Its Geometrical Origin and Its General Significance. [Boston Studies in the Philosophy of Science, Vol. XXV] 1974 ISBN 90-277-0532-1; Pb 90-277-0543-7
76. J. E. Murdoch and E. D. Sylla (eds.), *The Cultural Context of Medieval Learning*. [Boston Studies in the Philosophy of Science, Vol. XXVI] 1975 ISBN 90-277-0560-7; Pb 90-277-0587-9
77. S. Amsterdamski, *Between Experience and Metaphysics*. Philosophical Problems of the Evolution of Science. [Boston Studies in the Philosophy of Science, Vol. XXXV] 1975 ISBN 90-277-0568-2; Pb 90-277-0580-1
78. P. Suppes (ed.), *Logic and Probability in Quantum Mechanics*. 1976 ISBN 90-277-0570-4; Pb 90-277-1200-X
79. H. von Helmholtz: *Epistemological Writings. The Paul Hertz / Moritz Schlick Centenary Edition of 1921 with Notes and Commentary by the Editors*. Newly translated from German by M. F. Lowe. Edited, with an Introduction and Bibliography, by R. S. Cohen and Y. Elkana. [Boston Studies in the Philosophy of Science, Vol. XXXVII] 1975 ISBN 90-277-0290-X; Pb 90-277-0582-8
80. J. Agassi, *Science in Flux*. [Boston Studies in the Philosophy of Science, Vol. XXVIII] 1975 ISBN 90-277-0584-4; Pb 90-277-0612-2
81. S. G. Harding (ed.), *Can Theories Be Refuted? Essays on the Duhem-Quine Thesis*. 1976 ISBN 90-277-0629-8; Pb 90-277-0630-1
82. S. Nowak, *Methodology of Sociological Research*. General Problems. 1977 ISBN 90-277-0486-4
83. J. Piaget, J.-B. Grize, A. Szeminska and V. Bang, *Epistemology and Psychology of Functions*. Translated from French. 1977 ISBN 90-277-0804-5
84. M. Grene and E. Mendelsohn (eds.), *Topics in the Philosophy of Biology*. [Boston Studies in the Philosophy of Science, Vol. XXVII] 1976 ISBN 90-277-0595-X; Pb 90-277-0596-8
85. E. Fischbein, *The Intuitive Sources of Probabilistic Thinking in Children*. 1975 ISBN 90-277-0626-3; Pb 90-277-1190-9
86. E. W. Adams, *The Logic of Conditionals*. An Application of Probability to Deductive Logic. 1975 ISBN 90-277-0631-X
87. M. Przełęcki and R. Wójcicki (eds.), *Twenty-Five Years of Logical Methodology in Poland*. Translated from Polish. 1976 ISBN 90-277-0601-8
88. J. Topolski, *The Methodology of History*. Translated from Polish by O. Wojtasiewicz. 1976 ISBN 90-277-0550-X
89. A. Kasher (ed.), *Language in Focus: Foundations, Methods and Systems*. Essays dedicated to Yehoshua Bar-Hillel. [Boston Studies in the Philosophy of Science, Vol. XLIII] 1976 ISBN 90-277-0644-1; Pb 90-277-0645-X

SYNTHESE LIBRARY

90. J. Hintikka, *The Intentions of Intentionality and Other New Models for Modalities*. 1975
ISBN 90-277-0633-6; Pb 90-277-0634-4
91. W. Stegmüller, *Collected Papers on Epistemology, Philosophy of Science and History of Philosophy*. 2 Volumes. 1977
Set ISBN 90-277-0767-7
92. D. M. Gabbay, *Investigations in Modal and Tense Logics with Applications to Problems in Philosophy and Linguistics*. 1976
ISBN 90-277-0656-5
93. R. J. Bogdan, *Local Induction*. 1976
ISBN 90-277-0649-2
94. S. Nowak, *Understanding and Prediction*. Essays in the Methodology of Social and Behavioral Theories. 1976
ISBN 90-277-0558-5; Pb 90-277-1199-2
95. P. Mittelstaedt, *Philosophical Problems of Modern Physics*. [Boston Studies in the Philosophy of Science, Vol. XVIII] 1976
ISBN 90-277-0285-3; Pb 90-277-0506-2
96. G. Holton and W. A. Blanpied (eds.), *Science and Its Public: The Changing Relationship*. [Boston Studies in the Philosophy of Science, Vol. XXXIII] 1976
ISBN 90-277-0657-3; Pb 90-277-0658-1
97. M. Brand and D. Walton (eds.), *Action Theory*. 1976
ISBN 90-277-0671-9
98. P. Gochet, *Outline of a Nominalist Theory of Propositions*. An Essay in the Theory of Meaning and in the Philosophy of Logic. 1980
ISBN 90-277-1031-7
99. R. S. Cohen, P. K. Feyerabend, and M. W. Wartofsky (eds.), *Essays in Memory of Imre Lakatos*. [Boston Studies in the Philosophy of Science, Vol. XXXIX] 1976
ISBN 90-277-0654-9; Pb 90-277-0655-7
100. R. S. Cohen and J. J. Stachel (eds.), *Selected Papers of Léon Rosenfield*. [Boston Studies in the Philosophy of Science, Vol. XXI] 1979
ISBN 90-277-0651-4; Pb 90-277-0652-2
101. R. S. Cohen, C. A. Hooker, A. C. Michalos and J. W. van Evra (eds.), *PSA 1974. Proceedings of the 1974 Biennial Meeting of the Philosophy of Science Association*. [Boston Studies in the Philosophy of Science, Vol. XXXII] 1976
ISBN 90-277-0647-6; Pb 90-277-0648-4
102. Y. Fried and J. Agassi, *Paranoia*. A Study in Diagnosis. [Boston Studies in the Philosophy of Science, Vol. L] 1976
ISBN 90-277-0704-9; Pb 90-277-0705-7
103. M. Przełęcki, K. Szaniawski and R. Wójcicki (eds.), *Formal Methods in the Methodology of Empirical Sciences*. 1976
ISBN 90-277-0698-0
104. J. M. Vickers, *Belief and Probability*. 1976
ISBN 90-277-0744-8
105. K. H. Wolff, *Surrender and Catch*. Experience and Inquiry Today. [Boston Studies in the Philosophy of Science, Vol. LI] 1976
ISBN 90-277-0758-8; Pb 90-277-0765-0
106. K. Kosik, *Dialectics of the Concrete*. A Study on Problems of Man and World. [Boston Studies in the Philosophy of Science, Vol. LII] 1976
ISBN 90-277-0761-8; Pb 90-277-0764-2
107. N. Goodman, *The Structure of Appearance*. 3rd ed. with an Introduction by G. Hellman. [Boston Studies in the Philosophy of Science, Vol. LIII] 1977
ISBN 90-277-0773-1; Pb 90-277-0774-X
108. K. Ajdukiewicz, *The Scientific World-Perspective and Other Essays, 1931-1963*. Translated from Polish. Edited and with an Introduction by J. Giedymin. 1978
ISBN 90-277-0527-5
109. R. L. Causey, *Unity of Science*. 1977
ISBN 90-277-0779-0
110. R. E. Grandy, *Advanced Logic for Applications*. 1977
ISBN 90-277-0781-2
111. R. P. McArthur, *Tense Logic*. 1976
ISBN 90-277-0697-2
112. L. Lindahl, *Position and Change*. A Study in Law and Logic. Translated from Swedish by P. Needham. 1977
ISBN 90-277-0787-1
113. R. Tuomela, *Dispositions*. 1978
ISBN 90-277-0810-X
114. H. A. Simon, *Models of Discovery and Other Topics in the Methods of Science*. [Boston Studies in the Philosophy of Science, Vol. LIV] 1977
ISBN 90-277-0812-6; Pb 90-277-0858-4

SYNTHESE LIBRARY

115. R. D. Rosenkrantz, *Inference, Method and Decision*. Towards a Bayesian Philosophy of Science. 1977 ISBN 90-277-0817-7; Pb 90-277-0818-5
116. R. Tuomela, *Human Action and Its Explanation*. A Study on the Philosophical Foundations of Psychology. 1977 ISBN 90-277-0824-X
117. M. Lazerowitz, *The Language of Philosophy*. Freud and Wittgenstein. [Boston Studies in the Philosophy of Science, Vol. LV] 1977 ISBN 90-277-0826-6; Pb 90-277-0862-2
118. Not published 119. J. Pelc (ed.), *Semiotics in Poland, 1894–1969*. Translated from Polish. 1979 ISBN 90-277-0811-8
120. I. Pörn, *Action Theory and Social Science*. Some Formal Models. 1977 ISBN 90-277-0846-0
121. J. Margolis, *Persons and Mind*. The Prospects of Nonreductive Materialism. [Boston Studies in the Philosophy of Science, Vol. LVII] 1977 ISBN 90-277-0854-1; Pb 90-277-0863-0
122. J. Hintikka, I. Niiniluoto, and E. Saarinen (eds.), *Essays on Mathematical and Philosophical Logic*. 1979 ISBN 90-277-0879-7
123. T. A. F. Kuipers, *Studies in Inductive Probability and Rational Expectation*. 1978 ISBN 90-277-0882-7
124. E. Saarinen, R. Hilpinen, I. Niiniluoto and M. P. Hintikka (eds.), *Essays in Honour of Jaakko Hintikka on the Occasion of His 50th Birthday*. 1979 ISBN 90-277-0916-5
125. G. Radnitzky and G. Andersson (eds.), *Progress and Rationality in Science*. [Boston Studies in the Philosophy of Science, Vol. LVIII] 1978 ISBN 90-277-0921-1; Pb 90-277-0922-X
126. P. Mittelstaedt, *Quantum Logic*. 1978 ISBN 90-277-0925-4
127. K. A. Bowen, *Model Theory for Modal Logic*. Kripke Models for Modal Predicate Calculi. 1979 ISBN 90-277-0929-7
128. H. A. Bursen, *Dismantling the Memory Machine*. A Philosophical Investigation of Machine Theories of Memory. 1978 ISBN 90-277-0933-5
129. M. W. Wartofsky, *Models*. Representation and the Scientific Understanding. [Boston Studies in the Philosophy of Science, Vol. XLVIII] 1979 ISBN 90-277-0736-7; Pb 90-277-0947-5
130. D. Ihde, *Technics and Praxis*. A Philosophy of Technology. [Boston Studies in the Philosophy of Science, Vol. XXIV] 1979 ISBN 90-277-0953-X; Pb 90-277-0954-8
131. J. J. Wiatr (ed.), *Polish Essays in the Methodology of the Social Sciences*. [Boston Studies in the Philosophy of Science, Vol. XXIX] 1979 ISBN 90-277-0723-5; Pb 90-277-0956-4
132. W. C. Salmon (ed.), *Hans Reichenbach: Logical Empiricist*. 1979 ISBN 90-277-0958-0
133. P. Bieri, R.-P. Horstmann and L. Krüger (eds.), *Transcendental Arguments in Science*. Essays in Epistemology. 1979 ISBN 90-277-0963-7; Pb 90-277-0964-5
134. M. Marković and G. Petrović (eds.), *Praxis*. Yugoslav Essays in the Philosophy and Methodology of the Social Sciences. [Boston Studies in the Philosophy of Science, Vol. XXXVI] 1979 ISBN 90-277-0727-8; Pb 90-277-0968-8
135. R. Wójcicki, *Topics in the Formal Methodology of Empirical Sciences*. Translated from Polish. 1979 ISBN 90-277-1004-X
136. G. Radnitzky and G. Andersson (eds.), *The Structure and Development of Science*. [Boston Studies in the Philosophy of Science, Vol. LIX] 1979 ISBN 90-277-0994-7; Pb 90-277-0995-5
137. J. C. Webb, *Mechanism, Mentalism and Metamathematics*. An Essay on Finitism. 1980 ISBN 90-277-1046-5
138. D. F. Gustafson and B. L. Tapscott (eds.), *Body, Mind and Method*. Essays in Honor of Virgil C. Aldrich. 1979 ISBN 90-277-1013-9
139. L. Nowak, *The Structure of Idealization*. Towards a Systematic Interpretation of the Marxian Idea of Science. 1980 ISBN 90-277-1014-7

SYNTHESE LIBRARY

140. C. Perelman, *The New Rhetoric and the Humanities*. Essays on Rhetoric and Its Applications. Translated from French and German. With an Introduction by H. Zyskind. 1979
ISBN 90-277-1018-X; Pb 90-277-1019-8
141. W. Rabinowicz, *Universalizability*. A Study in Morals and Metaphysics. 1979
ISBN 90-277-1020-2
142. C. Perelman, *Justice, Law and Argument*. Essays on Moral and Legal Reasoning. Translated from French and German. With an Introduction by H.J. Berman. 1980
ISBN 90-277-1089-9; Pb 90-277-1090-2
143. S. Kanger and S. Öhman (eds.), *Philosophy and Grammar*. Papers on the Occasion of the Quincentennial of Uppsala University. 1981
ISBN 90-277-1091-0
144. T. Pawlowski, *Concept Formation in the Humanities and the Social Sciences*. 1980
ISBN 90-277-1096-1
145. J. Hintikka, D. Gruender and E. Agazzi (eds.), *Theory Change, Ancient Axiomatics and Galileo's Methodology*. Proceedings of the 1978 Pisa Conference on the History and Philosophy of Science, Volume I. 1981
ISBN 90-277-1126-7
146. J. Hintikka, D. Gruender and E. Agazzi (eds.), *Probabilistic Thinking, Thermodynamics, and the Interaction of the History and Philosophy of Science*. Proceedings of the 1978 Pisa Conference on the History and Philosophy of Science, Volume II. 1981
ISBN 90-277-1127-5
147. U. Mönnich (ed.), *Aspects of Philosophical Logic*. Some Logical Forays into Central Notions of Linguistics and Philosophy. 1981
ISBN 90-277-1201-8
148. D. M. Gabbay, *Semantical Investigations in Heyting's Intuitionistic Logic*. 1981
ISBN 90-277-1202-6
149. E. Agazzi (ed.), *Modern Logic – A Survey*. Historical, Philosophical, and Mathematical Aspects of Modern Logic and Its Applications. 1981
ISBN 90-277-1137-2
150. A. F. Parker-Rhodes, *The Theory of Indistinguishables*. A Search for Explanatory Principles below the Level of Physics. 1981
ISBN 90-277-1214-X
151. J. C. Pitt, *Pictures, Images, and Conceptual Change*. An Analysis of Wilfrid Sellars' Philosophy of Science. 1981
ISBN 90-277-1276-X; Pb 90-277-1277-8
152. R. Hilpinen (ed.), *New Studies in Deontic Logic*. Norms, Actions, and the Foundations of Ethics. 1981
ISBN 90-277-1278-6; Pb 90-277-1346-4
153. C. Dilworth, *Scientific Progress*. A Study Concerning the Nature of the Relation between Successive Scientific Theories. 4rd Rev. ed. 2007
ISBN 978-1-4020-6353-1; 978-1-4020-6354-1 (e-book)
154. D. Woodruff Smith and R. McIntyre, *Husserl and Intentionality*. A Study of Mind, Meaning, and Language. 1982
ISBN 90-277-1392-8; Pb 90-277-1730-3
155. R. J. Nelson, *The Logic of Mind*. 2nd. ed., 1989
ISBN 90-277-2819-4; Pb 90-277-2822-4
156. J. F. A. K. van Benthem, *The Logic of Time*. A Model-Theoretic Investigation into the Varieties of Temporal Ontology, and Temporal Discourse. 1983; 2nd ed., 1991
ISBN 0-7923-1081-0
157. R. Swinburne (ed.), *Space, Time and Causality*. 1983
ISBN 90-277-1437-1
158. E. T. Jaynes, *Papers on Probability, Statistics and Statistical Physics*. Ed. by R. D. Rozenkrantz. 1983
ISBN 90-277-1448-7; Pb (1989) 0-7923-0213-3
159. T. Chapman, *Time: A Philosophical Analysis*. 1982
ISBN 90-277-1465-7
160. E. N. Zalta, *Abstract Objects*. An Introduction to Axiomatic Metaphysics. 1983
ISBN 90-277-1474-6
161. S. Harding and M. B. Hintikka (eds.), *Discovering Reality*. Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science. 1983
ISBN 90-277-1496-7; Pb 90-277-1538-6
162. M. A. Stewart (ed.), *Law, Morality and Rights*. 1983
ISBN 90-277-1519-X

SYNTHESE LIBRARY

163. D. Mayr and G. Stüssmann (eds.), *Space, Time, and Mechanics*. Basic Structures of a Physical Theory. 1983 ISBN 90-277-1525-4
164. D. Gabbay and F. Guentner (eds.), *Handbook of Philosophical Logic*. Vol. I: Elements of Classical Logic. 1983 ISBN 90-277-1542-4
165. D. Gabbay and F. Guentner (eds.), *Handbook of Philosophical Logic*. Vol. II: Extensions of Classical Logic. 1984 ISBN 90-277-1604-8
166. D. Gabbay and F. Guentner (eds.), *Handbook of Philosophical Logic*. Vol. III: Alternative to Classical Logic. 1986 ISBN 90-277-1605-6
167. D. Gabbay and F. Guentner (eds.), *Handbook of Philosophical Logic*. Vol. IV: Topics in the Philosophy of Language. 1989 ISBN 90-277-1606-4
168. A. J. I. Jones, *Communication and Meaning*. An Essay in Applied Modal Logic. 1983 ISBN 90-277-1543-2
169. M. Fitting, *Proof Methods for Modal and Intuitionistic Logics*. 1983 ISBN 90-277-1573-4
170. J. Margolis, *Culture and Cultural Entities*. Toward a New Unity of Science. 1984 ISBN 90-277-1574-2
171. R. Tuomela, *A Theory of Social Action*. 1984 ISBN 90-277-1703-6
172. J. J. E. Gracia, E. Rabossi, E. Villanueva and M. Dascal (eds.), *Philosophical Analysis in Latin America*. 1984 ISBN 90-277-1749-4
173. P. Ziff, *Epistemic Analysis*. A Coherence Theory of Knowledge. 1984 ISBN 90-277-1751-7
174. P. Ziff, *Antiaesthetics*. An Appreciation of the Cow with the Subtile Nose. 1984 ISBN 90-277-1773-7
175. W. Balzer, D. A. Pearce, and H.-J. Schmidt (eds.), *Reduction in Science*. Structure, Examples, Philosophical Problems. 1984 ISBN 90-277-1811-3
176. A. Peczenik, L. Lindahl and B. van Roermund (eds.), *Theory of Legal Science*. Proceedings of the Conference on Legal Theory and Philosophy of Science (Lund, Sweden, December 1983). 1984 ISBN 90-277-1834-2
177. I. Niiniluoto, *Is Science Progressive?* 1984 ISBN 90-277-1835-0
178. B. K. Matilal and J. L. Shaw (eds.), *Analytical Philosophy in Comparative Perspective*. Exploratory Essays in Current Theories and Classical Indian Theories of Meaning and Reference. 1985 ISBN 90-277-1870-9
179. P. Kroes, *Time: Its Structure and Role in Physical Theories*. 1985 ISBN 90-277-1894-6
180. J. H. Fetzer, *Sociobiology and Epistemology*. 1985 ISBN 90-277-2005-3; Pb 90-277-2006-1
181. L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized*. Essays on the Philosophical and Foundational Work of Gottlob Frege. 1986 ISBN 90-277-2126-2
182. M. Detlefsen, *Hilbert's Program*. An Essay on Mathematical Instrumentalism. 1986 ISBN 90-277-2151-3
183. J. L. Golden and J. J. Pilotta (eds.), *Practical Reasoning in Human Affairs*. Studies in Honor of Chaim Perelman. 1986 ISBN 90-277-2255-2
184. H. Zandvoort, *Models of Scientific Development and the Case of Nuclear Magnetic Resonance*. 1986 ISBN 90-277-2351-6
185. I. Niiniluoto, *Truthlikeness*. 1987 ISBN 90-277-2354-0
186. W. Balzer, C. U. Moulines and J. D. Sneed, *An Architectonic for Science*. The Structuralist Program. 1987 ISBN 90-277-2403-2
187. D. Pearce, *Roads to Commensurability*. 1987 ISBN 90-277-2414-8
188. L. M. Vaina (ed.), *Matters of Intelligence*. Conceptual Structures in Cognitive Neuroscience. 1987 ISBN 90-277-2460-1

SYNTHESE LIBRARY

189. H. Siegel, *Relativism Refuted*. A Critique of Contemporary Epistemological Relativism. 1987
ISBN 90-277-2469-5
190. W. Callebaut and R. Pinxten, *Evolutionary Epistemology*. A Multiparadigm Program, with a Complete Evolutionary Epistemology Bibliograph. 1987
ISBN 90-277-2582-9
191. J. Kmita, *Problems in Historical Epistemology*. 1988
ISBN 90-277-2199-8
192. J. H. Fetzer (ed.), *Probability and Causality*. Essays in Honor of Wesley C. Salmon, with an Annotated Bibliography. 1988
ISBN 90-277-2607-8; Pb 1-5560-8052-2
193. A. Donovan, L. Laudan and R. Laudan (eds.), *Scrutinizing Science*. Empirical Studies of Scientific Change. 1988
ISBN 90-277-2608-6
194. H.R. Otto and J.A. Tuedio (eds.), *Perspectives on Mind*. 1988
ISBN 90-277-2640-X
195. D. Batens and J.P. van Bendegem (eds.), *Theory and Experiment*. Recent Insights and New Perspectives on Their Relation. 1988
ISBN 90-277-2645-0
196. J. Österberg, *Self and Others*. A Study of Ethical Egoism. 1988
ISBN 90-277-2648-5
197. D.H. Helman (ed.), *Analogical Reasoning*. Perspectives of Artificial Intelligence, Cognitive Science, and Philosophy. 1988
ISBN 90-277-2711-2
198. J. Woleński, *Logic and Philosophy in the Lvov-Warsaw School*. 1989
ISBN 90-277-2749-X
199. R. Wójcicki, *Theory of Logical Calculi*. Basic Theory of Consequence Operations. 1988
ISBN 90-277-2785-6
200. J. Hintikka and M.B. Hintikka, *The Logic of Epistemology and the Epistemology of Logic*. Selected Essays. 1989
ISBN 0-7923-0040-8; Pb 0-7923-0041-6
201. E. Agazzi (ed.), *Probability in the Sciences*. 1988
ISBN 90-277-2808-9
202. M. Meyer (ed.), *From Metaphysics to Rhetoric*. 1989
ISBN 90-277-2814-3
203. R.L. Tieszen, *Mathematical Intuition*. Phenomenology and Mathematical Knowledge. 1989
ISBN 0-7923-0131-5
204. A. Melnick, *Space, Time, and Thought in Kant*. 1989
ISBN 0-7923-0135-8
205. D.W. Smith, *The Circle of Acquaintance*. Perception, Consciousness, and Empathy. 1989
ISBN 0-7923-0252-4
206. M.H. Salmon (ed.), *The Philosophy of Logical Mechanism*. Essays in Honor of Arthur W. Burks. With his Responses, and with a Bibliography of Burk's Work. 1990
ISBN 0-7923-0325-3
207. M. Kusch, *Language as Calculus vs. Language as Universal Medium*. A Study in Husserl, Heidegger, and Gadamer. 1989
ISBN 0-7923-0333-4
208. T.C. Meyering, *Historical Roots of Cognitive Science*. The Rise of a Cognitive Theory of Perception from Antiquity to the Nineteenth Century. 1989
ISBN 0-7923-0349-0
209. P. Kosso, *Observability and Observation in Physical Science*. 1989
ISBN 0-7923-0389-X
210. J. Kmita, *Essays on the Theory of Scientific Cognition*. 1990
ISBN 0-7923-0441-1
211. W. Sieg (ed.), *Acting and Reflecting*. The Interdisciplinary Turn in Philosophy. 1990
ISBN 0-7923-0512-4
212. J. Karpinśki, *Causality in Sociological Research*. 1990
ISBN 0-7923-0546-9
213. H.A. Lewis (ed.), *Peter Geach: Philosophical Encounters*. 1991
ISBN 0-7923-0823-9
214. M. Ter Hark, *Beyond the Inner and the Outer*. Wittgenstein's Philosophy of Psychology. 1990
ISBN 0-7923-0850-6
215. M. Gosselin, *Nominalism and Contemporary Nominalism*. Ontological and Epistemological Implications of the Work of W.V.O. Quine and of N. Goodman. 1990
ISBN 0-7923-0904-9
216. J.H. Fetzer, D. Shatz and G. Schlesinger (eds.), *Definitions and Definability*. Philosophical Perspectives. 1991
ISBN 0-7923-1046-2
217. E. Agazzi and A. Cordero (eds.), *Philosophy and the Origin and Evolution of the Universe*. 1991
ISBN 0-7923-1322-4

SYNTHESE LIBRARY

218. M. Kusch, *Foucault's Strata and Fields*. An Investigation into Archaeological and Genealogical Science Studies. 1991 ISBN 0-7923-1462-X
219. C.J. Posy, *Kant's Philosophy of Mathematics*. Modern Essays. 1992 ISBN 0-7923-1495-6
220. G. Van de Vijver, *New Perspectives on Cybernetics*. Self-Organization, Autonomy and Connectionism. 1992 ISBN 0-7923-1519-7
221. J.C. Nyíri, *Tradition and Individuality*. Essays. 1992 ISBN 0-7923-1566-9
222. R. Howell, *Kant's Transcendental Deduction*. An Analysis of Main Themes in His Critical Philosophy. 1992 ISBN 0-7923-1571-5
223. A. García de la Sienra, *The Logical Foundations of the Marxian Theory of Value*. 1992 ISBN 0-7923-1778-5
224. D.S. Shwayder, *Statement and Referent*. An Inquiry into the Foundations of Our Conceptual Order. 1992 ISBN 0-7923-1803-X
225. M. Rosen, *Problems of the Hegelian Dialectic*. Dialectic Reconstructed as a Logic of Human Reality. 1993 ISBN 0-7923-2047-6
226. P. Suppes, *Models and Methods in the Philosophy of Science: Selected Essays*. 1993 ISBN 0-7923-2211-8
227. R. M. Dancy (ed.), *Kant and Critique: New Essays in Honor of W. H. Werkmeister*. 1993 ISBN 0-7923-2244-4
228. J. Woleński (ed.), *Philosophical Logic in Poland*. 1993 ISBN 0-7923-2293-2
229. M. De Rijke (ed.), *Diamonds and Defaults*. Studies in Pure and Applied Intensional Logic. 1993 ISBN 0-7923-2342-4
230. B.K. Matilal and A. Chakrabarti (eds.), *Knowing from Words*. Western and Indian Philosophical Analysis of Understanding and Testimony. 1994 ISBN 0-7923-2345-9
231. S.A. Kleiner, *The Logic of Discovery*. A Theory of the Rationality of Scientific Research. 1993 ISBN 0-7923-2371-8
232. R. Festa, *Optimum Inductive Methods*. A Study in Inductive Probability, Bayesian Statistics, and Verisimilitude. 1993 ISBN 0-7923-2460-9
233. P. Humphreys (ed.), *Patrick Suppes: Scientific Philosopher*. Vol. 1: Probability and Probabilistic Causality. 1994 ISBN 0-7923-2552-4
234. P. Humphreys (ed.), *Patrick Suppes: Scientific Philosopher*. Vol. 2: Philosophy of Physics, Theory Structure, and Measurement Theory. 1994 ISBN 0-7923-2553-2
235. P. Humphreys (ed.), *Patrick Suppes: Scientific Philosopher*. Vol. 3: Language, Logic, and Psychology. 1994 ISBN 0-7923-2862-0
- Set ISBN (Vols 233–235) 0-7923-2554-0
236. D. Prawitz and D. Westerståhl (eds.), *Logic and Philosophy of Science in Uppsala*. Papers from the 9th International Congress of Logic, Methodology, and Philosophy of Science. 1994 ISBN 0-7923-2702-0
237. L. Haaparanta (ed.), *Mind, Meaning and Mathematics*. Essays on the Philosophical Views of Husserl and Frege. 1994 ISBN 0-7923-2703-9
238. J. Hintikka (ed.), *Aspects of Metaphor*. 1994 ISBN 0-7923-2786-1
239. B. McGuinness and G. Oliveri (eds.), *The Philosophy of Michael Dummett*. With Replies from Michael Dummett. 1994 ISBN 0-7923-2804-3
240. D. Jamieson (ed.), *Language, Mind, and Art*. Essays in Appreciation and Analysis, In Honor of Paul Ziff. 1994 ISBN 0-7923-2810-8
241. G. Preyer, F. Siebelt and A. Ulfig (eds.), *Language, Mind and Epistemology*. On Donald Davidson's Philosophy. 1994 ISBN 0-7923-2811-6
242. P. Ehrlich (ed.), *Real Numbers, Generalizations of the Reals, and Theories of Continua*. 1994 ISBN 0-7923-2689-X

SYNTHESE LIBRARY

243. G. Debrock and M. Hulswit (eds.), *Living Doubt*. Essays concerning the epistemology of Charles Sanders Peirce. 1994 ISBN 0-7923-2898-1
244. J. Szrednicki, *To Know or Not to Know*. Beyond Realism and Anti-Realism. 1994 ISBN 0-7923-2909-0
245. R. Egidi (ed.), *Wittgenstein: Mind and Language*. 1995 ISBN 0-7923-3171-0
246. A. Hyslop, *Other Minds*. 1995 ISBN 0-7923-3245-8
247. L. Pólos and M. Masuch (eds.), *Applied Logic: How, What and Why*. Logical Approaches to Natural Language. 1995 ISBN 0-7923-3432-9
248. M. Krynicki, M. Mostowski and L.M. Szczerba (eds.), *Quantifiers: Logics, Models and Computation*. Volume One: Surveys. 1995 ISBN 0-7923-3448-5
249. M. Krynicki, M. Mostowski and L.M. Szczerba (eds.), *Quantifiers: Logics, Models and Computation*. Volume Two: Contributions. 1995 ISBN 0-7923-3449-3
Set ISBN (Vols 248 + 249) 0-7923-3450-7
250. R.A. Watson, *Representational Ideas from Plato to Patricia Churchland*. 1995 ISBN 0-7923-3453-1
251. J. Hintikka (ed.), *From Dedekind to Gödel*. Essays on the Development of the Foundations of Mathematics. 1995 ISBN 0-7923-3484-1
252. A. Wiśniewski, *The Posing of Questions*. Logical Foundations of Erotetic Inferences. 1995 ISBN 0-7923-3637-2
253. J. Peregrin, *Doing Worlds with Words*. Formal Semantics without Formal Metaphysics. 1995 ISBN 0-7923-3742-5
254. I.A. Kieseppä, *Truthlikeness for Multidimensional, Quantitative Cognitive Problems*. 1996 ISBN 0-7923-4005-1
255. P. Hugly and C. Sayward: *Intensionality and Truth*. An Essay on the Philosophy of A.N. Prior. 1996 ISBN 0-7923-4119-8
256. L. Hankinson Nelson and J. Nelson (eds.): *Feminism, Science, and the Philosophy of Science*. 1997 ISBN 0-7923-4162-7
257. P.I. Bystrov and V.N. Sadovsky (eds.): *Philosophical Logic and Logical Philosophy*. Essays in Honour of Vladimir A. Smirnov. 1996 ISBN 0-7923-4270-4
258. Å.E. Andersson and N-E. Sahlin (eds.): *The Complexity of Creativity*. 1996 ISBN 0-7923-4346-8
259. M.L. Dalla Chiara, K. Doets, D. Mundici and J. van Benthem (eds.): *Logic and Scientific Methods*. Volume One of the Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence, August 1995. 1997 ISBN 0-7923-4383-2
260. M.L. Dalla Chiara, K. Doets, D. Mundici and J. van Benthem (eds.): *Structures and Norms in Science*. Volume Two of the Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence, August 1995. 1997 ISBN 0-7923-4384-0
Set ISBN (Vols 259 + 260) 0-7923-4385-9
261. A. Chakrabarti: *Denying Existence*. The Logic, Epistemology and Pragmatics of Negative Existentials and Fictional Discourse. 1997 ISBN 0-7923-4388-3
262. A. Biletzki: *Talking Wolves*. Thomas Hobbes on the Language of Politics and the Politics of Language. 1997 ISBN 0-7923-4425-1
263. D. Nute (ed.): *Defeasible Deontic Logic*. 1997 ISBN 0-7923-4630-0
264. U. Meixner: *Axiomatic Formal Ontology*. 1997 ISBN 0-7923-4747-X
265. I. Brinck: *The Indexical 'I'*. The First Person in Thought and Language. 1997 ISBN 0-7923-4741-2
266. G. Hölmström-Hintikka and R. Tuomela (eds.): *Contemporary Action Theory*. Volume 1: Individual Action. 1997 ISBN 0-7923-4753-6; Set: 0-7923-4754-4

SYNTHESE LIBRARY

267. G. Hölmström-Hintikka and R. Tuomela (eds.): *Contemporary Action Theory*. Volume 2: Social Action. 1997 ISBN 0-7923-4752-8; Set: 0-7923-4754-4
268. B.-C. Park: *Phenomenological Aspects of Wittgenstein's Philosophy*. 1998 ISBN 0-7923-4813-3
269. J. Paśniczek: *The Logic of Intentional Objects*. A Meinongian Version of Classical Logic. 1998 Hb ISBN 0-7923-4880-X; Pb ISBN 0-7923-5578-4
270. P.W. Humphreys and J.H. Fetzer (eds.): *The New Theory of Reference*. Kripke, Marcus, and Its Origins. 1998 ISBN 0-7923-4898-2
271. K. Szaniawski, A. Chmielewski and J. Woleński (eds.): *On Science, Inference, Information and Decision Making*. Selected Essays in the Philosophy of Science. 1998 ISBN 0-7923-4922-9
272. G.H. von Wright: *In the Shadow of Descartes*. Essays in the Philosophy of Mind. 1998 ISBN 0-7923-4992-X
273. K. Kijania-Placek and J. Woleński (eds.): *The Lvov-Warsaw School and Contemporary Philosophy*. 1998 ISBN 0-7923-5105-3
274. D. Dedrick: *Naming the Rainbow*. Colour Language, Colour Science, and Culture. 1998 ISBN 0-7923-5239-4
275. L. Albertazzi (ed.): *Shapes of Forms*. From Gestalt Psychology and Phenomenology to Ontology and Mathematics. 1999 ISBN 0-7923-5246-7
276. P. Fletcher: *Truth, Proof and Infinity*. A Theory of Constructions and Constructive Reasoning. 1998 ISBN 0-7923-5262-9
277. M. Fitting and R.L. Mendelsohn (eds.): *First-Order Modal Logic*. 1998 Hb ISBN 0-7923-5334-X; Pb ISBN 0-7923-5335-8
278. J.N. Mohanty: *Logic, Truth and the Modalities from a Phenomenological Perspective*. 1999 ISBN 0-7923-5550-4
279. T. Placek: *Mathematical Intuitionism and Intersubjectivity*. A Critical Exposition of Arguments for Intuitionism. 1999 ISBN 0-7923-5630-6
280. A. Cantini, E. Casari and P. Minari (eds.): *Logic and Foundations of Mathematics*. 1999 ISBN 0-7923-5659-4 set ISBN 0-7923-5867-8
281. M.L. Dalla Chiara, R. Giuntini and F. Laudisa (eds.): *Language, Quantum, Music*. 1999 ISBN 0-7923-5727-2; set ISBN 0-7923-5867-8
282. R. Egidi (ed.): *In Search of a New Humanism*. The Philosophy of Georg Hendrik von Wright. 1999 ISBN 0-7923-5810-4
283. F. Vollmer: *Agent Causality*. 1999 ISBN 0-7923-5848-1
284. J. Peregrin (ed.): *Truth and Its Nature (if Any)*. 1999 ISBN 0-7923-5865-1
285. M. De Caro (ed.): *Interpretations and Causes*. New Perspectives on Donald Davidson's Philosophy. 1999 ISBN 0-7923-5869-4
286. R. Murawski: *Recursive Functions and Metamathematics*. Problems of Completeness and Decidability, Gödel's Theorems. 1999 ISBN 0-7923-5904-6
287. T.A.F. Kuipers: *From Instrumentalism to Constructive Realism*. On Some Relations between Confirmation, Empirical Progress, and Truth Approximation. 2000 ISBN 0-7923-6086-9
288. G. Holmström-Hintikka (ed.): *Medieval Philosophy and Modern Times*. 2000 ISBN 0-7923-6102-4
289. E. Grosholz and H. Breger (eds.): *The Growth of Mathematical Knowledge*. 2000 ISBN 0-7923-6151-2

SYNTHESE LIBRARY

290. G. Sommaruga: *History and Philosophy of Constructive Type Theory*. 2000
ISBN 0-7923-6180-6
291. J. Gasser (ed.): *A Boole Anthology*. Recent and Classical Studies in the Logic of George Boole. 2000
ISBN 0-7923-6380-9
292. V.F. Hendricks, S.A. Pedersen and K.F. Jørgensen (eds.): *Proof Theory*. History and Philosophical Significance. 2000
ISBN 0-7923-6544-5
293. W.L. Craig: *The Tensed Theory of Time*. A Critical Examination. 2000
ISBN 0-7923-6634-4
294. W.L. Craig: *The Tenseless Theory of Time*. A Critical Examination. 2000
ISBN 0-7923-6635-2
295. L. Albertazzi (ed.): *The Dawn of Cognitive Science*. Early European Contributors. 2001
ISBN 0-7923-6799-5
296. G. Forrai: *Reference, Truth and Conceptual Schemes*. A Defense of Internal Realism. 2001
ISBN 0-7923-6885-1
297. V.F. Hendricks, S.A. Pedersen and K.F. Jørgensen (eds.): *Probability Theory*. Philosophy, Recent History and Relations to Science. 2001
ISBN 0-7923-6952-1
298. M. Esfeld: *Holism in Philosophy of Mind and Philosophy of Physics*. 2001
ISBN 0-7923-7003-1
299. E.C. Steinhart: *The Logic of Metaphor*. Analogous Parts of Possible Worlds. 2001
ISBN 0-7923-7004-X
300. P. Gärdenfors: *The Dynamics of Thought*. 2005
ISBN 1-4020-3398-2
301. T.A.F. Kuipers: *Structures in Science Heuristic Patterns Based on Cognitive Structures*. An Advanced Textbook in Neo-Classical Philosophy of Science. 2001
ISBN 0-7923-7117-8
302. G. Hon and S.S. Rakover (eds.): *Explanation*. Theoretical Approaches and Applications. 2001
ISBN 1-4020-0017-0
303. G. Holmström-Hintikka, S. Lindström and R. Sliwinski (eds.): *Collected Papers of Stig Kanger with Essays on his Life and Work*. Vol. I. 2001
ISBN 1-4020-0021-9; Pb ISBN 1-4020-0022-7
304. G. Holmström-Hintikka, S. Lindström and R. Sliwinski (eds.): *Collected Papers of Stig Kanger with Essays on his Life and Work*. Vol. II. 2001
ISBN 1-4020-0111-8; Pb ISBN 1-4020-0112-6
305. C.A. Anderson and M. Zelény (eds.): *Logic, Meaning and Computation*. Essays in Memory of Alonzo Church. 2001
ISBN 1-4020-0141-X
306. P. Schuster, U. Berger and H. Osswald (eds.): *Reuniting the Antipodes – Constructive and Nonstandard Views of the Continuum*. 2001
ISBN 1-4020-0152-5
307. S.D. Zwart: *Refined Verisimilitude*. 2001
ISBN 1-4020-0268-8
308. A.-S. Maurin: *If Tropes*. 2002
ISBN 1-4020-0656-X
309. H. Eilstein (ed.): *A Collection of Polish Works on Philosophical Problems of Time and Space-time*. 2002
ISBN 1-4020-0670-5
310. Y. Gauthier: *Internal Logic*. Foundations of Mathematics from Kronecker to Hilbert. 2002
ISBN 1-4020-0689-6
311. E. Ruttkamp: *A Model-Theoretic Realist Interpretation of Science*. 2002
ISBN 1-4020-0729-9
312. V. Rantala: *Explanatory Translation*. Beyond the Kuhnian Model of Conceptual Change. 2002
ISBN 1-4020-0827-9
313. L. Decock: *Trading Ontology for Ideology*. 2002
ISBN 1-4020-0865-1

SYNTHESE LIBRARY

314. O. Ezra: *The Withdrawal of Rights*. Rights from a Different Perspective. 2002
ISBN 1-4020-0886-4
315. P. Gärdenfors, J. Woleński and K. Kijania-Placek: *In the Scope of Logic, Methodology and Philosophy of Science*. Volume One of the 11th International Congress of Logic, Methodology and Philosophy of Science, Cracow, August 1999. 2002
ISBN 1-4020-0929-1; Pb 1-4020-0931-3
316. P. Gärdenfors, J. Woleński and K. Kijania-Placek: *In the Scope of Logic, Methodology and Philosophy of Science*. Volume Two of the 11th International Congress of Logic, Methodology and Philosophy of Science, Cracow, August 1999. 2002
ISBN 1-4020-0930-5; Pb 1-4020-0931-3
317. M.A. Changizi: *The Brain from 25,000 Feet*. High Level Explorations of Brain Complexity, Perception, Induction and Vagueness. 2003
ISBN 1-4020-1176-8
318. D.O. Dahlstrom (ed.): *Husserl's Logical Investigations*. 2003
ISBN 1-4020-1325-6
319. A. Biletzki: *(Over)Interpreting Wittgenstein*. 2003
ISBN Hb 1-4020-1326-4; Pb 1-4020-1327-2
320. A. Rojszczak, J. Cachro and G. Kurczewski (eds.): *Philosophical Dimensions of Logic and Science*. Selected Contributed Papers from the 11th International Congress of Logic, Methodology, and Philosophy of Science, Kraków, 1999. 2003
ISBN 1-4020-1645-X
321. M. Sintonen, P. Ylikoski and K. Miller (eds.): *Realism in Action*. Essays in the Philosophy of the Social Sciences. 2003
ISBN 1-4020-1667-0
322. V.F. Hendricks, K.F. Jørgensen and S.A. Pedersen (eds.): *Knowledge Contributors*. 2003
ISBN Hb 1-4020-1747-2; Pb 1-4020-1748-0
323. J. Hintikka, T. Czarnecki, K. Kijania-Placek, T. Placek and A. Rojszczak † (eds.): *Philosophy and Logic In Search of the Polish Tradition*. Essays in Honour of Jan Woleński on the Occasion of his 60th Birthday. 2003
ISBN 1-4020-1721-9
324. L.M. Vaina, S.A. Beardsley and S.K. Rushton (eds.): *Optic Flow and Beyond*. 2004
ISBN 1-4020-2091-0
325. D. Kolak (ed.): *I Am You*. The Metaphysical Foundations of Global Ethics. 2004
ISBN 1-4020-2999-3
326. V. Stepin: *Theoretical Knowledge*. 2005
ISBN 1-4020-3045-2
327. P. Mancosu, K.F. Jørgensen and S.A. Pedersen (eds.): *Visualization, Explanation and Reasoning Styles in Mathematics*. 2005
ISBN 1-4020-3334-6
328. A. Rojszczak (author) and J. Wolenski (ed.): *From the Act of Judging to the Sentence*. The Problem of Truth Bearers from Bolzano to Tarski. 2005
ISBN 1-4020-3396-6
329. A. Pietarinen: *Signs of Logic*. Peircean Themes on the Philosophy of Language, Games, and Communication. 2005
ISBN 1-4020-3728-7
330. A. Aliseda: *Abductive Reasoning*. Logical Investigations into Discovery and Explanation. 2005
ISBN 1-4020-3906-9
331. B. Feltz, M. Crommelinck and P. Goujon (eds.): *Self-organization and Emergence in Life Sciences*. 2005
ISBN 1-4020-3916-6
- 332.
333. L. Albertazzi: *Immanent Realism*. An Introduction to Brentano. 2006
ISBN 1-4020-4201-9
334. A. Keupink and S. Shieh (eds.): *The Limits of Logical Empiricism*. Selected Papers of Arthur Pap. 2006
ISBN 1-4020-4298-1
335. M. van Atten: *Brouwer meets Husserl*. On the Phenomenology of Choice Sequences. 2006
ISBN 1-4020-5086-0
336. F. Stjernfelt: *Diagrammatology*. An Investigation on the Borderlines of Phenomenology, Ontology, and Semiotics. 2007
ISBN 978-1-4020-5651-2

SYNTHESE LIBRARY

337. H. van Ditmarsch, W. van der Hoek and B. Kooi (eds.): *Dynamic Epistemic Logic*. 2007
ISBN 978-1-4020-5838-7; Pb 978-1-4020-6908-6
338. N. Nottelmann: *Blameworthy Belief*. A Study in Epistemic Deontology. 2007
ISBN 978-1-4020-5960-5
339. N.B. Cocchiarella: *Formal Ontology and Conceptual Realism*. 2007
ISBN 978-1-4020-6203-2
340. A. Mura (ed.): *Bruno de Finetti: Philosophical Lectures on Probability*. Collected, edited,
and annotated by Alberto Mura. 2008
ISBN 978-1-4020-8202-3