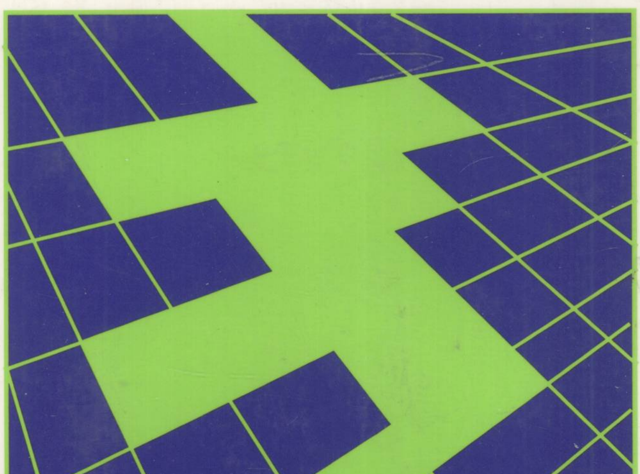
Cambridge Studies in Probability, Induction, and Decision Theory

FOUNDATIONS OF PROBABILITY WITH APPLICATIONS

Selected Papers 1974-1995



PATRICK SUPPES

MARIO ZANOTTI

Foundations of probability with applications

Selected papers, 1974–1995

Patrick Suppes
Stanford University

Mario Zanotti
Computer Curriculum Corporation

CAMBRIDGE
UNIVERSITY PRESS

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1996

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Suppes, Patrick, 1922–

Foundations of probability with applications: selected papers
1974–1995 / Patrick Suppes, Mario Zanotti.

p. cm. – (Cambridge Studies in probability, induction, and decision theory)

ISBN 0-521-43012-7. – ISBN 0-521-56835-8 (pbk.)

1. Probabilities. I. Zanotti, Mario. II. Title. III. Series.
QA273.18.S86 1996
519.2 – dc20 96-7895
CIP

A catalog record for this book is available from the British Library.

ISBN 0-521-43012-7 Hardback
ISBN 0-521-56835-8 Paperback

This is an important collection of papers on the foundations of probability that will be of value to philosophers of science, mathematicians, statisticians, psychologists, and educators.

The collection falls into three parts. Part I comprises five papers on the axiomatic foundations of probability. Part II contains seven articles on probabilistic causality and quantum mechanics, with an emphasis on the existence of hidden variables. The third part consists of a single extended essay applying probabilistic theories of learning to practical questions of education: it incorporates extensive data analysis.

Patrick Suppes in one of the world's foremost philosophers in the area of probability and has made many contributions to both the theoretical and the practical side of education. The statistician Mario Zanotti is a long-time collaborator.

Foundations of probability with applications

Cambridge Studies in Probability, Induction, and Decision Theory

General editor: Brian Skyrms

Ellery Eells, *Probabilistic Causality*
Cristina Bicchieri and Maria Luisa Dalla Chiara (eds.), *Knowledge, Belief, and Strategic Interaction*
Richard Jeffrey, *Probability and the Art of Judgment*
Robert Koons, *Paradoxes of Strategic Rationality*
Cristina Bicchieri, *Rationality and Coordination*
Patrick Maher, *Betting on Theories*
J. Howard Sobel, *Taking Chances*
Jan Von Plato, *Creating Modern Probability: Its Mathematics, Physics, and Philosophy in Historical Perspective*
Ellery Eells and Brian Skyrms (eds.), *Probability and Conditionals*
Christina Bicchieri, Richard Jeffrey, and Brian Skyrms (eds.), *The Dynamics of Norms*

Preface

The joint articles that we have written over the past twenty years fall into three different areas and consequently there is a natural division of this collection of our papers into three parts. Part I is concerned with our work in the foundations of probability, Part II with causality and quantum mechanics, and Part III with application of probabilistic models in education.

The five papers in Part I represent our joint efforts to clarify and extend the qualitative foundations of probability first given impetus in the important work of Bruno de Finetti. Both of us, but separately over a good many years, had extensive conversations with de Finetti about the foundations of probability, and several of our papers were inspired by questions he raised. In fact, the first two articles on necessary and sufficient conditions for existence of a strictly agreeing measure of a qualitative probability ordering were a response to de Finetti's early qualitative axioms and the subsequent search for necessary and sufficient axioms. The important point about our work is that in order to get simple axioms we had to go beyond the usual event structure of a probability space to elementary random variables. But we felt at the time, and continue to feel, that this is an extension that is very much in de Finetti's own line of thought, as reflected for example, in his introduction at an early stage of random quantities without an underlying probability space to represent them. It is also worth noting that characterizing the expectation of elementary random variables axiomatically builds on a long tradition in the theory of probability to define probability in terms of expectation. This is already to be found in Bayes's eighteenth century treatise.

Another direction of representation is to generalize from probabilities that subjectively seem sometimes difficult to manage to the weaker concept of having upper and lower probabilities. The third and fourth papers deal with such problems, and particularly the fourth paper is related to questions raised by de Finetti. In particular we examine in this article the conditions on upper and lower probabilities for them to imply the existence of probabilities. The fifth and final paper in Part I addresses a somewhat different but related problem, namely, that of giving random-variable rather than numerical representation for extensive quantities in the theory of measurement. Here the methods are somewhat different and in fact we use in a central way the well-known Hausdorf moment theorem to give qualitative conditions for numerical random variables whose distributions have finite support, that is, are defined on a bounded set. The central problem addressed in this paper is very much related to the long history of the study of the theory of error in probability theory going back to the early work of Simpson in the eighteenth century and later work by many others. This problem of finding appropriate qualitative axioms for errors or variability in measurements remains one of the least satisfactorily solved problems in the foundations of measurement.

Part II contains seven papers on the foundations of the theory of probabilistic causality and more particularly on the foundations of quantum mechanics. The first three papers deal directly with problems in quantum mechanics, the first with the stochastic incompleteness of quantum mechanics looked at from the standpoint of stochastic processes, the second and third with the problems that arise in connection with hidden-variable theories and in particular with the problem of such theories in the context of Bell's theorem. Our 1976 paper on these matters, which is the second paper in this section, was among the earliest to introduce in a formal way precise probabilistic statements about the independence conditions required for the derivation of Bell's theorem.

In the third paper we move away from specific questions in quantum mechanics to showing the impossibility of hidden variables when we impose natural conditions such as exchangeability and symmetry. It is important for positive scientific work regarding hidden variables to understand that rather natural conditions are sufficient to show that hidden variables cannot exist. The fourth paper moves in the opposite direction. Here we show that without such conditions probabilistic explanations in terms of hidden variables, in fact deterministic ones, are always possible whenever the observable random variables have a joint probability distribution. This is an interesting twist on the old and mistaken tale of the conflict between determinism and probability. Here they go hand in hand. Deterministic hidden variables can be found if and only if the phenomenological observables have a joint probability distribution. In the search for hidden variables that have conceptual meaning we need to impose much stronger conditions of the sort to be found in the third paper in this section. As the fourth paper shows, without such conditions a mathematical construction of hidden variables can always be given, even if this mathematical construction does not have any direct conceptual interpretation in terms of the scientific framework involved.

The fifth paper reviews and extends our earlier work on probabilistic causality and symmetry with special reference to quantum mechanics. The sixth paper returns to hidden variables in the context of Bell's theorem and gives a necessary condition for existence of a hidden variable when the number of hidden variables is greater than four. Recall that in Bell's theorem only four observables make up the inequalities, but it is natural to go beyond four and to ask more general questions. Finally, in the seventh paper in this section, written to celebrate the 90th birthday of Karl Popper, we show that if we weaken the conditions in the Bell-type situation to the existence, not of a probability distribution, but of an upper probability measure, then such a measure can always be found that is also compatible with the results of quantum mechanics. However, as we show in the paper, this upper probability measure is not monotonic, which means that there are events A and B such that even though A must occur if B occurs, the upper probability measure assigns a lower measure to A than to B. We extend in that framework our earlier results on causality to the existence of common causes in the framework of such upper measures. The results turn out to be rather satisfactory from the standpoint of their neatness of formulation, but whether or not they will end up having any application scientifically is yet to be seen.

Part III on applications in education of probabilistic models contains only one paper. It has not been previously published, but in fact over the past twenty years we have probably devoted as much of our joint effort to this work as to any other. Much of what we have done remains unpublished, although our first publication in this domain appeared in 1976 and, in fact, even earlier as a technical report in 1973.

References to our earlier work in education are given at the end of the long article that constitutes Part III. This paper deals with work that we are continuing and will continue beyond the point at which the present volume appears, namely, work on stochastic models of mastery learning. Here we try to give a sample of the kind of detailed empirical work we think is possible in education in the application of specific probabilistic models of learning to actual instruction when it takes place in the framework of computer-based education. It would be our hope that there will be in the future much further development of the kind of models we describe and develop, which are, as can be seen in the complete formulation at the end of the paper, rather intricate.

Acknowledgments for permission to reproduce the various articles are given at the bottom of the first page of each paper, but thanks are extended here to the many editors and publishers who generously agreed to publication. Finally, we acknowledge the help and patience of Laura von Kampen in preparing this volume for publication.

Patrick Suppes
Mario Zanotti

Stanford, California

Contents

Part III: Applications in education

I

Foundations of probability

1

Necessary and sufficient conditions for existence of a unique measure strictly agreeing with a qualitative probability ordering

1. Conceptual Background

Let Ω be a nonempty set and let $\mathcal{F}$ be an algebra of events on Ω, i.e., an algebra of sets on Ω. Let $\succcurlyeq$ be a qualitative ordering on $\mathcal{F}$. The interpretation of $A \succcurlyeq B$ for two events A and B is that A is *at least as probable* as B. A (finitely additive) probability measure P on $\mathcal{F}$ is *strictly agreeing* with the relation $\succcurlyeq$ if and only if, for any two events A and B in $\mathcal{F}$,

$$P(A) \geqslant P(B) \text{ iff } A \succcurlyeq B.$$

A variety of conditions that guarantee the existence of a strictly agreeing measure is known. Without attempting a precise classification, the sets of conditions are of the following sorts: (i) sufficient but not necessary conditions for existence of a unique measure when the algebra of events is infinite (Koopman, 1940; Savage, 1954; Suppes, 1956); (ii) sufficient but not necessary conditions for uniqueness when the algebra of events is finite or infinite (Luce, 1967); sufficient but not necessary conditions for uniqueness when the algebra of events is finite (Suppes, 1969); (iv) necessary and sufficient conditions for existence of a not necessarily unique measure when the algebra of events is finite (Kraft, Pratt, & Seidenberg, 1959; Scott, 1964; Tversky, 1967). A rather detailed discussion of these various sets of conditions is to be found in Chapters 5 and 9 of Krantz, Luce, Suppes, and Tversky (1971).

The difficulties of giving reasonably simple conditions in terms of the qualitative ordering of events are exemplified by Luce's axiom, which is weaker than Koopman's equipartition axiom or Savage's related but somewhat stronger axiom. Luce's axiom is the following (Krantz *et al.*, 1971, p. 207):

Reprinted from *Journal of Philosophical Logic* **5** (1976), 431–438.

For any events A, B, C, and D such that $A \cap B = \emptyset$, $A \succ C$, and $B \succsim D$, there exist events C', D', and E such that

(i) $E \approx A \cup B$;
(ii) $C' \cap D' = \emptyset$;
(iii) $C' \cup D' \subset E$;
(iv) $C' \approx C$ and $D' \approx D$.

Here $\succ$ is the strict ordering relation and $\approx$ the equivalence relation defined in terms of the weak ordering $\succsim$. The meaning of this axiom is complex and not easy to state in words. As we search for weaker axioms, closer to being necessary and not merely sufficient, the situation seems likely to get worse. The moral of the effort is that events are the wrong objects to consider. Some slightly richer concept is needed. Extension from one set of objects to a larger and richer set is a characteristic move in mathematics. The most familiar examples are extension of the rational numbers to the real numbers and extension of the real numbers to the complex numbers. As Georg Kreisel has emphasized in several conversations, the introduction of auxiliary concepts is an indispensable practical move in solving significant problems in many domains of mathematics and science.

The main result of this article exemplifies how easily simplification can follow from the introduction of auxiliary concepts. In the present case the move is from an algebra of events to an algebra of extended indicator functions for the events. By this latter concept we mean the following. As before, let Ω be the set of possible outcomes and let $\mathcal{F}$ be an algebra of events on Ω, i.e., $\mathcal{F}$ is a nonempty family of subsets of Ω, and is closed under complementation and union, i.e., if A is in $\mathcal{F}$, $\neg A$, the complement of A with respect to Ω, is in $\mathcal{F}$, and if A and B are in $\mathcal{F}$ then $A \cup B$ is in $\mathcal{F}$. Let $\boldsymbol{A}^c$ be the indicator function (or characteristic function) of event A. This means that $\boldsymbol{A}^c$ is a function defined on Ω such that for any ω in Ω,

$$\boldsymbol{A}^c(\omega) = \begin{cases} 1 \text{ if } \omega \in A \\ 0 \text{ if } \omega \notin A. \end{cases}$$

The algebra $\mathcal{F}^*$ of *extended* indicator functions relative to $\mathcal{F}$ is then just the smallest semigroup (under function addition) containing the indicator functions of all events in $\mathcal{F}$. In other words, $\mathcal{F}^*$ is the intersection of all sets with the property that if A is in $\mathcal{F}$ then $\boldsymbol{A}^c$ is in $\mathcal{F}^*$, and if $\boldsymbol{A}^*$ and $\boldsymbol{B}^*$ are in $\mathcal{F}^*$, then $\boldsymbol{A}^* + \boldsymbol{B}^*$ is in $\mathcal{F}^*$. It is easy to show that any function $\boldsymbol{A}^*$ in $\mathcal{F}^*$ is an integer-valued function defined on Ω. It is the extension from indicator functions to integer-valued functions that justifies calling the elements of $\mathcal{F}^*$ extended indicator functions.

The qualitative probability ordering must be extended from $\mathcal{F}$ to $\mathcal{F}^*$, and the intuitive justification of this extension must be considered. Let $\boldsymbol{A}^*$ and $\boldsymbol{B}^*$ be two extended indicator functions in $\mathcal{F}^*$. Then, to have $\boldsymbol{A}^* \succsim \boldsymbol{B}^*$ is to have the expected value of $\boldsymbol{A}^*$ equal to or greater than the expected value of $\boldsymbol{B}^*$. As should be clear, extended indicator functions are just random variables of a restricted sort. The qualitative comparision is now not one about the probable occurrences of events, but about the expected value of certain restricted random variables. The indicator functions themselves form, of course, a still more restricted class of random variables, but qualitative comparison of their expected values is conceptually identical to qualitative comparison of the probable occurrences of events.

There is more than one way to think about the qualitative comparison of the expected value of extended indicator functions, and so it is useful to consider several examples.

(i) Suppose Smith is considering two locations to fly to for a weekend vacation. Let A_i be the event of sunny weather at location i and B_i be the event of warm weather at location i. The qualitative comparision Smith is interested in is the expected value of $\boldsymbol{A}_1^c + \boldsymbol{B}_1^c$ *versus* the expected value of $\boldsymbol{A}_2^c + \boldsymbol{B}_2^c$. It is natural to insist that the utility of the outcomes has been too simplified by the sums $\boldsymbol{A}_i^c + \boldsymbol{B}_i^c$. The proper response is that the expected values of the two functions are being compared as a matter of belief, not value or utility. Thus it would seem quite natural to bet that the expected value of $\boldsymbol{A}_1^c + \boldsymbol{B}_1^c$ will be greater than that of $\boldsymbol{A}_2^c + \boldsymbol{B}_2^c$, no matter how one feels about the relative desirability of sunny *versus* warm weather. Put another way, within the context of decision theory, extended indicator functions are being used to construct the subjective probability measure, not the measurement of utility. In this context it is worth recalling the importance of certain special decision functions – the gambles – in Savage's theory.

(ii) Consider a particular population of n individuals, numbered $1, \ldots, n$. Let A_i be the event of individual i going to Hawaii for a vacation this year, and let B_i be the event of individual i going to Acapulco. Then define

$$\boldsymbol{A}^* = \sum_{i=1}^{n} \boldsymbol{A}_i^c \qquad \text{and} \qquad \boldsymbol{B}^* = \sum_{i=1}^{n} \boldsymbol{B}_i^c .$$

Obviously $\boldsymbol{A}^*$ and $\boldsymbol{B}^*$ are extended indicator functions – we have left implicit the underlying set Ω. It is meaningful and quite natural to qualitatively compare the expected values of $\boldsymbol{A}^*$ and $\boldsymbol{B}^*$. Presumably such comparisons are in fact of definite significance to travel agents, airlines, and the like.

We believe that such qualitative comparisons of expected value are natural in many other contexts as well. What the main theorem of this article shows is that very simple necessary and sufficient conditions on the qualitative comparison of extended indicator functions guarantee existence of a strictly agreeing, finitely additive measure, whether the set Ω of possible outcomes is finite or infinite. Moreover, when it is required that the measure also be an expectation function for the extended indicator functions, it is unique. The proof of the theorem, it should be mentioned, depends directly upon the theory of extensive measurement developed in Chapter 3 of Krantz *et al.* (1971).

2. Formal developments

The axioms are embodied in the definition of a qualitative algebra of extended indicator functions. Several points of notation need to be noted. First, $\boldsymbol{\Omega}^c$ and $\boldsymbol{\emptyset}^c$ are the indicator or characteristic functions of the set Ω of possible outcomes and the empty set $\emptyset$, respectively. Second, the notation $n\boldsymbol{A}^*$ for a function in $\mathcal{F}^*$ is just the standard notation for the (functional) sum of $\boldsymbol{A}^*$ with itself n times. Third, the same notation is used for the ordering relation on $\mathcal{F}$ and $\mathcal{F}^*$, because the one on $\mathcal{F}^*$ is an extension of the one on $\mathcal{F}$: for A and B in $\mathcal{F}$,

$$A \succcurlyeq B \text{ iff } \boldsymbol{A}^c \succcurlyeq \boldsymbol{B}^c.$$

Finally, the strict ordering relation $\succ$ is defined in the usual way: $\boldsymbol{A}^* \succ \boldsymbol{B}^*$ iff $\boldsymbol{A}^* \succcurlyeq \boldsymbol{B}^*$ and not $\boldsymbol{B}^* \succcurlyeq \boldsymbol{A}^*$.

DEFINITION *Let Ω be a nonempty set, let $\mathcal{F}$ be an algebra of sets on Ω, and let $\succcurlyeq$ be a binary relation on $\mathcal{F}^*$, the algebra of extended indicator functions relative to $\mathcal{F}$. Then the qualitative algebra $(\Omega, \mathcal{F}^*, \succcurlyeq)$ is qualitatively satisfactory if and only if the following axioms are satisfied for every $\boldsymbol{A}^*$, $\boldsymbol{B}^*$, and $\boldsymbol{C}^*$ in $\mathcal{F}^*$:*

Axiom 1. *The relation $\succcurlyeq$ is a weak ordering of $\mathcal{F}^*$;*
Axiom 2. $\boldsymbol{\Omega}^c \succcurlyeq \boldsymbol{\emptyset}^c$;
Axiom 3. $\boldsymbol{A}^* \succcurlyeq \boldsymbol{\emptyset}^c$;
Axiom 4. $\boldsymbol{A}^* \succcurlyeq \boldsymbol{B}^*$ *iff* $\boldsymbol{A}^* + \boldsymbol{C}^* \succcurlyeq \boldsymbol{B}^* + \boldsymbol{C}^*$;
Axiom 5. *If $\boldsymbol{A}^* \succ \boldsymbol{B}^*$ then for every $\boldsymbol{C}^*$ and $\boldsymbol{D}^*$ in $\mathcal{F}^*$ there is a positive integer n such that*

$$n\boldsymbol{A}^* + \boldsymbol{C}^* \succcurlyeq n\boldsymbol{B}^* + \boldsymbol{D}^*.$$

These axioms should seem familiar from the literature on qualitative probability. Note that Axiom 4 is the additivity axiom that closely resembles de Finetti's additivity axiom for events: *If* $A \cap C = B \cap C = \emptyset$, then $A \succcurlyeq B$ *iff* $A \cup C \succcurlyeq B \cup C$. As we move from events to extended indicator functions, functional addition replaces union of sets. What is formally of importance about this move is seen already in the exact formulation of Axiom 4. The additivity of the extended indicator functions is unconditional – there is no restriction corresponding to $A \cap C = B \cap C = \emptyset$. The absence of this restriction has far-reaching formal consequences in permitting us to apply without any real