
Bayesian and Non-Bayesian Approaches to Statistical Inference: A Personal View

Bruno Chiandotto

Abstract

Bayesian and non-bayesian approaches to statistical inference are compared giving particular attention to the emerging field of causal statistical inference and causal statistical decision theory. After a brief review of the evolution of statistical inference, as extraction of information and identification of models from data, the problematic issues of causal inference and causal decision theory will be reviewed. The aim is to provide some basic ideas for unifying the different approaches and for strengthening the future of statistics as a discipline.

Prologue

Hume (2003) argued that induction is irrational. This view, often called Humean irrationalism, conflicts with the empiricist view that affirms that science proceeds in a rational and inductive way. Many attempts have been made to refute Hume. One of the earliest is due to Bayes (1763) and Laplace (1812). According to Bayes, rational learning proceeds by assigning probabilities Keynes (1921), usually called prior probabilities, to hypotheses. Using Bayes's theorem, these prior probabilities are then updated in the light of experience. To determine these probabilities, Laplace used what is often called the *principle of insufficient reason*.

Subsequently, the Laplacian account of rational learning was criticized as applying the same intuition to a different representation of the problem often yields different probabilities. Keynes (1921) and Carnap (1950) tried to improve Laplace's approach by interpreting the prior probabilities as a measure of quantifying logical relations between statements. Fisher (1930, 1935, 1956) and Popper (1959) sharply

B. Chiandotto (✉)

Department of Statistics, University of Florence, Florence, Italy

e-mail: chiandot@ds.unifi.it

rejected the Bayes–Laplace tradition and proposed other solutions to the problem of rational learning. With his theory of significance testing, Fisher revolutionized statistical theory and practice. Meanwhile, Popper developed the falsificationist methodology and had a similar influence on the philosophy of science. Both solutions to the problem of rational learning are based on the same principle, namely, that it is rational to accept hypotheses if they have survived rigorous testing. In Popper’s terminology, such hypotheses are called corroborated. A similar approach is due to [Gini \(1943\)](#)¹ and [Pompilj \(1952, 1961\)](#).

1 Introduction

The history of statistical inference is marked by controversies about its fundamental principles. Historically, one can consider roughly four principal approaches to statistical inference.

The first approach is called Fisherian. Fisher has emphasized the need for a variety of approaches for different problems; he was dismissive of axiomatic arguments. A second approach due to [Neyman and Pearson \(1928\)](#), initially developed to explain Fisher’s ideas more concretely, is strongly based on the frequency theory of probability and emphasize operational concepts. A third approach, where probability represents a rational degree of belief, in which different people faced by the same evidence share the same probability, goes back to Laplace and his predecessors and in its modern form it is associated with Carnap and, especially, with [Jeffreys \(1931\)](#). This (*objective*) approach has been extended by specific characterization of probability in which the degree of belief is constrained only by the requirement of self-consistency. In this fourth approach, (*personalistic or subjective*), associated with [Ramsey \(1931\)](#), [Good \(1960\)](#), [De Finetti \(1937\)](#) and [Savage \(1951, 1954\)](#), there is no assumption that different people with the same knowledge express the same probability on a specified event.

In the first two approaches, usually referred to as *classical theory of statistical inference*, the procedures are justified by their performance under hypothetical repetitions of the experiment, i.e. frequency properties. The differences between the two are minor and are essentially the following: (a) in the Fisherian approach, emphasis is placed on the simple test of significance, on the likelihood function and principles as sufficiency; (b) in the Neyman–Pearson approach, operational requirements, such as power and other explicit indicators of sensitivity, are emphasized and confidence interval and acceptance and rejection of hypotheses terminology are introduced.

Jeffreys’s approach to inference has the same target as Fisher’s: what can be reasonably learned about a parameter of the hypothesized model from the data? But, in contrast to Fisher, Jeffreys argues that a different notion of probability is needed to achieve this, specifically, a reasonable degree of belief computed by means of

¹On the contributions of Gini to the foundations of probability and statistical inference I strongly recommend a forthcoming paper by [Piccinato \(2011\)](#).

Bayes's rule; the a priori distribution is taken, in accordance with Laplace, to be dispersed, representing lack of knowledge.

Jeffreys's and the personalistic approach are often referred to as *Bayesian* (or *neo-bayesian*) *approaches to statistical inference*. Although they are formally the same, there are some fundamental and philosophical differences: the personalistic degree of belief, in contrast to the reasonable degree of belief, measures how strongly you believe in something in the light of the model for the data; the direct consequence is that the choice of the prior is substantially different.

There are other approaches to statistical inference. The most relevant are: fiducial inference, likelihood inference, plausibility inference, structural inference, pivotal inference, prequential inference, and predictive inference.

All the approaches to statistical inference utilize some kind of information to obtain a description (through a statistical model) of the phenomenon under study. In my view, every approach based on mathematical models should accommodate all the different approaches and provide tools for making comparative analyses. Such an approach is the *decision approach* substantially already present in the Fisher and Neyman–Pearson theories. Moreover, the decision approach gives a satisfactory solution so far, at the so-called *pragmatic problem of induction*.

Many authors (Cox 1958; Smith 1961) affirm that a distinction must be made between statistical inference and statistical decision theory. But other authors such as Lindley (2006), and this is also my opinion, consider statistical decision theory as one of the possible extensions of statistical inference. Moreover, the decision approach, combining various theories of statistical inference, avoids dogmatism that can lead to paradoxical situations. It is free from logical error, is more effective in applications, and treats successfully a broader range of problems than competing approaches.

2 Bayesianism

Even if the most influential version of neo-bayesianism has been proposed by Savage, the term Bayesianism is used in a wider sense than Savage's approach. It includes the logical probability, frequentist probability, and some other attempts to objectify prior probabilities. Savage showed that a reasonable preference order over the set of all conceivable strategies can be represented by expected utilities of strategies, where now not only the utilities but also the probabilities for computing the expectations can be derived from the preference order. Substantially, Savage provided a general theory of rational learning and decision making. The relevance of neo-bayesianism, where all probabilities are the subjective degrees of belief, lies in the fact that it is a very general philosophy that seamlessly covers science and decision making starting from the problem of induction. Bayesian rationality constitutes progress beyond Humean irrationalism. Even if Bayesianism is not helpful when nothing is known, it might be helpful in the case of partial rather than complete prior information (Joyce 2010).

Real-world decision problems often have to be simplified to become tractable. According to contemporary model-building wisdom, *finding the right simplifications is an art, not a science*; it involves knowledge and requires experts in the field, this conviction is widely shared by experts. Bayesianism, it seems, gives the experts a possibility to bring their experience to bear on the problem. They can choose a prior probability measure in the light of their experience. Given this choice, which can be communicated to others, decision making can proceed, if the computations are feasible; if not, one can try to find an approximation. Indeed, model building is itself a matter of approximation; Bayesian experts might construct simplified models by excluding possibilities that they assign, in the light of their experience, a low prior probability. Thus, it could be argued that Bayesianism describes a rational way of expressing partial expert knowledge that cannot easily be expressed in another way. However, Bayesianism leaves in the dark how experts proceed when trying to transform experience into a prior. On the other hand, experts might learn from experience in a rational fashion. In this case, we already know how ideal Bayesian experts proceed. They start with a prior probability before making experiences, updating their prior, and when after some time they are viewed as experts, the prior they bring to a new problem is actually a posterior probability measure embodying their experience. The problem with this analysis is, however, that the everything-goes theorem implies that the expert's posterior is arbitrary. According to Bayesianism, all conclusions drawn from experience are equally reasonable or unreasonable.

3 Decision Theory and Utility

The foundations of the (*normative*) modern *statistical decision theory* is due to [Von Neumann and Morgenstern \(1947\)](#), for the so-called Expected Utility (EU) and Savage for the Subjective Expected Utility (SEU). These authors, on the basis of a series of postulates, or rational axioms of behavior of the decision maker, prove the existence of a real-valued utility function that can be derived from the betting rule.

Decision theory recommends an act that maximizes utility, that is, an act whose utility equals or exceeds the utility of every other act. It evaluates an act utility by calculating the act expected utility. It uses probabilities and utilities of an act possible outcomes to define an act expected utility.

Since people usually do not behave in ways consistent with the *axiomatic* rules and hence lead to violations of optimality, there is a related area of study, called a *descriptive decision theory*, attempting to describe what people actually do.

A series of criticisms (particularly [Allais 1953](#) and, for an up-to-date and reasonably extended review, [Chiandotto and Bacci 2004](#)) have been made against EU and SEU. The criticisms regards, mostly, the empirical relevance of the rational axioms of behavior.

Even if the problem of the importance of the axioms on the behavior of the decision maker has to be viewed not in the sense of a good description, but in that

of a good rule (i.e., it concerns identifying the best decision to take, assuming an ideal decision maker who is fully informed, able to compute with perfect accuracy and full rationality) different authors have proposed alternative systems of axioms less restrictive and more compatible with the actual behavior of decision makers.

To generalize the normative decision theory, some authors adopted different terminology like *prescriptive decision theory* (Bell et al. 1988), *constructive decision theory* (Roy 1993; Tsoukiàs 2007). These approaches hypothesize weaker axioms than the classic ones; in particular, since the more frequently violated axiom is independence, the new theories release the property of linearity in the probability. Machina (1982) develops a utility theory without the presence of the independence axiom. Other theories, instead, do not include the axiom of transitivity (Fishburn 1973). Among the more interesting theoretical proposals (generalization of utility theories) we should include the rank-dependent utility (Quiggin 1993), the prospect theory (Kahneman and Tversky 1979), and cumulative prospect theory (Tversky and Kahneman 1992). Aiming at giving to decision theory useful operating tools, it must be considered the so-called causal approach to the theory of the decisions. This approach, although mainly developed in the context of the philosophical reflection, results of large interest for his statistical implications.

4 Causality

In spite of the innumerable developments, generalized utility theories are still not able to solve in a satisfactory way operative decision-making problems. In fact such theories discuss situations in which the consequences of acts are dependent on the *state of the world* whenever the action chosen has no effect on such state. This hypothesis in many contexts is not satisfied. In fact, in many situations the choice made by the decision maker has a, sometime, relevant effect on the state of nature (*the act causes the state*). Therefore, to solve decision problems, the analysis of causality becomes relevant in its theoretical aspects and in its operative implications.

Regarding causality, the paper of Freedman (1999) and three contributions of Mealli et al. (2011), Cox and Wermuth (2004), and of Frosini (2006) are especially useful. This latter author presents a synthetic but exhaustive panorama of the developments of the concept of causality: starting from the Aristotelian doctrine of causation he arrives to the more recent developments on relevant aspects to statistical modeling and, particularly, on acyclic graphical models (*Directed Acyclic Graphs—DAGs*). Also Cox and Wermuth, after an interesting close examination of three different definitions of causality, analyze graphical models focusing on the concepts of statistical independence and particularly on the difference between conditioning and intervention.

The paper of Mealli, Pacini, and Rubin gives a complete and up-to-date account of the so-called Neyman–Rubin–Holland model of causality. The framework proposed especially by (Holland and Rubin 1988; Rubin 1974, 2004) is very powerful and general, it provides a definition of causal effects in terms of potential outcomes, as well as a general statement of the assumptions, sufficient to make

causal inferences possible, even with observational data. Unfortunately, because of its generality the standard Neyman–Rubin–Holland model operates at a level of abstraction that is far away from the underlying mechanisms and processes that account for how observational data are generated. While such generality makes the model very powerful, its agnosticism about the underlying causal mechanisms can make it difficult to be applied in settings that are not close to a well-designed experiment.

Graphical models (Lauritzen and Richardson 2002) represent a generalization of the graphs of influence (Dawid 2002; Howard and Matheson 1984) that represent an extension of the path diagrams proposed by Wright (1921). In path analysis, the connections among the variables of interest are expressed in a graphical form, allowing to distinguish spurious from causal, direct and indirect effects, of variables. Other very interesting contributions to the statistical analysis of causality are Dawid (2000); Holland (1986); Pearl (1995, 2009); Spirtes et al. (2000) and, above all, Woodward (2003).

Woodward collects in his volume a 30-year of research activity presenting a new theory of causality that he considers superior to the counterfactual theory of causality developed by Lewis. The contribution of Woodward is placed in line with the studies of Spirtes, Glymour, and Scheines and of Pearl. While these latter authors concentrate their attention on the theoretical–methodological aspects, Woodward deals particularly with the philosophical foundations of the reasoning introducing a simple, but clear, definition of causality: C causes A if and only if the value of A is modified by an intervention on C. Woodward presents the tools for the analysis, graphics, and equations, for proceeding to the development of its *theory of manipulation*.

The different approaches to causality outlined above are characterized by specificities that are considered by the authors themselves not compatible: each author considers his own approach to be superior to the others. In my opinion, this position does not appear acceptable, as many of them are compatible at least in some fundamental aspects. Regarding superiority, there does not exist a statistical tool of universal validity able to give a satisfactory solutions in all research frameworks. The combined use of different approaches (Lauritzen 2004; White and Chalak 2006) seems the correct route to pursue for achieving the more interesting and significant results.

5 Causal Decision Theory

How much what we have said about causality can be relevant in the decision-making context? Causal decision theory adopts principles of rational choice that attend to an act consequences. It maintains that an account of rational choice must use causality to identify the considerations that make a choice rational. An act expected utility is a probability-weighted average of its possible outcome utilities. Possible states of the world that are mutually exclusive and jointly exhaustive, and so form a partition, generate an act possible outcomes. An act-state pair specifies an

outcome. Each product specifies the probability and utility of a possible outcome. The sum is a probability-weighted average of the possible outcomes utilities, where the probabilities depend casually on the act, probability are causal rather than merely evidential.

Joyce (1999) gives an account of rational decision making and probabilistic theories of evidence and confirmation. This author begins with an historical introduction to the topic of decision theory, including a critical discussion of Savage's theory, followed by a treatment of the modern *evidential theory* of decision making. Two chapters are deal with causal decision theory. The final chapter reports a unified representation theorem that simultaneously provides a firm foundation for both evidential and causal decision theory.

The accounts of rational decision discussed by Joyce presuppose that a rational agent should act so as to maximize some sort of "expected utility," which is a sort of weighted average of the utilities of the outcomes of a decision. What's at issue in the foundational disputes is which kind of expected utility should be maximized, and, consequently, which weights should be used in the weighted average of the values of the outcomes. All parties seem to agree that the weights should be set according to the probabilities of the outcomes given that the act is performed. The disagreement concerns how to unpack this subtle conditional-like expression for the purpose at hand. Evidential decision theory recommends performing that act which provides the *best evidence* for the good outcomes (on average). On the other hand, causal decision theorists propose a different way of unpacking. They suggest that we unpack this as the degree to which the act causally promotes the state. Several interpretations of causal probability have been proposed in the literature, and the connections between the various kinds of conditionals have been studied extensively in recent decades.

Armendt (1986), in a paper on the foundations of causal decision theory, distinguishes three different approaches to causal decision theory, similar in the contents but philosophically different, that go back, respectively, to Gibbard and Harper (1976), Skyrms (1979) and Lewis (1981). Gibbard and Harper distinguished causal decision theory, which uses probabilities of subjunctive conditionals, from evidential decision theory, which uses conditional probabilities. As in decision problems probabilities of subjunctive conditionals track causal relations, using them to calculate an option expected utility makes decision theory causal. They argued that expected utility, calculated with probabilities of conditionals, yields genuine expected utility. Skyrms presented a version of causal decision theory that dispenses with probabilities of subjunctive conditionals. His theory separates factors that the agent's act may influence from factors that the agent's act may not influence. Lewis defines the expected utility of an option and his formula for an option expected utility that is the same as Skyrms. The handy interpretation of the probability of a state if one performs an act, however, is not completely satisfactory. A good decision aims to produce a good outcome rather than evidence of a good outcome. Causal decision theory interprets the probability of a state, if one performs an act, as a certain type of causal probability rather than as a standard conditional probability. This aspect makes expected utility track efficacy, rather than auspiciousness.

As already outlined, Pearl, Spirtes, Glymour, and Scheines and Woodward present methods of inferring causal relations from statistical data. They use DAGs and associated probability distributions to construct causal models. In a decision problem, a causal model yields a way of calculating an act effect. A causal graph and its probability distribution express a dependency hypothesis and yield each act causal influence given that hypothesis. They specify the causal probability of a state under supposition of an act. An act expected utility is a probability-weighted average of its utilities according to the dependency hypotheses that candidate causal models represent.

Heckerman and Shachter (1995) proposed a version of Pearl's causality definition in the decision-making framework. This formulation has been rejected by Pearl himself. Heckerman and Shachter (2003) some years later, discussing the work "Statistics and Causal Inference" of Pearl, say: "... Unfortunately, Pearl has downplayed the strong connections between his work and decision theory as well as the suitability of the influence diagram as a representation of causal interactions. On the contrary, we believe that people who are familiar with decision theory will find comfort, as we have, in these connections ..."

6 Conclusions

The importance of Bayesianism, in which all probabilities are subjective degrees of belief, lies in the fact that it is a very general philosophy that seamlessly covers science and decision making from the problem of induction, which provides the context where it originated, to the theoretical and practical problems of statistical inference. Bayesian rationality constitutes a progress beyond Humean irrationalism. Even if Bayesianism is not helpful when nothing is known, it might be helpful in the case of partial rather than complete knowledge (Joyce 2010).

Causal knowledge plays an important role in everyday reasoning, it enables to predict future outcomes, explain past events, control the environment. Correlations among events can often be good indicators of the presence of some causal relation, but it is well known that observed associations are insufficient to disambiguate causal structure. For this reason much of causal learning takes place in the context of intervention that, in the real world often involves learning a complex network of relations among many events (Pearl 2011). To learn from interventions one must first decide which intervention to make.

Intervention is the central subject of the contributions of Pearl on causality. This author, in my opinion, has given the more interesting and innovative contributions to the analysis of causality, but his contributions, to become really useful from an empirical point of view, must be reinterpreted, as suggested by Heckerman and Schachter, in a decision theoretic framework. The decision-making process allows learners to use interventions to disambiguate particular causal structures, namely, those that they have in mind as potential models of the causal system.

References

- Allais, M.: Le comportement de l'homme rationel devant le risque: critique des axiom et postulates de l'ecole americane. *Econometrica* **21**, 503–546 (1953)
- Armendt, B.: A foundations of causal decision theory. *Topoi* **5**, 3–19 (1986)
- Bayes, T.: Essay towards solving a problem in the doctrine of chances. *Philos. Trans. R. Soc. Lond.* **53**, 370–418 (1763)
- Bell, D.E., Raiffa, H., Tversky, A.: *Decision Making*. Cambridge University Press, Cambridge (1988)
- Carnap, R.: *Logical Foundations of Probability*. Chicago University Press, Chicago (1950)
- Chiantotto, B., Bacci, S.: Decisioni razionali per il governo dell'università, un prerequisite essenziale: la teoria dell'utilità. Università degli Studi di Firenze (2004)
- Cox, D.R.: Some problems connected with statistical inference. *Ann. Math. Stat.* **29**, 357–372 (1958)
- Cox, D.R., Wermuth, N.: Causality: a statistical view. *Int. Stat. Rev.* **72**, 285–305 (2004)
- Dawid, A.P.: Causal inference without counterfactuals. *J. Am. Stat. Assoc.* **95**, 407–448 (2000)
- Dawid, A.P.: Influence diagrams for causal modelling and inference. *Int. Stat. Rev.* **70**, 161–189 (2002)
- De Finetti, B.: La prévision: ses lois logiques, ses source subjectives. *Ann. del'Institut Henri Poincaré* **24**, 17–24 (1937)
- Fishburn, P.: A mixture-set axiomatization of conditional subjective expected utility. *Econometrica* **41**, 1–25 (1973)
- Fisher, R.A.: Inverse probability. *Math Proc. Camb. Philos. Soc.* **26**, 528–535 (1930)
- Fisher, R.A.: The logic of Inductive inference (with discussion). *J. R. Stat. Soc.* **98**, 39–82 (1935)
- Fisher, R.A.: *Statistical Method and Scientific Inference*. Oliver and Boyd, Edinburgh (1956)
- Freedman, D.: From association to causation: some remarks on the history of statistics. *Stat. Sci.* **14**, 243–258 (1999)
- Frosini, B.V.: Causality and causal models: a conceptual perspective. *Int. Stat. Rev.* **74**, 305–334 (2006)
- Gibbard, A., Harper, W.: Counterfactuals and two kinds of expected utility. In: Harper, W., Stalnaker, R., Pearce, G. (eds.) *Conditionals, Belief, Decision, Chance, and Time*, pp. 153–190. Dordrecht-Reidel, Dordrecht (1976)
- Gini, C.: I test di significatività, *Atti della VII Riunione della Società Italiana di Statistica*, Roma (1943)
- Good, I.J.: Weights of evidence, corroboration, explanatory power, information and utility of experiment. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **22**, 319–331 (1960)
- Heckerman, D., Shachter, R.: Decision-theoretic foundations for causal reasoning. *J. Artif. Intell. Res.* **3**, 405–430 (1995)
- Heckerman, D., Shachter, R.: Discussion in Pearl, J.: Statistics and causal inference: a review. *Test* **12**, 101–165 (2003)
- Holland, P.: Statistics and causal inference. *J. Am. Stat. Assoc.* **81**, 945–960 (1986)
- Holland, P., Rubin, D.: Causal inference in retrospective studies. *Eval. Rev.* **13**, 203–231 (1988)
- Howard, R.A., Matheson, J.E.: Influence diagrams. In: Howard, R.A., Matheson, J.E. (eds.) *Readings in the Principles and Applications of Decision Analysis*. Strategic Decision Group, Menlo Park (1984)
- Hume, D.: *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*. Book 1: Of the Understanding (2003). The Project Gutenberg eBook. Release Date: December, 2003
- Jeffreys, H.: *Scientific Inference*. Cambridge University Press, Cambridge (1931)
- Joyce, J.: *The Foundations of Causal Decision Theory*. Cambridge University Press, Cambridge (1999)
- Joyce, J.: A defense of imprecise credence and decision making. *Philos. Perspect.* **24**, 281–323 (2010)

- Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. *Econometrica* **47**, 263–293 (1979)
- Keynes, J.M.: *A Treatise on Probability*. MacMillan & Co. London (1921)
- Laplace, P.S.: *Théorie analytique des probabilités*, 3rd edn. (1820). Courcier, Paris (1812)
- Lauritzen, S.L.: Discussion on causality in Rubin, D.B. (2004): Direct and indirect causal effects via potential outcomes. *Scand. J. Stat.* **31**, 161–170 (2004)
- Lauritzen, S., Richardson, T.: Chain graph models and their causal interpretations (with discussion). *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**, 321–361 (2002)
- Lewis, D.: Causal decision theory. *Australas. J. Philos.* **59**, 5–30 (1981)
- Lindley, D.V.: *Understanding Uncertainty*. Wiley, Hoboken (2006)
- Machina, M.J.: Expected utility analysis without the independence axiom. *Econometrica* **50**, 227–323 (1982)
- Mealli, F., Pacini, B., Rubin, D.B.: Statistical inference for causal effects. In: Kenett, R., Salini, S. (eds.) *Modern Analysis of Customer Satisfaction Surveys*. Wiley, Chichester (2011)
- Neyman, J., Pearson, E.S.: On the use and interpretation of certain test criteria for purposes of statistical inference. Part. I and II. *Biometrika* **20A**, 175–240 and 263–294 (1928)
- Pearl, J.: Causal diagrams for empirical research. *Biometrika* **82**(4), 669–710 (1995)
- Pearl, J.: *Causality: Models, Reasoning, and Inference*, 2nd edn. Cambridge University Press, Cambridge (2009)
- Pearl, J.: The causal foundations of structural equation modeling. In: Hoyle, R.H. (ed) *Handbook of Structural Equation Modeling*. Guilford Press, New York (2011)
- Piccinato, L.: Gini's criticism to the theory of inference: a missed opportunity. *Metron* **69**, 101–117 (2011)
- Pompij, G.: Logica della conformità. *Archimede* **4**, 22–28 (1952)
- Pompij, G.: *Teoria dei campioni, Applicazioni alla sperimentazione, alla produttività e alle rilevazioni campionarie*. Partial reprint (1956), Veschi, Roma (1961)
- Popper, K.R.: *The Logic of Scientific Discovery*. First version of this book appeared as *Logik der Forschung*, 1934. Hutchison Education, London (1959)
- Quiggin, J.: *Generalized Expected Utility the Rank-Dependent Model*. Kluwer, Boston (1993)
- Ramsey, F.P.: Truth and probability. In: Braithwaite, R.B. (ed.) *The Foundations of Mathematics and Other Logical Essays*, pp. 56–198. Routledge & Kegan Paul, London (1931)
- Roy, B.: Decision science or decision-aid science? *Eur. J. Oper. Res.* **66**, 184–203 (1993)
- Rubin, D.B.: Estimating causal effects of treatments in randomized and non-randomized studies. *J. Educ. Psychol.* **66**, 688–701 (1974)
- Rubin, D.B.: Direct and indirect causal effects via potential outcomes. *Scand. J. Stat.* **31**, 161–170 (2004)
- Savage, L.J.: The theory of statistical decision. *J. Am. Stat. Assoc.* **46**, 55–67 (1951)
- Savage, L.J.: *The Foundations of Statistics*. Wiley, New York (1954)
- Skyrms, B.: *Causal Necessity*. Yale University Press, New Haven (1979)
- Smith, C.A.B.: Consistency in statistical inference and decision. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **25**, 1–37 (1961)
- Spirtes, P., Glymour, C., Scheines, R.: *Causation, Prediction and Search*, 2nd edn. MIT Press, Cambridge (2000)
- Tsoukiàs, A.: On the concept of decision aiding process: an operational perspective. *Ann. Oper. Res.* **154**, 3–27 (2007)
- Tversky, A., Kahneman, D.: Advances in prospect theory: cumulative representation of uncertainty. *J. Risk. Uncertain.* **5**, 297–323 (1992)
- Von Neumann, J., Morgenstern, O.: *Theory of Games and Economic Behavior*, 2nd edn. Princeton University Press, Princeton (1947)

-
- Woodward, J.: Making things happen: a theory of causal explanation. In: Oxford Studies in the Philosophy of Science. Oxford University Press, New York (2003)
- White, H., Chalak, K.: A Unified Framework for Defining and Identifying causal Effects, UCSD Department of Economics Working Paper, pp. 1–53 (2006)
- Wright, S.: Correlation and causation. *J. Agric. Res.* **20**, 557–585 (1921)