International Encyclopedia of Statistical Science
Springer 2010
http://www.springer.com/statistics/book/978-3-642-04897-5

# Imprecise Probability*

**Frank P.A. Coolen, Matthias C.M. Troffaes, Thomas Augustin**
Durham University, UK and Ludwig Maximilians University Munich, Germany
*frank.coolen@durham.ac.uk, matthias.troffaes@durham.ac.uk, thomas@stat.uni-muenchen.de*

## 1 Overview

Quantification of uncertainty is mostly done by the use of precise probabilities: for each event $A$, a single (classical, precise) probability $P(A)$ is used, typically satisfying Kolmogorov's axioms [4]. Whilst this has been very successful in many applications, it has long been recognized to have severe limitations. Classical probability requires a very high level of precision and consistency of information, and thus it is often too restrictive to cope carefully with the multi-dimensional nature of uncertainty. Perhaps the most straightforward restriction is that the quality of underlying knowledge cannot be adequately represented using a single probability measure. An increasingly popular and successful generalization is available through the use of *lower and upper probabilities*, denoted by $\underline{P}(A)$ and $\overline{P}(A)$ respectively, with $0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1$, or, more generally, by lower and upper expectations (previsions) [33, 36, 41]. The special case with $\underline{P}(A) = \overline{P}(A)$ for all events $A$ provides precise probability, whilst $\underline{P}(A) = 0$ and $\overline{P}(A) = 1$ represents complete lack of knowledge about $A$, with a flexible continuum in between. Some approaches, summarized under the name *nonadditive probabilities* [18], directly use one of these set-functions, assuming the other one to be naturally defined such that $\underline{P}(A^c) = 1 - \overline{P}(A)$ , with $A^c$ the complement of $A$. Other related concepts understand the corresponding intervals $[\underline{P}(A), \overline{P}(A)]$ for all events as the basic entity [38, 39]. Informally, $\underline{P}(A)$ can be interpreted as reflecting the evidence certainly in favour of event $A$, and $1 - \overline{P}(A)$ as reflecting the evidence against $A$ hence in favour of $A^c$.

The idea to use imprecise probability, and related concepts, is quite natural and has a long history (see [22] for an extensive historical overview of nonadditive probabilities), and the first formal treatment dates back at least to the middle of the nineteenth century [9]. In the last twenty years the theory has gathered strong momentum, initiated by comprehensive foundations put forward by Walley [36] (see [30] for a recent survey), who coined the term *imprecise probability*, by Kuznetsov [27], and by Weichselberger [38, 39], who uses the term *interval probability*. Walley's theory extends the traditional subjective probability theory via buying and selling prices for gambles, whereas Weichselberger's approach generalizes Kolmogorov's axioms without imposing an interpretation. Usually assumed consistency conditions relate imprecise probability assignments to non-empty closed convex sets of classical probability distributions. Therefore, as a welcome by-product, the theory also provides a formal framework for models used in frequentist robust statistics [6] and robust Bayesian approaches [31]. Included are also concepts based on so-called two-monotone [24] and totally monotone capacities, which have become very popular in artificial intelligence under the name (Dempster-Shafer) belief functions [17, 32]. Moreover, there is a strong connection [14] to Shafer and Vovk's notion of game-theoretic probability [20].

*Contribution to *International Encyclopedia of Statistical Sciences (Springer)*

The term ‘imprecise probability’—although an unfortunate misnomer as lower and upper probability enable more accurate quantification of uncertainty than precise probability—appears to have been established over the last two decades, and actually brings together a variety of different theories. In applications, clear advantages over the established theory of precise probability have been demonstrated (see Section 2). This justifies the further development of imprecise probability, particularly towards building a complete methodological framework for applications in statistics, decision support, and related fields. Imprecise probability provides important new methods that promise greater flexibility for uncertainty quantification. Its advantages include the possibility to deal with conflicting evidence, to base inferences on weaker assumptions than needed for precise probabilistic methods, and to allow for simpler and more realistic elicitation of subjective information, as imprecise probability does not require experts to represent their judgements through a full probability distribution, which often does not reflect their beliefs appropriately.

The Society for Imprecise Probability: Theories and Applications (`www.sipta.org`) organises conferences, workshops and summer schools, and provides useful introductory information sources and contacts through its web-page.

# 2 Applications

The increased attention to imprecise probability during the last two decades has led to many new methods for statistical inference and decision support, with applications in a wide variety of areas.

## 2.1 Imprecise Probabilities in Statistics and Decision Theory

Following Walley [36], many of the imprecise probability-based contributions to statistics follow a generalized Bayesian approach, using a standard precise parametric sampling model with a set of prior distributions. In particular, the use of models from the exponential family is popular in conjunction with classes of conjugate priors. Walley’s Imprecise Dirichlet Model (IDM) for inference in case of multinomial data [37] has attracted particular attention [8]. One successful application area for the IDM is classification [42], where the use of lower and upper probabilities makes the learning process more stable and enables in a quite natural way an item to be explicitly not classified into a unique category, indicating that no clear decision for a single category can be made on the basis of the information available. In these models, updating to take new information into account is effectively done by updating all elements of the set of prior distributions as in Bayesian statistics with precise prior distributions, leading to a set of posterior distributions which forms the basis for inferences. From the technical perspective this procedure is therefore closely related to robust Bayesian inference, but, by reporting the indeterminacy resulting from limited information, use and interpretation of the resulting imprecise posterior goes far beyond a simple sensitivity and robustness analysis.

Other approaches to statistical inference with imprecise probabilities have been developed, which tend to move further away from the precise probabilistic approaches. Examples of such approaches are Nonparametric Predictive Inference [11], generalizations of the frequentist approach, see e.g. [3] for hypotheses testing and [21] for estimation, as well as several approaches based on logical probability [26, 28, 40].

Imprecise probabilities have also proven their use in decision support [34], where, in tradition of Ellsberg’s experiments [19], ambiguity (or non-stochastic uncertainty) plays a crucial role [23]. If only little information is available about a variable, then it is often more natural to refuse to determine a unique optimal decision, when gains and losses depend on that variable. Imprecise probability theory grasps this in a rigorous manner, resulting in a set of possibly optimal decisions, rather than providing only a single, seemingly arbitrarily chosen, optimal decision from this set. Imprecise probability theory is especially useful in critical decision problems where gains and losses heavily depend on variables which are not completely known, such as for instance in pollution control [10] and medical diagnosis [43].

### 2.2 Further Applications

Recent collections of papers [5, 13, 16] give an impression of the huge variety of fields of potential application. In artificial intelligence, for example in pattern recognition [29] and information fusion [7], uncertain expert knowledge can be represented more accurately by means of imprecise probability. Because imprecise probability methods can process information without having to add unjustified assumptions, they are of great importance in risk and safety evaluations, design engineering [2] and reliability [12]. The ongoing intensive debate on bounded rationality makes reliable decision theory based on imprecise probability particularly attractive in microeconomics and in social choice theory. In finance, imprecise probability is gaining strong influence given its very close connection to risk measures [1, 35]. Imprecise probability also yields deeper insight into asset pricing [25]. The study of Markov chains with imprecise transition probabilities [15] is also important for many areas of application.

## 3 Challenges

Imprecise probability and its applications in statistical inference and decision support offer a wide range of research challenges. On foundations, key aspects such as updating have not yet been fully explored, and different approaches have different conditioning rules. The relation between imprecision and information requires further study, and many of the most frequently used statistical methods (such as complex regression models) have not yet been fully generalized to deal with imprecise probability. In cases where generalizations are easily found, it may be unclear which of many possible approaches is most suitable. Of course, early developments of new theoretic frameworks tend to include illustrative applications to mostly text-book style problems. The next stage required towards widely applicable methods involves upscaling, where in particular computational aspects provide many challenges. Even methods such as simulation, mostly straightforward with precise probabilities, become non-trivial with imprecise probabilities.

For applications which require the use of subjective information, elicitation of expert judgements is less demanding when lower and upper probabilities are used, but while practical aspects of elicitation have been widely studied this has, thus far, only included very few studies involving imprecise probabilities.

In decision making, algorithms to find optimal solutions need to be improved and implemented for large-scale applications. As many problems have a sequential nature, ways of representing sequential solutions efficiently also need to be developed, the more so as classical techniques such as backward induction and dynamic programming often cannot be extended directly.

## References

[1] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[2] J.M. Aughenbaugh and C.J.J. Paredis. The value of using imprecise probabilities in engineering design. *Journal of Mechanical Design*, 128(4):969–979, 2006.

[3] Thomas Augustin. Neyman-Pearson testing under interval probability by globally least favorable pairs—reviewing Huber-Strassen theory and extending it to general interval probability. *Journal of Statistical Planning and Inference*, 105(1):149–173, June 2002.

[4] Thomas Augustin and Marco Cattaneo. *International Encyclopedia of Statistical Sciences*, chapter Foundations of probability. Springer, 2010.

[5] Thomas Augustin, Frank P.A. Coolen, Serafin Moral, and Matthias C.M. Troffaes, editors. *ISIPTA'09: Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Applications, Durham University, Durham, UK, July 2009*. SIPTA, 2009.

[6] Thomas Augustin and Robert Hable. On the impact of robust statistics on imprecise probability models: a review. *Structural Safety*, 2010. To appear.

[7] Alessio Benavoli and Alessandro Antonucci. Aggregating imprecise probabilistic knowledge: application to zadeh's paradox and sensor networks. *International Journal of Approximate Reasoning*, 2010. To appear.

[8] Jean-Marc Bernard. Special issue on the Imprecise Dirichlet Model. *International Journal of Approximate Reasoning*, 50:201–268, 2009.

[9] George Boole. *An investigation of the laws of thought on which are founded the mathematical theories of logic and probabilities*. Walton and Maberly, London, 1854.

[10] Morgane Chevé and Ronan Congar. Optimal pollution control under imprecise environmental risk and irreversibility. *Risk Decision and Policy*, 5:151–164, 2000.

[11] F. P. A. Coolen. *International Encyclopedia of Statistical Sciences*, chapter Nonparametric predictive inference. Springer, 2010.

[12] F. P. A. Coolen and L. V. Utkin. *International Encyclopedia of Statistical Sciences*, chapter Imprecise reliability. Springer, 2010.

[13] P. Coolen-Schrijner, F. Coolen, M. Troffaes, and T. Augustin. Special issue on statistical theory and practice with imprecision. *Journal of Statistical Theory and Practice*, 3(1):1–303, 2009. Appeared also with Sat Gupta as editional editor under the title *Imprecision in Statistical Theory and Practice*, Grace, Greensboro.

[14] Gert de Cooman and Filip Hermans. Imprecise probability trees: Bridging two theories of imprecise probability. *Artificial Intelligence*, 172(11):1400–1427, July 2008.

[15] Gert de Cooman, Filip Hermans, and Erik Quaeghebeur. Imprecise markov chains and their limit behavior. *Probability in the Engineering and Informational Sciences*, 23(4):597–635, 2009.

[16] Gert de Cooman, J. Vejnarová, and Marco Zaffalon, editors. *ISIPTA'07: Proceedings of the Fifth International Symposium on Imprecise Probabilities and Their Applications, Charles University, Prague, Czech Republic, July 2007*. SIPTA, 2007.

[17] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Statist.*, 38:325–339, 1967.

[18] Dieter Denneberg. *Non-additive Measure and Integral*. Kluwer, Dordrecht, 1994.

[19] Daniel Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75:643–669, 1961.

[20] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!* Wiley, 2001.

[21] Robert Hable. Minimum distance estimation in imprecise probability models. *Journal of Statistical Planning and Inference*, 2009. to appear.

[22] F. Hampel. Nonadditive probabilities in statistics. *Journal of Statistical Theory and Practice*, 3(1):11–23, 2009.

[23] Ming Hsu, Meghana Bhatt, Ralph Adolphs, Daniel Tranel, and Colin F. Camerer. Neural systems responding to degrees of uncertainty in human decision-making. *Science*, 310:1680–1683, 2005.

[24] P. J. Huber and V. Strassen. Minimax tests and the Neyman-Pearson lemma for capacities". *The Annals of Statistics*, 1:251–263, 1973.

[25] Victor Richmond R. Jose, Robert F. Nau, and Robert L. Winkler. Scoring rules, generalized entropy, and utility maximization. *Operations Research*, 56(5):1146–1157, Sep-Oct 2008.

[26] Henry E. Kyburg Jr. *The Logical Foundations of Statistical Inference*. Springer, 1974. Synthese Library.

[27] Vladimir P. Kuznetsov. *Interval Statistical Models*. Radio i Svyaz Publ., Moscow, 1991. In Russian.

[28] Isaac Levi. Enterprise of knowledge: Essay on knowledge, credal probability and chance. *MIT Press*, 1980.

[29] Kevin Loquin and Olivier Strauss. Noise quantization via possibilistic filtering. *International Journal of Approximate Reasoning*, 2010. To appear.

[30] Enrique Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48(2):628–658, June 2008.

[31] D. Ríos Insua and Fabrizio Ruggeri, editors. *Robust Bayesian Analysis*. Springer, New York, 2000.

[32] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.

[33] Cedric A. B. Smith. Consistency in statistical inference and decision. *Journal of the Royal Statistical Society*, B(23):1–37, 1961.

[34] Matthias C. M. Troffaes. Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning*, 45(1):17–29, May 2007.

[35] Paolo Vicig. Imprecise probabilities in finance and economics. *International Journal of Approximate Reasoning*, 49(1):99–100, September 2008.

[36] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[37] Peter Walley. Inferences from multinomial data: Learning about a bag of marbles. *Journal of the Royal Statistical Society, Series B*, 58(1):3–34, 1996.

[38] Kurt Weichselberger. The theory of interval probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24:149–170, 2000.

[39] Kurt Weichselberger. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I — Intervallwahrscheinlichkeit als umfassendes Konzept*. Physica, Heidelberg, 2001. In cooperation with T. Augustin and A. Wallner.

[40] Kurt Weichselberger. The logical concept of probability and statistical inference. In Fabio G. Cozman, Robert Nau, and Teddy Seidenfeld, editors, *ISIPTA '05: Proc. 4th Int. Symp. on Imprecise Probabilities and Their Applications*, pages 396–405, July 2005.

[41] Peter M. Williams. Notes on conditional previsions. Technical report, School of Math. and Phys. Sci., Univ. of Sussex, 1975. Republished in: International Journal of Approximate Reasoning, 44(3):366-383, 2007.

[42] Marco Zaffalon. The naive credal classifier. *Journal of Statistical Planning and Inference*, 105(1):5–21, June 2002.

[43] Marco Zaffalon, Keith Wesnes, and Orlando Petrini. Reliable diagnoses of dementia by the naive credal classifier inferred from incomplete cognitive data. *Artificial Intelligence in Medicine*, 29(1–2):61–79, 2003.