Journal of Mathematical Psychology 64–65 (2015) 44–57

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Extreme points of the credal sets generated by comparative probabilities

Enrique Miranda [a,*], Sébastien Destercke [b]

[a] University of Oviedo, Department of Statistics and O.R., C-Calvo Sotelo, s/n, Oviedo, Spain

[b] CNRS, HEUDIASYC Joint Research Unit, UMR 7253, Centre de recherche de Royallieu, UTC, F-60205 Compiegne cedex, France

HIGHLIGHTS

- We consider comparative probability models on the singletons of a finite space.
- We study the set of probability measures compatible with such a model.
- We characterize its extreme points by means of a graphical representation.
- We investigate the properties of the lower probability induced by this set.
- We provide tight bounds on the number of extreme points.

ARTICLE INFO

Article history:
Received 15 October 2013
Received in revised form
25 November 2014
Available online 27 December 2014

Keywords:
Comparative probabilities
Credal sets
2-monotone capacities
Belief functions
Extreme points
Imprecise probability masses

ABSTRACT

When using convex probability sets (or, equivalently, lower previsions) as uncertainty models, identifying extreme points can help simplifying various computations or the use of some algorithms. In general, sets induced by specific models such as possibility distributions, linear vacuous mixtures or 2-monotone measures may have extreme points easier to compute than generic convex sets. In this paper, we study extreme points of another specific model: comparative probability orderings between the singletons of a finite space. We characterize these extreme points by means of a graphical representation of the comparative model, and use them to study the properties of the lower probability induced by this set. By doing so, we show that 2-monotone capacities are not informative enough to handle this type of comparisons without a loss of information. In addition, we connect comparative probabilities with other uncertainty models, such as imprecise probability masses.

1. Introduction

In the last decades, there has been a growing interest in imprecise probability models as alternative models to probability in situations where the available information is vague or scarce. These types of models include for instance possibility measures (Dubois & Prade, 1988), belief functions (Shafer, 1976), 2- and n-monotone capacities (Choquet, 1953; Denneberg, 1994) and probability boxes (Destercke, Dubois, & Chojnacki, 2008; Ferson, Kreinovich, Ginzburg, Myers, & Sentz, 2003). All the examples above can be seen as particular cases of coherent lower and upper previsions (Walley, 1991).

The adequacy of each of these models for a particular problem depends, among other things, on the interpretation we are giving to our uncertainty. In this paper, we consider a *robust Bayesian* interpretation (Berger, 1994; Good, 1962): we assume the existence of a precise, but unknown, probability model, and work with the set of probability measures that are compatible with the available information. This gives rise to a *credal set*, as considered by Levi (1980).

Here, we consider the case where the information is expressed by means of a *comparative probability model* (de Finetti, 1931; Koopman, 1940a,b): we consider a finite possibility space $\mathcal{X}$ and assume that we are given judgements of the type "the probability of A is at least as great as that of B". Comparative probabilities have been deemed of particular interest within the context of subjective probability theory (Fine, 1973, 1979; Fishburn, 1986; Suppes, 1974; Walley & Fine, 1979); we also refer to Capotorti and Formisano (2008), Christian, Conder, and Slinko (2007) and Nehring (2009) for some recent work and to Walley (1991, Section 4.5) for a study from the point of view of coherent lower previsions. One of their advantages is that they seem well suited for modelling qualitative judgements (e.g., expert opinions). More-

* Corresponding author.
E-mail addresses: mirandaenrique@uniovi.es (E. Miranda), sebastien.destercke@hds.utc.fr (S. Destercke).

http://dx.doi.org/10.1016/j.jmp.2014.11.004

over, they have been shown to be more general than classificatory probabilities (Walley & Fine, 1979), and they can also be used to compare random quantities (Cohen, 1991).

In spite of this, there are only few works dealing with the numerical and practical aspects of comparative probabilities (Regoli, 1996). One reason for this is that it is not easy to summarize the set of probabilities associated with the comparative assessments, for instance by means of a lower and an upper probability, and this renders it difficult to handle the information about the probability of the events of interest. In this paper, we solve this problem for the specific case of comparisons between the probabilities of singletons. We do so by characterizing the comparative probability models by means of the extreme points of their associated sets of probabilities. Characterizing such extreme points is instrumental in a number of applications of imprecise probabilities, including inferences for graphical (Cano & Moral, 2000) and statistical models (Walley, 1991, Section 8.5), and also to compute bounds of some functionals such as entropy (Abellán & Moral, 2005). It is a problem that has been studied for other types of imprecise probability models, such as 2-monotone capacities (Chateauneuf & Jaffray, 1989), possibility measures (Miranda, Couso, & Gil, 2003), probability intervals (de Campos, Huete, & Moral, 1994) and belief functions (Dempster, 1967); however, as we shall detail later, there is only one partial result for the case we shall consider in this paper (Gulordava, 2010).

After giving some preliminary results in Section 2, we show in Section 3 that, when the comparison judgements are made on the probabilities of the singletons, a graphical representation of these judgements makes it easy to derive the extreme points of the associated credal sets. In Section 4, we use this result to discuss some practical aspects of these models: we establish tight lower and upper bounds of the number of extreme points; investigate their relationship with other imprecise probability models; provide algorithms for the extraction of these extreme points; and discuss the computation of conditional lower probabilities and the merging of multiple comparison judgements. Some additional remarks related to the practical use of these models and their extensions are provided in Section 5.

2. Preliminaries

Consider a finite space $\mathscr{X} = \{x_1, \ldots, x_n\}$, modelling the set of outcomes of some experiment. In this paper, we assume that our information about these outcomes can be modelled by means of *comparative probability orderings of the states*, i.e., statements of the type "the probability of x_i is at least as great as that of x_j". Hence, we shall represent the available information by means of a subset $\mathscr{L}$ of $\{1, \ldots, n\} \times \{1, \ldots, n\}$ modelling the (pre)order relation between the states.

The set of probability measures compatible with this information is given by

$$\mathscr{P}(\mathscr{L}) = \{P \in \mathbb{P}_{\mathscr{X}} : \forall (i,j) \in \mathscr{L}, P(x_i) \geq P(x_j)\}, \tag{1}$$

where $\mathbb{P}_{\mathscr{X}}$ denotes the set of all probability measures on $\mathscr{X}$. This set is called the *natural extension* of the ordering by Walley (1991, Section 4.5.1). This can be equivalently stated by saying that we consider a preorder $\succeq$ between the singletons and we want to characterize the set of probability measures P that agree with this order, in the sense that

$$x_i \succeq x_j \Rightarrow P(x_i) \geq P(x_j).$$

For the purposes of this paper, it shall be useful to represent these assessments by means of a graph $\mathscr{G} = (\mathscr{X}, \mathscr{L})$ where the nodes are the elements of $\mathscr{X}$ and we draw an edge between x_i and x_j when $(i,j) \in \mathscr{L}$.

Example 1. Consider the space $\mathscr{X} = \{x_1, \ldots, x_5\}$ and the assessments

$$\mathscr{L} = \{(1,3), (1,4), (2,5), (4,5)\}.$$

The acyclic graph $\mathscr{G}$ associated with $\mathscr{L}$ is given by Fig. 1. ◊

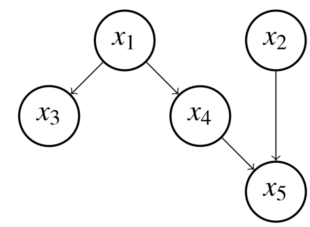

Fig. 1. Graph $\mathscr{G}$ of Example 1.

Note that the credal set $\mathscr{P}(\mathscr{L})$ determined by Eq. (1) is always non-empty, because it includes for instance the uniform probability distribution. Further, the set $\mathscr{P}(\mathscr{L})$ coincides with the set $\mathscr{P}(\mathscr{C}(\mathscr{L}))$ determined by the *transitive closure* $\mathscr{C}(\mathscr{L})$ of $\mathscr{L}$, as the additional constraints of $\mathscr{C}(\mathscr{L})$ are redundant with those of $\mathscr{L}$: since any model in $\mathscr{P}(\mathscr{L})$ (a probability measure) is transitive, so should be the relationship $\succeq$ associated with $\mathscr{L}$.

It is interesting to compare $\mathscr{P}(\mathscr{L})$ with the set

$$\widetilde{\mathscr{P}}(\mathscr{L}) = \{P \in \mathbb{P}_{\mathscr{X}} : \forall (i,j) \in \mathscr{L}, P(x_i) > P(x_j)\},$$

i.e., with the credal set associated with *strict* probability comparisons, which also appear sometimes in the literature (Fishburn, 1986). Since $\mathscr{P}(\mathscr{L})$ is a convex polytope in $\mathbb{R}^n$, it follows from basic convex analysis that $\mathscr{P}(\mathscr{L})$ corresponds to the closure of $\widetilde{\mathscr{P}}(\mathscr{L})$ when the latter set is non-empty, and this non-emptiness is easy to characterize.

Lemma 1. *$\widetilde{\mathscr{P}}(\mathscr{L}) \neq \emptyset$ if and only if its associated graph $\widetilde{\mathscr{G}}$ is acyclic.*

Proof. "Only if": $\widetilde{\mathscr{G}}$ cyclic means that there are at least two indices i, j such that $P(x_i) > P(x_j)$ and $P(x_i) < P(x_j)$, leading to an inconsistency.

"If": if $\widetilde{\mathscr{G}}$ is acyclic, then it can be associated with a preorder over the probability masses $P(x_i)$. We can then take a linear extension of this preorder and associate it with a permutation σ of $\{1, \ldots, n\}$ such that

$$P(x_{\sigma(1)}) < P(x_{\sigma(2)}) < \cdots < P(x_{\sigma(n)});$$

then, it is easy to see that there exists a probability satisfying all these constraints: we may for instance consider the probability measure associated with the probability mass $P(x_{\sigma(i)}) = 1/n - (n-i)\epsilon$ for $i = 1, \ldots, n-1$ and $P(x_{\sigma(n)}) = 1/n + \epsilon + \cdots + (n-1)\epsilon$ with $\epsilon \in (0, \frac{1}{n^2})$. □

Indeed, a cyclic graph $\widetilde{\mathscr{G}}$ is incompatible with the irreflexive property that strict comparative assessments must satisfy. Nevertheless, in this paper we shall focus on non-strict comparative assessments, and for those the associated graph $\mathscr{G}$ may possess cycles, as we shall discuss later.

Note also that we can straightforwardly connect our current model with the axiomatic view of comparative probabilities (de Finetti, 1931). From $\mathscr{L}$, we can obtain a probability ordering $\succeq$ over subsets of $\mathscr{X}$ such that $A \succeq B$ whenever $P(A) \geq P(B)$ for all $P \in \mathscr{P}(\mathscr{L})$. Using (Walley, 1991, Section 4.5.3), this probability ordering satisfies a number of properties, in particular all axioms required by de Finetti (1931) except for the one of completeness. Hence, while we focus in this paper on the numerical aspects associated with specific comparative probabilities, we are completely in-line with the axiomatic view.