European Journal of Operational Research 234 (2014) 209–220

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision Support

Decision making with imprecise probabilities and utilities by means of statistical preference and stochastic dominance

Ignacio Montes *, Enrique Miranda, Susana Montes

University of Oviedo, Dep. of Statistics and O.R, Spain

ARTICLE INFO

Article history:
Received 27 March 2013
Accepted 12 September 2013
Available online 21 September 2013

Keywords:
Preference learning
Binary relations
Stochastic dominance
Statistical preference
Random sets
Imprecise probabilities

ABSTRACT

A problem of decision making under uncertainty in which the choice must be made between two sets of alternatives instead of two single ones is considered. A number of choice rules are proposed and their main properties are investigated, focusing particularly on the generalizations of stochastic dominance and statistical preference. The particular cases where imprecision is present in the utilities or in the beliefs associated to two alternatives are considered.

1. Introduction

In decision making under uncertainty, it is not uncommon to encounter situations with vague or conflicting information about the probabilities or the utilities associated to the different alternatives. We may think for instance of conflicts among the opinions of several experts, limits or errors in the observational process, or simply partial or total ignorance about the process underlying the alternatives. In any such case, the elicitation of a unique probability/utility model for each of the alternatives may be difficult and its use, questionable.

One of the solutions that have been proposed for situations like this is to consider a robust approach, by means of a set of probabilities and utilities. The use of this approach to compare two alternatives is formally equivalent to the comparison of two sets of alternatives, those associated to each possible probability-utility pair. Hence, it becomes useful to consider comparison methods that allow us to deal with sets of alternatives instead of single ones. However, the way to compare of sets of alternatives is no longer immediate: we may compare all possibilities within each of the sets, or also select some particular elements of each set, to take into account phenomena of risk aversion, for instance. This gives rise to a number of possibilities. Moreover, even in the simpler case where we choose one alternative from each set, we must still decide which criterion we shall consider to determine the preferred one.

There is quite an extensive literature on how to deal with imprecise beliefs and utilities when our choice is made by means of an expected utility model (Aumann, 1962; Nau, 2006; Ríos Insua, 1992; Seidenfeld, Schervish, & Kadane, 1995). However, the problem has almost remained unexplored for other choice functions. In this paper we focus mostly on two different optimality criteria that serve as an alternative to the expected utility model: stochastic dominance and statistical preference. The former is based on the comparison of the distribution functions associated to the alternatives, and has been applied in economics (Denuit, Dhaene, Goovaerts, & Kaas, 2005; Goovaerts, Kaas, Van Heerwaarden, & Bauwelinckx, 1990); the latter can be seen as a robust alternative to expected utility which is based on the median instead of the mean as a location parameter, and was introduced in De Schuymer, De Meyer, De Baets, and Jenei (2003), as an equivalent and graded version of the method presented in Boland, Hollander, Joag-Kev, and Kochar (1996); it is also a counterpart of the expected utility model when the rewards of the different alternatives are expressed in a qualitative scale (Dubois, Fargier, & Perny, 2003). We shall recall the basic aspects of these two criteria in Section 2.

In Section 3, we define a number of choice models for sets of alternatives starting from some binary relation, based on earlier work on this problem carried out in Montes, Miranda, and Montes (2013), and apply them to the particular cases where this relation is the one associated to stochastic dominance or statistical preference. Then we consider two particular cases: first, in Section 4 we deal with the case where we have precise information about the beliefs but imprecise one about the utilities. We model this situation by means of multi-valued mappings, or random sets

* Corresponding author. Tel.: +34 985186466; fax: +34 985103353.
E-mail addresses: imontes@uniovi.es (I. Montes), mirandaenrique@uniovi.es (E. Miranda), montes@uniovi.es (S. Montes).

0377-2217/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2013.09.013

(Dempster, 1967) which, under the epistemic interpretation considered in Kruse and Meyer (1987), can be seen as collections of random variables imprecisely specified. We show that under some conditions the comparison can be simplified using the lower and upper probabilities induced by the random set. Secondly, we consider in Section 5 the case where we have precise utilities but imprecise beliefs, and show that there are two additional notions that may be more useful in such a scenario.

The different conditions and their properties are illustrated by means of an example in Section 6. We conclude the paper by giving a number of additional remarks in Section 7.

2. Preliminary concepts

Let us review the basics about the two optimality criteria for decision making under uncertainty we shall consider in this paper. To clarify our set up, we consider a problem where we must choose between alternatives X, Y whose utilities depend on the values ω of the states of nature. We assume that we have probabilistic information about these states of nature, so that X, Y are defined as variables from a probability space $(\Omega, \mathcal{A}, P)$ and taking values on an utility space Ω'. For the most part, we shall assume that Ω' is a bounded subset of the reals; however, in the case of statistical preference we may have qualitative utilities, and then Ω' may correspond to an ordered qualitative scale.

2.1. Stochastic dominance

The notion of stochastic dominance between random variables is based on the comparison of their corresponding distribution functions. Assume that our utility scale is $\Omega' = [0,1]$ (the results in this section generalize immediately to the case where Ω' is any bounded interval of real numbers). Distribution functions are thus defined in the following way:

Definition 1. A cumulative distribution function on $[0,1]$ is a function F: $[0,1] \to [0,1]$ satisfying the following properties:

- $x \leqslant y \Rightarrow F(x) \leqslant F(y)\ \forall x,y$ [Monotonicity].
- $F(1) = 1$ [Normalization].
- $F(x) = \lim_{\epsilon \downarrow 0} F(x+\epsilon)\ \forall x < 1$ [Right-continuity].

Any F satisfying the properties of monotonicity and normalization is associated to a finite additive probability measure, and we shall call it a *finitely additive distribution function*.

One of the most popular methods for the comparison of cumulative distribution functions is stochastic dominance (Levy, 1998):

Definition 2. Given two cumulative distribution functions F and G, we say that F *stochastically dominates* G, and denote it $F \succeq_{\mathrm{FSD}} G$, if $F(t) \leqslant G(t)$ for every t in $[0,1]$, and given two random variables X, Y taking values on $[0,1]$, we say that X stochastically dominates Y, and denote it $X \succeq_{\mathrm{FSD}} Y$ when its associated distribution function F_X stochastically dominates F_Y, where

$F_X(t) = P(X \leqslant t)$ and $F_Y(t) = P(Y \leqslant t) \quad \forall t \in [0,1]$.

In the literature, this notion is sometimes called *first degree stochastic dominance*, in order to distinguish it from a number of weaker conditions called *second, third,...* degree stochastic dominance (Levy, 1998). This is the reason of the notation $\succeq_{\mathrm{FSD}}$. Occasionally the notation $\succeq_{st}$ is also employed (see for instance Müller & Stoyan, 2002).

This definition induces a partial order in the space $\mathbb{F}$ of cumulative distribution functions, from which we can derive the notions of strict stochastic dominance, indifference and incomparability:

- We say that F stochastically dominates G *strictly*, and denote it by $F \succ_{\mathrm{FSD}} G$, if $F \succeq_{\mathrm{FSD}} G$ but $G \not\succeq_{\mathrm{FSD}} F$. This holds if and only if $F \leqslant G$ and there is some $t \in [0,1]$ such that $F(t) < G(t)$.
- F and G are *stochastically indifferent*, and denote it by $F \equiv_{\mathrm{FSD}} G$, if $F \succeq_{\mathrm{FSD}} G$ and $G \succeq_{\mathrm{FSD}} F$, or equivalently, if $F = G$.
- F and G are *stochastically incomparable*, and denote it by $F \nsim_{\mathrm{FSD}} G$, if $F \not\succeq_{\mathrm{FSD}} G$ and $G \not\succeq_{\mathrm{FSD}} F$.

Thus, $(\mathbb{F}, \succ_{\mathrm{FSD}}, \equiv_{\mathrm{FSD}}, \nsim_{\mathrm{FSD}})$ constitutes a preference structure (Roubens & Vincke, 1985).

Stochastic dominance is commonly used in economics and finance (Denuit et al., 2005; Goovaerts et al., 1990) and can be given the following interpretation: $F \succeq_{\mathrm{FSD}} G$ means that the choice of F over G is rational, in the sense that we prefer the alternative with greater probability of providing a utility above a certain threshold t, and this for all possible t. The notion has also been used in other frameworks such as reliability theory, statistical physics, and epidemiology. We refer to Levy (1998), Müller and Stoyan (2002) for more information, and to Batur and Choobineh (2012), Dupacova and Kopa (2013) for recent works in the context of decision making. It is characterized by the following property:

Theorem 1. (Levy, 1998). *Given two random variables X and Y it holds that: $X \succeq_{FSD} Y$ if and only if $E(u(X)) \geqslant E(u(Y))$ for every non-decreasing u.*

The result is based on the equivalence between $X \succeq_{\mathrm{FSD}} Y$ and the inequality $E(I_{[t,\infty)}(X)) \geqslant E(I_{[t,\infty)}(Y))\ \forall t \in \mathbb{R}$, so if we denote by $\mathcal{U}^*$ the set of non-decreasing and bounded maps from $\mathbb{R}$ to $\mathbb{R}$, we may also characterize stochastic dominance by:

$$X \succeq_{\mathrm{FSD}} Y \iff E(u(X)) \geqslant E(u(Y)) \quad \text{for every } u \in \mathcal{U}^*. \tag{1}$$

2.2. Statistical preference

We next introduce the notion of statistical preference. One of its advantages is that it is applicable to variables X, Y taking values on any ordered qualitative scale Ω', which need not be numerical. It is based on the notion of probabilistic relation.

Definition 3. (Bezdek, Spillman, and Spillman, 1978). Given a set of alternatives $\mathcal{D}$, a *probabilistic relation* is a map $Q : \mathcal{D} \times \mathcal{D} \to [0,1]$ satisfying $Q(a,b) + Q(b,a) = 1$ for all a, b in $\mathcal{D}$.

Consider two variables X, Y from a probability space $(\Omega, \mathcal{A}, P)$ to an ordered utility space Ω' and define

$$Q(X,Y) = P(X > Y) + \frac{1}{2} P(X = Y); \tag{2}$$

then it is easy to see that Q is a probabilistic relation. The value $Q(X,Y)$ can be interpreted as a measure of the strength of our preference of X over Y. Statistical preference can then be introduced as a decision criterion based on this probabilistic relation:

Definition 4 (*De Schuymer, De Meyer, and De Baets, 2003; De Schuymer et al., 2003*). We say that the random variable X is:

- statistically preferred to Y, and denote it by $X \succeq_{\mathrm{SP}} Y$, if $Q(X,Y) \geqslant \frac{1}{2}$;
- strictly statistically preferred to $Y (X \succ_{\mathrm{SP}} Y)$ if $Q(X,Y) > \frac{1}{2}$;
- statistically indifferent to $Y (X \equiv_{\mathrm{SP}} Y)$ when $Q(X,Y) = \frac{1}{2}$.

Note that, if $\mathcal{D}$ denote a set of random variables defined on the probability space, $(\mathcal{D}, \succ_{\mathrm{SP}}, \equiv_{\mathrm{SP}})$ constitutes a preference structure without incomparable elements.