Preference Relations and Families of Probabilities: Different Sides of the Same Coin

Inés Couso*

Department of Statistics and O.R.
University of Oviedo
couso@uniovi.es

Abstract. The notion of preference is reviewed from different perspectives, including the Imprecise Probabilities' approach. Formal connections between different streams of the literature are provided, and new definitions are proposed.

Keywords: Imprecise probabilities, stochastic orderings, Walley's desirability, fuzzy rankings, possibility theory.

1 Introduction

The problem of finding appropriate rankings to evaluate and/or sort different alternatives appears in a variety of disciplines such as Economics ([2]), Political Sciences ([5]) or Artificial Intelligence ([16]), among many others. We can usually identify four main ingredients in a decision making problem, namely:

1. The set of alternatives (also referred to as "gambles" ([36]), "actions" ([4,32]),"options" ([3,24]) or "candidates" ([5,39]).
2. The "set of states of nature" –also called "criteria" in multi-criteria decision problems ([4,20,25,32]), "experts" ([13]) or "individuals" in group decision making ([22]), or "voters" in political decision making ([5,26])–.
3. An evaluation on every alternative and/or a preference ordering between different alternatives for each state of nature.
4. A merging method in order to combine evaluations or preference relations for different states of nature, that allows us to select the best (or a set of non dominated) alternative(s).

Different streams of the literature make different assumptions about the two last ingredient. With respect to the preference ordering over the set of alternatives, for each particular state of the nature, we face at least two different approaches:

* The author is very thankful to the Executive Directors and the Organizers for their kind invitation to participate in the conference. The research in this work has been supported by the Spanish grant TIN2011-24302. This financial support is gratefully acknowledged.

A. Laurent et al. (Eds.): IPMU 2014, Part I, CCIS 442, pp. 1–9, 2014.

3a. In the first one, a quantitative [12,10,20,35,36] or a qualitative (see [31], for instance) assessment is associated to every single alternative. In the case of quantitative assessments, the usual ordering between numbers can be a natural choice to sort the set of alternatives, although it is not the only option ([21]). Recently, (fuzzy) interval-valued assessments instead of crisp numerical values have been considered in some problems where imprecision is involved in this assessment process [1,34]), and therefore specific partial or total pre-orderings between (fuzzy) intervals must be selected in order to sort the different alternatives, for every particular state of nature.

3b. According to the second approach, preference statements over the set of alternatives are made, instead of considering value functions providing single assessments. Those preference statements can be given by means of graphical representations, for instance, and may lead to a partial ordering over the set of alternatives, for each state of nature or criterion (see [18,25], for instance.)

Regarding the combination of our preference orderings under the different states of nature into a single (partial) ordering, we can distinguish at least three approaches:

4a. The expert initial information is assessed by means of comparative preference statements between options. A (family of) probability measures(s) over the states of nature is derived from it. This is the "behavioral" approach initiated by Savage in a precise probabilities environment ([11,33]) and followed afterwards in the general theory of imprecise probabilities (see [8,36]).

4b. A (family of) probability measure(s) over the class of states of nature is initially considered. A preference relation on the set of alternatives ("random variables" or "gambles") is derived from it. This is the approach considered for instance in [21,9,12], for the case of precise probabilities and in [14,29,34,35]. for the case of partially known probabilities or weights. Such a preference relation, derived from the initial (sometimes partial) information about the weights of the different states of nature (or "criteria", in multi-criteria decision), is closely related to the notion of "almost preference" in Walley's framework. This approach will be reviewed in Subsection 3.2 ("credal set approach").

4c. This approach usually appears in multi-criteria decision problems. Traditionally, a weight function, formally equivalent to a probability measure, is defined over the (finite) set of criteria, in order to determine the relative importance of each of them. The global score of an alternative is therefore calculated according to the weighted mean, that can be seen as a (discrete) Lebesgue expectation. Lately ([27]) some authors have generalized this approach in order to include the possibility of interaction between different criteria. In order to do so, the weights associated to some families of criteria do not necessarily coincide with the sum of the weights of each of them. This kind of complex information can be represented by means of a non additive set-function defined over the set of criteria. According to this approach, the global score of each alternative may be determined by the Choquet integral of the partial scores with respect to such a set function. We will briefly review of this approach in Subsection 3.3 ("aggregation operators-based approach").

Once we have described the four usual ingredients in decision making problems, we need to take into account an additional issue. The decision-maker is usually interested by two questions: on the one hand, (s)he may wish to know what decisions are preferred to others; on the other hand, (s)he may be interested in determining whether single alternatives are "satisfactory" ([28]) or "desirable" ([36]) or not. P. Walley ([36]) established a formal link between both issues.

The goals in this paper are threefold:

- First, we aim to highlight what are the formal connections and commonalities between different combination methods (classical stochastic orderings, Walley's preferences, etc.) from the literature.
- Second, we will show Walley's partial orderings as a generalization of the expected utility criterion ([33]), and we will explore other possible generalizations of stochastic orderings within the Imprecise Probabilities setting.
- Finally, we will face the problem of ranking fuzzy intervals from the perspective of Imprecise Probabilities, assuming that each fuzzy interval characterizes a possibility distribution, considered as an upper probability.

Sections 2 and 3 deal with the two first goals, and Section 4, with the third one.

2 Preference Modeling in a Probabilistic Setting: Stochastic Orderings

Let us first consider a probability space $(\Omega, \mathcal{F}, P)$ determining the "weights" of the different states of the nature. Let us consider a pair of random variables (alternatives) defined on Ω, $X, Y : \Omega \to \mathbb{R}$. This section reviews three well known stochastic orderings in the literature:

SO1. *Dominance in the sense of expected utility* [33].- Given an increasing function $u : \mathbb{R} \to \mathbb{R}$, X dominates Y wrt u if $E_P(u(X)) \geq E_P(u(Y))$. We will denote it $X \geq_u Y$. A special case is *Dominance in Expectation*: X dominates Y if $E_P(X) \geq E_P(Y)$. This relation represents the particular case of the previous one, when the utility function u is the identity function $u(x) = x$, $\forall\, x \in \mathbb{R}$.

SO2. *Statistical preference* [9].- X is statistically preferred to Y if $P(X > Y) + 0.5 \cdot P(X = Y) \geq 0.5$. We will denote it $X \geq_{SP} Y$.

SO3. *First order stochastic dominance* [21].- X dominates Y if $P(X > a) \geq P(Y > a), \forall\, a \in \mathbb{R}$. It is well known that $X \geq_{1st} Y$ if and only if $X \geq_u Y$, for all increasing utility functions $u : \mathbb{R} \to \mathbb{R}$. We will denote it $X \geq_{1st} Y$.

According to [7], all the above stochastic orderings can be put into a common formulation. In fact, we can express each of them as follows:

$$X \text{ is preferred to } Y \text{ if and only if } E_P[f(X,Y)] \geq E_P[f(Y,X)], \tag{1}$$

$$\text{or equivalently, if } E_P[f(X,Y) - f(Y,X)] \geq 0, \tag{2}$$

for a certain function $f : \mathbb{R}^2 \to \mathbb{R}$, increasing in the first component and decreasing in the second one. Below, we provide the specific expression of $f(\cdot,\cdot)$ for each of the above orderings:

- *Dominance in the sense of expected utility.*- Let us consider the function $f_u(x,y) = u(x), \forall\, x \in \mathbb{R}$ (which is a constant wrt the second argument). We can easily check that $X \geq_u Y$ if and only if $E_P[f_u(X,Y)] \geq E_P[f_u(Y,X)]$.
- *Statistical preference.*- Let us consider the mapping $f(x,y) = \operatorname{sgn}(x-y)$, where sgn denotes the "sign" function taking the values $1, 0$ or -1, depending on the sign of the argument. It is checked in [7,8] that $X \geq_{SP} Y$ if and only if $E[f(X,Y)] \geq E[f(Y,X)]$.
- *First stochastic dominance.*- Let us now consider the family of functions $f^a(x,y) = 1_{x>a} - 1_{y>a},\ \forall\, a \in \mathbb{R}$. X dominates Y if and only if $E[f^a(X,Y)] \geq E[f^a(Y,X)],\ \forall\, a \in \mathbb{R}$. Equivalently, and according to the formal existing relation between dominance in the sense of expected utility and first stochastic dominance, we can say that X dominates Y if and only if $E[f_u(X,Y)] \geq E[f_u(Y,X)]$, for every (increasing) utility function $u : \mathbb{R} \to \mathbb{R}$.

3 Partial Orderings in Imprecise Probabilities

3.1 The Behavioral Approach

Peter Walley [37] establishes a list of axioms of coherence for preference relations between gambles. A preference relation $\succeq$ defined on a linear space of gambles[1] $\mathcal{K}$ is said to be *coherent* when it satisfies the following properties:

P1. Not $X \succeq X$.
P2. If $X(\omega) \geq Y(\omega),\ \forall\, \omega \in \Omega$ and $X(\omega) > Y(\omega)$, for some $\omega \in \Omega$, then $X \succeq Y$.
P3. If $X \succeq Y$ and $c > 0$ then $cX \succeq cY$.
P4. If $X \succeq Y$ and $Y \succeq Z$ then $X \succeq Z$.
P5. $X \succeq Y$ if and only if $X - Y \succeq 0$.

He also establishes a duality between the notions of *desirability* and *preference*: X is desirable if and only if it is preferred to the null gamble (the gamble that provides no reward, no matter the state of the nature, $\omega \in \Omega$.) Conversely, X is preferred to Y when their difference $X - Y$ is a desirable gamble. In such a case, one is willing to give up Y in return for X. Any coherent preference relation

[1] A gamble $X : \Omega \to \mathbb{R}$ is a bounded mapping defined on the space of states of nature. If you were to accept gamble X and ω turned to be true, then you would gain $X(\omega)$. (This reward can be negative, and then it will represent a loss.)

determines a *lower prevision*, $\underline{P}$, representing the supremum of acceptable buying prices for gambles, that is formally defined as follows:

$$\underline{P}(X) = \sup\{c \in \mathbb{R} : X \succeq c\}, \ \forall X.$$

The lower prevision $\underline{P}$, determines the credal set[2], $M(\underline{P})$, formed by the set of linear previsions that dominate $\underline{P}$: $M(\underline{P}) = \{P : P(X) \geq \underline{P}(X), \ \forall X\}$. According to the last formula, it can be interpreted as the minimum of the expectation operators associated to a convex family of (finitely additive) probabilities. Furthermore, the following implications hold for any pair of gambles, X, Y:

$$\underline{E}(X - Y) > 0 \Rightarrow X \text{ is preferred to } Y \Rightarrow \underline{E}(X - Y) \geq 0. \tag{3}$$

According to Equation 3, Walley's preference clearly generalizes the "dominance in expectation" criterion reviewed in Section 2.

We have recently explored ([8]) the generalization of statistical preference, defining the preference of X over Y as the desirability of the sign of their difference, $\mathrm{sgn}(X - Y)$. Under this criterion, we just take into account whether the consequent $X(\omega)$ is greater than (or preferred to) $Y(\omega)$ or not, but not the magnitude of their difference. Thus, it does not require the existence of a numerical scale, unlike Walley's criterion. The following equation reminds Eq. 3, where the lower expectation has been replaced by the median:

$$\underline{\mathrm{Me}}(X - Y) > 0 \Rightarrow X \text{ is signed preferred to } Y \Rightarrow \underline{\mathrm{Me}}(X - Y) \geq 0.$$

Taking into account the last section, and some of the ideas suggested in [7], the notion of first stochastic dominance could be easily generalized to the imprecise probabilities setting, if we consider that X is preferred to Y whenever the gamble $1_{X>a}$ is preferred to $1_{Y>a}$, for every $a \in \mathbb{R}$.

3.2 The Credal Set-Based Approach

Lately, different authors have independently generalized some stochastic orderings to the case where our imprecise information about the underlying probability distribution, P, over the set of states of nature is determined by means of a set of probability measures $\mathcal{P}$. Taking into account the general formulation of stochastic orderings proposed in Section 2, many of those new definitions can be seen from a general perspective, where the values $E_P[f(X,Y)]$, $E_P[f(Y,X)]$ and $E_P[f(X,Y)] - E_P[f(Y,X)]$ are replaced by their respective sets of admissible values. More specifically, they can be seen as generalizations of Eq. 1 and 2, where the inequality $\geq$ is replaced by a particular (sometimes partial) preorder between intervals (or, more generally, between arbitrary sets of numbers). In this respect, the criteria considered in [35] generalize the "dominance in expectation"

[2] A credal set is a convex and closed family of linear previsions, i.e., of linear functionals $P : \mathcal{L} \to \mathbb{R}$ satisfying the constraint $P(1) = 1$. The last concept generalizes the notion of expectation operator to the case of non necessarily $\sigma-$additive probabilities.

criterion (SO1). On the other hand, the criterion considered in [34] generalizes the notion of statistical preference (SO2), to the case where the usual ordering between numbers is replaced by an interval ordering ([19]). As a side remark, let us notice that there exist a formal connection between this last criterion and the notion of *fuzzy preference* ([30]). In fact, it can be easily checked that the mapping f defined as $f(X,Y) = \overline{P}(X,Y)$ is a *connected* [3] or *complete* fuzzy preference, since it satisfies the general axioms of fuzzy preferences and the inequality $f(X,Y) + f(Y,X) \geq 1$. If, in addition, $\overline{P}$ is a possibility measure, then the fuzzy preference relation derived from it is *strongly connected.* Finally, the four generalizations of "first stochastic dominance" (SO3) proposed in [14] and the six ones proposed in [29] also follow the general procedure described in this section. A more detailed explanation about how the criteria reviewed in this section fit this general formulation is provided in [7].

3.3 The Aggregation Operators-Based Approach

The weighted mean is a very often used aggregation criterion in multicriteria decision problems, and it is formally equivalent to "dominance in expectation", where the probability masses play the role of "weights of importance" of the different criteria instead of being interpreted in terms of stochastic uncertainty. During the last decades (see [27] and references therein), generalizations of the weighted mean operator, like the discrete Choquet integral (that includes OWA operators [38] as particular cases) have been considered. Within this more general setting, the degree of importance associated to a particular set of criteria is not forced to coincide with the sum of the weights assigned to the particular criteria included in the set, allowing the possibility of modeling the effect of interaction between different criteria. The discrete Choquet integral wrt the non-additive monotone set-function, μ, that assigns the weight $\mu(C)$ to each particular set of criteria, C, seems to be the natural generalization of the weighted mean under this general approach. If the non-additive set function satisfies some additional properties, like *submodularity*, for instance, the Choquet integral plays the role of a lower prevision or a lower expectation. In those cases, the resulting aggregation method is formally linked to the preference criterion defined by Walley reviewed in Section 3.1, as well as to the generalizations of the "dominance in expectation" criteria considered in Section 3.2.

Labreuche et al. [25] also interpret the "weights" as relative degrees of importance of criteria. In their paper, partial preference relations between the options, instead of single evaluations are considered for each particular criterion (see 3a). Those preference relations are combined afterwards according to "statistical preference" (SO2). Weights on sets of criteria are not directly provided. Instead, preference relations between indicator functions over the set of criteria are given, and the set of feasible probabilities (weight vectors) is derived from them. So the approach in this paper is very much connected to the procedure described in (4a). It would be interesting to explore the formal connection between this approach and the preference criterion proposed in [8].

4 An Imprecise Probabilities' Approach to the Notion of Fuzzy Ranking

According to the possibilistic interpretation of fuzzy sets, the problem of ranking fuzzy numbers can be seen from an Imprecise Probabilities' perspective ([6,17]. In fact, a pair of fuzzy sets can be seen as an incomplete description of a joint probability measure $P_{(X,Y)}$, and therefore, any of the criteria reviewed in Sections 2 and 3 could be applied. For instance, Detyniecki et al. [15] apply statistical preference to the joint probability distribution induced by a pair of independent random variables whose respective density functions are proportional to the respective fuzzy sets membership functions, considered as possibility distributions. Sánchez et al. ([34]) also generalize statistical preference, but this time, they consider the whole set of probability measures dominated by a joint upper probability determined by both possibility distributions. The criterion of dominance of expectation has been also generalized in the recent literature (see [10] for instance). A deep analysis studying well-known fuzzy rankings from the perspective of Imprecise Probabilities has been developed in [6].

References

1. Aiche, F., Dubois, D.: An extension of stochastic dominance to fuzzy random variables. In: Hüllermeier, E., Kruse, R., Hoffmann, F. (eds.) IPMU 2010. LNCS, vol. 6178, pp. 159–168. Springer, Heidelberg (2010)
2. Arrow, K.J.: Social Choice and Individual Values, 2nd edn. Wiley, New York (1963)
3. Barret, R., Salles, M.: Social choice with fuzzy preferences. In: Handbook of Social Choice and Welfare, pp. 367–389. Elsevier (2004)
4. Chiclana, F., Herrera, F.: Integrating three representation models in fuzzy multipurpose decision making based on fuzzy preference relations. Fuzzy Sets Syst. 97, 33–48 (1998)
5. Condorcet, M.: Essay sur le application de l'analyse à la probabilité des decisions rendues à la pluralité des voix. Imprimirie Royale, Paris (1785)
6. Couso, I., Destercke, I.: Ranking of fuzzy numbers seen through the imprecise probabilistic lense. In: Eurofuse 2013 (2013)
7. Couso, I., Dubois, D.: An imprecise probability approach to joint extensions of stochastic and interval orderings. In: Greco, S., Bouchon-Meunier, B., Coletti, G., Fedrizzi, M., Matarazzo, B., Yager, R.R. (eds.) IPMU 2012, Part III. CCIS, vol. 299, pp. 388–399. Springer, Heidelberg (2012)
8. Couso, I., Sánchez, L.: The behavioral meaning of the median. In: Borgelt, C., González-Rodríguez, G., Trutschnig, W., Lubiano, M.A., Gil, M.Á., Grzegorzewski, P., Hryniewicz, O. (eds.) Combining Soft Computing and Statistical Methods in Data Analysis. AISC, vol. 77, pp. 115–122. Springer, Heidelberg (2010)
9. David, H.: The method of paired comparisons, Griffin's Statistical Monographs & Courses, vol. 12. Charles Griffin & D. Ltd., London (1963)
10. De Cooman, G.: Further thoughts on possibilistic previsions: A rejoinder. Fuzzy Sets Syst. 153, 375–385 (2005)
11. De Finetti, B.: Theory of Probability, vol. 1 & 2. John Wiley & Sons (1974-1975)
12. De Schuymer, B., De Meyer, H., De Baets, B., Jenei, S.: On the cycle-transitivity of the dice model. Theory and Decision 54, 261–285 (2003)

13. Delgado, M., Herrera, F., Herrera-Viedma, E., Martínez, L.: Combining numerical and linguistic information in group decision making. Inform. Sciences 107, 177–194 (1998)
14. Denœux, T.: Extending stochastic ordering to belief functions on the real line. Inform. Sciences 179, 1362–1376 (2009)
15. Detyniecki, M., Yager, R.R., Bouchon-Meunier, B.: A context-dependent method for ordering fuzzy numbers using probabilities. Inform. Science 138, 237–255 (2001)
16. Domshlak, C., Hüllermeier, E., Kaci, S., Prade, H.: Preferences in AI: An overview. Artificial Intelligence 175, 1037–1052 (2011)
17. Dubois, D.: The role of fuzzy sets in decision sciences: Old techniques and new directions. Fuzzy Sets and Systems 184, 3–28 (2011)
18. Dubois, D., Fargier, H., Prade, H.: Decision making under ordinal preferences and comparative uncertainty. In: Proceedings of International Conference on Uncertainty in Artificial Intelligence (UAI 1997), pp. 157–164 (1997)
19. Fishburn, P.C.: Interval Orderings and interval graphs. Wiley, New-York (1985)
20. Grabisch, M., Labreuche, C.: A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. Ann. Oper. Res. 175, 247–286 (2010)
21. Hadar, J., Russell, W.: Rules for Ordering Uncertain Prospects. American Economic Review 59, 25–34 (1969)
22. Herrera, F., Herrera-Viedma, E., Verdegay, J.L.: A model of consensus in group decision making under linguistic assessments. Fuzzy Sets Syst. 78, 73–87 (1996)
23. Kaci, S., Labreuche, C.: Argumentation framework with fuzzy preference relations. In: Hüllermeier, E., Kruse, R., Hoffmann, F. (eds.) IPMU 2010. LNCS, vol. 6178, pp. 554–563. Springer, Heidelberg (2010)
24. Kacprzyk, J.: Group decision making with a fuzzy linguistic majority. Fuzzy Sets Syst. 18, 105–118 (1986)
25. Labreuche, C., Maudet, N., Ouerdane, W.: Justifying Dominating Options when Preferential Information is Incomplete. In: ECAI 2012, pp. 486–491 (2012)
26. Ladha, K.K.: The Condorcet Jury Theorem, Free Speech, and Correlated Votes. American Journal of Political Science 36, 617–634 (1992)
27. Marichal, J.L.: An Axiomatic Approach of the Discrete Choquet Integral as a Tool to Aggregate Interacting Criteria. IEEE Trans. Fuzzy Syst. 8, 800–807 (2000)
28. Marichal, J.L.: Aggregation Functions for Decision Making. In: Bouysou, et al. (eds.) Decision-making Process. Concepts and Methods. Wiley (2009)
29. Montes, I., Miranda, E., Montes, S.: Stochastic dominance with imprecise information. Computational Statistics & Data Analysis 71, 868–886 (2014)
30. Orlovsky, S.: Decision-making with a fuzzy preference relation. Fuzzy Sets Syst. 1, 155–168 (1978)
31. Rico, A., Grabisch, M., Labreuche, C., Chateauneuf, A.: Preference modeling on totally ordered sets by the Sugeno integral. Discrete Appl. Math. 147, 113–124 (2005)
32. Roubens, M.: Preference relations on actions and criteria in multicriteria decision making. Eur. J. Oper. Res. 10, 51–55 (1982)
33. Savage, L.J.: The Foundations of Statistics. Wiley (1954)
34. Sánchez, L., Couso, I., Casillas, J.: Genetic Learning of Fuzzy Rules based on Low Quality Data. Fuzzy Sets and Systems 160, 2524–2552 (2009)
35. Troffaes, M.: Decision making under uncertainty using imprecise probabilities. Int. Journal Appr. Reas. 45, 17–29 (2007)

36. Walley, P.: Statistical reasoning with imprecise probabilities. Chapman and Hall, London (1991)
37. Walley, P.: Towards a unified theory of imprecise probability. Int. J. Appr. Reas. 24, 125–148 (2000)
38. Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Trans. Syst., Man, Cybern. 18, 183–190 (1988)
39. Young, H.P.: Condorcet's Theory of Voting. American Political Science Review 82, 1231–1244 (1988)