Theory of Sets of Probabilities (and related models) in a Nutshell

Fabio Cozman

December 30, 1999

This is an attempt to *very* briefly cover essential aspects of the theory of sets of probabilities and its cousins, lower previsions, lower probability, lower envelopes, Choquet capacities, etc. I concentrate on the theory of sets of probabilities exposed and axiomatized by Giron and Rios [5], as the others can be derived from it, or at least can be explained from it.

Usually, a theory of sets of probabilities is a *normative theory of decision making.* So the purpose is to explain how an agent should make decisions. Making a decision is deciding which possible act to follow. Theories are *normative* because they only offer guidelines; they are not supposed to be a description of how real agents work. But they offer some sensible guidelines that have been discussed and used for ages.

The theory of sets of probabilities advocates that a rational agent chooses an act based on expected loss considerations. Expected loss is defined by two entities:

1. A loss function, which translates the preferences of the agent. Suppose an agent prefers a banana as much as the agent prefers two apples plus an orange. Then the loss function l(.) must be such that:

   $$\mathrm{l(banana)} = 2\ \mathrm{l(apple)} + \mathrm{l(orange)}.$$

   Some use the term utility for the reverse of loss; i.e., loss with a minus sign. I will loss as the basic concept, because most of the literature in Statistics and Electrical Engineering use it this way. But loss and utility are essentially the same thing.

2. A set of probability distributions. If you are used to Bayesian theory, you have heard about a single probability distribution. Here we talk about a set of probability distributions. Using the banana example: in usual probability theory you say: "I think the

chance I'll have a banana is 40%, so I intend to bet in bananas at 2:5 odds". Note the imprecision of the probability statements; from them you can say: "I intend to bet in bananas at odds that go from 1:5 up to 3:5 (20% to 60%)".

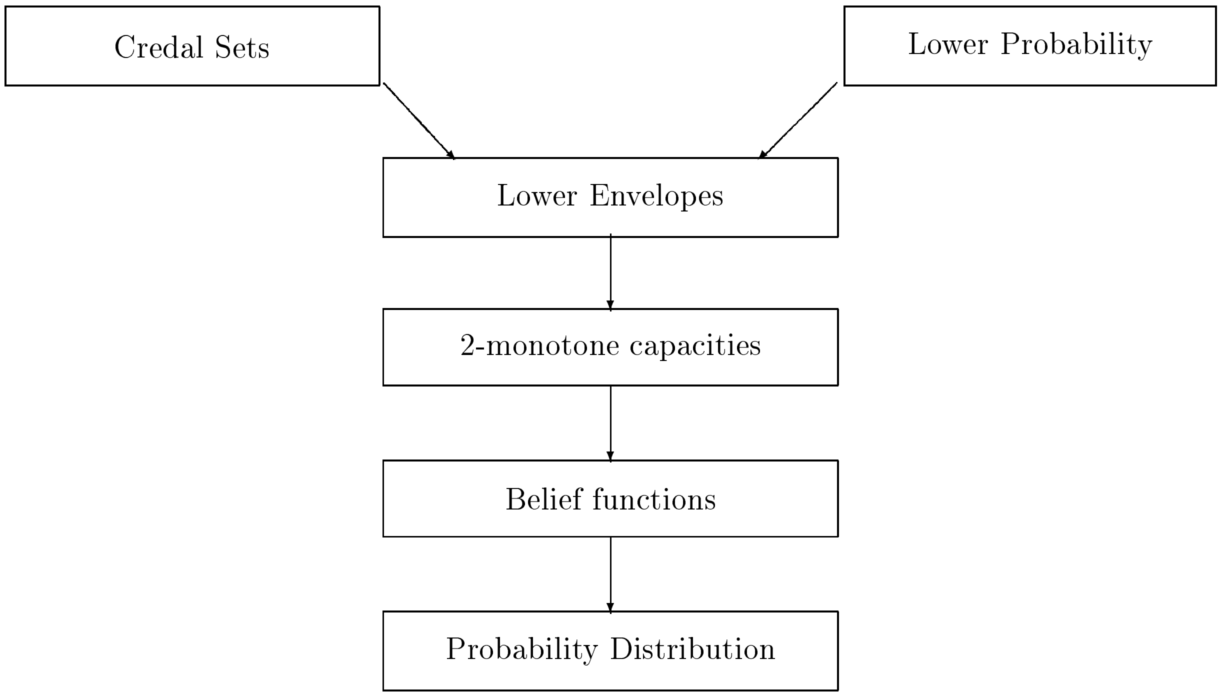
The set of probability distributions is called the *credal set* [7, 8]. Usually a credal set is assumed to be a convex set of probability distributions. A credal set conveys the beliefs of an agent about the possible states of the world.

So you see that the theory is strongly subjectivist, just like Bayesian theory [2, 3, 10]. The subjective beliefs and preferences of the agent are represented by mathematical entities; those are the credal set and the loss function respectively. And the agent tries to pick the option with lower expected loss. But notice that, since there is a *set* of distributions, there may be a *set* of possible options.

In real inference problems, the agent starts with a prior credal set, uses a likelihood credal set and reaches a posterior credal set. Then the agent picks an option that minimizes expected loss (there may be more than one option that is admissible). Note that Bayesian theory is the particular case of a credal set with a single element. People that like sets of distributions point out that real experts are more likely to specify probability values in intervals: "Well, I guess the probability of rain is between 20% and 25%...", and so on [11]. Also, people on the sets of probabilities camp can use a credal set to investigate how sensitive decisions are with respect to an *extended* range of models; that's mostly what is done in Robust Bayesian Statistics [1, 6]. Others simply point out that even if experts give sharp numbers, real agents are more faithfully represented by a theory that does not *require* that a single number be produced for every possible act, specially if a group of agents is to be considered [9]. And others say that, for an agent with finite computational resources, it may be advantageous to manipulate sets of distributions as a cruder (abstracted) model of a particular decision problem.

There are many theories, more or less related to Bayesian theory, that also use sets or intervals of probabilities to represent an agent's preferences or beliefs. Here are some you may have heard: Lower Probabilities [4], Choquet Capacities, Monotone Capacities, 2-monotone Capacities, Infinitely monotone Capacities (belief functions), Lower Envelopes [6], Lower Expectations, Lower Previsions [11]. The last five names are essentially identical to "Convex Sets of Probabilities of Some Sort". Quasi-Bayesian theory is just a possible axiomatization of these theories, which is very general (it actually emcompasses all of them) and which is based on decision theory. The first six names come from somewhat different axiomatizations which emphasize the use of *intervals* of probabilities, rather than sets of probability distributions. There are several definitions for each one of these things, so my rough classification may be a bit too rough for some cases.

The following figure gives you a simplified family tree. The down arrows indicate that the upper family contains the lower family (e.g., Lower Envelopes can always be represented as convex sets of probability distributions); but no arrow means equality. Note that I used Lower Probabilities to mean "dominated or undominated lower probabilities" and Lower Envelopes to mean "lower bounds of sets of probabilities" (that is, Lower Envelopes are dominated Lower Probabilities) — this is not entirely standard terminology as there is no complete agreement on the what exactly are Lower Probabilities.



References

[1] J. O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, 1985.

[2] M. DeGroot. *Optimal Statistical Decisions*. McGraw-Hill, New York, 1970.

[3] J. Earman. *Bayes or Bust?* The MIT Press, Cambridge, MA, 1992.

[4] T. L. Fine. Lower probability models for uncertainty and nondeterministic processes. *Journal of Statistical Planning and Inference*, 20:389–411, 1988.

[5] F. J. Giron and S. Rios. Quasi-Bayesian behaviour: A more realistic approach to decision making? In J. M. Bernardo, J. H. DeGroot, D. V. Lindley, and A. F. M. Smith, editors, *Bayesian Statistics*, pages 17–38. University Press, Valencia, Spain, 1980.

[6] P. J. Huber. *Robust Statistics.* Wiley, New York, 1980.

[7] I. Levi. *The Enterprise of Knowledge.* MIT Press, Cambridge, Massachusetts, 1980.

[8] I. Levi. *Hard Choices: Decision Making under Unresolved Conflict.* Cambridge University Press, Cambridge, 1986.

[9] I. Levi. Compromising Bayesianism: A plea for indeterminacy. *Journal of Statistical Planning and Inference*, 25:347–362, 1990.

[10] J. Pearl. On probability intervals. *International Journal of Approximate Reasoning*, 2:211–216, 1988.

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