Some results about unbounded expected utilities

based on joint work [SeScK 2009 and ScSeK 2009)

Mark Schervish, Teddy Seidenfeld, and Jay Kadane – CMU

This presentation engages several challenges for an *Expected Utility* theory of *coherent* preferences over *random quantities* when:

1. **Utilities are (for random variables that are) *unbounded*.**

2. ***Coherence* – that is, avoidance of uniform dominance in the partition by *states* – is the liberal standard for rational preference afforded by de Finetti's (1974) theory.**

- ***Part 1*, I review de Finetti's (finitely additive) *coherence* criterion.**
- ***Part 2*, I present some fresh challenges that confront a theory of coherent preference for unbounded quantities.**
- **Part 3, I give a progress report towards a *de Finetti*-style theory of *Expected Utility* for unbounded quantities.**

Part 1: A review of de Finetti's theory of *coherent previsions*.

An important and historically early application of strict dominance in decision making is de Finetti's criterion of *coherence of previsions*.

De Finetti uses (1) a privileged *partition* Ω and
(2) a *class of variables* defined on Ω.

(1) A privileged partition into *states*, $\Omega = \{\omega_i: i \in I$ an index set$\}$.

→ (uniform, strict) Dominance applies in the privileged partition Ω.
That is what defines *states* in de Finetti's theory of coherent preference, in contrast with other partitions.

***Conceptual aside*: For validity of dominance, there are no moral hazards –**
There is no act/state dependence.

(2) A class of bounded, real-valued variables, $\chi = \{X_j: j \in J\}$, defined on Ω.

***Technical aside*: Since de Finetti's theory (as well as Savage's) accommodates finitely additive expectations, one may use the powerset of Ω as the algebra, thereby sidestepping issues about measurability of the variables.**

For each random variable $X \in \chi$, the rational agent has a (2-sided) *prevision* P(X) which is to be interpreted as a *fair* price (both for *buying* and *selling*) gambles on the outcomes of the variable X.

For all real $\beta > 0$, small enough so that the agent is willing to pay the possible losses, the agent is willing

to pay βP(X) in order *to buy* (i.e., to receive) βX in return.

and, is willing

to accept βP(X) in order *to sell* (i.e., to pay) βX in return.

That is, the agent will accept the *fair gamble*

$$\beta[X - \mathrm{P}(X)]$$

as a change in fortune, for all sufficiently small (positive or negative) β.

The agent is required to accept all <u>finite</u> sums of fair gambles of the preceding form. That is, for all finite n and all small, real $\beta_1, \ldots, \beta_n$ and all $X_1, \ldots, X_n \in \chi$, the agent will accept the combination of fair gambles as fair:

$$\sum_{i=1}^{n} \beta_i[X_i - P(X_i)].$$

Where β_i is positive, the agent buys β_i-units of X_i for a price of $\beta_i P(X_i)$.

Where β_i is negative, the agent sells β_i-units of X_i for a price of $\beta_i P(X_i)$.

- **The previsions are *<u>incoherent</u>* if there is a *uniformly negative* (“fair”) finite combination of acceptable gambles.**

That is, the agent’s 2-sided previsions are incoherent if there exists a sum of the form above and $\varepsilon > 0$ such that, for each $\omega \in \Omega$,

$$\sum_{i=1}^{n} \beta_i[X_i(\omega) - P(X_i)] < -\varepsilon.$$

Otherwise the agent’s previsions are *<u>coherent</u>*.

- **Note the special role played by the privileged partition Ω in defining *coherence*.**

With incoherent previsions, the sure-loss *book* constitutes a combination of gambles that is uniformly, strictly dominated by *not-betting* (= 0).

De Finetti's *Coherence Theorem*:

- **A set of previsions are coherent if and only if they are the expected values for the respective random variables under a (finitely additive) probability distribution over Ω.**

- **When variables are indicator functions for events (subsets of Ω), coherent previsions are the values of a finitely additive probability. And then the $|\beta_i|$ are the stakes in winner-take-all bets, where the previsions fix betting rates, $P(X_i) : 1-P(X_i)$.**

De Finetti's result applies to *called-off previsions*, given an event *B*.
These use called-off gambles of the form

$$B\,\beta[X - \mathrm{P}(X)]$$

where *B* is the indicator function for the conditioning event *B*.

Then, with a proviso about using only *non-null* conditioning events, coherence of all previsions assures that

- **Coherent called off (2-sided) previsions are finitely additive conditional expectations, given the conditioning event.**

***Note well*: Called-off previsions correspond to *normal form* decisions, and not to *extensive form* decisions.**

There is no *dynamical coherence* in de Finetti's theory. His theory covers merely *static* aspects of coherence, since conditional expectations are matched with called-off previsions.

- **Thus, de Finetti's theory of coherence does not require updating/learning by Bayesian conditional probabilities.**

A second, and equally important result is de Finetti's

Fundamental Theorem of Coherent Previsions

Suppose that coherent (2-sided) previsions P(·) are given for all variables in a set χ defined with respect to Ω.

Let *Y* be a real-valued function defined on Ω but not in χ.

Define: $\underline{A} = \{X: X(\omega) \leq Y(\omega)$ **and** X **is in the linear span of** $\chi\}$

$\overline{A} = \{X: X(\omega) \geq Y(\omega)$ **and** X **is in the linear span of** $\chi\}$

Let

$$\underline{P}(Y) = \sup_{X \in \underline{A}} P(X) \quad \text{and} \quad \overline{P}(Y) = \inf_{X \in \overline{A}} P(X)$$

Then **the 2-sided prevision, P(*Y*), may be any finite number from $\underline{P}(Y)$ to $\overline{P}(Y)$ and the resulting enlarged set of previsions is coherent.**

Outside this closed interval, the enlarged set of previsions are incoherent.

Part 2: Dominance for unbounded random variables (SeScK, 2009)

Challenge: With unbounded utilities, _coherent_ preferences,

(i.e., preferences that respect simple dominance in the partition Ω)

do not also respect indifferences between _equivalent variables_.

Definition: Two variables are _equivalent_ if they have the same probability distribution over outcomes.

Example: Consider a fair coin toss with P(_H_) = P(_T_) = ½
Let _X_ be the variable _X_(_H_) = 1 and _X_(_T_) = 0
Let _Y_ be the variable _Y_(_H_) = 0 and _Y_(_T_) = 1.
_X_ and _Y_ are equivalent as P(_X_=1) = P(_Y_=1) = ½, etc.

- **In canonical _EU_-theories, utility is over the outcomes of variables: the decision maker is _indifferent_ between equivalent variables.**

  See: von Neumann-Morgenstern (1947); Savage (1954); Anscombe-Aumann (1963).

- **In these theories, preference is defined over _lotteries_ (aka _gambles_), which are the equivalence classes of equivalent variables.**

Two Heuristic Examples illustrating the *Challenge*

Each of the following two examples provides a collection of unbounded but equivalent variables that cannot all be indifferent to each other, on pain of incoherence.

If variables *X* and Y have equal expected utility, EU(*X*) = EU(*Y*), then their difference, *Z* = *X*-*Y*, is indifferent to the *status quo* – EU(*Z*) = 0.

But particular combinations of the following equivalent random variables have differences that are (uniformly) bounded away from 0.

Hence, they cannot all be indifferent to one another.

Common structure for both heuristic examples

- **Let events E_n (n = 1, …) form a partition $\pi_E = \{E_n\}$ with a Geometric(½) probability distribution: P(E_n) = 2^{-n} (n = 1, 2, …).**

  Flip a fair coin until the first head. E_n = first head on flip #n.

- **Let $\pi_A = \{A_H, A_T\}$ be the outcome of another fair-coin flip, independent of the events E_n. $P(A_H|E_n) = P(A_H) = ½$.**

- **Consider the countable state-space $\pi_A \times \pi_E$.**

Heuristic Example 1: St. Petersburg variables

Define three (equivalent) St. Petersburg random variables, *X*, *Y*, and *Z*, as follows.

	E_1	E_2	....	E_n	....
A_H	$Z = 2$	$Z = 4$		$Z = 2^n$	
	$X = 4$	$X = 8$		$X = 2^{n+1}$	
	$Y = 2$	$Y = 2$		$Y = 2$	
A_T	$Z = 2$	$Z = 4$		$Z = 2^n$	
	$X = 2$	$X = 2$		$X = 2$	
	$Y = 4$	$Y = 8$		$Y = 2^{n+1}$	

For each state in $\pi_A \times \pi_E$,

$$X + Y - 2Z = 2, \text{ a constant quantity.}$$

This situation contradicts indifference between all 3 pairs of these equivalent variables. Such indifference requires that the expected utility EU(·) for the difference between two equivalent variables is 0. In this example, that entails,

$$EU(X - Z) + EU(Y - Z) = EU(X + Y - 2Z) = 0.$$

But the utility of a constant is that constant.

So, $EU(X + Y - 2Z) = 2$ **a contradiction.**

Thus, coherent preferences, here, are <u>not</u> defined merely by the probability distribution of utility outcomes.

Aside: ***Heuristic Example*** **1 uses non-Archimedean preference. The St. Petersburg variables do not have finite utility.** ***Heuristic Example 2*** **uses Archimedean preferences.**

***Heuristic Example* 2 – Coherent *boost* for unbounded variables.**

As before, consider the countable state-space $\pi_A \times \pi_E$,

with the Geometric(½) probability distribution on π_E,

and with an independent "fair coin" distribution on π_A.

Define the three equivalent (Geometric) random variables *X*, *Y*, and *Z*.

	E_1	E_2	....	E_n	....
	$X = 1$	$X = 2$		$X = n$	
A_H	$Y = 2$	$Y = 3$		$Y = n+1$	
	$Z = 1$	$Z = 1$		$Z = 1$	
	$X = 1$	$X = 2$		$X = n$	
A_T	$Y = 1$	$Y = 1$		$Y = 1$	
	$Z = 2$	$Z = 3$		$Z = n+1$	

- **X, Y, and Z are equivalent Geometric(½) variables.**

  But for each state in $\pi_A \times \pi_E$, $\quad Y + Z - X = 2.$

If all equivalent variables have equal Expected Utility

$$\mathbf{EU}(Y - X) + \mathbf{EU}(Z - X) = 0 \qquad \textit{if and only if}$$

$$\mathbf{EU}(Y) = \mathbf{EU}(Z) = \mathbf{EU}(X) = 2.$$

Then Expected Utility for a Geometric(½) variable X is its *countably additive* expectation, 2, and Expected Utility is continuous from below.

Specifically, **if a sequence of variables $< X_n > \rightarrow X$ (pointwise convergence)**
and for each state ω, $X_n(\omega) \leq X(\omega)$,
then $\lim_{n\to\infty} \mathbf{EU}(X_n) = \mathbf{EU}(X)$.

That is, in order to have indifference over equivalent Geometric(½) random variables, preferences must be continuous from below.

However, de Finetti's theory of *coherence* requires only that preference respects (uniform) dominance in the partition by *states*.

This entails respecting *bounds* from sequences of bounded random variables without requiring continuity from below.

Consider, the unbounded Geometric(½) variable X from the example, where

$$X(\{A_T, E_n\}) = X(\{A_T, E_n\}) = n; \text{ with } P(E_n) = 2^{-n}.$$

Let X_n be the bounded, truncated variable:

$$X_n(\{A_T, E_m\}) = X(\{A_T, E_m\}) = m \text{ for } m \leq n$$

and $$X_n(\{A_H, E_m\}) = X(\{A_H, E_m\}) = 0 \text{ for } m > n.$$

So, for each n = 1, 2, …, and for each state ω,

$$X_n(\omega) \leq X(\omega).$$

Also, $$\langle X_n \rangle \rightarrow X.$$

Respect for (uniform) dominance in the partition by states entails that

$$lim_{n\rightarrow\infty}\, \mathbf{EU}(X_n) \leq \mathbf{EU}(X).$$

Thus, if we start with the class of bounded variables and extend to included X, Y and Z, there is no sure-loss that results from the values EU(X) = 10, EU(Y) = 4, and EU(Z) = 8; when, X has *boost* 8, Y has *boost* 2, and Z has *boost* 6.

Principal Lesson from this Heuristic Example 2

Unless preferences are continuous from below,

which entails that probability is countably additive, then

Expected Utility for unbounded variables

***is not a function (solely) of the probability distribution of outcomes*!**

…. And there is the parallel problem of what to do with the

(non-Archimedean) coherent preference for St. Petersburg variables.

Part 3 *A progress report on (finitely additive) Expected Utilities for unbounded random variables*

From a mathematical perspective, an expectation of a real-valued function f is an integral, taken with respect to some set function or measure, μ.

$$\mathbf{E}_{\mu}(f) = \int_{\Omega} f(\omega)\, \mu(d\omega)$$

The integral may even be allowed to be merely finitely additive:

$$\mathbf{E}_{\mu}(f + g) = \mathbf{E}_{\mu}(f) + \mathbf{E}_{\mu}(g)$$

<u>without</u> requiring continuity, *aka* countable additivity for events.

That is, though a sequence of functions $< f_i >$ is suitably convergent to a function g, $< f_i > \Rightarrow g$, nonetheless it may be that

$$\mathbf{E}_{\mu}(g) \neq lim_i\, \mathbf{E}_{\mu}(f_i).$$

- **We see that this approach to a theory of expectations, based on a set function or measure μ, cannot serve our purposes.**

We face a challenging situation where equivalent variables may be required to have unequal expectations.

In order to accommodate this aspect of a coherent, finitely additive preference we require an integral that is *not* based on a measure.

- **Such an expectation-concept is available based on the central idea in the *Daniell* integral. (See Royden, 1968.)**

Begin with integrals given on a class *C* of *elementary functions* (including constants), which are closed under linear operations, so the integral is a (positive) linear operator. Then, the integral can be extended to a larger class of functions $D \supset C$ by using the functions from *C* to bound the values of the integral on functions in *D*, without requiring continuity and without basing integrals on a measure.

Two results obtained using this concept of expectation.

1) **When the integrals for a class of *elementary functions C* are *coherent* in de Finetti's sense, i.e., if these integrals respect uniform dominance in the privileged partition by states, then the values allowed for extending the integral to a larger class *D* match exactly the range of coherent extensions under de Finetti's *Fundamental Theorem of Previsions*.**

***Example*: Let *C* be the class of bounded random variables on a state-space Ω. Let their integrals be their finitely additive expectations under some probability *P*. Let class *D* include the unbounded Geometric variables of Heuristic Example 2. Then their coherent "boosted" expectations are permitted values of the extended integrals.**

2) **What is distinctive about merely finitely additive probability is revealed more clearly in its theory of conditional expectations than in its theory of unconditional expectations.**

***(de Finetti) Conglomerability in a partition*:**

Probability P is *conglomerable in a partition* $\pi = \{h_1, h_2, \ldots\}$ provided that, for each event E in the algebra, the unconditional probability $P(E)$ lies inside the closed interval of conditional probabilities $\{P(E \mid h)\}$.

$$inf_{h\in\pi}\, P(E \mid h) \leq P(E) \leq sup_{h\in\pi}\, P(E \mid h)$$

Example A non-conglomerable f.a. probability (Dubins, 1975).

Let <Ω, ℬ, P> be a finitely additive measure space such that:

- **Countable partition $\Omega = \pi_E \times \pi_N$. $\pi_E = \{E_C, E_F\}$. $\pi_N = \{1, 2, \ldots\}$.**
- **Algebra ℬ is the powerset of Ω.**
- **Unconditional probability P satisfies: $P(E_C) = P(E_F) = ½$.**
- **Conditional probabililty P(·|·) satisfies:**

  $P(N \mid E_C)$ is Geometric(½)

  $P(N \mid E_F)$ is *purely finitely additive* – pick a "random" integer.

	N=1	N=2	....	N=m	....
E_C	$1/2^2$	$1/2^3$	....	$1/2^{(m+1)}$	....
E_F	0	0	....	0	....

Table of unconditional probabilities for states in Dubins' example.

$P(N{=}m) = 2^{-(m+1)} > 0$. So conditional probability given N is determined by unconditional probability.

$$P(E_C) = ½ < 1 = P(E_C \mid N{=}m).$$

and P fails to be conglomerable in the partition π_N.

- ***Theorem* (*SSK*, 1984): Each finitely but not countably additive probability fails to be conglomerable in some countable partition.**

In a partition where probability for an event is not conglomerable, there probability is not an average of its conditional probabilities.

When probability is not conglomerable for an event *E* in partition π, then P is not *disintegrable* in π either:

$$\mathbf{P}(E) \neq \int_{h\in\pi} \mathbf{P}(E|h)\, d\mathbf{P}(h).$$

But Probability is merely the special case of Expectation restricted to indicator functions: P(*E*) = *EU*(*E*(ω))

So, for bounded variables, the *conglomerability* and *disintegration* apply to Expectations and Conditional Expectations.

For bounded random variables in a class $\chi = \{X\}$

an Expected Utility function is *disintegrable* over χ in partition π if

$$\forall X \in \chi \quad EU(X) = \int_{h\in\pi} EU(X|h)\, dP(h)$$

and

Expected Utility is *conglomerable* over χ in π

if $\forall X \in \chi \quad inf_{h\in\pi}\, EU(X \mid h) \leq EU(X) \leq sup_{h\in\pi}\, EU(X \mid h).$

One of Dubins' (1975) important results is that with respect to the class χ of all bounded variables, these are equivalent!

A finitely additive expectation is conglomerable over χ in a partition π

just in case

it is disintegrable over its conditional expectations given elements of π.

Under the following finiteness conditions on unbounded variables, using the finitely additive, Daniell-styled integral described before, we extend Dubins' result that conglomerable and disintegrable expectations are coextensive. And we show somewhat more.

Finiteness conditions on (unbounded) random variables

- **The variables are real-valued – no St. Petersburg variables.**
- **The variables have finite absolute expectations:** $EU(|X|) < \infty$.
- **Each conditional expectation is finite:** $EU(X \mid h) < \infty$.
- **Expectation of conditional expectation is finite:** $EU(\,EU(X|h)\,) < \infty$.

***Note*: The set of all variables that satisfy these conditions forms a linear space.**

Let *EU*(·) be a (de Finetti) coherent expectation as represented by a finitely additive Daniell-styled integral, and π be a partition.

Let *W* be a class of variables that meet the finiteness conditions.

- **Say that *W* is of *Class-0* relative to an *EU*(·) and a partition π if *EU*(·) is not conglomerable (hence, also not disintegrable) in π over *W*.**

***Aside*: Let $W \subseteq Z$. Non-conglomerability is inherited by the larger class *Z*.**
So, if *W* is of *Class-0* and then *Z* also is of *Class 0*.

- **Say that *W* is of *Class-1* relative to an *EU*(·) and a partition π if *EU*(·) is conglomerable but *not* disintegrable in π over *W*.**

- **Say that *W* is of *Class-2* relative to an *EU*(·) and a partition π if *EU*(·) is *both* conglomerable and disintegrable in π over *W*.**

- **Dubins' (1975) result, applied all bounded random variables says, relative to an *EU*(·) and a π either it is of Class-0 or of Class-2.**
- **We show the same for classes of unbounded random variables that satisfy the finiteness conditions, and which form a linear space.**
- **Also, we give an example of a coherent f.a. expectation, a partition π, and a class of variables (but which do not form a linear space) that includes all the bounded random variables, which is of Class-1:**

  The expectations are conglomerable in π but not disintegrable under the Daniell-styled integral.

Summary of our progress

Goal 1 **We have adapted an existing theory of integrals – the *Daniell* integral – so that it matches de Finetti's coherence criterion for a class of functions (including constants) that form a linear space.**

- **This class includes the unbounded variables from the 2nd heuristic example.**

We are able to incorporate finite *boost* into our integral theory of expectations. The finitely additive *Daniell* integral is not required to be a function of the distribution of outcomes.

Goal 2 **Under the finiteness conditions on unbounded variables, we extend Dubins' result: *conglomerable-in-π* and *disintegrable-in-π* expectations are coextensive. And we show somewhat more.**

***However*, the first heuristic example, the one involving St. Petersburg variables, requires infinite (non-Archimedean) expectations. These do not satisfy our finiteness conditions.**

- **There is work to be done on an integral representation for *non-Archimedean*, finitely additive expected utility!**

References

Anscombe, F.J. and Aumann, R. (1963) A definition of subjective probability. *Annals of Math. Stat*. 34: 199-205.

de Finetti, B. (1974) *Theory of Probability* (2 volumes). New York, John Wiley.

Dubins, L. (1975) Finitely additive conditional probabilities, conglomerability, and disintegrations. *Ann. Prob*. 3: 89-99.

Royden, H. L. (1968) *Real Analysis*. London, Macmillan.

Savage, L.J. (1954) *Foundations of Statistics*. New York, John Wiley.

Scherivsh, M.J., Seidenfeld, T. and Kadane, J.B. (1984) The Extent of non-conglomerability of finitely additive probabilities. *Z. f. Wahr.* 66: 205-226.

Schervish, M.J., Seidenfeld, T. and Kadane, J.B. (2009) Disintegrability …

Seidenfeld, T., Schervish, M.J., and Kadane, J.B. (2009) Preference for equivalent random variables: a price for unbounded utilities. *J. Math. Econ.* 45: 329-340.

Von Neumann, J. and Morgenster, O. (1947) Theory of Games and Economic Behavior (2nd ed.) Princeton, Princeton University Press.