Lewis on chance
2 October 2006

1. Conceptual truths about chance

Kolmogorov's axioms:

P(A) >= 0.
P(A) = 1 whenever A is a logical truth.
P(A v B) = P(A) + P(B) whenever A and B are logically incompatible.

(Updating) Chances at later times are derived from chances at earlier times by conditionalising on the complete truth about intervening history.

- What if intervening history was 'infinitely unlikely'? Two options for making sense of this: (i) primitive conditional chances; (ii) infinitesimal chances.

2. The best-system analysis

We have various *candidate systems*: functions from the possible worlds in some set and times to probability distributions, which obey (Updating).

At any given possible world, they can be rated in three ways:

(i) simplicity
(ii) strength (of the set of propositions which are assigned probability 1 relative to all times and worlds in the set)
(iii) fit (= probability assigned [at the first instant?] to the actual course of history).

The chance of a proposition at a time is the value assigned to it by whichever candidate system does the best job of jointly maximising these factors.

4. Time-independent chance?

Proposed analysis: the chance of A at t is its time-independent chance conditional on the complete truth about history up to *t*.

Given this, 'candidate systems' can simply be *probability distributions* (over worlds-propositions). The true time-independent chances are the probabilities in the best candidate system.

3. Worries about "fit" in infinite worlds (Elga)

4. How must belief about chance constrain belief about other matters?

5. Lewis's 'Bayesian' framework

Subjects believe propositions [or rather, *centred* propositions, which Lewis elsewhere identifies with properties of temporal stages] to various degrees, which can be measured on a scale from 0 to 1.

The degrees of belief of an *ideally rational* subject are a probability distribution (conform to the the Kolmogorov axioms):
(I) All logical truths are believed to degree 1.
(II) When P and Q are logically incompatible, the degree to which one believes that P or Q is equal to the sum of the degree to which one believes that P and the degree to which one believes that Q.

At any time, an ideally rational subject has certain *evidence* (a set of [centred] propositions).

There are some probability distributions called the *'reasonable initial credence functions'*, such that at any time, the degrees of belief of an ideally rational subject are derived from one of them by conditionalising on the subject's evidence.

Regularity: reasonable initial credence functions give zero probability only to impossible propositions.

6. The Principal Principle

Lewis's first version of the Principal Principle:

Let A be any proposition. Let X be the proposition that the chance at t that A is x. Let E be any proposition that is *admissible at t*. Let C be any reasonable initial credence function. Then $C(A \mid EX) = x$.

- E needs to be in there for two main reasons: (i) we want a principle which entails something about the degrees of belief of rational subjects who have evidence; (ii) we want to constrain their conditional degrees of belief too.
- Disadvantage of this formulation: lack of a theory of admissibility. Lewis gets by with sufficient conditions: (i) propositions entirely about history up to *t* are admissible at *t*, (ii) propositions about 'how history determines chance' are admissible at all times. (Not clear how exactly these are to be understood).
- Notice that being admissible at a time entails having a chance of 1 at that time.

Second version:

Let A be any proposiiton. Let C be any reasonable initial credence function. Let H_{tw} be any consistent proposition about history up to some time *t*. Let T_w be any fully informative 'theory of chance', i.e. a proposition about "how history determines chance". Then $C(A \mid H_{tw}T_w) = P_{tw}(A)$, where $P_{tw}(A)$ is the number x (if there is one) such that T_w entails that if H_{tw} is true, then the chance of A at t is x.

- This still needs us to specify what exactly the 'theories of chance' are.
- Follows from the first version, given that $H_{tw}T_w$ is admissible at *t* and entails that the chance of A is $P_{tw}(A)$.

- Weaker than the first version, even if we assume that the disjunction of Lewis's two sufficient conditions for admissibility is a necessary condition. Many EXs will be equivalent only to *infinite* disjunctions of $H_{tw}T_w$s. And *conglomerability* fails in general when we're dealing with infinite partitions: a proposition can have a certain probability conditional on each of infinitely many mutually incompatible propositions, and have a different probability conditional on the disjunction of all of them.
  - For example: consider a partition of [0,1] into three-membered sets, each of which has just one member < ½ . One could, I think, assign 'the dart landed in [0, ½] probability 1/3 conditional on 'the dart landed in three-membered set S' for each such three-membered S, while assigning it probability ½ conditional on 'the dart landed in [0,1].

Why do we need all this talk about 'theories of chance'? What would be the matter with the following formulation?

Let C be any reasonable initial credence function. Let P_t be the proposition that the chance distribution at *t* is P. Let H_t be any proposition entirely about history up to *t* that is consistent with P_t. Then $C(A \mid H_tP_t) = P(A)$.

- This no longer entails that the chances at later times are derived from the chances at earlier times by conditionalising on the intervening history (unless facts about chance at t themselves counted as facts about 'history').

Lewis's formulations are less perspicuous than they might be because he wants his principles to entail 'everything we know about chance'—or at least everything of a 'formal' character that we know about chance. If we give up this aspiration, we can happily take the claim that the chances at t2 are derived from those at t1 by conditionalising on history from t1 to t2 as axiomatic, in which case we can simplify our chance-credence principle even further, as follows:

Let C be any reasonable initial credence function. Let P_t be the proposition that the chance distribution at *t* is P. Then $C(A \mid P_t) = P(A)$.

or if we prefer to do things in term of time-independent chance:

Let C be any reasonable initial credence function. Let P_i be the proposition that the time-independent chance distribution is P. Then $C(A \mid P_i) = P(A)$.

7. Undermining

There are two distinct probability distributions P1 and P2 such that

(i) it's possible for either P1 and P2 to be the chance distribution at some time t.
(ii) There is some total history H to which both P1 and P2 both assign nonzero chance.

By (i), (ii) and PP, $C(H \mid \text{the chances at t are P1}) \neq 0$ and $C(H \mid \text{the chances at t are P2}) \neq 0$. So C(H and the chances at t are P1) ≠ 0 and C(H and the chances at t are P2) ≠ 0. But at least one of these two propositions is impossible if Humeanism is true, since H determines what the chances are. Therefore at least one of them must have prior credence equal to zero: contradiction.

8. The necessary a posteriori to the rescue?

Objection: what's impossible often does not deserve a prior credence of zero, e.g. 'water = XYZ'.

Response: even if it's not a priori in general what it is for a certain probability distribution to be the chance distribution, what we need for the argument is much less. Consider worlds that consist of a finite sequence of coin-tosses. Let P1 be the system on which the tosses are independent and each has a probability of 1/3 of coming up Heads; let P2 be the system on which they are independent and on which each has a probability of 2/3 of coming up Heads.

Surely it's a priori that *if* P1 could be the chance-distribution in a world where there are no extra non-Humean facts, and there is no chance of there being any such facts, then P2 could. What in our usage of words like 'chance' could break the tie?

But by the argument of the previous section, it's not the case that both could, and we can know this a priori. So it's a priori that neither P1 nor P2 could be the chance distribution in a Humean world.

9. Desparate remedies

Lewis considers a view on which (in our terms) it's necessary that the ur-chances are what they are.

This is a special case of a more general view: no two distribution either of which could be the ur-chance distribution gives positive probability to the same world.

Problem: this goes against our belief that simple hypotheses about the chances (e.g. Bernouilli distributions) are consistent.