6-1-2016

Imprecise Probability and Chance

Anthony F. Peressini
Marquette University, anthony.peressini@marquette.edu

Accepted version. *Erkenntnis,* Vol. 81, No. 3 (June 2016): 561-586. The final publication is available at Springer via http://dx.doi.org/10.1007/s10670-015-9755-9.

Imprecise Probability and Chance

Anthony F. Peressini
Philosophy Department, Marquette University
Milwaukee, Wisconsin

Abstract: Understanding probabilities as something other than point values (e.g., as intervals) has often been motivated by the need to find more realistic models for degree of belief, and in particular the idea that degree of belief should have an objective basis in "statistical knowledge of the world." I offer here another motivation growing out of efforts to understand how chance evolves as a function of time. If the world is "chancy" in that there are non-trivial, objective, physical probabilities at the macro-level, then the chance of an event e that happens at a given time is *<1* until it happens. But whether the chance of e goes to one continuously or not is left open. Discontinuities in such chance trajectories can have surprising and troubling consequences for probabilistic analyses of causation and accounts of how events occur in time. This, coupled with the compelling evidence for quantum discontinuities in chance's evolution, gives rise to a "(dis)continuity bind" with respect to chance probability trajectories. I argue that a viable option for circumventing the (dis)continuity bind is to understand the probabilities "imprecisely," that is, as intervals rather than point values. I then develop and motivate an alternative kind of continuity appropriate for interval-valued chance probability trajectories.

1 Introduction

Understanding probabilities as something other than points, perhaps as intervals or more general sets, has often been motivated by the need to find more realistic models for degree of belief, and in

[*Erkenntnis*, Vol 81, No. 3 (June 2016): pg. 561-586. DOI.

particular the idea that degree of belief should have an objective basis in "statistical knowledge of the world" (Kyburg 1999, 2). I offer here another motivation growing out of efforts to understand how chance evolves as a function of time. If the world is "chancy" in that there are non-trivial, objective, physical probabilities at the macro-level, then the chance of an event e that happens at a given time is <1 until it happens. But whether the chance of e as a function of time, $P_e(t)$, continuously approaches 1 or not is left open. Discontinuities in $P_e(t)$ have surprising and troubling consequences for probabilistic analyses of causation and our understanding of how events occur in time. This, coupled with the compelling evidence for quantum discontinuities in chance's evolution, gives rise to a "(dis)continuity bind" with respect to chance probability trajectories. I argue that a viable option for circumventing the (dis)continuity bind is to understand the probabilities "imprecisely," that is, as intervals rather than point values.

Imprecise (non-point-valued) probability has been studied for some time in applied and subjective probability settings, e.g., Walley (1991), Kyburg (1999) and Weichselberger (2000), and are of renewed interest of late; see Augustin et al. (2014). And most recently imprecise probabilities have been extended to objective understandings of chance by Glynn (2014). While within the setting of point probabilities the pull toward and away from continuity does indeed constitute a bind, this is not so in imprecise probability settings.

The advantage of interval-valued probability is that the notion of a continuous function opens up when the function in question is not a point-valued function. It turns out that there are multiple ways to generalize the standard (point-valued) definition of continuous. Thus one can find kinds of continuity that stabilize causally salient inequality claims between probability trajectories without being so restrictive as to decide substantive philosophical questions by definition. In particular, such kinds of continuity retain the possibility of "jumps" in chance to capture quantum or other theoretically motivated "discontinuity."

The plan of the paper will be to begin by introducing chance probability trajectories and the (dis)continuity bind they give rise to (Sects. 2, 3). I then present (Sect. 4) interval-valued trajectories as an alternative to point-valued trajectories. Finally, I develop and motivate an alternative kind of continuity (Sect. 5), appropriate for interval-valued chance trajectories, that I argue alleviates the bind (Sect. 6).

2 Chance as a Function of Time: Continuous or Discontinuous?

To introduce the continuity question I will present it in a vague causal setting, which I will firm up and connect to specific accounts below in Sect. 3. Let x and y denote token events, where *x* takes place at time and place (t_x, s_x) and y takes place at (t_y, s_y). Suppose further in some plausible way that *x's* being *Xcausedy's* being *Y*, where *x* is of type *X* and *y* is of type *Y*. The focus for now will be on how the chance of token event *y's* being *Y* evolves between t_x and t_y, that is, how the chance of *y's* being *Y* changes as a function of time. (From here on I will abbreviate the token events of "*x being X*" and "*y being Y*" by just writing the properties exemplified, *X* and *Y*, respectively.)

It is a starting assumption for what follows that chance be understood as a single-case time-dependent probability akin to what are sometimes called "physical probabilities." I begin with the standard assumption for such discussions that the chance probability function P is part of a probability space triple $<\Omega, \mathcal{F}, P>$where Ω is a set, $\mathcal{F}$ is a σ-field over Ω, and P is a probability function on F$\mathcal{F}$ that obeys the standard (Kolmogorov) axioms of the probability calculus. These physical probabilities (chances) apply to particular events, ones that occur or fail to occur at a particular time and place, and hence have values defined relative to a time of evaluation. To make explicit the temporal index, t, involved in evaluating the chance of event $Y \in \mathcal{F}$ at time t, I use the notation $P_Y(t)$. Thus in general if an event *Y* occurs at a time t_y, $P_Y(t)$ is strictly between 0 and 1 prior to t_y, and 1 at time t_y and all later times. I will use the terms "chance trajectory" and "probability trajectory" interchangeably to refer to chance analyzed in terms probability in this way.1

[*Erkenntnis*, Vol 81, No. 3 (June 2016): pg. 561-586. DOI. This article is © Springer and permission has been granted for this version to appear in e-Publications@Marquette. Springer does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Springer.]

I am roughly following Ismael (2011, 419–420) with my pre-theoretic understanding of chance, taking it to be "...the link between the fundamental level of physical description in quantum mechanics and the measurement results that mark the points of empirical contact between theory and world." I follow her in that my understanding is that chance is objective and non-trivial (not everywhere zero or one), though I remain agnostic with respect to her ultimate analysis of it and especially whether its grounding is at the quantum level or some higher level as in Glynn (2010, 2014) or Sober (2010).

To illustrate a chance probability trajectory further, consider the following commonly discussed example originally from Rosen (1978) and modified here from Eells (1991):

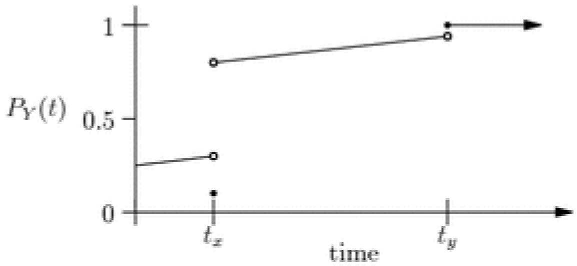
Example 1

A poorly putted golf ball is rolling roughly in the direction of the cup when a squirrel runs by and bumps it in such a way that its resulting trajectory is directly toward the cup and it continues right into the cup.

Again, I am assuming that the probability values of $P_Y(t)$ in this example reflect the objective chance of the event (ball going in the hole) and that its chance is strictly <1 until it happens. To fill out the relevant probabilities in this example, suppose that the probability of the ball going into the cup given its initial trajectory, velocity, etc. is 0.4. Suppose further that, in general, the (type) probability of balls going in when squirrels bump them is very low (say 0.05), however, in this (token) case the particular trajectory of the ball immediately following the bump made the probability of the ball falling in the cup likely, say 0.8. Denoting the event of the squirrel bumping the ball as *x* being *X* and the event of the ball going into the cup as *y* being *Y*, we can depict the probability trajectory of Y as in the graph in Fig. 1.

Fig. 1

[*Erkenntnis*, Vol 81, No. 3 (June 2016): pg. 561-586. DOI.



Chance trajectory with discontinuous jump at occurring event

The usual (intuitive) causal verdict in this example is that the squirrel's kick *X* caused the ball to drop into the cup *Y*, even though in general squirrel kick's in such situations almost never result in the ball going in the hole. The salient features of the graph for causal considerations are that the chance of *Y* takes an immediate point drop in probability at t_x, corresponding to the type-level fact that *X*-type events generally decrease the chance of Y-type events, and that the chance of *Y* recovers immediately after the ball is bumped at t_x to a higher value than it had before because of the favorable trajectory/velocity actually imparted by the token event *X*. While hopefully this causal story seems plausible enough, its details are not of particular concern here. For present purposes, the important feature of the graph is the discontinuity at t_y, that is, the fact that the chance of *Y* "jumps" to 1 at moment the ball falls into the cup.2

This "jump" is perhaps a natural way to incorporate the assumption that the world is chancy or indeterministic at the macro-level. This "occurring event discontinuity" assumption is made and discussed *explicitly* in, for example, Eells' (1991, 294) account of singular causation, but the question arises in any setting in which probability trajectories and chance are involved. I refer to this assumption as

> DJP (Discontinuous Jump Principle) The chance trajectory of an event *e* that occurs at t_e jumps discontinuously to 1 at time t_e.

It is important to note that in order for there to be a discontinuous jump (jump discontinuity) as *Y* occurs (or at any other significant time, e.g., t_x), it is necessary that the trajectory be

[*Erkenntnis*, Vol 81, No. 3 (June 2016): pg. 561-586. DOI.

continuous in some (perhaps small) interval to the left of the jump discontinuity—this will become significant below.

Notice that the assumption that the probability of an event is not one until the event occurs is also consistent with the graph continuously approaching one. While it may be more natural to require the graph to "jump" to one, this "jump" is not entailed by a chance or indeterminism assumption.4 Consider the alternate graph of Example 1 depicted in Fig. 2, in which the chance trajectory continuously approaches one at t_y. It is equally true in this graph that the probability of *Y* is strictly *<1* until it actually occurs at t_y. The difference between this graph and the graph in Fig. 1 is that in Fig. 1 the value of $P_Y(t)$ is bound away from one prior to t_y, while in Fig. 2 the value of $P_Y(t)$ becomes arbitrarily close to, but always *<1* as *t* approaches t_y.

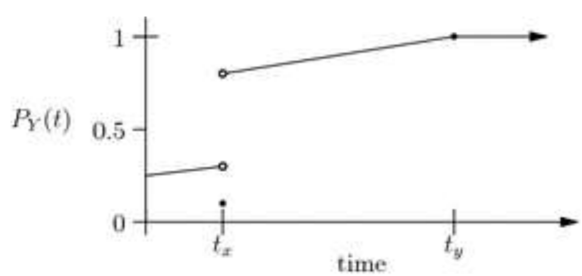


Fig. 2
Chance trajectory with continuous chance at occurring event

This question of the continuity of the chance trajectory of an occurring event is but one of a host of issues that come up concerning the continuity of chance trajectories. In the next section I argue that these general continuity issues have no straightforward resolution and in fact present something of a bind.

3 The Continuity Bind

In this section I argue that there are compelling reasons both for and against the possibility of discontinuities in chance trajectories, and that consequently we are faced with a "continuity bind." The bind arises from the tension between the following three considerations: (1) systematic discontinuities in chance trajectories like those required by DJP are problematic, but (2) assuming continuous chance trajectories runs afoul the compelling evidence for discontinuous chance from quantum phenomena, and yet (3) many probabilistic accounts of causality depend in one way or another on continuity assumptions.

3.1 Discontinuity Problems

Return now to Example 1. Consider the period from after the time the squirrel bumps the ball to the time it enters the cup. The instant the ball comes off the bump it has a certain trajectory and speed, one that will take it directly into the cup, and this helps make the chance of the ball going in as high as it is after that instant. As time gets closer to t_y and the ball gets closer to the cup, the number of eventualities that could prevent the fall into the cup decreases, and hence its chance continues to increase. That is, as the ball passes by points on the green closer and closer to the cup with the same favorable trajectory and speed, the chance of its going in the cup would naturally be expected to continue to get closer and closer to one. These considerations alone would seem to favor a continuous increasing of $P_Y(t)$ toward one, but there are more compelling reasons for rejecting the discontinuous version.5

Suppose that DJP is correct, namely, that the chance trajectories for events that occur do jump at the instant they occur. It follows then that the chance trajectory for any of the occurring events leading up to the event under consideration would also have a jump discontinuity at the time at which they occur. It is clear that the chance (trajectory) of the original event is not independent of the chance (trajectories) of certain of the events leading up to it, i.e., its chance depends on those events that need to "fall into place" in order for it to happen. And this leads to problems. Consider again Example 1: between the time of the cause-event t_x and the time of the effect-event t_y, both version's graph (Figs. 1, 2) depict the chance trajectory as continuous in the interval just to the left of t_y. But this does not accord with the "jumpy" nature of the probabilistically (causally) relevant prior events falling into place.6 If all the events involved in

[*Erkenntnis*, Vol 81, No. 3 (June 2016): pg. 561-586. DOI. This article is © Springer and permission has been granted for this version to appear in e-Publications@Marquette. Springer does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Springer.]

the ball traversing the points on the green after being bumped and before entering the cup have chance trajectories that have a jump discontinuity at the time they occur, then it seems that the chance trajectory of *Y* (ball falling into the cup), which depends upon such events falling into place, should reflect this discontinuous "jumping" at the times these prior events occur in the interval before t_y.

This reasoning suggests that the discontinuous version of how chance trajectories increase to one is inconsistent with the chance trajectory $P_Y(t)$ being continuous in the interval just before t_y, as it must be in order to have a jump discontinuity. If this is right, then assuming something like DJP in such settings is inconsistent, since $P_Y(t)$ is required (as depicted) to be continuous in at least some small interval to the left of t_y. The details of the formal argument and discussion can be found elsewhere (Peressini forthcoming), so I will just sketch it here. It begins by constructing a series of probabilistically relevant events "converging" to the time t_y, the moment *Y* occurs. Consider a sequence of moments $\{t_i\}$converging to t_y and a sequence of events $\{X_i\}$occurring at these times and upon which *Y's* chance (probabilistically) depends: in the setting of Example 1, these events and moments might be where the ball was (with its same favorable trajectory) at half of a second before it went in, and a fourth of a second before, an eighth of a second before, ... More formally we might put this as $t_1 = t_y - \frac{1}{2^i}$ and X_i= the event of the ball being where it was at t_i with the particular favorable trajectory it had.7 One then makes use of Bayes' Theorem to formalize how Y's chance trajectory is dependent on the chance trajectories of the X_i. The argument takes the form of an inconsistent/incoherent dilemma, namely that DJP in this setting entails either that

1. the chance trajectory, $P_Y(t)$ is discontinuous from the left at t_y (has no left hand limit), which is inconsistent with there being a jump discontinuity at t_y, or that

2. the certainty (distance chance is from 1) of antecedent events upon which *Y* depends becomes arbitrarily larger than the certainty of *Y* itself, which will be shown to be an *incoherent* result.

The relevant detail for purposes here is that employing a discontinuity principle like DJP has the unexpected consequence that chance trajectories are radically discontinuous or otherwise incoherent. A related result is that employing something parallel to the DJP in analyzing causation, i.e., utilizing discontinuous jumps as in the graphs of Example 1 whenever causally relevant events fall into place is similarly problematic. I will call this causal version of DJP:

CDJP (Causal Discontinuous Jump Principle) The chance trajectory of an event *e* that occurs at t_e jumps discontinuously at times when events causally relevant to e occur.

It should be clear that CDJP runs the risk of the same problems as DJP. If a similar construction of a sequence of temporally converging probabilistically relevant antecedent events can be found, then CDJP will face the same inconsistency/incoherency.8

In summary, while it seems natural in some causal contexts to understand the chance of an event as discontinuously "jumping" as its cause(s) occur, such discontinuities lead to problems: when jump discontinuities are required in general as occurring events occur (DJP) or as causally relevant events occur (CDJP), then chance functions will suffer from the essential or radical discontinuity problems (or be incoherent). But it would be rash to rule out such discontinuity in general, since quantum theory so notoriously invokes just such irreducible chance.

3.2 Irreducible Quantum Chance

As mentioned above, an obvious way out of the discontinuity problem is to understand chance trajectories as everywhere continuous, but this runs afoul the very real possibility of irreducible quantum "chance," both at the quantum level and "percolating up" to our macro-level.

Quantum events, say the decay of a U-238 atom, may well have a non-trivial chance of occurring that does not change through time. If so, then at the instant the event occurs, the chance trajectory will

[*Erkenntnis*, Vol 81, No. 3 (June 2016): pg. 561-586. DOI. This article is © Springer and permission has been granted for this version to appear in e-Publications@Marquette. Springer does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Springer.]

jump from its constant value to one. See Fig. 3. Such quantum events seem to be a singular kind of event that does not depend probabilistically (or causally) on any other factors, and hence has a chance trajectory that does not "evolve through time" until it jumps to 1. And while it may be initially tempting to conclude from this that such discontinuous behavior is isolated to the quantum-level, this is not plausible. Various examples have been developed to show that this jumpiness can be made to "percolate up," even if it does not do so on a regular basis. A simple one involves nothing more than a Geiger counter that emits a clicking sound (macro-level event) when a micro-level decay event is detected.

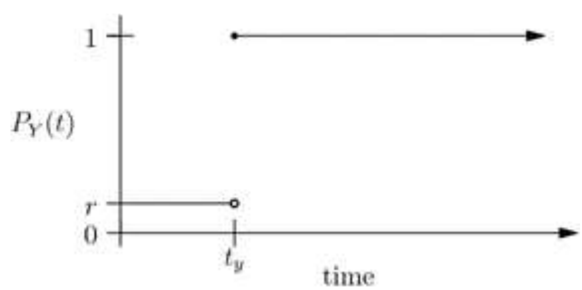


Fig. 3
The discontinuous chance trajectory of a quantum-level event

A possible response might be to maintain that while at the quantum-level such quantum events have discontinuous chance trajectories, macro-level events that involve them "dampen out" the discontinuity. Such a view might insist that events at the macro-level always have duration; they consist of intervals of time (and space). If so, then the discontinuity is avoided at the macro-level because the detection event and/or the ensuing clicking event have temporal duration during which the chance of the click (detection event) can increase sharply but continuously to 1. But at best, understanding macro- and quantum-level chance as distinct kinds such that temporal duration at the macro-level "smooths out" the quantum discontinuities, saves only macro-level continuity—and at the cost of assuming a bifurcated view of chance that involves a significant assumption about how empirical theory will ultimately go. Countenancing the possibility of discontinuous chance trajectories at all levels seems unavoidable.

3.3 Implicit Continuity Needs

How should one understand the continuity of the chance of an event as the event occurs? The upshot of Sect. 3.1 is that employing a discontinuity principle like DJP or CDJP has the unexpected consequence that chance trajectories are radically discontinuous (or otherwise incoherent). So accounts of causation like those of Eells (1991) that depend on such discontinuities are immediately problematic.

But other probabilistic analyses of causation have an opposite problem: they depend (at least implicitly) on continuity in the chance trajectories. Accounts most obviously affected by continuity are those like Menzies (1989), who utilizes "temporally dense" chains of (counter-factual) probability increases, and Kvart (2004), whose account looks for "stable screeners" and "causal relevance neutralizers" in temporally intermediate events between cause and effect. If chance cannot be assumed to be continuous, however, this undercuts such accounts by rendering the probability in the interval potentially unstable in that it may "jump" between values that may or may not preserve the presence of the relevant "probability increases" or the absence of "stable screeners" (probability decreasers).9

Even in probabilistic accounts of causation that lack explicit reference to the evolution of chance through time, there are potential complications. For example, in Noordhof (1999), Hitchcock (2004), Northcott (2010) and Glynn (2011) one finds reference to probabilities (and probability inequalities) assessed "shortly before" the time of the cause and/or effect.10 These accounts in one way or another compare the probability of an event e before it occurs at time t_e conditional on the presence and absence of a putative cause *c* at time t_c. Depending on the details of the account, one evaluates inequalities involving conditional probabilities at moments "just before" the time of the cause $t_c{-}\epsilon$ or "shortly before" the time of effect $t_e{-}\epsilon$, and perhaps at times in between. But such inequalities are stable, that is, one can safely ignore the $\epsilon>0$ magnitude expressed by "just before," thereby assuming that if ϵ is sufficiently small the inequality will hold for all smaller, only if the probability functions are continuous to the left of t_c

[Erkenntnis, Vol 81, No. 3 (June 2016): pg. 561-586. DOI.

or t_e. In general, when an inequality of the form $P_{t-\epsilon}(e|c)>P_{t-\epsilon}(e|{\sim}c)$ is employed as it is in these accounts, its stability is dependent in this sense on a continuity assumption.

Glynn's (2011) careful account helps reveal that even when utilizing variables instead of events for the relevant probability assessments, there typically remains a dependence on time, and hence continuity. In my original copy of Glynn's "A Probabilistic Analysis of Causation," he makes use of a "just before" ϵϵ-inequality, but the version published as Glynn (2011) has removed such explicit reference to time, opting to express causal conditions in terms of conditional probabilities of variables attaining a value.11 Nonetheless, a temporal index plays a role in the definition of Glynn's (2011) Revealer of Positive Evidence Set, which is to "include only variables representing events occurring no later than t_E" (p. 358, my italics). And in his discussion of the "Hiker Ducking Boulder" example, Glynn proceeds by "interpolating a variable" along the route of the boulder by which time it is too late for the Hiker to duck (p. 382). That one can (and must at times) interpolate such a variable defined in terms of time reveals that the temporal index and its attendant continuity issues are still present in such accounts despite the use of "variables attaining values" instead of events with temporal indices.

In summary, if chance trajectories cannot be assumed to be continuous, accounts of causation may be undercut by rendering the chance around the times of interest potentially "unstable" in the sense of jumping between values that may or may not preserve the relevant features of the probabilistic analysis—typically probability increases in the presence of the putative cause expressed in the form of an inequality.

3.4 Neither Continuous Nor Discontinuous

I hope at this point to have made a case for a "continuity bind" in how one understands probability trajectories in the context of chance. The above considerations push both toward and away from discontinuities in chance trajectories; there are pressing needs for both continuity and discontinuity. In particular:

- requiring systematic discontinuities like (DJP) or (CDJP) is problematic,
- requiring continuous probability trajectories is too restrictive, and yet
- probabilistic causal analyses require some continuity assumptions.

I now consider how an imprecise account of chance trajectories may be able to help with this bind.

4 Interval-Valued Imprecise Probability Trajectories

In section I present the theory of interval-valued functions and discuss their ordering and continuity properties, which will prove important in the next section. The idea behind an "imprecise probability" is that probability ought to be measured by something other than a point value, and the natural choice is a set of values from $\mathbb{R}[0,1]$. Often the sets of values are assumed to be closed intervals, as opposed to more general sets of numbers, because intervals naturally capture uncertainty with respect to a precise value that is often assumed to be lurking behind our ignorance.12 Additionally, intervals have a structure that makes them simpler to deal with than generalized sets. I too will focus on closed intervals here, though most of what I will say has a straightforward extension to generalized sets.

4.1 Interval Analysis Basics

My concern here is not so much with issues internal to imprecise probabilities themselves, but rather with how an imprecise (interval) framework might provide a needed alternative to point probabilities with respect to the temporal evolution of chance. Thus, I will simply assume that there is in the background an account of imprecise or, better yet, "interval probability," perhaps along the lines of Weichselberger (2000). My focus will be on how to think about the interval probabilities as they evolve through time. To this end, I will deal primarily with the interval probability trajectory of a given event, A as a function, P:$\mathbb{R}\rightarrow\mathbb{IR}$([0,1]) which maps $\mathbb{R}$ (time) to the set of all closed subintervals of the unit interval, defined by: $\mathbb{IR}([0,1]) = \{[\bar{a}, \underline{a}] \mid 0 \leq \bar{a} \leq \underline{a} \leq 1\}$. This allows the results of interval analysis to be brought to bear, especially Ramon Moore's pioneering work (Moore 1966, 1979; Moore et al. 2009).

4.2 Ordering Intervals

Deciding on an order relationship in $\mathbb{IR}$([0,1]) is a problem of considerable interest, especially in applied settings such as linear programming and optimization, approximation theory, and artificial intelligence. The complications stem from the fact that the standard order relationship on $\mathbb{R}$, "<", has multiple "natural" extensions to $\mathbb{IR}$([0,1]). For example, Moore (1966) introduces the following ordering which preserves transitivity:

$$[a, \bar{a}] < T[\underline{b}, \bar{b}] \ \ if\ and\ only\ if\ \ \bar{a} < \underline{b}.$$

In addition to transitivity, this order retains the property from < in $\mathbb{R}$ that if *A*<*B* then there exists a *C* such that *A*<*C*<*B*. See Fig. 4. Another common ordering is an "end point" order:

$$[a, \overline{a}] < E\left[\underline{b}, \overline{b}\right] if\ and\ only\ if\ \ a\ <\ \underline{b}\ and\ \bar{a}\ < \bar{b}$$

A virtue of this order is that it is weaker than <*T* and still entails that each of the end points satisfy the < relationship in $\mathbb{R}$, and it preserves the property from $\mathbb{R}$ that $A<A+\epsilon$ for any $\epsilon>0$.

These interval orderings differ from their real counterparts in that they are only partial orderings as opposed to a total or linear orderings, that is, ordering such that for all *A, B* ∈ $\mathbb{IR}$ ([0,1]). either *A*<*B* or *B*<*A*. For example, nested intervals such as

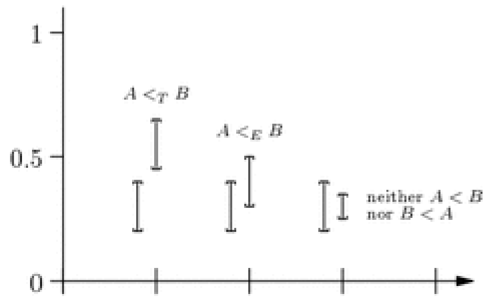


Fig. 4
Ordering intervals with $<T$ and $<E$ partial order relationships Fig. 4

$A= [.2,.4]$ and $B=[.25,.35]$ are such that neither $A<B$ nor $B<A$. (See Fig. 4.) Within applied work in interval analysis, studies of the properties of different orderings is a lively area (Li and Li 2010; Guerra and Stefanini 2011).13 For purposes here, the transitive "$<T$" generally will be utilized, though again, how to analyze "less than" in the context of interval valued chance should be considered a "site of contention."

4.3 Convergence and Continuity

Convergence and continuity are among the most central concepts in analysis—and both of these notions depend on the ability measure distance. Recalling the standard definition in $\mathbb{R}$, for convergence of a sequence $\{a_i\}$to a: for every $\epsilon>0$ there is a natural number $N=N(\epsilon)$ such that

$$|a_i-a|<\epsilon$$

for all $I \geq N(\epsilon)$. And a function $f(x)$ is continuous at a point x_0 if for every $\epsilon>0$ there is a positive number $\delta=\delta(\epsilon)>0$ such that

$$|f(x)-f(x_0)|<\epsilon$$

whenever $|x-x_0|<\delta$. The distance function in the setting of real analysis is of course the familiar $D_{\mathbb{R}}(x, y) =|x-y|$, which is a special case of the general notion of a *Hausdorff metric*.

It is well known that the space of bounded closed intervals $\mathbb{IR}$ with Hausdorff metric H defined by $d(A,B)=\max\{|\underline{b}-\underline{a}|,|\overline{b}-\overline{a}|\}$ where $A=[\underline{a}, \overline{a}]$ and $B=[\overline{b}, \underline{b}]$ is a complete metric space (Aubin and Frankowska 1990). Thus the closed and bounded subset $\mathbb{IR}([0,1])$ of

[Erkenntnis, Vol 81, No. 3 (June 2016): pg. 561-586. DOI. This article is © Springer and permission has been granted for this version to appear in e-Publications@Marquette. Springer does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Springer.]

IRIR is also a complete metric space under the inherited metric HIRHIR.

In $\mathbb{IR}([0,1])$ convergence may be defined similarly with $D_{\mathbb{R}}$ replaced by $D_{\mathbb{IR}}$: if $A = [\underline{a}, \overline{a}]$ and $\{A_i\} = \{[\underline{a}_i, \overline{a}_i]\}$ then {Ai}→A if for every *ϵ>0* there is a natural number *N=N(ϵ)* such that

$$D_{\mathbb{IR}}(A_i - A) < \epsilon$$

for all *i≥N(ϵ)* A straightforward consequence of this definition (Moore et al. 2009, Sec. 6.1) is that

$$\{A_i\} \rightarrow A \text{ if and only if } \{\text{ai})\rightarrow \text{a and} \{\text{ai}\} \rightarrow \text{a}\{\underline{a}_i\} \rightarrow \underline{a} \text{ and } \{\overline{a}_i\} \rightarrow \overline{a}.$$

In a like way for continuity, a function $F: \mathbb{R} \rightarrow \mathbb{IR}([0,1])$is continuous at x_0 if for every ϵ>0 there is a positive number δ=δ(ϵ)> such that

$$D_{\mathbb{IR}}(F(x) - F(x_0)) < \epsilon,$$

whenever x−x_0|<δ. Again it is straightforward consequence that if $F(x) = [\underline{F}(x), \overline{F}(x)]$ then at $x \in \mathbb{R}$

$F(x)$is continuous if and only if $\underline{F}(x)$ and $\overline{F}(x)$ are continuous.

As is the case with point functions in $\mathbb{R}$, *F's* continuity at x_0 is equivalent to $\lim_{x \rightarrow x_0} F(x) = F(x_0)$ see e.g., Flores-Franulic et al. (2013, 1460).

This inherited "ordinary" continuity for interval-valued functions is so tightly tied to the continuity of the real-valued endpoint functions that it will not be of use here, since it consequently inherits directly the ordinary version's incompatibility with any "jumpiness." Thus no additional "maneuvering room" is gained to help work free of the continuity bind. Fortunately, other conceptions of continuity are possible. Below I consider weaker notions of continuity for interval functions that preserve crucial aspects of the notion without entailing ordinary continuity. This will free analyses based on interval probability trajectories from the dichotomous continuity of precise probability settings.

5 Specialized Continuity for IP Trajectories

[Erkenntnis, Vol 81, No. 3 (June 2016): pg. 561-586. DOI.

Given that different notions of continuity for interval functions are possible, how should one arrive at one appropriate for IP trajectories understood as functions of the form $F: \mathbb{R} \rightarrow \mathbb{IR}([0,1])$? Recalling the continuity bind for point-valued trajectories, any such requirement ought to:

- stabilize inequality claims between causally salient probability trajectories,
- retain the possibility of "jumpiness" to capture quantum or other theoretically motivated "discontinuity,"
- not be so restrictive as to decide substantive philosophical or empirical questions by definition.

That there is no requirement satisfying these three desiderata in the context of point-valued functions is what gives rise to the continuity bind, and as we have seen, the inherited ordinary version of continuity in the context of interval-valued functions will not work either. Thus, a weaker specialized IP continuity must be found that allows for the endpoint functions, $\underline{F}$ and $\overline{F}$ to be discontinuous in the ordinary sense.

At this point there are two general strategies for broadening the ordinary notion of continuity for IP functions:

1. Focus first on more general set-valued functions (of which interval-valued functions are a special case) and then apply such general insights to develop a distinct kind of continuity for interval-valued functions, or

2. Focus first on the general functional space of interval-valued functions, and then explore the properties of particular subspaces of interval-valued functions generated by weaker notions of continuity.

In what follows, I primarily employ the first strategy, though I briefly discuss the second in Sect. 5.3.

5.1 Set-Valued Analysis and Continuity

Once one generalizes functions to entities more complicated than point-values and their Cartesian products, the ways to theorize continuity multiply. I next develop an informal notion of continuity appropriate for IP trajectories and then work out a formal definition of it utilizing the idea of semicontinuity from generalized set-valued analysis.

In the early thirties, Bouligand, Kuratowski, and Wilson formalized the notions of upper semicontinuous and lower semicontinuous maps on generalized sets with metrics. These two notions were required to capture the ordinary sense of continuity from real analysis, because in set-valued analysis, the ϵ-δ formulation, which requires that arbitrarily small neighborhoods in the range be mapped into by sufficiently small neighborhoods in the domain, is independent of the (equivalent in real analysis) formulation that "continuous functions map converging series to converging series" (Aubin and Frankowska 1990, 38–40). Upper semicontinuity formalizes the standard ϵ-δ definition as follows:

Upper Semicontinuity (Set-Valued Functions)

A set valued function $F : X \rightsquigarrow Y$ is *upper semicontinuous at* $x_0 \in X$ if and only if for any neighborhood $\mathcal{U}$ of $F(x_0)$,

$$\exists \eta > 0 \text{ such that } \forall x \in B_X(x_0, \eta), F(x) \subset \mathcal{U}.$$

The function F is *upper semicontinuous* if it is upper semicontinuous at every point in its domain.

Its parallel with the standard ϵ-δ definition should be clear with *U* playing the role of the ϵ-neighborhood and the η-ball about x0 playing the role of the δ-neighborhood. To capture the notion of mapping converging sequences onto converging sequences we have lower semicontinuity:

[Erkenntnis, Vol 81, No. 3 (June 2016): pg. 561-586. DOI.

Lower SemiContinuity (Set-Valued Functions)	A set-valued function $F : X \rightsquigarrow Y$ is *lower semicontinuous at* x_0 if and only if for any $y_0 \in F(x_0)$ and for any sequence $x_i \in X$ with $\{x_i\} \to x_0$, $\exists$ a sequence of elements $y_i \in F(x_i)$ with $\{y_i\} \to y_0$. The function F is *lower semicontinuous* if it is lower semicontinuous at every point in its domain.

Again, this definition requires that there be the requisite converging sequence in *Y* for any point in *F(x_0)* and any sequence in *X* converging to x_0.

It is instructive to see how a function may be upper semicontinuous (USC) and fail to be lower semicontinuous (LSC) and *vice versa*. For example, consider

$$F(x) = \begin{cases} [0,1] & \text{if} x = 1, \\ [\frac{1}{2}, \frac{1}{2}] & \text{otherwise.} \end{cases}$$

This is graphed in Fig. 5. It is USC at *x*=1 because any neighborhood *U* about *F(1)*=[0,1] will contain *F(x)* for all the *x* in any neighborhood about 1 since $F(x) = [\frac{1}{2}, \frac{1}{2}] \subset [0,1] \subset \mathcal{U}$ everywhere except at *F(1)* which is $[0,1] \subset \mathcal{U}$. And it fails to be LSC at *x=1* because there is a $y_0 \in F(1)$ say $y_0 = \frac{1}{4} \in F(1)$ such that the elements of a (any) sequence of x_i converging to 1 do not have F(x_i) converging to y_0—this is because all such $F(x_i) = \frac{1}{2} \neq \frac{1}{4} = y_0$. As an example of a LSC function that is not USC consider

$$G(x) = \begin{cases} [\frac{1}{2}, \frac{1}{2}] & \text{if} x = 1, \\ [0,1] & \text{otherwise.} \end{cases}$$

This is graphed in Fig. 6. It is LSC at *x=1* because only $\frac{1}{2}$ is in *G(1)* and every $x_i \neq 1$ has $\frac{1}{2} \in G(x_i)$, so for every sequence of x_i converging to 1, $y_i = \frac{1}{2} \in G(x_i)$ is such that y_i converges to $\frac{1}{2}$. It is not USC because the small neighborhood about $\frac{1}{2} \in G(1)$ given by (.4, .6) is such that any neighborhood about 1 has an x in it such that $G(x) = [0,1] \not\subset (.4, .6)$.

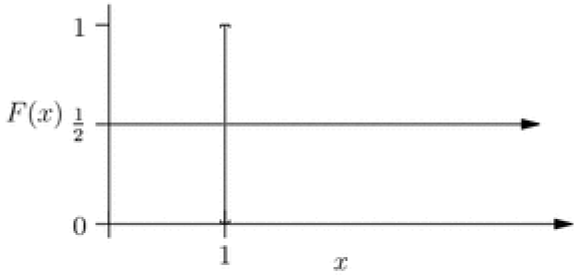


Fig. 5
Graph of upper but not lower semicontinuous function

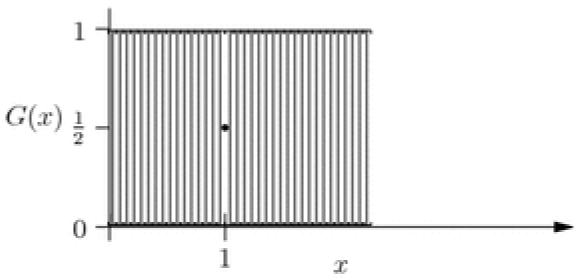


Fig. 6
Graph of lower but not upper semicontinuous function

With the notions of USC and LSC in hand, the ordinary notion of continuity for set-valued functions can then be defined as:

Continuity (Set-Valued Functions) A set-valued function $F : X \rightsquigarrow Y$ is *continuous at* $x_0 \in X$ if and only if it is both upper semicontinuous **and** lower semicontinuous at x_0. The function F is *continuous* if and only if it is continuous at every point in its domain.

Now once again this ordinary continuity for set-valued functions is (as intended by its developers) tightly tied to continuity for real-valued functions, in particular, it brings together in the set-valued context the ϵ-δ definition and the "converging sequences" definition. So, for purposes here, which involve working out weaker, less dichotomous notions of continuity, the ordinary conception is unhelpful. There is now, however, a clear way forward by which to open up such room, namely, by not requiring both lower and upper semicontinuity.

5.2 A Proposal: "Gapless" Interval Functions

With the formal tools now in place to explore weaker variations of continuity, it will be helpful to reconsider desiderata for continuity for IP trajectories. First, such continuity should preserve the possibility of "jumps" in chances because some causal and quantum phenomena may be such that the intervals representing the chance of effects "jump" discontinuously. But chance at a given moment need not (and perhaps should not) be "completely disconnected" from chance temporally near by. That is, given the interdependent nature of events, both in kind (causal and constitutive) and level (micro, meso, macro), chances are not "completely disconnected" from one moment to another. So while the chance interval of an event evolves through time, perhaps "jumping" (discontinuously), it should at no point in time "jump" in such a way that there is an actual gap between the chance intervals. Call this the:

Gapless Jump Criterion	The IP trajectory of an event *A* should be such that any discontinuous jumps in chance intervals are not such that there is a gap from "one moment to the next."

The idea is to rule out cases like that of Fig. 7 below, while still allowing discontinuous IP trajectories as in Fig. 8. A bit more precisely, the idea behind this is that while the chance intervals $[\underline{P}(t), \overline{P}(t)]$may be discontinuous, they should not be "jumping" in a way that there is a time t_x such that the chance interval at that time $[\underline{P}(t_x), \overline{P}(t_x)]$is bound away from chance intervals at times arbitrarily close to t_x

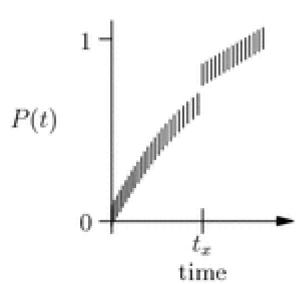


Fig. 7
Interval function with "gap" discontinuity

[Erkenntnis, Vol 81, No. 3 (June 2016): pg. 561-586. DOI.

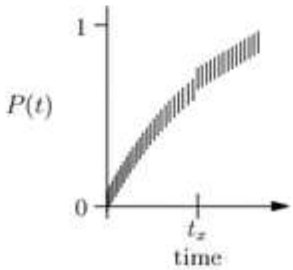


Fig. 8
Interval function with "non-gap" discontinuity

An alternative way of fleshing out the requirement that chance at a given moment should not be completely disconnected from chance temporally near by is that there is always an "unbroken" (continuous) path through the graph of the temporally evolving chance intervals. Call this the:

Path Possibility Criterion	The IP trajectory of an event *A* should be such that there is a *continuous path* the chance values can take.

In other words, P should not be "jumping" in a way that rules out the possibility that there could be at each time t in the domain of *P* a value c_t contained in the chance interval $[\underline{P}(t), \overline{P}(t)]$such that the function, $v: \mathbb{R} \rightarrow \mathbb{R}$, defined as $v(t) = c_t$is itself continuous.

As will become evident below, these are not equivalent conditions: gaplessness is independent of continuous path possibility. And as is clear from the graphs of quantum-style discontinuities (Fig. 3), which are not "path continuous" in this sense, gaplessness is the better choice in order to countenance quantum chance.14

The way to formalize this gaplessness returns to the definition of set-valued continuity in terms of upper and lower semicontinuity. As suggested by the graphs of examples of USC without LSC and LSC without USC (Figs. 5, 6), both of which are gapless in the desired sense, each of LSC and USC are sufficient for "gapless" continuity. The proposal is that G(apless)-continuity be characterized as

G-Continuity (Interval-Valued Maps)	An interval function $F : \mathbb{R} \to \mathbb{IR}[0,1]$ is *G-continuous* at $x_0 \in \mathbb{R} \leftrightarrow$ it is lower **or** upper semicontinuous at x_0. The function F is *G-continuous* if it is G-continuous at all $x \in \mathbb{R}$.

It is clear that if F is continuous in the ordinary sense, then F is G-continuous, and the functions in Figs. 5, 6 are each examples of G-continuity without ordinary continuity. As another more perspicuous example of a function that is G-continuous but not continuous, consider the following function:

$$F(x) = \begin{cases} [0.4,0.6] & \text{if} \quad x < 1, \\ [0.4,0.7] & \text{if} \quad x = 1, \\ [0.5,0.7] & \text{if} \quad x > 1. \end{cases}$$

This is graphed in Fig. 9. It is discontinuous in the ordinary sense (the end point functions are discontinuous) at x=1x=1 and it also fail to be LSC at x=1x=1. It is however USC at x=1x=1. Notice also that G-continuity is a generalization of ordinary continuity in that the point functions in $\mathbb{R}[0,1]$ that are G-continuous are precisely the continuous functions in $\mathbb{R}[0,1]$.

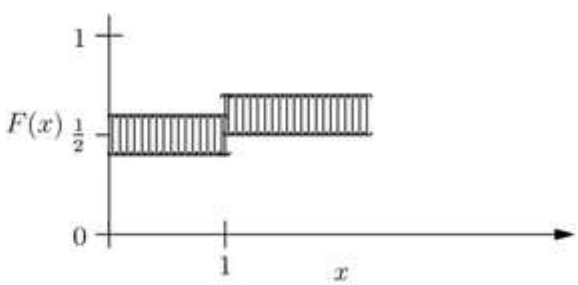


Fig. 9
G-continuous but not (ordinarily) continuous interval function

Finally, while G-continuity is sufficient for "gaplessness," it is not necessary. There are functions that are not "gappy" but are neither LSC nor USC. Consider the function:

$$F(x) = \begin{cases} [0.4,0.8] & \text{if} \quad x \in \mathbb{Q}, \\ [0.2,0.6] & \text{if} \quad x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$$

where $\mathbb{Q}$ denotes the rational numbers and $\mathbb{R} \setminus \mathbb{Q}$ the irrational numbers. This function (graphed in Fig. 10) is not gappy and there is a continuous path through it (e.g., $v(x) = \frac{1}{2}$, and yet it is neither USC nor LSC, and hence not G-continuous. While functions like these are not so pathological as to to be outside the realm of possible physical theory (e.g., they are Lebesgue integrable), there does not seem to be reason from physical theory to be concerned with them at this point. Though were that to change, there are ways of weakening semicontinuity to include such functions.

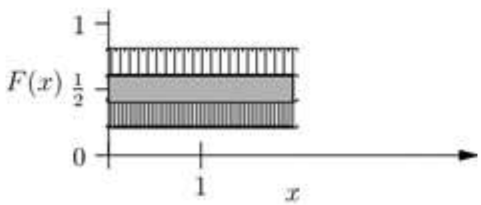


Fig. 10
"Gappless" but neither upper nor lower semicontinuous function

The approach to continuity for IP trajectories developed so far, G-continuity, began from the more general setting of set-valued functions. It is worthwhile at this point to examine briefly the second approach, one focusing on particular subspaces of the interval-valued function space.

5.3 Interval-Valued Function Spaces and Continuity

When the interval functions $\mathbb{IR}[0,1]$are treated a function space, the way is opened to characterize and explore subspaces distinct from the the subspace of ordinary continuous functions of $\mathbb{IR}[0,1]$. The subspaces of interest are ones that both properly contain contain the ordinary continuous functions of $\mathbb{IR}[0,1]$but are also still generalizations of ordinary continuity in the sense that the only point valued functions they contain are precisely the (ordinary) continuous functions in $\mathbb{IR}[0,1]$.

In this vein, Roumen Anguelov et al. (2006) develops three distinct notions of continuity: S-continuity, D-continuity, and H-continuity that apply to interval functions. The class of S-continuous functions are of particular interest for IP trajectories as they are the weakest of the three. The class of S-continuous functions are defined in terms of the lower and upper Baire operators, a corollary of which is that an interval function $F = [\underline{f}, \overline{f}]$is S-continuous if and only if $\underline{f}$is lower semicontinuous and $\overline{f}$is upper semicontinuous.16 Despite this relationship to its endpoints functions, an S-continuous function is a completely novel entity from both algebraic and topological points of view. Such functions can be quite "jumpy" (discontinuous in the ordinary sense) with the primary restriction being that the upper endpoint function can only jump up and the lower endpoint function can only jump down, and hence do not have the "gaps" of discontinuous point-valued functions.

The S-continuous functions are strictly contained in the G-continuous functions, since S-continuity entails upper SC and hence G-continuity. But the example of lower continuity without upper (Fig. 6) is not S-continuous (its upper endpoint function is not USC and its lower endpoint function is not LSC), hence the strict containment. One potential advantage of S-continuity over something like G-continuity is its connection to continuous functions and the fact that its structure is well understood and characterizable in ways that make connections to other kinds of continuity. The class of S-continuous functions contain the completed graphs of all point-wise infima and suprema of sets of continuous functions (Anguelov et al. 2006, 18). Also, S-continuous functions can be characterized as the set of interval functions whose graph is a closed subset of the Cartesian product of its domain and $\mathbb{R}$ (Anguelov and Markov 2007, 280).

In any case, the generalized continuity work based in functional analysis offers another viable route for the continuity of IP trajectories.

6 Putting IP Continuity to Work

Having introduced the above more "open" varieties of continuity for interval-valued functions, I return now to the original motivating concerns for IP trajectories, namely, continuity questions concerning occurring macro-level events, quantum discontinuities, and inequality instability for probabilistic causation.

6.1 IP Trajectories and Occurring Events

The issue of how to think of the IP trajectory *P(t)* of event A as it occurs at time t_A raises several distinct questions. The first question is whether $P(t_A)$is [1, 1] or whether it is [a, 1] for some 0≤a<1, that is, whether the value is a proper interval with 1 as its upper endpoint or whether it is a point interval. A second question is how precisely the endpoint functions $\underline{P}(t)$and $\overline{P}(t)$converge to their values, since this is left open by G-continuity. These questions undoubtedly depend on broader theoretical (and likely empirical) considerations that cannot be resolved here, but it is important to note how G-continuity can accommodate various renderings of the IP trajectories of occurring events.

In the context of the Golf Ball example (Example 1), consider the following interval-valued chance trajectory of the events, given in Fig. 11. In this rendering, the upper endpoint function jumps discontinuously to 1, while the lower endpoint function converges (from the left) continuously to a value <1, thus $P(t_y) = [a, 1]$for some *0≤a<1*. Another feature of *P(t)* to note is that at t_x, when the squirrel kick occurs, the value of the chance trajectory "jumps" (discontinuously) to an interval that contains both the limit intervals from the left and right. An interpretation of this would be, again, that the instant of the kick "brings together" the higher chance that ensues immediately after the moment of the kick with the lower chance associated with the original trajectory up until the moment of the kick. This rendering and interpretation involve understanding the kick as a point event. And again, this IP trajectory is G-continuous (at t_x and t_y) because it is USC there.

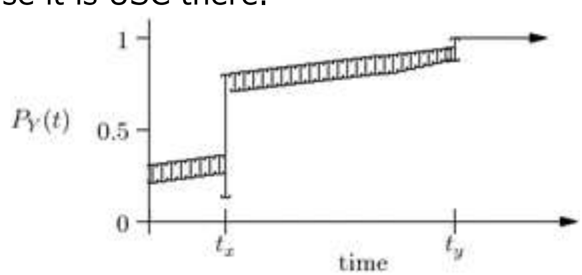


Fig. 11
Graph of Example 1 with IP Trajectory jumping to 1

Now consider an alternative rendering, one in which *P(*t_y*)* is the point interval *[1, 1]* as depicted in Fig. 12. Notice that at t_x, when the squirrel kick occurs, the chance trajectory immediately "jumps" (discontinuously) to an interval that again contains both the limit intervals from the left and right, but it remains "wide" for an interval of time before (discontinuously) decreasing in size to the relatively tight interval it is a bit later in time. In this rendering, the kick event is interpreted as a temporally extended event. Finally, this IP trajectory is discontinuous in the ordinary sense and still G-continuous because it is USC at t_x and LSC at t_y.

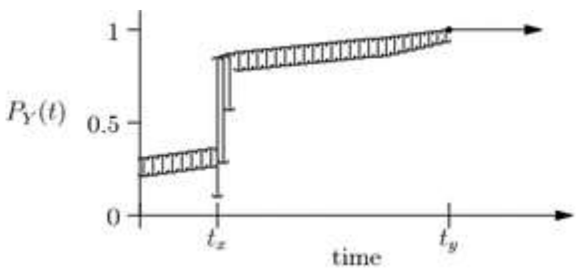


Fig. 12
Graph of Example 1 with IP Trajectory converging to 1

It should be clear as well that G-continuity can also accommodate quantum-style events, namely, ones in which the probability of *A* is *r* until it happens at t_A. There is an obvious G-continuous IP function corresponding to the situation:

$$P(t) = \begin{cases} [r,r] & \text{if} \quad t < t_A, \\ [r,1] & \text{if} \quad t = t_A, \\ [1,1] & \text{if} \quad t > t_A. \end{cases}$$

The function P is graphed in Fig. 13. Under the prevailing interpretations of quantum theory, such an event is "uncaused" or "irreducibly probabilistic." If such events are indeed qualitatively different from macro-events, as there is good reason to think, then the IP framework with G-continuity is particularly apt because it provides multiple ways to formalize the difference, e.g., quantum events trajectories take on interval values only at the moment of the event, or alternatively quantum event trajectories are G-continuous but fail to satisfy the Path Possibility Criterion (p. 17).

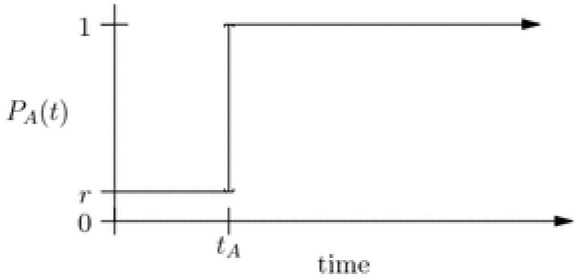


Fig. 13
G-continuous Quantum Event

6.2 Causality and Continuity

Recall the "inequality instability" issue from Sect. 3.3, namely, that many probabilistic analyses of causation compare the chance of an event e conditional on the presence and absence of a putative cause *c* at time $t{-}\epsilon$ "shortly before" the time of the cause and/or effect. I argued above that this comparison is typically assumed to be stable, that is, that one can ignore the precise $\epsilon>0$ magnitude involved in "shortly before," safely assuming that if ϵ is sufficiently small, the values of the chances will be "indicative" of the property of interest (the inequality in this case). But this kind of stability can be assumed only if the chance trajectories are continuous with respect to time to the left of t_c and/or t_e.18 For clarity, in what follows I will focus on time t_e, but the same follows for time t_c or any other point between them. The general form of the inequality of interest is:

$$P_{t_e-\epsilon}(e|c) > P_{t_e-\epsilon}(e|\sim c). \qquad (1)$$

The chances that are of interest are described as being "shortly before" the time of the putative effect e because at the precise time of e the relevant chances (conditional probabilities) are trivial. If the chance trajectories involved are not assumed to be continuous to the left of t_e, then the mere fact that the inequality holds at a given time "shortly before t_e" is insufficient to guarantee that it will hold (to the left) in any interval about t_e. And if the inequality could be "flipping" in the neighborhood $(t_e-\epsilon, t_e)$then it holding at $t_e-\epsilon$ is not going to be decisive for the causal efficacy of c, since such inequality reversing would undercut the understanding that c was decisive in the sense of ruling out the possibility of there being further factors that could act as "stable screeners" or "neutralizer" or "failure sets."

It might appear that some thinkers are indirectly working in a "continuity condition" with respect to such inequalities. Kvart (2004), for example, supplements the inequality with conditions that effectively rule out reversals of the inequality in the interval (t_c,t_e). Glynn (2011) may be seeking to do the same thing by requiring only that there be the right combinations "increasers" (supporters of the inequality) and "decreasers" (under-cutters of the inequality) in the interval (t_c,t_e). But these conditions are intended to make sure (1) is preserved (or violated) by holding fixed or allowing to vary appropriate causal background factors. This should not be confused with the issue I am urging consideration of, namely, that once all of the factors are included and the inequality is asserted/denied, it is still asserted relative to a temporal index, whether explicitly stated or not, and that the decisiveness of the inequality as the "last word" on c's causal relevance is dependent on the behavior of the chance functions not just at a point in time just prior to t_e but also in a neighborhood around it—because an interval contains an uncountable number of points, but a requirement involving a sequential specification of comparisons (Kvart) or the right combination of increasers/decreasers (Glynn) can only be effective on a countable (Kvart) or finite (Glynn) set of points. Only if such a neighborhood exists can one be sure that the probability of the effect (just before it happens) is higher in the presence of the cause and that it remain higher until it occurs. Or put more intuitively, that the probability increase that qualifiesc as cause has "the last say" in the evolution of the probability trajectory from some appropriate time before the effect until the time it occurs.

As seen above, interval-valued functions can provide an alternative to requiring ordinary (overly strong) continuity. Consider the properties of inequalities like (1) in the context of G-continuous IP trajectories. Making use of the transitive interval ordering $<T$, notice that for interval-valued G-continuous functions, F and G, the inequality

$$F(x_0) >_T G(x_0), \qquad (2)$$

has implications for the behavior of the functions in an interval around x_0. In the case of ordinary continuity, one could infer from (2) that the inequality holds in an appropriately small interval about x_0. While G-

[Erkenntnis, Vol 81, No. 3 (June 2016): pg. 561-586. DOI.

continuity does not quite support this entailment, it does yield something almost as strong, namely that

$$F(x) \nless_T G(x) \quad (3)$$

in some neighborhood $(x_0 - \epsilon, x_0 + \epsilon)$ about x_0. The proof of this is straightforward and instructive.
Claim 1w (weak):

Claim 1w (weak): Let F and G be G-continuous interval-valued functions and x_0 a point in their domain such that $F(x_0) >_T G(x_0)$, then there exists a neighborhood $(x_0 - \epsilon, x_0 + \epsilon)$ about x_0 such that $F(x) \nless_T G(x)$ for all x in $(x_0 - \epsilon, x_0 + \epsilon)$.

Proof

The definition of G-continuous requires that the functions F and G be either USC or LSC at x_0. The four possibilities are (1) both F and G are USC, (2) both F and G are LSC, (3) F is USC and G is LSC, or (4) G is USC and F is LSC. The proof for each of these four cases is as follows:

Case 1: Both USC If both are USC, then it is an immediate consequence of the definition of USC that a neighborhood of x_0 can be found so that each of $F(x)$ and $G(x)$ are less than half the distance between $F(x_0)$ and $G(x_0)$ for all x in that neighborhood, which actually entails that $F(x) >_T G(x)$ in the given neighborhood, so $F(x) \nless_T G(x)$ there too.

[Erkenntnis, Vol 81, No. 3 (June 2016): pg. 561-586. DOI. This article is © Springer and permission has been granted for this version to appear in e-Publications@Marquette. Springer does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Springer.]

Case 2: Both LSC	Proceed indirectly by assuming otherwise: if no neighborhood about x_0 existed such that $F(x) \not<_T G(x)$, then one could construct a sequence of $\{x_i\}$ converging to x_0 such that $\underline{G}(x_i) \geq \overline{F}(x_i)$. But LSC entails that we can find sequences $\{f_i\}$ and $\{g_i\}$ with $f_i \in F(x_i)$ and $g_i \in G(x_i)$ such that $\{f_i\} \to \underline{F}(x_0)$ and $\{g_i\} \to \overline{G}(x_0)$. But this means that for sufficiently large i, f_i and g_i must be within $\epsilon = \frac{1}{2}(\underline{F}(x_0) - \overline{G}(x_0)) > 0$ of $\underline{F}(x_0)$ and $\overline{G}(x_0)$, respectively, from which it follows that $f_i > g_i$, but this entails that $\underline{G}(x_i) < \overline{F}(x_i)$, which contradicts the construction of $\{x_i\}$.
Case 3: One of each	F is USC and G is LSC. The proof proceeds as in Case 2, constructing the $\{g_i\}$ as before, but construct the $\{f_i\}$ in this case using the USC of F in order to ensure that $\lvert f_i - \underline{F}(x_0)\rvert < \frac{1}{i}$, by picking f_i from $F(x_j)$ for large enough j, which is guaranteed by USC to be within $\frac{1}{i}$ of $\underline{F}(x_0)$. Then by the same reasoning as in Case 2, the $\{f_i\}$ and $\{g_i\}$ will lead to the same contradiction.
Case 4: Other of each	G is USC and F is LSC. Again proceed as in Case 2, constructing the $\{f_i\}$ as before, but this time construct the $\{g_i\}$ using the USC of G to ensure that $\lvert g_i - \overline{G}(x_0)\rvert < \frac{1}{i}$, by picking g_i from $G(x_j)$ for large enough j, which is guaranteed by USC to be within $\frac{1}{i}$ of $\overline{G}(x_0)$. Then by the reasoning of Case 2, the $\{f_i\}$ and $\{g_i\}$ will lead to the same contradiction. □

From this it follows that in the setting of G-continuous interval valued probability functions, *P(e|c)* and $P(e|\sim c)$, if the inequality $P_{t_e-\epsilon}(e|c) >_T P_{t_e-\epsilon}(e|\sim c)$holds at some time $t_e - \epsilon$prior to and sufficiently close to t_e, then one has that it does not reverse itself in the interval from $t_e - \epsilon$ to t_e. That is, the inequality is in fact stable in that it cannot "flip" in the neighborhood $(t_e - \epsilon, t_e)$, and so can be decisive for the causal efficacy of *c*.

And further, the stability given by (3), translated back into the terms of the IP inequality (1) yields that the putative cause c is such that the chance of e is higher, given c at some appropriate time, *and that it remains at least as high (not less than) for some interval of time after*. This degree of stability is considerably more than is present in the point probability setting without a continuity requirement. What is more, as the proof of Claim 1w makes clear, it is the LSC cases (Cases 2, 3, and 4) that necessitate the weaker result. Thus in order to obtain the stronger result *(>T instead of ≮T)*, one could require USC instead of G-continuity's weaker "USC or LSC", and thereby ensure the full stability of (2) obtaining in an interval about x_0.21 This stronger version would be:

Claim 1s (strong):

Claim 1s (strong): Let F and G be USC interval-valued functions and x_0 a point in their domain such that $F(x_0) >_T G(x_0)$, then there exists a neighborhood $(x_0 - \epsilon, x_0 + \epsilon)$ about x_0 such that $F(x) >_T G(x)$ for all x in $(x_0 - \epsilon, x_0 + \epsilon)$.

Thus, in the IP setting, different levels (strengths) of stability would be available (depending on the kind of continuity employed) that could mediate between the particular needs of causal analyses and the kinds of "jumpiness" (ordinary discontinuity) required by empirical or other theoretical constraints.

7 Conclusion

I hope to have made a case for the advantages of interval-valued probability in settings where objective chance trajectories as a function of time are of interest. In such settings issues of the continuity of chance trajectories become pressing: discontinuities have surprising and troubling consequences for probabilistic analyses of causation and how events occur in time, and yet there is compelling reason to retain the possibility of discontinuities in chance's evolution.

In the imprecise setting of interval-valued probability, the notion of a continuous function opens up, and it turns out that there are multiple ways to generalize the standard point function definition of continuous. This yields kinds of continuity that can both stabilize probability inequality claims between trajectories and still retain the possibility of "jumpiness" that can capture quantum or other theoretically motivated discontinuity. And equally important, having such a repertoire of continuity alleviates the need to decide substantive empirical and/or philosophical questions by "definitions."

Footnotes

1. I will assume that the basic form of these chance probabilities is unconditional; this is in contrast to general probability, which applies to classes of event and whose basic forms is conditional. I assume this

for clarity and convenience only: the continuity issues I deal with here are not sensitive to whether the physical probabilities of chance are analyzed in the standard Kolmogorovian way or some other way, with a different conditionalization and/or with conditional probabilities as the basic form; see for example Hájek (2003).

2. As an anonymous reviewer points out, there are other equally (perhaps more) plausible ways of understanding the token squirrel kick's effect on the probability trajectory, e.g., it might be understood as "immediately" raising the probability if focusing on how it "immediately" improves the balls trajectory, or understood as smoothly lowering it if focusing on the chance of the squirrel collision becoming more and more likely. But nothing here turns on these particulars—as long as some sort of discontinuity is plausible in some setting, which defensible understandings of some quantum examples provide. The intent of the example here is only to illustrate clearly a chance discontinuity. The point drop rendering above (following Eells) is particularly helpful (though not essential) for my purposes because it exhibits two different discontinuities. I note too that since all most all (excepting Eells) probabilistic analyses of causation are explicitly neutral with respect to the continuity question, the mere possibility of discontinuities needs to be explicitly accommodated or ruled out, since the possibility itself undercuts such analyses. Both of these points will be taken up at length below.

3. A jump discontinuity is one in which the left- and right-hand limits exist, but are not equal. The other two possibilities, that the left and right hand limits exist and are equal, or that one (or both) fail to exist are called removable and essential discontinuities, respectively. The essential discontinuity case will come up again below.

4. Eells, for one, does recognize that chance could also be represented in a continuous fashion, with the probability continuously approaching one from below. But he writes that his analysis does not "pay attention" to whether the trajectory is continuous at the time the event occurs (Eells 1991, 294, note 6). See Peressini (forthcoming) for an argument to the contrary.

5. A possible exception to this might be an irreducibly probabilistic (point) event, e.g., whether element U-238 will emit an electron by time t. According to prevalent interpretations of quantum physics this probability will be bound away from one right up to the instant it happens, at which point it will "jump" to one. I discuss this case below.

6. I will not distinguish in what follows between causally and probabilistically relevance. The questions of if and how these notions coincide is of course at the center of the debate about whether causation can analyzed probabilistically. For the purposes of this paper, probabilistic relevance is sufficient, since the concern here is with probabilistic analyses of causation.

7. It is important to stress that since the argument requires the construction of a series of events upon which Y probabilistically depends, it most obviously succeeds when there is a space-time process leading up to or constituting the event Y, as there is in Example 1. And as a consequence, the argument does not necessarily apply to certain classes of quantum events, which (under certain interpretations) fail to have such probabilistically relevant antecedent events; this is as it should be as there is nothing incoherent about such quantum-level examples. While there is debate about whether all macro-level examples of causation need to have such an intermediate process, even accepting a pluralistic view, e.g., Hall (2004), it is sufficient for my argument here that it work for the large class of macro-level cases (like Example 1) in which there is such a mediating process. I owe thanks to an anonymous reviewer for help with this point.

8. I note that CDJP does not give rise to any novel problems from those that follow from DJP, and in fact may be seen as following from DJP, since DJP entails that there be discontinuous "jumps" in an event's trajectory at each moment its "causes" occur. But while intuitive, the actual argument to establish this entailment is far from trivial; see Peressini (forthcoming). Furthermore, it is important to distinguish between the two principles because the rationales for introducing each are different, namely causal concerns versus more general ontological concerns regarding determinism, chance and event ontology.

9. Menzies' account, like Kvart's, while not explicitly addressing continuity, does implicitly constrain discontinuities. He builds on Lewis' (1986) counter-factual analysis in terms of unconditional probabilities. Menzies requires that causally related events c and e be probabilistic dependent—which amounts to there being intermediate events corresponding to any finite set of intervening times between the times of c and e such that the actual probability of each of the intervening events is significantly higher than it would have been had the immediately preceding event in the set not happened. This effectively requires the chance function to be monotonically increasing, and turns out to be an implausibly strong condition; Menzies (1996) himself disavows even an amended version of this theory. As I draw out below, the point probability framework and this continuity bind often force one to choose between stability in the chance function and such overly strong constraints on it.

10. Hitchcock (2004, 414) reports Ned Hall's suggestion that one evaluate the probability of an effect shortly before the time at which the effect occurs; Hitchcock also outlines there a related proposal of his own.

11. The original version is still available online at http://web.mit.edu/gradphilconf/2008/A%20Probabilistic%20Analysis%20of%20Causation.pdf.

12. The term "imprecise probability" traces back to Walley's (1991) foundational work in the area.

13. Other prominent orderings are center-point and radius less than, center-point less than and radius greater than, lower point and center point less than, upper point and center point less than. See Guerra and Stefanini (2011).

14. Even if the quantum jump from *r* to 1 at time t_y in Fig. 3 is defined to be the interval value *[r, 1]*, the function is still not path-continuous. Were there compelling motivations, there are ways to accommodate such jumps within a path continuous framework, e.g., by relaxing it to require only left or right path continuity or by defining *P(t)* to be *[r, 1]*

at t_y and all subsequent times, but as things stand gaplessness works equally well without the complications.

15. For example, one may weaken the definition of Lower SC by requiring only that one (rather than all) elements in the domain set at a point have converging sequences. So an interval function $F: \mathbb{R} \to \mathbb{IR}[0,1]$is Weak LS Continuous at $x_0 \in \mathbb{R}$if and only if there is a $y_0 \in F(x_0)$such that for any sequence $x_i \in \mathbb{R}$ and $\{x_i\} \to x_0$there exists a sequence of elements $y_i \in F(x_i)$ with $\{y_i\} \to x_0$.

16. It should be noted that semicontinuity for real valued functions like $\underline{f}$ and $\overline{f}$ is distinct from, though not unrelated to, semicontinuity for generalized set-valued functions and interval-valued functions like F. In particular, semicontinuous real-valued functions may well be gappy in a way that is precluded in $\mathbb{IR}[0,1]$ or more general set-valued spaces.

17. An interesting question for further work is whether the G-continuity of $F = [\underline{f}, \overline{f}]$is equivalent to $[((\underline{f})$ being LSC) and ($\overline{f}$ being USC)] or [($\underline{f}$ being USC) and ($\overline{f}$ being LSC)], but not both of $\underline{f}$ and $\overline{f}$ being one of LSC or USC.

18. I stress that it can be ignored only if P is continuous; it is not true that the inequality holds only if P is continuous.

19. It would be an interesting project in itself to recast all of the particular idiosyncratic details of the various competing probabilistic accounts of causation in term of imprecise probabilities, including reassessing each of the examples and arguments they employ.

20. Even when the temporal index is explicitly expressed as in (1) (as opposed to placed out of sight within a "variables taking on values" approach), as far as I can tell the temporal index is simply "carried along," that is, the continuity properties of chance as a function of time are not addressed. I note too that as mentioned above, Menzies' (1989) account does indirectly rule out the possibility of any (and therefore any discontinuous) drops in chance, but at the cost of an

implausibly strong monotonicity requirement, which in part leads him to disavow the account altogether (Menzies 1996).

21. Of course the tradeoff with this move is that it would rule out certain kinds of "gapless" functions, i.e., those that are LSC and not USC. Recall Sect. 5.2.

Acknowledgments

Key parts of this work were done on sabbatical in Berlin 2012-13; I thank Dörte and Frieder Middelhauve and Jodi Melamed for helping make the sabbatical possible (and fun) along with Michael Pauen and the Berlin School of Mind and Brain for providing a place to work, and a stimulating and gemütlich community. A version of this paper was presented at the Imprecise Probabilities in Statistics and Philosophy Conference at Ludwig-Maximilians-Universität München, June 28, 2014. I am grateful for the comments and suggestions I received from participants. I thank also this journal's anonymous reviewers for their helpful comments.

References

Anguelov, R., & Markov, S. (2007). Numerical computations with hausdorff continuous functions. In T. Boyanov, S. Dimova, K. Georgiev, & G. Nikolov (Eds.), *Numerical methods and applications, lecture notes in computer science* (Vol. 4310, pp. 279–286). Berlin: Springer.

Anguelov, R., Markov, S., & Sendov, B. (2006). The set of hausdorff continuous functions–The largest linear space of interval functions. *Reliable Computing*, 12(5), 337–363.

Aubin, J., & Frankowska, H. (1990). *Set-valued analysis*. Birkhäuser: Systems & Control.

Augustin, T., Coolen, F. P., de Cooman, G., & Troffaes, M. C. (2014). *Introduction to imprecise probabilities*. New York: Wiley.

Eells, E. (1991). *Probabilistic causality*. New York: Cambridge University Press.

Flores-Franulic, A., Chalco-Cano, Y., & Roman-Flores, H. (2013). An ostrowski type inequality for interval-valued functions. *In IFSA world congress and NAFIPS annual meeting* (IFSA/NAFIPS), 2013 Joint, pp. 1459–1462.

Glynn, L. (2010). Deterministic chance. *British Journal for the Philosophy of Science*, 61(1), 51–80.

Glynn, L. (2011). A probabilistic analysis of causation. British Journal for the Philosophy of Science, 62(2), 343–392.

Glynn, L. (2014). *Unsharp best system chances*. http://philsci-archive.pitt.edu/10239/.

Guerra, M., & Stefanini, L. (2011). A comparison index for interval ordering. *In IEEE symposium on foundations of computational intelligence* (FOCI), 2011, pp. 53–58.

Hájek, A. (2003). What conditional probability could not be. *Synthese*, 137(3), 273–323.

Hall, N. (2004). Two concepts of causation. In J. D. Collins, E. J. Hall, & L. A. Paul (Eds.), *Causation and counterfactuals* (pp. 181–276). Cambridge, MA: MIT Press.

Hitchcock, C. (2004). Do all and only causes raise the probabilities of effects? In J. D. Collins, E. J. Hall, & L. A. Paul (Eds.), *Causation and counterfactuals* (pp. 403–418). Cambridge, MA: MIT Press.

Ismael, J. (2011). A modest proposal about chance. *Journal of Philosophy*, 108(8), 416–442.

Kvart, I. (2004). Causation: Probabilistic and counterfactual analyses. In J. D. Collins, E. J. Hall, & L. A. Paul (Eds.), *Causation and counterfactuals* (pp. 359–386). Cambridge, MA: MIT Press.

Kyburg, H.E. (1999). Interval-valued probabilities. *In The society for imprecise probability: Theories and applications.* http://www.sipta.org/documentation/interval_prob/kyburg.

Lewis, D. (1986). Postscripts to 'causation'. In D. Lewis (Ed.), *Philosophical papers* (Vol. II, pp. 172–213). Oxford: Oxford University Press.

Li, F.C., & Li, J. (2010). Ordering method of interval numbers based on synthesizing effect. *In International conference on machine learning and cybernetics* (ICMLC), 2010, (vol. 1, pp. 108–112).

Menzies, P. (1989). Probabilistic causation and causal processes: A critique of Lewis. *Philosophy of Science*, 56(4), 642–663.

Menzies, P. (1996). Probabilistic causation and the pre-emption problem. *Mind*, 105(417), 85–117.

Moore, R. (1966). *Interval analysis*. Prentice-Hall series in automatic computation. Englewood Cliffs: Prentice-Hall.

Moore, R. (1979). *Methods and Applications of Interval Analysis*. Studies in Applied and Numerical Mathematics. Society for Industrial and Applied Mathematics.

Moore, R., Kearfott, R., & Cloud, M. (2009). *Introduction to Interval Analysis*. SIAM e-books. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104).

Noordhof, P. (1999). Probabilistic causation, preemption and counterfactuals. *Mind*, 108(429), 95–125.

Northcott, R. (2010). Natural-born determinists: A new defense of causation as probability-raising. *Philosophical Studies*, 150(1), 1–20.

Peressini, A. (forthcoming) Causation, probability, and the continuity bind.

Rosen, D. A. (1978). In defense of a probabilistic theory of causality. *Philosophy of Science*, 45(4), 604–613.

Sober, E. (2010). Evolutionary theory and the reality of macro probabilities. In E. Eells & J. Fetzer (Eds.), *The place of probability in science* (pp. 133–162). Netherlands: Springer.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. New York: Chapman and Hall.

Weichselberger, K. (2000). The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24(2–3), 149–170.