Question 1. Continuity*

Glenn Shafer and Vladimir Vovk

January 27, 2003

Question

Kolmogorov's measure-theoretic framework for probability includes an axiom of continuity, which is equivalent to countable additivity once finite additivity is assumed. Does the game-theoretic framework require a similar axiom?

Brief Answer

No axiom of continuity was assumed in our book, *Probability and Finance: It's Only a Game!* Because the axiom of continuity for measure-theoretic probabilities is not well motivated, our being able to prove the classical limit theorems of probability without it should count as an advantage of our framework over the measure-theoretic framework. On the other hand, as we have shown in Working Paper #5 for the **Game-Theoretic Probability and Finance Project**, an alternative definition of game-theoretic upper probability that does satisfy the axiom of continuity can play a useful simplifying role in the game-theoretic treatment of continuous-time processes.

*This document can be downloaded from www.probabilityandfinance.com, the web site for the **Game-Theoretic Probability and Finance Project**. Look for Question 1 in the list of Frequently Asked Questions.

Contents

1 Fuller Statement of the Question

In his celebrated 1933 treatise on the foundations of probability, *Grundbegriffe der Wahrscheinlichkeitsrechnung* [4], Kolmogorov began with a field $\mathcal{F}$ of subsets of a set Ω. He considered a real-valued function P on $\mathcal{F}$ that satisfies two axioms:

- $\mathrm{P}(\Omega) = 1$.
- If A and B are in $\mathcal{F}$ and $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.

These axioms imply that $P(\emptyset) = 0$ and that

$$P(A_1 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n) \tag{1}$$

whenever $A_1, \ldots, A_n$ are disjoint sets in $\mathcal{F}$. Equation (1) expresses the condition that the set function P be *finitely additive*.

For infinite Ω, Kolmogorov further required that $\mathcal{F}$ be a σ-field (i.e., that it be closed under countably infinite unions and intersections) and that P satisfy an *axiom of continuity*:

- If $A_1, A_2, \ldots$ is a decreasing sequence of elements of $\mathcal{F}$, and $\cap_{i=1}^{\infty} A_i = \emptyset$, then $\lim_{i\to\infty} P(A_i) = 0$.

In the presence of the other axioms and assumptions, this axiom is equivalent to each of the following statements:

- If $A_1, A_2, \ldots$ is a decreasing sequence of elements of $\mathcal{F}$, then

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P(A_i). \tag{2}$$

- If $A_1, A_2, \ldots$ is an increasing sequence of elements of $\mathcal{F}$, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P(A_i). \tag{3}$$

- If $A_1, A_2, \ldots$ is a sequence of disjoint elements of $\mathcal{F}$, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right). \tag{4}$$

Equation (4) expresses the condition that P be *countably additive*.

The game-theoretic framework does not necessarily produce probabilities in the usual sense. In general, it produces only lower and upper probabilities, pairs $\underline{P}(A)$ and $\overline{P}(A)$ satisfying $0 \le \underline{P}(A) \le \overline{P}(A) \le 1$ and

$$\underline{P}(A) + \overline{P}(A^{\mathrm{c}}) = 1, \tag{5}$$

where A^c is the complement of A. In general, $\underline{P}(\emptyset) = \overline{P}(\emptyset) = 0$, and $\underline{P}(\Omega) = \overline{P}(\Omega) = 1$, but neither $\underline{P}$ nor $\overline{P}$ need be finitely additive. So there is no question of requiring that they be countably additive. It does make sense, however, to ask about condition (2) for $\underline{P}$ or condition (3) for $\overline{P}$. Should $\underline{P}$ satisfy

$$\underline{P}\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \underline{P}(A_i) \tag{6}$$

when $A_1, A_2, \ldots$ is a decreasing sequence of sets? Equivalently, should $\overline{P}$ satisfy

$$\overline{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \overline{P}(A_i) \tag{7}$$

when $A_1, A_2, \ldots$ is a increasing sequence of sets? These are the questions we are considering here.

A reader who comes to this discussion with no sense whatsoever of the intuitive meaning of $\underline{P}$ and $\overline{P}$ might ask why we are considering the condition with decreasing sets for $\underline{P}$ and the condition for increasing sets with $\overline{P}$. Why not the other way around? The reason why the conditions make no sense the other way around is that $\underline{P}(A)$ measures the degree to which we expect A to happen, while $\overline{P}(A)$ measures the plausibility of its happening. In the extreme case where we have no evidence at all, we have no reason to expect anything in particular: $\underline{P}(A) = 0$ for all $A \neq \Omega$. And everything is plausible: $\overline{P}(A) = 1$ for all $A \neq \emptyset$. Since $\underline{P}(\Omega) = 1$ and $\overline{P}(\emptyset) = 0$, this example violates (3) for $\underline{P}$ and (2) for $\overline{P}$.

Equation (5) tells us that $\underline{P}$ and $\overline{P}$ can each be defined in terms of the other, and hence that it is unnecessary to study them both. Here, as in our book, we single out $\overline{P}$ for study. So the question we discuss is whether $\overline{P}$ should satisfy (7) when the sequence $A_1, A_2, \ldots$ is increasing.

2 Why is Countable Additivity Controversial?

Countable additivity for probability has always been controversial. Émile Borel, who introduced it, and Andrei Kolmogorov, who confirmed its role in measure-theoretic probability, were both ambivalent about it. They saw no conceptual argument for requiring probabilities to be countably additive. It is merely mathematically convenient to assume they are. As Kolmogorov explained in his *Grundbegriffe*, countable additivity has no meaning for empirical experience, which is always finite, but it is mathematically useful. Kolmogorov did not list examples where it is useful, but many such examples were already known at the time, the foremost being the strong law of large numbers. Kolmogorov did not know how to prove this theorem without first assuming countable additivity.

We can elaborate Kolmogorov's explanation by pointing out that infinities enter into applied mathematics not as representation but as simplification. Though finite, our experience of reality is often too complicated to understand easily. We move to infinity to smooth away some of the complications, leaving structures that are easier to understand. Kolmogorov's axiom of continuity can be thought of as part of this simplifying or smoothing process.

We would like, of course, as much justification as possible for the particular ways we use infinities to simplify. That an infinitary axiom enables us to prove theorems we cannot otherwise prove is one argument for adopting it, but not by itself a fully satisfying argument when the mathematics we are doing is supposed to be applied. We would also like to understand as well as possible the nature of the simplification imposed by the axiom. If the axiom suppresses certain messy details, we may want to know something about how these details re-emerge when we translate the theorems the axiom enables back into finitary terms. If the axiom is completely irrelevant to finitary reality, we may be puzzled about how the theorems it enables can be useful. These questions seem not to have been answered in a fully satisfying way within the measure-theoretic framework, and this may be why debate over countable additivity still continues.

Kolmogorov was a frequentist. He justified finite additivity by the fact that it is satisfied by relative frequencies in a real and therefore finite experiment. He was hesitant about countable additivity because it is irrelevant to such a finite experiment. The most criticism of countable additivity has come, however, from scholars who see probability as an idealization of belief rather than as an idealization of frequency. The subjectivist Bruno De Finetti was already objecting to countable additivity in the early 1930s [3].

3 Questions for the Game-Theoretic Framework

Instead of starting with an assignment of numbers to sets, the game-theoretic framework starts with a perfect-information game between two players, Skeptic and Reality. It then defines upper probabilities in terms of Skeptic's opportunities to increase his capital in this game. We believe that this deeper starting point provides opportunities to see more clearly when and how continuity is useful.

Because game-theoretic upper probabilities are defined quantities, not primitives, it makes no sense to impose (7) on them as axiom. But we can ask two questions:

- Can we impose simple and natural conditions on the details of the game (the move spaces of Skeptic and Reality and the rule for calculating Skeptic's capital at each step) or on the rule for defining the upper probabilities that guarantee (7)?
- Are such conditions useful? Do they help us prove interesting theorems or simplify the statement of theorems in useful ways?

This gives us a lot to talk about.

4 Definitions of Upper Probability

Although we cannot completely review the game-theoretic framework here, it may be useful to give some basic terminology:

- The *sample space*, for which we write Ω, is the set of all permitted sequences of moves by Reality.
- An *event* is a subset of Ω. In other words, an event is something Reality does.
- A strategy for Skeptic is *risk-free* if Skeptic does not risk bankruptcy when he follows the strategy starting with unit capital. This means his capital will never become negative no matter what Reality does.
- Given an event A and a constant $C > 1$, *Skeptic can refute A at level C* if he has a risk-free strategy that guarantees him at least C at the end of the game if he starts with unit capital and Reality does A.

In our book, we define the upper probability of an event A by

$$\overline{P}(A) := \inf\left\{\frac{1}{C} \,|\, \text{Skeptic can refute } A \text{ at level } C\right\}.$$

In Working Paper #5, where we are concerned with continuous-time processes, we define it instead by

$$\overline{P}^{\dagger}(A) := \inf\left\{\frac{1}{C} \,|\, \text{there is a sequence } A_1 \subseteq A_2 \ldots \text{ of events such that}\right.$$

$$\left.\text{(a) } A \subseteq \textstyle\bigcup_{i=1}^{\infty} A_i \text{ and (b) for each } i \text{ Skeptic can refute } A_i \text{ at level } C\right\}$$

$$= \inf\left\{\lim_{i\to\infty} \overline{P}(A_i) \,|\, A_1 \subseteq A_2 \subseteq \cdots \;\text{ and }\; A \subseteq \bigcup_{i=1}^{\infty} A_i\right\}.$$

If Ω is finite, then $\overline{P}^{\dagger}(A) = \overline{P}(A)$. But if Ω is infinite, we can only say that $\overline{P}^{\dagger}(A) \le \overline{P}(A)$.

It is easy to check that the set function $\overline{P}^{\dagger}$ satisfies the axiom of continuity (7). But nothing equally general can be said about the set function $\overline{P}$; whether it satisfies (7) depends on the details of the probability game—the particular move spaces available to Skeptic and Reality and the rule for calculating Skeptics gain from their moves.

We can offer the following comments:

- In the simple bounded forecasting game we study in Chapter 3 of *Probability and Finance* (this chapter may be downloaded from www.probabilityandfinance.com), the axiom (7) is satisfied by $\overline{P}$. This is because the possible moves by Skeptic are bounded in each situation in the game, and hence we can always form a convex linear combination of a countable number of strategies for Skeptic, without worrying about the convergence of the linear combination. (See the proof of Lemma 3.2 on page 68.)

- If we define the general concept of a probability game as in §8.3 of *Probability and Finance*, then it is easy to construct artificial examples of probability games that do not satisfy (7).
- We do not know interesting examples where (7) is not satisfied.
- As for the other interesting examples of probability games in our book (interesting because they correspond to classical theorems), we do not know whether these examples satisfy (7). We have not studied the question exhaustively, because nothing seems to turn on it.

5 Definitions of "Almost Surely"

To gain more insight, let us leave upper probabilities aside for a moment and merely consider some alternative definitions of "almost surely":

1. A happens *almost surely in our first sense* if Skeptic has a risk-free strategy that guarantees him infinite capital at the end of the game if he begins with unit capital and Reality does not make A happen (our events are always statements about what Reality does).
2. A happens *almost surely in our second sense* if $\overline{P}(A^{c}) = 0$. This is equivalent to saying that Skeptic can refute A^{c} at level C for every $C > 1$, no matter how large C is. In other words, for every $C > 1$, Skeptic has a risk-free strategy that guarantees him capital C at the end of the game if he begins with unit capital and Reality does not make A happen.
3. A happens *almost surely in our third sense* if $\overline{P}^{\dagger}(A^{c}) = 0$. This is equivalent to saying that there is a sequence of events $A_1 \subseteq A_2 \subseteq \cdots$ such that (1) $A \subseteq \cup_{i=1}^{\infty} A_i$ and (2) for any $C > 1$ and any i, $i = 1, 2, \ldots$, Skeptic has a risk-free strategy that guarantees him capital C at the end of the game if he begins with unit capital and Reality does not make A_i happen.

These are successively weaker definitions:

$$
\begin{aligned}
&A \text{ happens almost surely in our first sense} \\
&\quad\Longrightarrow A \text{ happens almost surely in our second sense} \\
&\quad\Longrightarrow A \text{ happens almost surely in our third sense.}
\end{aligned}
$$

When we consider only what happens "almost surely", the question of whether upper probabilities are continuous is replaced by the question of whether

$$A_i \text{ happens almost surely}, i = 1, 2, \ldots \Longrightarrow \bigcap_{i=1}^{\infty} A_i \text{ happens almost surely.} \quad (8)$$

Because $\overline{P}^{\dagger}$ is continuous, the implication (8) holds for the third sense of "almost surely". It does not always hold for the other two senses.

As it turns out, the first sense is most useful for studying the classical strong laws of probability, where one considers an infinitely long sequence of trials. The third sense is most useful, on the other hand, for studying processes that are continuous in time, where the asymptotics is concerned not with successive trials but with successively finer descriptions of what happens over a possibly finite period of time.

6 The Classical Strong Laws

The strong law of strong numbers, as we formulated it in our book, says that the average error of certain predictions in a certain game tends to zero almost surely. It turns out that we can prove this in the first sense of almost surely: Skeptic has a risk-free strategy that makes him infinitely rich if the convergence to zero does not take place.

Notice how this statement is weakened when we use the second sense of almost surely: now we are saying that for any wealth C Skeptic has a risk-free strategy for getting C if the convergence to zero does not take place. If this were all we knew, we could say that Skeptic can get as rich as he wants if the convergence does not take place, but we would not know that there a single strategy that works for him no matter what goal C he chooses.

In the simple case of coin tossing, the game-theoretic strong law of large numbers reduces to a theorem that Borel first formulated in 1909 [1]. Borel's proof of this theorem, which tried to avoid the use of countable additivity as an axiom, was not convincing to later mathematicians, who replaced it with proofs that do use countable additivity [6]. The fact that we have been able to prove the theorem without an appeal to countable additivity can be regarded as a vindication, after over 80 years, of Borel's intuition.

We can do without countable additivity as an axiom in our proofs of the strong laws because we can replace it with an argument using the convex combination of a countable number of strategies. This may not always work; in general convergence needs to be checked. (If the strategies $\mathcal{P}_i$, $i = 1, 2, \ldots$, call for Skeptic to make the move M_i in a particular situation, and the magnitude of M_i increases without bound, then an effort to form the strategy $\alpha_1\mathcal{P}_1 + \alpha_2\mathcal{P}_2 + \ldots$, where $\alpha_1, \alpha_2, \ldots$ is a sequence of nonnegative numbers adding to one, will fail if the series $\alpha_1 M_1 + \alpha_2 M_2 + \ldots$ diverges.) But as we show in our book, strategies that converge as required can be found to prove all the classical strong laws. No artificial appeal to countable additivity or continuity is needed.

All the classical limit theorems concern what happens in the limit in a countable number of successive trials. Borel called the study of such countable sequences of trials the theory of "denumerable probability." Our results can be said to vindicate Borel's intuition not just for coin tossing but for the whole topic of denumerable probability.

7 Continuous-Time Processes

The study of continuous-time processes goes beyond denumerable probability; as we have already explained, it concerns a limit of successively finer descriptions of a process over a possibly finite time period rather than the limit of a successively longer sequences of trials.

Continuous-time processes are studied game-theoretically in Part II of our book. As the reader of the book will see, we put continuous-time into the game-theoretic framework using an ultraproduct of finitary games, analogous to the ultraproduct of copies of the real numbers that is used in nonstandard analysis. This is a relatively unfamiliar but powerful way of producing a smooth infinitary picture. Most of the continuous-time results in Part II are concerned with hedging, however, not with probability. Continuous-time probability appears only in Chapter 14, where we discuss diffusion processes from the game-theoretic point of view.

We have now studied game-theoretic continuous-time probability further, in Working Paper #5 [7]. This paper establishes a result that seems fundamental and prototypical: if a speculator in a continuous-time market has no risk-free strategy that will make him infinitely rich in a finite period of time, then the price of a security being traded in the market must either (1) be constant or (2) have the same fractal dimension as Brownian motion (this means its increments must have the order of magnitude $\sqrt{dt}$). This result is well known in the measure-theoretic literature, where it is proven under particular stochastic assumptions (usually one assumes that the process is fractal Brownian motion). It also has messy but well known nonstochastic finitary counterparts (in connection with $\mathcal{R}/\mathcal{S}$ analysis). Like these finitary counterparts, our game-theoretic version makes no stochastic assumptions.

What we want to explain here is how the statement of our game-theoretic result depends on how we define "almost surely". If we use the second sense of infinitesimal, we obtain the result that the price of the security, if it does not have increments of order $\sqrt{dt}$, can vary only by an amount that we can make as small as we want. But if we use the third sense, we obtain the simpler statement that the variation must be infinitesimal. Simplicity is what we want here; messiness belongs in the finitary version of the story. So we are led to adopt the third definition of "almost surely", along with the corresponding definition of upper probability, denoted by $\overline{P}^{\dagger}$ in our preceding discussion. (Although this adoption is, at this stage, of experimental nature; we have not as yet convinced ourselves that the axiom of continuity will not lead to contradictions or awkward implications in the game-theoretic framework.)

8 Summary

As we have explained, continuity is not a condition we can simply impose on game-theoretic upper probabilities. Whether it is satisfied depends on how the probability game is constructed and how upper probabilities are then defined.

Upper probabilities in fully finitary games are necessarily continuous, though in an uninteresting way. So the question of continuity arises only for infinitary games, which are constructed to simplify finitary reality. As it turns out, the infinitary games that interest us do satisfy the condition of continuity, either because infinitely many strategies can be combined or because the continuity is assured by the definition of upper probability that is most convenient. This is hardly surprising, because a story where such continuity holds is simpler than one where it does not. Continuity of upper probabilities is not an important end in itself, however, and there seems to be no reason to go out of our way merely to make it hold.

References

[1] Émile Borel. Les probabilités dénombrables et leurs applications arithmétiques. *Rendiconti del Circolo Matematico di Palermo*, 27:247–270, 1909. Reprinted in [2], Volume 2, pp. 1055–1079.

[2] Émile Borel. *Œuvres de Émile Borel*. Éditions du Centre National de la Recherche Scientifique, Paris, 1972. Four volumes.

[3] Donato Michele Cifarelli and Eugenio Regazzini. De Finetti's contribution to probability and statistics. *Statistical Science*, 11:253-282, 1996.

[4] Andrei N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin, 1933. An English translation by Nathan Morrison appeared under the title *Foundations of the Theory of Probability* (Chelsea, New York) in 1950, with a second edition in 1956. A Russian translation, by G. M. Bavli, appeared under the title Основные понятия теории вероятностей (Наука, Moscow) in 1936, with a second edition, slightly expanded by Kolmogorov with the assistance of Albert N. Shiryaev, in 1974.

[5] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!* Wiley, 2001.

[6] Glenn Shafer and Vladimir Vovk. The sources of Kolmogorov's *Grundbegriffe*. Working Paper #4. The Game-Theoretic Probability and Finance Project. www.probabilityandfinance.com.

[7] Vladimir Vovk and Glenn Shafer. A game-theoretic explanation of the $\sqrt{dt}$ effect. Working Paper #5. The Game-Theoretic Probability and Finance Project. www.probabilityandfinance.com.