Concept and Intuition in Abstract Probability Theory

Kolmogorov's axioms present an abstract conceptual formalisation of probability that runs counter to our intuitive image of randomness and its concrete instances. But are the relations between concept and intuition, concrete and abstract, so straightforward? And does the revolutionary historical sequence leading from set theory and measure theory to abstract probability occlude a deeper, conceptual order of priority? Elie Ayache takes the true measure of this revolution in our understanding of randomness and probability, and its as yet unthought ramifications.

The Concrete and the Real

When abstract probability theory makes a distinction between the concrete sample ω (also known as a random outcome or trial) and the event A that is realised if ω ∈ A, it does something entirely new: this is essentially a distinction between the *concrete* and the *real*. When probability no longer pertains to the random outcome as such, but only to the event, then probability is literally *separated* from randomness. The great foundational gesture of abstract probability theory was to shatter our image of randomness. There is no random *generator* any longer, and it is no longer a matter of *expecting* the random outcomes. Once it is understood that the random outcome matters only in so far as it is the set-theoretic element of an event, then set theory becomes the foundation of probability theory, and everything relating to expectation and to the concrete field of randomness is reduced to the sole measurement of sets. And when we examine the strong law of large numbers, which is what lends tense to the notion of probability and gives us the impression of expecting something to happen with some probability, we realise that measure theory has only been extended to sets of non-denumerable cardinality, and that we now only measure the set of typical (infinite) random sequences, which is of measure 1 and in which no sequence is distinguished in particular.

We must resist the thought that, if the sample ω is an element of the event A, then the event A is realised. The event is realised because of probability and because of the *interpretation* conferred upon a probability equal to 1; it is realised outside of the formalism. Indeed, the formalism of measure theory only measures sets, and never distinguishes the elements that 'realise' them. Under no circumstance can we begin with the random trial ω, and then go on to seek the event to which it belongs, and which it will therefore have 'realised'. It is not because ω

URBANOMIC / DOCUMENTS

URBANOMIC.COM

belongs to A that A is realised; ω belongs to A anyway. We never have ω—it has never been identified as a distinct entity—even though the formal construction begins with this infrastructure. Our intuitive image of randomness and of the random trial is that of drawing balls from an urn; it is that of the *materialisation* of the random trial, of the manifestation of the concrete; but in the formalism of probability, everything points in the opposite direction, that of the measure of sets alone, that of the infinite and non-constructive limit where, precisely, individual trials are indistinguishable and lose their identity.

Our intuitive image of randomness and of the random trial is that of the [materialisation](#) of the random trial; but in the formalism of probability, everything points in the opposite direction

Measure theory needs the elements ω contained in events like A. Otherwise, how could it distinguish between an empty set and a set of measure zero? And how could it establish the distinction between the concrete and the manifest, whose other name is the real? There must exist a situation where it would be impossible to track back from the observed and measured event to its concrete cause; measurement must stop at a certain threshold. Thought must distinguish its two elements: the one that can measure and express things, or that builds up a point of view, elaborates a language and thus expectations, that is to say objects; and the one that imagines, unfathomably, that that which is within the event and brings it about—namely, the random trial—is different from the expression of the event and from the point of view that the event represents; that that which happens and is thrown—this concrete trial—is a mute thing and is different from what is given to the understanding. Thus the concrete, the infrastructure, would be subjacent to the real and to the algebra of events. With the total separation of the concrete and the abstract, with the masterly gesture of Kolmogorov who succeeded in translating this into set theory by seeing the random sample ω as an element of the event, the whole category of thought relating to randomness and to the intuition of randomness is affected.

We have not yet sufficiently rethought our intuition of randomness and of probability in the light of Kolmogorov's formalism

It seems to us that we have not yet sufficiently rethought our intuition of randomness and of probability in the light of Kolmogorov's formalism, and that we have not yet drawn all the conclusions concerning the strong law of large numbers. Doubtless we haven't yet reflected upon it deeply enough, we haven't yet found the new mode of expression and language, the sequence of words that would have to be arranged in the right order, following the formalism. We aren't yet entirely sure how to write phrases like the one we have written above: 'The concrete is distinct from the real'.

The strong law of large numbers did well to dissolve into the non-constructive continuum of sets the feeling of tense and expectation which, for a thinker such as von Mises, remained attached to randomness, and it did well to dissolve the corresponding intuition along with it. We must take very seriously von Mises's desperate attempts to connect the intuitive formalism of collectives that he proposed with measure theory, which alone offers the means to prove the strong law of large numbers.[1] It is precisely here that the difficulty lies. Yes, probability sounds intuitive and we all feel like we understand it, but we must accept that, in order to truly understand probability and see the proof of the strong law of large numbers through to the end—a proof which can only be universal and infinite—we must *lose* the thread of intuition as we pass into the non-constructive logic of set-theory and its non-constructive theorems of (absolute) existence.

The weak law of large numbers tells us that the probability that the average of (independent and identically distributed) random variables will deviate from the common expected value by more than a given tolerance converges to zero as the number of random variables increases, whereas the strong law of large numbers says that their average converges to the expected value with a probability equal to 1.

1. R. von Mises, *Probability, Statistics and Truth* (New York: Dover, 1981).

Concrete cases where the average of random variables does not converge to the expected value thus form a set of *measure zero*. This is all that the theorem of convergence tells us, thanks to the non-constructive paradise that set theory opens up for it; but in any case, this set of measure zero is not identified term by term. Certainly, we can single out concrete cases where the average of random variables converges toward something else—for example, if all the random variables produce a value different from their expected value; but we cannot explicitly identify every such atypical concrete sequence. All we can do is to measure the set that they form.

Inversely, we cannot explicitly exhibit one single typical concrete sequence ω where the average of random variables converges to the expected value. How can an infinite concrete sequence of random numbers be explicitly exhibited?

What is characteristic of theorems of almost sure convergence (i.e. with a probability equal to 1) is precisely this: that we can measure sets, even dense sets, without identifying any particular element ω of them. The strong law of large numbers is proved by providing an upper bound to the measure of the set (the event) in which the average of random variables, computed beyond rank N, deviates from the expected value by more than a tolerance ε. We then observe that, for any given tolerance ε, we can make the upper bound of this measure as small as we wish for an appropriately chosen rank N. Given that the tolerance ε and the upper bound of the measure of the exceptional event can be varied independently, it is not that the concrete sequences ω in which the average converges are easier or more difficult to identify; it is just that we are controlling the measure of the set in which they are contained.

The theorem of almost sure convergence introduces new modes of identification. Identification no longer takes place in terms of the particular individual, or the concrete trial ω, but in terms of the whole set. And since we are dealing with a set of infinite cardinality, we no longer identify it by enumerating its elements, but only via its measure. The Cantor set proves the existence of sets of measure zero that are of the power of the continuum, meaning that the atypical sequences may themselves be non-denumerable. We therefore wonder why no one

Randomness, and even probability, in the sense of propensity, lose their intuitive meanings in the face of set theory's mode of identification

has taken the radical step of entirely subordinating the intuition of randomness to this non-constructive property of set-theoretical infinities and to this new manner of identification that we owe to probability. Randomness, in the sense of expecting a random generator to generate its outcomes, and even probability, in the sense of propensity, lose their intuitive meanings in the face of set theory's mode of identification (namely, that of non-individual identification).

If the objective meaning of probability is given by the strong law of large numbers, and if the latter now concerns only sets of random sequences (events) and not individual concrete sequences ω, then we ask why we would hesitate to finally declare that the mystery of probability has been dispelled. It is after all quite startling that the rigorous meaning of a phenomenon as intuitive as that of randomness, or the empirical law of large numbers, should ultimately emerge out of non-intuitive set theory and its treatment of the infinite. We even wish to invert the course of our intellectual history and say that, with our historical intuition of randomness and of the infinite sequence of dice-rolls, we already had the *intuition of set theory*. So the latter would be natural after all. Why would the strong law of large numbers be any different, once it is made rigorous, from the notion of the convergence of a function toward a limit? In that case also, we had to turn to set theory.

Inversely, we remain fascinated by the simplicity of Kolmogorov's solution. So it was enough to distinguish between random trial and event, and the benefit of this distinction was not only, as most textbooks indicate, the continuous nature of the sample space (geometrical probability). Above all, the benefit of the distinction is that it brings with it a clarification of the content of the strong law of large numbers, and of the way in which it is proved. Once again, what interests us is to arrange thought and the sequence of discourse in the right order. We believe absolutely that intuition must be subject to

revision by the concept—but we also believe that a new intuition can be invented, as well as a new matter, and therefore new words, *following the revolution of the concept*.

The Random Variable

The random variable X is defined as a function over an abstract space which, when we inspect it more closely, will be the very domain of definition of the concrete: the space Ω of concrete random trials ω. It is abstract because the concrete, in what is most proper to it and what is absolute in it, and in so far as the mathematical entity in question, X, depends upon it, appears as *abstraction itself*. The concrete represents withdrawal and depth, or what is most unfathomable for mathematical representation; so it appears here, inversely, as that which concretely *lacks* representation, and therefore as that which is most abstract. Some authors argue that the very term 'random variable' is unjustified, and may even lead to confusion, since the random variable X is defined in Kolmogorov's formalism as a *function* of the random sample ω and not as a variable: ω → X(ω). But we prefer to keep for it the term 'variable' which, in analysis, represents the lowest level of the hierarchy, because what this variable depends upon (and what disquiets those authors and leads them to label it as a function instead) is the *concrete domain*, or the famous 'abstract' space of random samples ω, which is in reality lower than the lowest level of 'concrete' representation in analysis, and constitutes a particular infrastructure for it. Moreover, true functions may be applied to the random variable, Y=f(X), in which case they will be functions of a variable in the sense of the usual mathematical representation of analysis.

The random variable X(ω) cannot be 'represented' as a function of ω, in the sense of analysis: its 'graph' cannot be traced out, because ω itself does not vary within a space of continuous or even representable variation. It is an 'abstraction' that cannot be represented by drawing, by a graph, or by a continuous train of thought. It is discontinuity itself, the bottomless pit, the horror of representation into which the whole concrete world plunges and retracts. What is more, we will speak of an 'abstract integral' when we wish to 'calculate' the average of the random variable X: $\int_{\Omega} X(\omega)P(d\omega)$. The abstract sample space is not a space of quantitative mathematical variation, in the sense of analysis; it is not the space of a variable. The only quantitative variation that applies to it is that of the *measure* of those aggregates of samples called *events*.

The isolated 'variable' ω evades all continuous mathematical representation, as such. Only the aggregate of samples—that is, the mass, the set of samples ω—has a real measure. So it would be tempting to turn the set of samples ω—the event—into the basic mathematical variable whose variation would be continuous and upon which the 'function' X would then depend. But this would not really solve our problem, since the variable X is really a function of the atomic sample ω, and not of the aggregate. Precisely, the measurable event will move over to the side of the random variable itself, so we will speak of the distribution function F(x) of the random variable X as the measure of the event { ω ∈ Ω: X(ω) ≤ x }.

Thus it is the values x of the random variable X that find a status in the usual mathematical representation, which requires a space of continuous variation for the variable. The distribution function F(x) is the first mathematical 'drawing' that can be made of the random variable X(ω). Instead of drawing the graph ω → X(ω), we draw the inverse graph which correlates the values x of X not with the elements ω but with the aggregates of elements ω, which, precisely, are measurable: the events that tell us that the variable X admits of values that are less than x. Since the measure of events satisfies the right additivity rules, it makes sense to correlate this measure with increments in the value of the variable X.

The random variable is a variable; but unlike ordinary variables, the concrete world, has intruded into its domain of variation

The random variable X has a peculiar status, then. It is certainly a function of a lower 'variable', the sample ω, but mathematical representation (analysis) cannot be grasped at this level, and so it is through an inversion that X recovers the role of the base variable, sending the mass of aggregates above it. X is therefore a variable, in the sense of analysis; but

unlike ordinary variables, the concrete trial ω, the concrete world, has intruded into its domain of variation. Something concrete but which, ironically, is called 'abstract' in the mathematical domain, will now rule whether or not the variable X 'presents' itself, and whether or not it is 'realised'. In the representation of *analysis*—the ordinary mathematical representation of a function and of its underlying variable, y=f(x)—all values of x are present and are realised at the same moment (which is a moment of thought). But in the case of the random variable X, something strange—the concrete world, 'realisation'—shatters this moment and this unity.

Has anyone thought of treating randomness and the intuition of randomness in the same way as Bolzano did the notion of the limit of a function in analysis (i.e. what is known as the *rigorisation* of analysis)? Why would the concept not equally prevail against the intuition of randomness? Why stubbornly maintain the intuition of the random generator and the image of the concrete random sequence ω, as von Mises would wish, when a formal treatment deals with the problem perfectly well—in a way that is non-intuitive, for sure, but that allows the semantic subject, whose 'knowledge' is not subjective but objective in the sense of *objective semantics*, to carry out the mental act that is necessary in order to perceive exactly what the concept or the formal script are trying to say?

The criteria (ε,η) of the rigorisation of analysis left the kinetic intuition of the limit of a function behind, and left to thought only the extensive logic of sets and their nested inclusions, the statics but also the infinity made available by set theory. Now, the infinity that is implicit, in the intuition of randomness, is twofold. On the inside, it is the infinity of the inherence of probability in the single die we are holding, and, on the outside, it is the infinity of the sequence of throws of that die. As to convergence, it is the convergence in limit theorems, such as the strong law of large numbers. We must understand that the intuition of randomness calls for another sort of abstraction, which furthers that of analysis. Not only should convergence toward the limit no longer concern the particular values of a function, as it used to be the case in analysis, and concern, instead, the complete function—in this case the random variable X (and doubtless it is this

difference from analysis that leads some authors to say that it is erroneous to call X a 'variable', and that it is really a function)—but, by the same token, the concrete random sequence ω (the argument of X) can no longer appear. From now on, the semantic mind must simply imagine it. We have to understand that, as the series of random variables X_n measuring the frequency of appearance of the face converge, qua functions, toward the limiting frequency (or as the strong law of large numbers singles out the convergence set), nothing is realised concretely and no particular sequence ω of concrete throws of the die appears. Convergence concerns only probability.

It was Kolmogorov who established the random variable, which would precisely remain 'variable' and wouldn't realise a particular value. This is how infinitistic theorems could be handled and the strong law of large numbers shown, keeping in mind that the statement of the strong law could be misleading at first as it might suggest that we had returned to the convergence in values. For the strong law does indeed state that, for all samples ω belonging to Ω, apart from a set of zero measure, $X_n(\omega) \to X(\omega)$. This gives us the impression that for each one of those particular random sequences ω, frequency, as a function, converges to a limiting value. But what returns once again to contradict intuition and to remind us that we are indeed in the domain of probability and not that of analysis, is the un-identifiability of those sequences ω. It is because absolute (Platonic) existence, authorised by set theory, allows us to reason about actual infinity and to operate infinite intersections and infinite unions of sets (the Borel-Cantelli lemma, typically) that we can, at the limit, affirm the existence of these elements ω, and even the existence of a set of measure 1 made up of them, without being able to identify any one of them in particular.

As a consequence of the limit theorem, the conception of randomness is not one of potential infinity, where the sequence would be drawn step by step without one being able to predict the next step, but rather one of actual infinity, where the sequence does indeed exist but resists imagination in a sense other than that of unpredictability. This is the deeper sense that brings us to the real critique and the real separation of the concrete and the real. There is a very subtle sense in which the concrete disappears

The conception of randomness is one of actual infinity, where the sequence does indeed exist but resists imagination in a sense other than that of unpredictability

at infinity, a sense that must, we argue, have repercussions at the very beginning and must already explain, already illuminate, the separation Kolmogorov made when he introduced a concrete that would no longer appear and would no longer ever manifest itself!

Repetition and Infinity

When we reproach measure theory for abolishing von Mises's intuition of randomness, on the pretext of its infinitistic and non-constructivist results, we forget that intuition is already neutralized from the very first step, with Kolmogorov's discovery (or brilliant intuition) of the separation $\omega \to X(\omega)$. In reality, the great non-intuitive mystery of randomness is largely announced from the very beginning. It is simplicity incarnate (although it potentially contains all of measure theory's use of the infinite and the non-constructive). It really consists in asking what the concrete is, what it means that something happens in the concrete world and that we observe something; what it means that we constitute language—that is to say, propositions open to verification—and constitute expectation and objectivity; what it means that the world should *trial* the dice from the inside and that the dice should present and *manifest* the face (the event) on the outside.

But if we say that the mystery of randomness is given from the beginning, we must mean that it concerns randomness as a whole—not just the randomness of tosses of the coin or throws of the die. Here, we are partisans of the idea that mathematical discovery drives intuition forward, or indicates to thought its *breakdown into its real elements*, which at first are not apparent to it. In fact we claim that nothing is non-intuitive, or rather, we ask that we redefine intuition and the power of thought in the light of abstract probability theory. Davidson is right: a theory of truth, and even a theory of thought's real access to the world, must be drawn from, inspired

by, the example of probability.[2] Doubtless probability, with the diagonal way in which it cuts simultaneously through the concrete and the real, indicates something really profound and even primitive—that is to say, something true. We must begin the programme of philosophy, which ultimately aims to say the simplest and most profound things, from what appears today as the least intuitive aspect of abstract probability theory.

Kolmogorov succeeded in squeezing the whole world into the separation $\omega \to X(\omega)$! From the start of the construction, Ω is simply set as the space of concrete samples (which we call abstract, by the way, thus indicating already the whole infinite divide of thought) and all that we superpose on it is the structure of the algebra of events, which is absolutely intuitive but to which we surreptitiously add the clause of passage to the infinite, or σ-additivity. The algebra becomes a σ-algebra and this gives us everything we need. In this preliminary structure, randomness is given in all its mystery, already surpassing every intuitive random sequence von Mises could imagine. For nothing more is needed to show the strong law of large numbers and to finally supply our thought with the elements it was missing (or at least their recognition) before it finally understood our first intuition of randomness.

The way in which the individual concrete sequence ω is not identified as such, in the set of measure 1 to which it is shown to belong 'at the end' of the theorem of infinite convergence, is linked to the power that measure theory wields by virtue of σ-algebra and σ-additivity. It has nothing to do with the non-constructive character of von Mises's random sequence. This is why we claim that randomness, here, is of another nature. It is at once more powerful, deeper and simpler. This randomness is the one deserving the true philosophical analysis that will show how counter-intuitive it seems at first and, subsequently, once the new intuition is acquired, how natural and intuitive it is on the contrary. The persistence of ω, both at the beginning and the end of the proof of the strong law of large numbers, the probability space that provides the outside frame both at the beginning and the end, show us that

2. 'An axiomatized theory of truth may be compared with, say, Kolmogorov's axiomatization of probability.' D. Davidson, *Truth and Predication* (London and Cambridge, MA: Harvard University Press, 2005), 32.

randomness, in all its difficulty and depth, is already found at the beginning, and that a search for its new intuition really has to start at the beginning.

What is more, are we so certain, when we say that the sequence of random variables converges except on a set of measure zero, that this really separates the set of measure zero, on one side, from the convergence set, on the other? A typical sequence that would belong to the set of measure 1 is never really considered, as in von Mises, in order to define probability as the limiting value of frequency. We don't say that convergence has taken place on the set of measure 1 with the aim of picking one of its elements, that is to say a concrete sequence of outcomes ω, and then evaluating probability on it. On the contrary, this typical sequence ω exists only as a representative of the convergence set of measure 1. What counts is not an individual sequence from which we could estimate or even imagine the limiting frequency (like von Mises), but the set of measure 1, or the *typicality* of the typical sequence. The outcome here is the measure 1, not the limiting frequency. We always 'normally' fall in a set of measure 1 and the sequence of random trials is 'normal', since what is normal is random and what is random is normal (why?), and therefore (and the implication goes in this direction) when one throws a die randomly, one is 'normally' sure that the frequency will converge.

What makes things even worse is that not only periodic sequences, or sequences exhibiting a recognizable pattern, must be excluded, but also random sequences in which the frequency would converge toward a value other than the probability—for example, sequences that would be produced by a biased coin. The set of measure zero that the strong law separates from the convergence set is, thus, far more 'inserted' into it and inextricably mixed with it than we think! It is really at the Cantorian infinite, the actual and not the potential infinite, that the full meaning of measure 1 is given (tail events). For instance, the set of sequences that would be produced by a coin whose probability of yielding FACE is 0.50001 and not 0.5 is a set of measure zero, and must therefore be separated from the set of measure 1 of sequences produced by a normal coin (keeping in mind that the measure in question is the one induced on the space 2^ω by the measure that assigns a probability of 0.5 to the event {FACE}). We are surprised that sequences which seem so similar a priori should suddenly be separated so violently into two sets that are so different from each other, one of measure 1 and the other of measure zero! Shouldn't there be continuity in the variation of probability? For the set of sequences produced by a coin whose probability of yielding FACE is 0.50000001 is also of measure zero, etc.

The strong law of large numbers takes place at actual infinity

In reality, the weak law of large numbers already establishes that the sequences will only be separated progressively, as the number N of trials increases, because of the tolerance that it sets over probability. We have to trial the coin further and further in order that the proportion of sequences in which the frequency differs from the probability by a certain given tolerance becomes as small as we like. At any number N of trials that we stop, the infinite sequences belonging to the set of measure 1 will share with the infinite sequences produced by a biased coin that this value N is unable to distinguish from the normal coin all the finite sequences that were drawn up to N. No matter how far we push N, there will always exist a tolerance ε such that a biased coin whose probability of yielding {FACE} is ½±ε will produce infinite sequences that begin with sequences drawn up to N and shared with the infinite sequences produced by the normal coin. Thus the separation between the set of measure 1 (or the set of random sequences produced by the normal coin) and the set of measure zero (or the set of random sequences produced by all biased coins, no matter how slight their bias), really takes place at the actual, final infinity. The strong law of large numbers takes place at actual infinity.

Thought and Matter

So what kind of conclusion could we draw from this—what could the subtlety of measure theory bring to philosophy? Has anyone already thought of the philosophy of probability no longer in the traditional sense of the reality of its inherence, or in the antirealist sense against such inherence, but as the analysis of the meaning of its infinitistic theorems

and the deep intuition that lies behind them (which seems so contrary to intuition, at first)? We have come to understand that probability did not exist in nature, but that it was linked to the fundamental elements of thought and to the way in which thought carves up the world into concrete and real, or into concrete and manifest; and we therefore conclude that the random objects or random phenomena, in which we thought at first the philosophy of probability should be read, are in reality only approximate illustrations of the formalism of random variables. This is why we say that the random variable, $\omega \rightarrow X(\omega)$, and the separation that it establishes between the concrete and the real, or between matter and computation, or again, between the *strike* of contingency and the *measure* of the event (which gives us the algebra of events and the combinatorial logic that correspond to the manifest aspect of thought), are in reality a discovery, not an invention or creation.

Curiously, with the extreme subtlety of measure theory and the impossibility of assigning probability except both at the beginning and the end, with the real meaning of the set of measure 1 that only emerges at actual infinity, and with the loss of identity of the concrete individual ω (even if this loss only takes place at actual infinity), we say that the meaning of probability is found and the path is finally open to conduct a real philosophy of probability. It is no longer a matter of placing probability within the physical object (objective probability) or within the subject (subjective probability), but at the deeper level of thought that makes Kolmogorov's brilliant intuition (the separation between the concrete and the real) and the 'non-intuition' of infinitistic theorems equally deep and equally simple. The philosophy of this equivalence remains to be written on a deeper plane than the plane of physics, in a register that is different from all those that have already been written.

In treatises on philosophy of probability, measure theory always appears as an appendix, as if its only contribution was to make computation rigorous or to cover the cases of continuous probability; and we are always astonished that, in order to complete the domain of computation and to make it consistent, it should be necessary to consider measure theory in all its infinite subtlety—as if probability should

apologize to us for that supplement. It is thought that the formalism and the whole of measure theory are just an accessory, a sort of abstraction whose only purpose it to provide a general mathematical frame. For many authors, abstract probability theory is abstract only because it is here to furnish the general framework of thought for different concrete phenomena all of which have randomness in common.

Certainly, it is randomness we are dealing with ultimately, but what we think is surprising and deserves philosophical wonder above all is that this randomness should call for infinitistic theorems, and their (apparent) non-intuitive character. Yes, we must consider randomness first and marvel at it first, but this is not because it is shared by all concrete random phenomena but because, being so common and so familiar to thought, it requires, in order to be formalised, something as deep for thought as measure theory.

It is as if randomness was a thread of thought that led us down to an unsuspected basement of thought—to its archaeology

Randomness is fundamental not because it is a common characteristic to natural phenomena, not because of the abstraction it represents, but because measure theory turns out to be necessary in order to capture it. It is as if randomness was a thread of thought that led us down to an unsuspected basement of thought—to its archaeology. As we have said, nothing is non-intuitive. Even when Bolzano destroys the kinetic intuition of the limit of a function in analysis, he delivers to thought the perception of a new domain, a deeper level in which it can comprehend and grasp things, that is to say, a new intuition. Just as Bolzano's critique of Kant's intuition is applied to analysis, we must now apply it to randomness. Precisely, the debate between intuition and concept in randomness (that is, between von Mises and Kolmogorov) should teach us that *the force of thought must be found and maintained in randomness*—and perhaps even in its absolute background, which is contingency. So it is ultimately an argument against the dissociation of thought and the absolute background that we are here contemplating.

Just as, with the weak law of large numbers, one does not escape from the interference of probability in any interpretation of probability that one would wish to extract from the law (the probability of an event is such that the probability that the relative frequency deviates from it by more than a certain threshold is smaller and smaller as the number of trials increases)—a matter which engenders an infinite regress—similarly, in the expression of the strong law of large numbers, which was supposed finally to give the interpretation of probability as the exact limiting frequency, one does not escape from the intervention of probability, since the equality of probability and the limiting frequency is only true with a probability equal to 1.

Now, I am convinced that the probability we are dealing with in the second case is different. It results from the power of measure theory and the infinite intersections of sets that occur in it, that is, from the σ-algebra and σ-additivity. It results from the non-constructive limits of set theory (the axiom of choice) and the infinitistic theorems whose real proof evaded Borel.

The weak law of large numbers is verifiable, whereas the strong law is not. I suspect that what the strong law presupposes and that the weak law does not have is the notion of random variable, a notion that was available neither to Bernoulli (who had already demonstrated a version of the weak law) nor to Borel (who believed he had demonstrated the strong law, or at least had correctly formulated it without correctly proving it). It is with the strong law of large numbers that the random variable takes on its full meaning (and its full power)—which lies in the separation of the concrete and the real and only in this. It is the strong law that confirms how essential the discovery of the random variable was.

It is in the infinite, non-constructive intersections which act on sets without enumerating or identifying their elements that the concrete sample ω takes on its full sense, which is that of being separated from the event A, and of precisely being an element of it. An element ω must be had; after we are done with all the infinite intersections, an element ω must remain in our hands; but at the same time the identity of this element ω must be lost. One could almost say, at this stage, that ω is no more than a symbol. Indeed, there

would be no need to create the notion of a set, if it were to enumerate its elements. The set only serves to be identified and measured in place of its elements, or above the collection of its elements, in an operation where, in spite of everything, one entirely retains the certainty that there are elements in this set.

The weak law of large numbers still corresponds to the sequence of concrete trials effectively carried out, whereas Kolmogorov, for instance, in preparation of the statement of the strong law, and even before saying what probability is going to be, declares that it still makes sense to speak of 'probability' after the infinite intersections that he will consider.[3] Now, the meaning of this probability, whose validity Kolmogorov is keen to check and check again, results exclusively from the pure abstraction of measure theory. For we might retort to Kolmogorov: 'What, then, is this meaning of probability?' Everyone was astonished by the theorem that Borel stated; everyone thought that the strong law of large numbers was an astonishing result; Borel had for the first time fallen upon infinitistic results, as von Plato said;[4] he had manipulated infinite sums of probability; but he still didn't have the power of measure theory at his disposal—don't forget that Borel was a constructivist. Certainly Borel's theorem is astonishing because of the exact result it expresses, but what is really astonishing in it, the real novelty here, is the manipulation of actual and not potential infinity: it is the element ω that is given at the end (where actual infinity is found) and that is, at the same time, given to our thought to grasp from the beginning.

It is this character, conferred specifically upon the element ω of the concrete sample space by the strong law of large numbers, precisely the character of being 'lost' rather than identified, that we wish to read as early as the first symbolic appearance of ω, as early as the first inscription that makes it subjacent to the random variable. The whole meaning of ω is to be situated underneath the level that

---

3. 'We may, therefore, speak of the probability of convergence of a sequence of random variables, for it always has a perfectly definite meaning.' A.N. Kolmogorov, *Foundations of the Theory of Probability* (New York: Chelsea Publishing Company, 1950), 33.

4. 'We encounter here a *genuinely infinitistic event*.' J. von Plato, *Creating Modern Probability: Its Mathematics, Physics and Philosophy in Historical Perspective* (Cambridge: Cambridge University Press, 1994), 48.

would be the lowest, in analysis, lower than the variable itself. The random sample ω 'runs through' the sample space in a very peculiar way, as we have said, and this non-representation of ω, this implicit character, is perfectly linked, we argue, to its primary assignment, which is that ω is not a variable but that it is unique, that it is the point of repetition of the world (which only gives the impression of variation). Recall the question that is attached to ω: if ω ∈ A, is A realised?

The sample ω is the indexical, this is why it is unique, always unique. But at the same time, abstract theory makes it the implicit element of a set. Precisely it cannot do otherwise, and therein lies its entire innovation. How to make the indexical vary?

Sets, or superstructure, are that which we manipulate in probability theory, and to which we apply probability calculus: addition, subtraction. They are the events that constitute an algebra. The laws of thought, which are laws of symmetry, proportion, and distribution—or the very laws of representation—are the ones that impose the algebra, on the surface. But what is new in probability is the infrastructure—that is, the sample ω that is an element of the event A. The so-called classical theory of probability had perfect access to the calculus and to the algebra; but it did not have the separation between the abstract and the concrete. Abstract probability theory is fundamentally linked to set theory because of the fundamental property that comes down to the separation between the element and the set; and abstract set theory is already fundamentally linked to the non-constructivist problematic, which is nowhere more apparent than in probability theory.

Probability theory is pressing. Hilbert explicitly asked that it be formalised. We say this to indicate that it is perhaps closer to us, ultimately, than set theory; that it is more familiar and more concrete; that it corresponds better than set theory to our intuition and to our access to the world, to the way in which our thought cuts the world up into concrete and real. Although it was created afterward, it is plausibly through probability theory that we can really access set theory. It is not surprising that most constructivists in set theory (Borel, Poincaré) ended up playing important roles in abstract probability theory.

This foundation, this crucial history, is hidden, literally abolished, from the formal exposition of probability: 'Shut up and calculate'

It seems to us that this foundation, this crucial history, is hidden, literally abolished, from the formal exposition of probability. Khrennikov is quite right to say that the slogan there tends to be: 'Shut up and calculate'.[5] But how to understand probability if we do not explain to the student of probability the extent to which what seem at first like the least intuitive and the least constructive results—the strong law of large numbers—are intimately linked to the deepest intuition, or to that element of thought that is even more deeply rooted than intuition?

It is certainly not intuitively that thought can grasp its own elements, since it can think nothing by decomposing itself into elements. We must teach those elements to thought, reveal thought to itself by way of the concept. Precisely, thought has the power of recognising, after the fact, the depth—that is to say the natural genetic character for thought—of this thing it has just learned about its own elements.

So, probability was formalised; but we are not certain that the intuitive thread was not lost in the meantime. Indeed, Kolmogorov's result is astonishing and magical, in many regards: precisely emerging out of silence, out of secrecy (from behind the iron curtain), out of nowhere! Kolmogorov did indeed formalise something: all of this stands beautifully on set theory and measure theory, but are we still dealing with the same intuitive probability?

One never knows: perhaps the strong law of large numbers is ultimately only a pure mathematical result! All of the language in which it is stated is extremely mathematical. And what is more, what should surprise us in the first place is that a phenomenon that is observed in the physical world, namely the stabilisation of the relative frequency of appearance of the faces of a die, could be shown mathematically. We know how to state Newton's law mathematically: $F = m\gamma$. This law is also manifested in the physical world;

5. A. Khrennikov, *Probability and Randomness: Quantum versus Classical* (London: Imperial College Press, 2016), 5.

but no one has ever thought of proving it! What is this thing, this particularity of the strong law of large numbers, which gives us in contrast the impression that we could—or even must—prove it? Did we intuitively feel that it was a consequence of certain properties of matter and not itself a primary given, and did we set out to seek these properties, from which then to deduce the strong law of large numbers? (This wouldn't be the first time, in physics, that the real cause is hidden—and therefore calls for discovery—and only linked to the manifest phenomenon by a mathematical derivation.)

Or were we first impressed, in the strong law of large numbers, by its 'oscillation', by the truth that it holds back from delivering to us straight away, by its stammering, by the grains of sand that seem little by little to disappear from the surface to reveal the hidden inscription—to be honest, impressed by its progressive aspect, as if this unsettled and hesitant law, almost a primitive law (in the sense of a primitive invention or machine), were itself in the process of trying to tell us something with great difficulty; as if it were itself in the process of showing us something with its own 'random' means, which were precisely limited; and as if thought, once it had divined the direction in which the truth lay, could get to the result faster than the apparent law? Presumably it is the characteristic of the laws of randomness that they should manifest themselves 'randomly' and to keep back, to keep hidden (waiting to be shown, literally) the certain principle. This mode of showing, this fundamental hesitation (are we dealing with an analytic or a synthetic proposition? are we dealing with a theorem or a law?) being doubtless the characteristic of randomness.

Rigorous and formalised probability, the only true probability, the only probability that could answer the philosophical question: 'What is probability?' does not exist in nature or in the laws of physics or dynamics (quantum theory is not probability, it is something else). How could it? Its generality alone should make it suspicious to our eyes that it may be inherent in nature (even allowing that nature may not have localised it in one place, or in one phenomenon) rather than being relative to another level, the level at which nature is *thought*, that is to say, thought semantically—thinking what nature or the material world *means* for thought, rather than what it truly is.

Certainly, probability has something to do with matter. But it has nothing to do with physics, or with the behaviour of matter in physics. Physically, nothing prevents the die we are rolling from following exactly the same trajectory as in the previously roll. No, it is before we even physically lift the die, while it is still at rest and thought explores, at rest, what matter can mean in the world and what it can mean that the concrete world trials it, it is already at that moment that matter presents the multiplicity of its faces on the outside and the one and only concrete world charges it from the inside. It is at this preliminary moment that randomness is conceived. And even the proof that this *semantic matter* is independent of the physical world, where physical experiments will be eventually conducted in physical time, is that this matter and this concrete have been formalised by Kolmogorov, under the form of $\omega \to X(\omega)$, which, despite its formal and explicit appearance, contains all the implicit already, that is to say, all that should implicitly refer to matter.

Rigorous probability, or probability as made rigorous by Kolmogorov, really begins with the formalism and really calls for the random variable. The random variable is not a matter of explaining or representing or modelling the probability that the coin or the die possesses. The inverse is the case. It is rather that the die, which has many faces that can be manifested as events (and obey the rules of combination of the corresponding algebra), *represents* formal probability and is an illustration of it. Like the random variable, the die has many faces; in this sense, we are really talking of two multiples that can correspond to one another. But then we must seek that which, in the die, might represent the element ω. This will be the unicity of the world that traverses the die and trials it. Yes, we say unicity; for it is this, the unique strike of a world, the unique trial, that characterises randomness and makes the random variable different from a general mathematical variable.

The world that breaks the continuous variable, the graph and the generality, is the irruption of repetition, of the difference that will become intensive

In other words, the world, the strike, *throws* the variable. The world that breaks the continuous variable,

the graph and the generality, is the irruption of repetition, of the difference that will become intensive. For the concrete that throws the die, this trial that connects itself to the world, has no sense apart from being unique and from repeating the experiment in the sense in which the world repeats itself— by no longer aligning copies but by making us explore its strike further and further, in a finer and finer repetition of that which makes the difference *in the unique concrete trial*. The irruption of the concrete into the algebra and into the calculus is the indication that something has happened, that something has been thrown: the coin, the die, the strike, the world. *Le sort en est jeté*. It seems to us that too many things happen, in thought, with the advent of probability, in its most extreme form and in its most advanced formalism which is that of measure theory, for it not to deserve a historical report, a true philosophical recognition.