Gerhard Ernst and Andreas Hüttemann (eds.): Time, Chance and Reduction. Philosophical Aspects of Statistical Mechanics. Cambridge University Press, Cambridge 2010, 71-91.

The natural-range conception of probability

Dr. Jacob Rosenthal

University of Bonn

Institut für Philosophie

Am Hof 1

D-53113 Bonn

Germany

Email: jacob.rosenthal@uni-bonn.de

1 *Objective interpretations of probability*

Objective interpretations claim that probability statements are made true or false by physical reality, and not by our state of mind or information. The task is to provide truth conditions for probability statements that are objective in this sense. Usually, two varieties of such interpretations are distinguished and discussed: frequency interpretations and propensity interpretations. Both face considerable problems,[1] the most serious of which I will briefly recall to motivate the search for an alternative.

Firstly, the frequency interpretations. Here the central problem is that it is very implausible (to say the least) to postulate a non-probabilistic connection between probabilities and relative frequencies. What a frequency approach claims seems either to be false or to presuppose the notion of probability. Take, for example, the repeated throwing of a fair die that has equal probabilities for each side. All you can say is that it is *very probable* that upon many repetitions each face will turn up with a relative frequency of approximately 1/6 (weak law of large numbers). Or that, *with probability* 1, the limiting relative frequency of each face would be 1/6 in an infinite number of repetitions (strong law of large numbers). You cannot drop the clauses “very probable” or “with probability 1” in these statements. There are no relative frequencies that the die *would* produce on repeated throwing, but it *could*, with varying probabilities, yield any frequency of a given outcome. This is not only clear from the standard mathematical treatment, according to which the throws are probabilistically independent, but one should expect it from the outset. After all, a chance experiment is

[1] Hájek (1997, 2009) gives a comprehensive list of those problems for frequency accounts, Eagle (2004) does the same for propensity accounts of probability.

repeated, so it should be possible to get the same result over and over again.[2] The idea of repeating a random experiment does not fit together with the claim that certain relative frequencies would or have to emerge. So, you cannot use the connection between probabilities and relative frequencies for an interpretation of the former. Any such connection is itself probabilistic and given by the laws of large numbers, which cannot plausibly be strengthened. Nor will it do to give a different interpretation to those second-order probabilities. They arise from a combination of the first-order probabilities, and thus there is no reason for interpreting them in another way.

Second, in the so-called single-case propensity theories, the entities called "propensities" are supposed to be fundamental entities in nature which somehow attach to individual physical processes and make their different possible outcomes more or less probable. So, propensities are, or create, irreducible objective single-case probabilities. But this implies that they constrain rational credence, because otherwise the term "probability" would be misplaced. Propensities yield the appropriate degrees of belief. There is no account of how they might be able to achieve this, they constrain rational credence "just so". They are assumed to be normative parts of reality, and their normativity is to be taken as a brute fact. If one tries to avoid this oddity by explicating propensities in terms of relative frequencies, one is saddled again with the central problem of the frequency approach. The so-called long-run propensity theories are actually varieties of the frequency interpretation and share most of its features and problems, whereas the single-case approaches don't have much to say about probabilities. They just say that there are entities out there which somehow adhere to single events and give us appropriate degrees of belief. This austere fact is concealed by the

---

[2] As usual in discussions of probability, I use the term (chance or random) "experiment" and related notions like "experimental arrangement" in a wide sense throughout, standing for "in principle repeatable physical process or constellation". It is not just to be thought of experiments that are or can be conducted by human experimentators.

introduction of various new notions, like “disposition of a certain strength”, “weighted physical possibility”, “causal link of a certain strength” or even “generalized force”. “Propensity” is one of these terms, too, which are in greater need of explication than “probability” itself, and the explication of which inevitably leads back to the very concept of probability they are supposed to explain.

In view of these difficulties it doesn’t seem superfluous to look for yet another objective interpretation of probability. And in fact there is a third, not actually new but to a large extent neglected, possibility for such an interpretation: viz., probabilities as deriving from ranges in suitably structured spaces of initial states. To put it very briefly, the probability of an event is the proportion of those initial states within the initial-state space attached to the underlying random experiment which lead to the event in question, provided that the space has a certain structure. I will call this the “range interpretation” of probability statements. It was already proposed by Johannes von Kries by the end of the 19th century, in his comprehensive book *Die Principien der Wahrscheinlichkeitsrechnung* (von Kries 1886, 2nd printing 1927). Where I will use the term “range”, von Kries spoke about “Spielraum”, which could also be translated as “leeway”, “room to move”, or “play space”.[3] I will present this interpretation by making a brief detour through the classical conception of probability, to which it is related.

2 From the classical interpretation to the range interpretation

According to the well-known classical conception, the probability of an event is the ratio of “favourable” (with regard to the event in question) to “possible” cases, where those possible

[3] See Heidelberger (2001).

cases must be “equally possible” and are judged so by the “principle of indifference” or “principle of insufficient reason”:

(CC) Let $\boldsymbol{E}$ be a random experiment. On every trial of $\boldsymbol{E}$ exactly one of the cases a_1, a_2, ..., a_n is realized. If we don’t see (subjectivist reading) or there isn’t (objectivist reading) any reason why a_i should occur rather than a_j for each i, j, then the cases a_1, a_2, ..., a_n are equally possible. If exactly k of them lead to the event A, then the probability of A on a trial of $\boldsymbol{E}$ is k / n.

This classical conception, which originated with the invention of the calculus of probabilities by Fermat and Pascal in the 17$^{\text{th}}$ century and culminated with Laplace, was abandoned in the course of the 19$^{\text{th}}$ century due to the following difficulties: First, equally possible cases are not always to be had, i.e., the principle of insufficient reason is not always applicable. Second, the opposite problem: sometimes there are different types of equally possible cases, i.e., the principle of insufficient reason is applicable in different ways which lead to contradicting probability ascriptions (Bertrand’s paradoxes). Third, since “equally possible” can mean nothing but “equally probable”, the classical interpretation provides no analysis of probability, but presupposes the concept. As will become clear, the range interpretation is able to overcome the first and the second problem. The third one proves harder and will turn out to be a touchstone for success or failure of the range approach.

Now, for example, in how far are the six sides of a symmetric die “equally possible” upon throwing? In how far *isn’t* there any reason (remember we are interested in an *objectivist* reading of the principle of insufficient reason) why one side should turn up rather than another? In how far is it equally easy to get a certain number a as the outcome and to get any other number? Take the outcome of a throw to be determined by its initial conditions. These can be represented by a vector of real numbers, which characterize things like the position of the die in space, its velocities in the different directions, its angular velocities in the different directions, etc., at the very beginning of its flight. Call any such vector, which represents a complete selection of initial conditions, an initial state. Thus, for some n, the n-dimensional real vector space $\mathbf{R}^n$, or a suitable subspace of it, can be viewed as the space of possible initial states for a throw of the die. Each point in the initial-state space $\boldsymbol{S}$ fixes the initial conditions completely, and thus fixes a certain outcome of the throw. Here it is assumed that other conditions, which one could call "boundary conditions", remain constant. These include the physical properties of the die and of the surface onto which it is thrown, the air pressure and temperature and other things like that. The difference between initial and boundary conditions is made in order to distinguish varying factors from constant ones. If, e.g., the air pressure might change from one throw to the next, it would also have to be included among the initial conditions. Strictly speaking, we would then face a different random experiment with a different initial-state space, but, at least in ordinary cases, with the same outcome probabilities.

If we could survey the initial-state space, we would find it consisting of very small patches where all the initial states contained lead to the same outcome. In the neighbourhood of a patch associated with the number a as outcome we would find patches of approximately equal size leading to any other number. Therefore, the outcome a is represented with a proportion of approximately 1/6 in any not-too-small segment of $\boldsymbol{S}$. Without loss of generality, we can lay down this condition for intervals: Let $\boldsymbol{S}_a \subseteq \boldsymbol{S}$ be the set of those initial states that lead to outcome a. Let $\boldsymbol{I} \subseteq \boldsymbol{S}$ be an interval. Let μ be the standard (Lebesgue-) measure. If $\boldsymbol{I}$ is not too small, we will find that

$$\frac{\mu(\boldsymbol{I} \cap \boldsymbol{S}_a)}{\mu(\boldsymbol{I})} \approx 1/6 .$$

The proportion of initial states that lead to outcome *a* in an arbitrary, not-too-small interval ***I*** of the initial-state space is about 1/6.

A loaded die can obviously be treated in exactly the same manner, i.e., the approximate probabilities attached to its sides can be found in any not-too-small segment of the space of initial states as the proportion of those initial states in that segment which lead to the respective outcome. This already overcomes the first shortcoming of the classical interpretation. The structure of the initial-state space explains the typical probabilistic patterns that emerge on repeated throwing of the die: why it is impossible to predict the outcome of a single throw, on the one hand, and why the different outcomes occur with certain characteristic frequencies in the long run, on the other hand. Of course, I have not *shown* that the initial-state space associated with the throwing of a certain die has the sketched structure. This would be a very demanding task. I have to assume that it is possible in principle to do so, using only classical mechanics, and hope that the reader finds this assumption plausible.

3 The natural-range interpretation of probability

The basic idea underlying the range conception is as follows: When we observe a random experiment that gives rise, on repetition, to probabilistic patterns, it is often plausible to suppose, as I did with the die, that this phenomenon can be attributed to an initial-state space with a structure of the indicated kind. On the one hand, in any (not too) small vicinity of an initial state leading to a given outcome *a* there are initial states leading to different outcomes, which explains why the outcome of a single trial cannot be predicted. On the other hand, for each outcome, the proportion of initial states leading to it is constant all over the initial-state space, i.e., it is approximately the same in any not-too-small segment of the space, which explains why there are certain stable characteristic relative frequencies with which the different outcomes occur. We cannot use those frequencies to define the probabilities of the respective outcomes – remember the main objection to frequency interpretations – but we can use the underlying initial-state space to do so.

Before setting up the definition, however, we have to specify the requirement that a certain outcome a is represented with approximately the same proportion in any not-too-small *segment* of the initial-state space. This proves a bit tricky, because if we demand too much here, the applicability of the range interpretation will be unnecessarily restricted, whereas if we demand too little, certain mathematically desirable features of the range approach are lost. First, we have to presuppose that the initial-state space $\boldsymbol{S}$, as well as the set $\boldsymbol{S}_a \subseteq \boldsymbol{S}$ of those initial states that lead to the outcome a, are Lebesgue-measurable subsets of $\mathbf{R}^n$ (for some n) of maybe infinite size. More precisely (to avoid problems with infinity), we assume that the intersection of $\boldsymbol{S}$ or $\boldsymbol{S}_a$ with any bounded measurable subset of $\mathbf{R}^n$ is again measurable. Second, we look at those measurable subsets of $\boldsymbol{S}$ whose size exceeds a certain minimum; let's call them "not too small" from now on. We certainly cannot assume that a is approximately equally represented in any such set, because $\boldsymbol{S}_a$ is itself one of these sets. A natural idea is to stick to bounded and connected measurable sets, or *areas*. As any area can be approximated by (n-dimensional) intervals, we could, without loss of generality, set up the proportion-condition for *intervals*. Unfortunately, even this proves too strong. Imagine a random experiment with a two-dimensional initial-state space, such that the "patches" the space consists of are extremely narrow but potentially infinitely long stripes. As the stripes are so narrow, the range conception of probability should be applicable, but as they are at the same time very long, there are quite large intervals within which all initial states give rise to the same outcome. So it seems that we ought to be even more restrictive and consider only *cube-shaped* or *equilateral intervals*. This is somewhat inelegant, but in this way we are able to capture exactly those cases in which the range approach is intuitively applicable, without giving up too much of the generality of the proportion-condition. Now we are in a position to state the range conception (or interpretation, or definition) of probability, (RC).

(RC) Let $\boldsymbol{E}$ be a random experiment and a a possible outcome of it. Let $\boldsymbol{S} \subseteq \mathbf{R}^n$ for some n be the initial-state space attached to $\boldsymbol{E}$, and $\boldsymbol{S}_a \subseteq \boldsymbol{S}$ be the set of those initial states that lead to the outcome a. Let μ be the standard (Lebesgue-) measure. If there is a number p such that for any not-too-small n-dimensional equilateral interval $\boldsymbol{I} \subseteq \boldsymbol{S}$, we have

$$\frac{\mu(\boldsymbol{I} \cap \boldsymbol{S}_a)}{\mu(\boldsymbol{I})} \approx p\,,$$

then there is an objective probability of a upon a trial of $\boldsymbol{E}$, and its value is p.

This is the first formulation of the range conception. It would be more accurate to speak of the "physical-" or "natural-range conception" (see comment K below), but I will stick to the short term. A second formulation of this approach to probabilities will be given in the next section, using so-called "arbitrary functions". Before turning to that, I am going to make several comments on (RC).

A. The range approach is meant to be a proposal for an objective interpretation of probability statements, i.e., as providing truth conditions for such statements that do not depend on our state of mind or our state of information and are therefore called "objective". The interpretation is reductive: probability is analysed in non-probabilistic terms. The axioms of the calculus of probability immediately follow from the definition, if one takes a small idealizing step and reads the approximate equation as an exact one. The central problem of frequency interpretations is avoided, because the relative frequency of a given outcome may, upon repetition of the random experiment, deviate as much as you like from the proportion with which the outcome is represented within the initial-state space. Such an event is not excluded, merely very improbable. Many of the problems of frequency approaches can be traced back to the confusion of assertibility conditions with truth conditions. What makes a statement about an objective probability true is not the frequencies that will or would arise or have arisen, but what explains those frequencies – in a probabilistic way, using the laws of large numbers. The statement is made true by the physical circumstances that give rise to the typical probabilistic pattern. The range approach claims to spell out what those circumstances are.

B. Second-order probabilities of the kind just mentioned, as they occur in the laws of large numbers, can also be given a range interpretation in a natural way. Considering k independent repetitions of $\boldsymbol{E}$ as a new, complex random experiment $\boldsymbol{E}^k$, the initial-state space attached to this new experiment is just the k-fold Cartesian product – call it $\boldsymbol{S}^k$ – of $\boldsymbol{S}$ with itself. The probabilities of the possible outcomes of $\boldsymbol{E}^k$ are the proportions with which these outcomes are represented in $\boldsymbol{S}^k$. Therefore, if the probabilities in a random experiment admit of a range interpretation, so do the probabilities associated with independent repetitions of the experiment, and the respective ranges are connected in a straightforward way. I will not discuss the problems connected with the notion of independence, in particular the very important question of how to justify the stochastic independence of different repetitions of $\boldsymbol{E}$, but refer the reader to Strevens (2003, ch. 3). Suffice it to say that the independence of the range probabilities need not be taken for granted, or simply assumed, but can be explained in much the same way as the existence of the probabilities themselves. It is a virtue of the approach that the question concerning the source of the objective probabilities and the one concerning the source of their independence on repeated trials admit of largely parallel answers.

C. The range interpretation works only in deterministic contexts, where the outcome of a chancy process is determined by initial conditions.[4] This means that terms like "random experiment" or "chance process" cannot be taken in a hard realist sense. What is random or chancy in this sense depends on our epistemic and computational abilities. So, according to the range interpretation, *that* there are probabilities depends on us, but *what* they are depends on the world. The structure of an initial-state space of the required kind, and in particular the proportions with which the various outcomes are represented within it, is perfectly objective. It only depends on the laws of nature and the dynamics of the process in question, but not on our state of mind or information. But that we call those proportions "probabilities", and call the underlying experiment a "random" or "chance experiment", is dependent on the epistemic and control capacities we possess (or are willing to exercise). There is an epistemic aspect to range probabilities, but this aspect does not concern their value. It concerns the fact that certain experiments are considered to be random or chancy.

[4] One could try to generalize the approach by considering processes in which an initial state does not fix an outcome, but only a probability distribution over the possible outcomes. But this distribution would have to be interpreted in turn, and if this were done according to the range approach, one could in principle eliminate these probability distributions by moving to another, higher-dimensional initial-state space. Therefore, the range approach is indeed restricted to deterministic contexts, or else becomes dependent on a different interpretation of probability.

D. There is a reservation to the range conception that may be addressed in connection with the example of a loaded die. One might speculate that with a die which is, e.g., biased in favour of “six”, we do not get the same proportion of “six” all over the initial-state space, but that this proportion increases when the die is thrown higher above the surface. If anything of this kind were true, there would neither be a characteristic relative frequency of “six” to be explained, nor an objective probability of this outcome. We would still face a random experiment with an initial-state space, to be sure, but we could not attach objective probabilities to the possible outcomes – at least not in the usual sense of point values. To get objective probabilities proper, we would have to constrain the method of throwing by demanding that the die is thrown in a certain fixed distance to the surface. What *could* rightly be said, however, is that the objective probability of “six” in the original random experiment is greater than 1/6. I do not think that suchlike occurs with a loaded die, but random experiments of this kind certainly exist. In these cases, the range approach does not provide definite, point-valued objective probabilities – even abstracting from the fact that (RC) merely states an approximate equality – nor are there any to be had. We may, however, get intervals for objective probabilities. It should not be too surprising that there may be random experiments which in this way give rise to interval-valued objective probabilities.

E. The definition (RC) is vague in two respects: The proportion of initial states leading to outcome a is *approximately* equal to p in any *not-too-small* cubic interval of the initial-state space $\boldsymbol{S}$. This is somewhat unsatisfactory; the first vagueness, in particular, has the consequence that range probabilities do not have exact point values. These are rather chosen for mathematical convenience. The ideal case would be the truly chaotic limiting case in which for *every* equilateral interval $\boldsymbol{I}$ we have $\frac{\mu(\boldsymbol{I} \cap \boldsymbol{S}_a)}{\mu(\boldsymbol{I})} = p$ *exactly*. This never occurs in reality, but often it is convenient to model chancy situations as if this were true. We approach the limiting case, for example, if in the throwing of a die we consider only throws that are made high above the surface with high initial angular velocities. That a situation appears chancy to us at all is due to the fact that it approximates the limiting case sufficiently; how good the approximation must be depends again on our epistemic and control abilities. They determine what the clause “not too small” in (RC) amounts to, i.e., what is the minimal interval size from which on the condition is required to hold.

F. The range definition of probability stands in-between two more clear-cut proposals. First: Just take the proportion of an outcome in the initial-state space as the probability of this outcome. Do not care how the initial states associated with this outcome are distributed over the space. Second: Only in the truly chaotic limiting case do we have objective probabilities. The equation has to hold *exactly* for *every* cube-shaped interval. The first proposal is untenable. If the objective probability of an event were just its proportion in the initial-state space, no matter how this space is structured, we would have to ascribe objective probabilities in cases that are not at all random to us, while in other cases we would get the wrong probabilities out of this proposal. This is because on repetition of the random experiment, we cannot expect the initial states that actually occur to be evenly distributed over the initial-state space. There will be regions of ***S*** that occur more often than others, and we have to guarantee that this does not influence the probabilities. Therefore, we have to require that the proportion of the initial states leading to a given outcome is (about) the same in any (not-too-small) segment of ***S***.

The second proposal is reasonable. One may constrain talk of objective probabilities to the limiting case. This has two advantages: It is impossible *in principle* to influence the probabilities by refinement of our abilities of measurement, computation and control. Nature may know the outcome in advance, but we will never know, nor can we improve in any way on the value p as an expectation whether or not the outcome a will occur on a single trial. Furthermore, the outcome probabilities are given precise numerical values by the structure of the initial-state space. The disadvantage, of course, is that this kind of situation never occurs in reality. But here one may simply say that many actual situations are very close to the limiting case, i.e., practically indistinguishable from it for us, given our epistemic capacities. Therefore we model them accordingly. As reality contains structures that can play the role of objective probability near enough, the ascription of objective probabilities is justified in such cases.

G. We do not get single-case probabilities out of the range approach. The initial-state space is something attached to a *type* of random experiment, whereas in a single trial, there is a definite initial state as well as a definite outcome, and no more. We neither have several possible outcomes, one of which is realized, nor an initial-state space, one element of which occurs – after all, a deterministic context is presupposed throughout. I have used the term “random experiment” in the type-sense, as I spoke about repetitions of a given experiment, and without a notion of the repetition or, equivalently, of the type of a given process, the talk of possible outcomes or possible initial states makes no sense. Consequently, there is no way to get anything like range probabilities for a single case considered in isolation. According to the range approach, an objective probability is never single-case, and a single-case probability requires an epistemic reading. This is not surprising, because objective single-case probabilities imply genuine indeterminism. It may be tempting to regard as indetermined which initial state is realized on a given occasion and to get objective single-case probabilities in this way, after all. That would also cut off the problem discussed in the next remark. But this move would make the occurrence of a particular initial state the outcome of a further, now genuinely indeterministic random experiment, and consequently suggest a further concept of probability. In order to preserve the point of the range approach, one has at least to remain agnostic regarding the question whether or not the occurrence of a certain initial state is itself determined.

H. The term “random experiment” is used in a broad sense, meaning “in-principle-repeatable physical process”, and exhibiting no particular connection to epistemic subjects that are thought to “intervene in nature”. It is therefore a matter of convention what one picks out as the initial state of a certain process. Take again the throwing of a die. The outcome is determined by initial conditions, but why take as initial conditions various physical parameters at the moment when the die leaves the gambler’s hand? Why not go further back in time and look, e.g., at the moment when he takes up the die? It seems that one could as well regard the physical circumstances at that moment, which include a physical characterization of the gambler’s brain and body, as the initial conditions of the throw. Obviously, there are in principle countless possibilities for what to take as the initial-state space of the experiment “throw of a die with such-and-such properties under normal circumstances”. Which one gives us the true probabilities?

In reply to this, one may simply say that these cases amount to several different random experiments with different initial-state spaces. In specifying a type of experiment or a type of physical process, one has to specify in particular what counts as the beginning of the process or experiment. That cuts off the problem, but at the cost of threatening to completely undermine the objectivity of probability ascriptions to event types like “getting six on a throw of a symmetric die under normal circumstances”. One can give a more substantial answer by observing that going further back in time presumably would not change the probabilities, and making such a feature a *requirement* for probability ascriptions according to the range conception. This is von Kries’s line (1886, ch. 2.3, 2.6). According to him, the ranges in an initial-state space provide objective probabilities only if they are “ursprünglich” (primordial, original), which means just that: if one regards earlier states as the initial states, this does not affect the respective proportions of the outcomes of the process. To be sure, one gets many different “initial”-state spaces for a given type of process, but always with the same properties with regard to the possible outcomes. Von Kries also provides arguments to the effect that normally we are justified in regarding as giving rise to original ranges the state spaces we are inclined to view as initial.

I leave open which route to take here, since the problem leads into very deep waters. The easy solution implies that range probabilities are relativized in the mentioned respect, and it thus becomes difficult to uphold the claim that they are objective. They depend on a cut in time that is conventionally made by epistemic subjects. The ambitious solution, on the other hand, seems to imply that whenever we ascribe objective probabilities to certain events, we have to make far-reaching assumptions. We have to be able to locate the “initial state” of the process – and with it the relevant state space – arbitrarily far in the past without changing the probabilities. Even if we are ready to make such assumptions, however, at some stage we will hit on the initial state of the universe, and it is doubtful whether talk about an “in-principle-repeatable physical process” still makes sense then. At this stage, the “repetition of the experiment” would amount to the creation of a new universe with the same laws of nature but, presumably, a different initial state out of the space of all possible initial states of the universe. And considering *this* state space, which no doubt rightly carries the name “primordial”, the proportions of the outcomes would *still* have to be the same. It would be quite embarrassing if one were (implicitly) committed to such considerations when ascribing range probabilities. It may however be the price for full objectivity, and if we are not prepared to pay it, we may have to be content with a pragmatically reduced range concept of probability.

I. In what sense is the range interpretation related to the classical one? One could say that the former treats the initial states in any small equilateral interval of ***S*** implicitly as "equally possible", because it identifies the probability of a certain outcome *a* with the proportion with which *a* appears in such an interval. One could interpret this as the implicit supposition of an approximately uniform distribution over any small cubic interval in ***S***. Therefore, the "equally possible" cases are in a weak sense always to be had, after all: initial states that are near to one another in ***S*** may be said to be treated implicitly as equally possible by (RC). This already points to the main problem of the range approach, to be discussed in section 5: as with the classical conception, "equally possible" means nothing but "equally probable" here, so the concept of probability seems to be presupposed by the range approach. But the approach is applicable to loaded dice and the like. Furthermore, it does not depend on a subjectively interpreted principle of indifference and consequently does not fall prey to Bertrand's paradoxes. So it can at least overcome the first two problems of the classical concept. Whether it is able to deal with the third problem, too, remains to be seen.

J. It is also instructive to compare the range interpretation to frequency and propensity accounts of objective probability. As I said in remark A, frequency approaches mistake the evidence for objective probabilities for the probabilities themselves. No doubt there is a close connection between objective probabilities and frequencies: the former can be used to explain and predict the latter, the latter to infer the former. But all these explanations, predictions and inferences are probabilistic in themselves, and what makes an objective probability statement true is not the observed or expected frequency pattern, but the physical features of the experimental set-up that give rise to this pattern (in a probabilistic way). Now, it would be quite natural to *define* objective probability by saying simply this: the objective probability of an event is that very feature, whatever it may be, of the experimental set-up that brings about the characteristic probabilistic or frequency patterns. What Popper and others say concerning propensities *may* be understood along these lines, which would turn the propensity conception into a very general, but also nearly empty account of objective probability. Understood in this way, it is not subject to the above-mentioned criticisms, but it is also uninformative: it says nothing more than that probability statements refer to such properties of experimental arrangements that can play the role objective probabilities are supposed to play. But *what* can play this role, what features of experimental set-ups can rightly be called "objective probabilities"? We do not yet have an interpretation of probability, but something more abstract. Given this reading, the propensity conception is, as it were, merely a promissory note that has yet to be cashed out. The range interpretation can be seen as obtaining this; it makes a proposal what those features of random experiments are. It is a reasonable conjecture, put forward by Strevens (2003, ch. 5), that it is precisely such constellations in which probabilistic patterns emerge and in which we can ascribe objective probabilities to certain events. Other proposals are made by the different propensity conceptions (this term now understood in the specific sense in which it was discussed at the beginning), notably by single-case theories. Whereas the range approach takes those features of experimental arrangements which can be identified as probabilities to be complex high-level properties, single-case propensity theories take them to be simple, irreducible and fundamental.

K. As is clear by now, the "ranges" we deal with here are physical or natural ranges. All examples in which the initial-state space can actually be surveyed come from physics, but in principle the approach is suited to all empirical sciences, a point argued for already by von Kries and again by Michael Strevens. What one needs to apply the range approach is a distinction between initial conditions and laws. Moreover, the initial conditions have to be viewed as forming a real vector space: as will become even clearer in the next section, we are in the realm of real analysis here. It may be possible to generalize the approach to more abstract topological (vector) spaces, but a notion of continuity is indispensable. I emphasize this, because von Kries's ideas received a rather one-sided reception. They were taken up by John Maynard Keynes as well as by the Vienna Circle, notably Friedrich Waismann and Ludwig Wittgenstein, who both used the term "Spielraum", but gave it a distinctly logical meaning.[5] Since then, the term "range", when it occurs in philosophical discussions of probability, usually refers to logical probabilities. But these face the same problems as classical ones, and are supposed to be *a priori* and uniquely singled out by requirements of rationality. As is clear by now from the heroic attempts of Rudolf Carnap, there is nothing of this kind to be had. One gives away the potential of von Kries's ideas if one develops them only in the direction of a logical concept of probability. To avoid misunderstandings, (RC) could be dubbed "physical-range conception", if this were not too narrow. Again, what is essential are laws of nature and continuous initial conditions, but not necessarily laws of physics. Therefore, the term "natural-range conception" seems to be most appropriate.

4 The method of arbitrary functions

Continuous functions can be characterized as functions that are approximately constant on any sufficiently small interval. It is enough for continuity if this holds for equilateral intervals. Hence, in the ideal limiting case of the range interpretation we have

[5] See von Wright (1982) and Heidelberger (2001).

$$\int_{S_a} \delta(\boldsymbol{x})\,\mathrm{d}\,\boldsymbol{x} = p$$

for any continuous density function $\delta : \boldsymbol{S} \rightarrow \mathbf{R}$ from the initial-state space $\boldsymbol{S}$ into the real numbers **R**. In reality, at best we approach the ideal case, and therefore the equation is neither exactly true nor does it hold for certain eccentric density functions. Two types of (continuous) densities are obviously apt to create problems: on the one hand, density functions that are periodic with a period exactly the size of the "patches" the initial-state space is divided into, and on the other hand, density functions that put most of the weight on a single patch (a patch being a measurable connected set of maximum size consisting of initial states that all lead to the same outcome). These densities fluctuate either with a high frequency, or else very strongly, on at least one relatively small area of the initial-state space. In order to exclude them, we have to restrict the set of admissible densities to those of appropriately bounded variation. Taking this into account, we can restate the range interpretation of probability thus:

(AF) Let ***E*** be a random experiment and *a* a possible outcome of ***E***. Let $\boldsymbol{S} \subseteq \mathbf{R}^n$ for some *n* be the initial-state space attached to ***E***, and $\boldsymbol{S}_a \subseteq \boldsymbol{S}$ be the set of those initial states that lead to outcome *a*. If there is a number *p* such that for any continuous density function $\delta : \boldsymbol{S} \rightarrow \mathbf{R}$ whose variation is appropriately bounded on small intervals of ***S***, we have $\int_{S_a} \delta(\boldsymbol{x})\,\mathrm{d}\,\boldsymbol{x} \approx p$, then there is an objective probability of *a* upon a trial of ***E***, and its value is *p*.

The bounded-variation-condition would have to be stated more precisely as follows: There must be positive real-valued constants *c*, *k* such that for any continuous density function the variation of which is less *c* on any interval of a size less than *k*, the approximate equation holds. In this way densities are excluded that fluctuate either with a very high frequency or very abruptly and strongly. Let us call such densities “irregular” or “eccentric”, the others “regular”, “ordinary”, or “well-behaved”. The central question for the range approach, to be discussed in the next section, concerns the justification of this constraint. Before addressing it, I will continue the list of remarks.

L. The approach to probabilities via continuous density functions is traditionally called “method of arbitrary functions”. A sketch of it can be found in von Kries (1886, ch. 3). It is, generally speaking, worth pointing out that von Kries addressed almost any aspect of the range conception, including, for example, its application to statistical physics (ch. 8). One can say with little exaggeration that the very first treatment of the approach was also essentially complete concerning the main ideas, but much less concerning their elaboration. Von Kries’s style is rather informal, even more than the one of the present paper, which is partly due to the fact that the requisite mathematical tools were not yet developed in his times. Henri Poincaré, in *Calcul des Probabilités* (1896) as well as in *Science and Hypothesis* (1902), was the first to apply the method of arbitrary functions rigorously, to the so-called wheel of fortune. This wheel consists of a disc which is divided into small segments of equal size whose colours alternate between red and black, and a spinning pointer which is pushed and eventually comes to rest on either a red or a black segment. It is a particularly simple example, because, provided that the pointer starts in a fixed initial position, its final position depends on its initial velocity only, and so the initial-state space has only one dimension. Poincaré shows that any ordinary distribution over initial velocities leads to a probability of approximately ½ for the outcomes “red” and “black”, respectively. Since then, the wheel of fortune has served as the standard example to introduce the method of arbitrary functions.

M. Later on, more demanding examples were treated by Eberhard Hopf (1934, 1936), who also gave the first systematic account of the method. The difficulty is, not just to make a more or less plausible *claim* to the effect that with a certain arrangement the assumptions of the range interpretation or the method of arbitrary functions are fulfilled and that, therefore, probabilities that are estimated, or known in advance from other sources, can be understood accordingly, but to *show* that those assumptions are fulfilled and to *derive* the probabilities mathematically by actually constructing and investigating the initial-state space. This proves too difficult even in comparatively simple examples, like the Galton-board, but Hopf treated some non-trivial examples at least, in which the probabilities are not known in advance from symmetry considerations. In general, the range conception is meant to provide *truth* conditions for probability statements, but it is not well suited to supply *assertibility* conditions. In most cases, one has to infer the probabilities from symmetries or relative frequencies instead of directly assessing the initial-state space of a random experiment.

N. Von Plato (1994, ch. 5) gives an historical overview of the method of arbitrary functions. A modern treatment of the mathematics is Engel (1992). Strevens (2003) is by far the most comprehensive contemporary investigation of the range approach. Remarkable about his treatment is that he sticks to a realistic modelling insofar as he avoids the transition to limiting cases that is typical for mathematically-oriented treatises. He sketches applications of the method to statistical physics (ch. 4.8) and population ecology (ch. 4.9) and investigates the presuppositions under which it could work in the social sciences (ch. 5.4). According to him, there is good reason to view probability as a complex high-level phenomenon that occurs if and only if initial-state spaces are structured in the indicated way. Consequently, he is inclined to reject the “simple probabilities” of the single-case propensity theory (ch. 5.6).

O. I have little to add to these comprehensive and thorough studies, except concerning the question of the interpretation of the emerging probabilities. While the early writers like Poincaré and Hopf adhered to a frequency theory of probability and accordingly understood the density functions as idealized representations of actual frequencies, Strevens deliberately remains neutral in view of interpretational questions (Strevens 2003, ch. 1.32, 1.33, 2.14, 2.3). He is interested in the "physics", not the "metaphysics", of probability. According to him, you may interpret the density functions in any way you like, whereupon your interpretation will simply carry over to the resulting outcome probabilities. The situation is: probabilities in – probabilities out; the density functions over the initial-state space represent probability distributions. In general, one can say that the phenomenon caught by the phrase "method of arbitrary functions" or "range approach", if discussed at all by writers on probability, was taken to be an *explanation* for the existence of probabilistic patterns, but not as an *interpretation* of the concept of probability. In contrast to that, I would like to take seriously the idea that the range approach provides an independent proposal for an interpretation of probability statements, in the sense that they are made true or false by the structure of the corresponding initial state spaces.

5 The main objection to the range interpretation

The range conception identifies probabilities with proportions of outcomes in suitably structured initial-state spaces. This is, furthermore, not just meant to be a contingent numerical identity, but a proposal for an objective interpretation of probability. What I am going to discuss now under the heading "main objection to the range interpretation" has two parts. The first does not yet challenge the numerical identification, but just denies that it is suitable for an interpretation of probability statements. The second calls into doubt the numerical identification, which is, of course, a minimal requirement without which you cannot even start to consider the range proposal. The objection goes like this:

> “As with the classical approach, we do not get an interpretation of probability, because the concept is in fact presupposed. The range approach has to assume that initial states in the same small equilateral interval are approximately “equally possible”, i.e., it must assume an approximately uniform probability distribution over any such interval. Otherwise, there would be no reason to identify the proportion with which a certain outcome is represented within an arbitrary (but not too small) such interval with the probability of that outcome. The point is even more obvious with the method-of-arbitrary-functions-formulation: What are those density functions supposed to be? Clearly, they represent probability distributions over initial states, and so the question of the interpretation of probability is merely shifted back, from probabilities of outcomes to probabilities of initial states.
>
> Furthermore, it is not even possible to claim that the outcome probabilities are *numerically* identical to the respective proportions in the initial-state space. The probabilities of the outcomes depend not only on this space, but also on the actual density function over it. An initial-state space plus a probability density on it fix outcome probabilities, the space by itself fixes nothing. There are, of course, cases in which all ordinary density functions lead to roughly the same outcome probabilities, and in which only quite eccentric densities would yield different ones. But this does not mean that in these cases you can fix the probabilities without taking into account densities over the initial-state space, at least implicitly. For the actual density might well be an eccentric one, in which case you would get the probabilities wrong.

> Therefore, it is not even possible to numerically identify outcome probabilities with proportions in initial-state spaces. They are not always identical, and if they are, this is contingently so."

This criticism is undoubtedly strong. But it underestimates the fact that the range approach refers precisely to situations in which the initial-state space is structured in a way that guarantees independence of the outcome probabilities from any particular density function. Well, not quite – we have to confine ourselves to not-too-eccentric densities. But the independence is very far-reaching. We need not assume a uniform distribution over the initial-state space, i.e. treat the initial states as equally possible cases, or something like that. By considering only continuous density functions with appropriately bounded variation, we treat adjacent small regions of the initial state space as approximately equally possible, but in fact this does not seem to be a severe restriction. The values of the probabilities we get out of the range approach would only be wrong if there were such a thing as the true probability distribution over the initial-state space and if this distribution were eccentric. In that case we would not dismiss the range approach, however, but rather conclude that we had overlooked some nomologically relevant factor, which means either that we had gotten the initial-state space wrong, or that this space is not primordial.

This observation is the key to answering the objection. The laws of nature determine the result, given the initial conditions, but they leave open what those initial conditions are. As *they* do not care for the initial conditions, and *we* can't control them sufficiently, it can only be by accident if on repeated trials of ***E*** an eccentric distribution over ***S*** emerges. This would not give us a reason to change our expectations concerning the possible outcomes, i.e., we would maintain our judgments concerning the objective probabilities. But *if*, on the contrary, we somehow convince ourselves that there is something behind the observed eccentric distribution, i.e. that it can be relied on for future predictions, we would conclude that there must be some nomological factor we have overlooked: the laws of nature actually *do* care, contrary to what we thought. This reaction shows, I suppose, that we are quite ready to analyse probability according to the range interpretation. We are not prepared to accept "just so" a falling-apart of objective probabilities and the respective proportions in the associated initial-state space. We would look for an explanation, which would then change our modelling of the experimental situation and thereby enable us to uphold the identification.

Furthermore, if we are able to maintain the *numerical* identity between objective probabilities and proportions of outcomes in initial-state spaces, we can analyse the concept of probability accordingly, without circularity, because in (RC) there is no talk about equally possible cases, nor about densities or distributions over initial-state spaces. The probabilities depend only on the space itself, after all. This discussion bears heavily on the distinction between laws of nature and initial conditions, or, in the wording of von Kries, between the "nomological" and the "ontological" (von Kries 1886, ch. 4.4). As in remark H above, I cannot do full justice to this topic here, but it seems to be clear that we tend to make the distinction in a way that supports the range conception.

Another reply to the main objection, put forward by Strevens (2003, ch. 2.53), considers perturbations of density functions. An eccentric density over the initial-state space would give us different probabilities, to be sure, but any perturbation of this density that is not very slight is liable to change the probabilities again. The probabilities that are due to an eccentric density are not resilient, their values depend on the fact that this very density, and no other, is the true one. In contrast to this, the probabilities given by the standard densities are insensitive to changes in the density: this is precisely the upshot of the method of arbitrary functions. Now, it is very doubtful if there ever is such a thing as "the true density" on the initial-state space. Why should nature, in its choice of initial states, follow some particular density function, after all (provided that we took all relevant laws of nature into account in modelling the initial-state space)? Thus we can say either that the objective outcome probabilities are given by the regular densities, which all lead to the same probabilities, or else that there are no objective probabilities at all. This confirms the foregoing point that if we could somehow rely on a certain eccentric density to be the true one, this would have to be due to some neglected nomological factor that guarantees the stability of this density. Without such a stabilizing factor the probabilities the density gives rise to could not properly be called "objective".

These answers to the objection are not yet completely satisfying, as shown by a second, closely related criticism. In considering the proportions with which the various possible outcomes are represented in the initial-state space, we assume a standard measure on this space. To be sure, any continuous transformation of this measure with appropriately bounded variation would lead to the same probabilities, but why not choose a devious or even completely gerrymandered measure? Take again the throwing of a die. As we have seen, if we were able to construct the initial-state space and to survey it in detail, we would find it consisting of very small regions, a kind of patches, each patch uniformly leading to a particular result. Now, in view of these patches, we might choose a measure that blows up just those patches which lead to the outcome "six", at the expense of all other patches. If one metricizes the initial-state space in such a non-standard way, there may well be a fixed proportion of "six" in any (not-too-) small equilateral interval, but this proportion would now exceed the true probability of "six". That probability has not changed, of course. The random experiment is the same as before, we merely represented the initial-state space attached to it in an unusual way, i.e. chose a strange measure. On what grounds can such a representation be excluded?

The connection with the foregoing discussion is that an appropriately distorted initial-state space, with the “six”-patches inflated and the others shrunk in relation to them, amounts to the same as the choice of the standard measure plus a density that is suitably periodic with a period of the size of the patches. Now, I said that *if* we should get the impression that nature’s choice of initial conditions could reliably be modelled by such a density function, we would conclude that we have overlooked some nomologically relevant factor and the initial-state space is not primordial. But couldn’t we base our considerations on the suitably distorted space as well? And if we did, wouldn’t we then be forced to dismiss all *ordinary* density functions as eccentric, because any such function has troughs over the “six”-patches of the distorted space? Wouldn’t we then, by an argument analogous to the foregoing one, be forced to conclude that we must have overlooked some nomologically relevant factor if we came to the conclusion that nature followed *such* a density in selecting initial conditions? This would be a *reductio* of the reasoning given above. In this scenario, *any* density can appear to be regular or irregular, to be ordinary or eccentric, depending on the measure on the initial-state space. Therefore, if we rule out certain densities as irregular or eccentric, this can only be in relation to the standard measure, or to a large class of well-behaved transformations of the standard measure, but not to just any old measure one happens to choose. Consequently, it has to be assumed that standard ways of measuring the distance between vectors of physical quantities are distinguished in an objective sense. Nature has to single out the measure (or at least a large class of suitably well-behaved measures that are equivalent with respect to the relevant proportions) together with the initial-state space, if we are to get objective probabilities out of the range approach. It is then only in relation to this “natural” measure (or class of measures) that a density counts as eccentric and thus as giving reason to re-model the physical situation by looking for a further, so far neglected nomological factor.

A distorting measure of the kind sketched above gives rise to a very unnatural representation of the possible initial states. The different components of the initial-state vector do not represent the usual physical magnitudes any longer, but related ones that are the result of complicated and entangled transformations of the former. (The components of the initial-state vector are not distorted one by one, but all together.) The new magnitudes we get in this way play no role in physics, and with good reason. For example, the usual conservation laws do not hold for them. Similar observations can be made for even more gerrymandered transformations of the initial-state space, if, e.g., the space is torn into pieces that would then be rearranged so as to yield different probabilities of the outcomes, or no probabilities at all.[6] The standard physical magnitudes would thereby be transformed in a not only complicated and entangled but even discontinuous way. In general, it should be possible to rule out as unnatural the magnitudes that are the result of such transformations, and with them the corresponding distance measures for initial states. This is sketchy and programmatic, of course. We cannot, however, be satisfied with just saying that if we represent the initial-state space in standard "reasonable" ways, we always get the same proportions of the possible outcomes. The laws of nature do not only have to tell us what the outcome of the experiment is, given certain initial conditions, but also which distance measure, or class of equivalent measures, to choose for initial states.[7] To put it the other way round: as far as the choice of the measure is a matter of convention, as far as the initial-state space has no "naturally-built-in" measure, or class of measures that are equivalent regarding the proportions of the outcomes, the probabilities we get out of the range approach are merely conventional as well.

[6] Strevens speaks of "racked" and "hacked" initial condition variables, respectively (Strevens 2003, ch. 2.53).

[7] See, however, Guttmann 1999, ch. 4.8, for a sceptical attitude concerning this program in the context of statistical mechanics.

That such an assumption is necessary can also be seen with Strevens's perturbation argument. The problem here is that a certain regular measure for the space of possible densities over the initial-state space, i.e. a distance measure for densities, is tacitly assumed. Otherwise we could not say that *small* perturbations of a density are likely to change the probabilities if the density is eccentric, but not if it is ordinary. It should be possible in principle to select a measure for densities that is itself eccentric in such a way so as to render the foregoing statement wrong. Again, we have to assume "natural" distance measures for densities to be able to assess the effects of "small" perturbations of densities.

To resume: The main objection to the range conception can be answered if and only if standard ways of measuring the distance between initial-state vectors are naturally distinguished. The probabilities we get out of the range conception are objective precisely to the extent in which this is the case. But in view of the serious difficulties of propensity and frequency theories, the range approach seems to be a viable option for an objective interpretation of probability.[8]

[8] Note of Acknowledgment: A shorter version of this paper was presented at the Munich Conference on Time, Chance, and Reduction (8–10 March 2006). I am very grateful to Gerhard Ernst and Andreas Hüttemann for giving me the opportunity, and to the conference participants for helpful discussion. Much of the paper was written about a year before during a stay at the University of Konstanz Center for Junior Research Fellows. Here, I am very grateful to Luc Bovens and Stephan Hartmann for inviting me, to the Alexander-von-Humboldt-Stiftung for providing a three-months scholarship, and to the members of the Philosophy, Probability, and Modeling Group for providing several instructive criticisms. Last but not least I am grateful to Christopher von Bülow for improving my English.

References

Eagle, Antony (2004): "Twenty-one arguments against propensity analyses of probability"; in: *Erkenntnis* 60, 371–416.

Engel, Eduardo (1992): *A Road to Randomness in Physical Systems*; Berlin: Springer.

Guttmann, Yair (1999): *The Concept of Probability in Statistical Physics*; Cambridge: Cambridge University Press.

Hájek, Alan (1997). "Mises redux" – redux: fifteen arguments against finite frequentism. *Erkenntnis*, **45**, 209–227.

Hájek, Alan (2009). Fifteen arguments against hypothetical frequentism. *Erkenntnis*, **70**, 211–235.

Heidelberger, Michael (2001): "Origins of the logical theory of probability: von Kries, Wittgenstein, Waismann"; in: *International Studies in the Philosophy of Science* 15, 177–188.

Hopf, Eberhard (1934): "On causality, statistics and probability"; in: *Journal of Mathematics and Physics* 13, 51–102.

Hopf, Eberhard (1936): "Über die Bedeutung der willkürlichen Funktionen für die Wahrscheinlichkeitstheorie"; in: *Jahresbericht der Deutschen Mathematiker-Vereinigung* 46, 179–195.

Poincaré, Henri (1896): *Calcul des Probabilités*; Paris.

Poincaré, Henri (1902): *La Science et l'Hypothèse*; Paris.

Strevens, Michael (2003): *Bigger than Chaos*; Cambridge (Mass.): Harvard University Press.

Von Kries, Johannes (1886): *Die Principien der Wahrscheinlichkeitsrechnung*; Tübingen: Mohr Siebeck. (2nd unchanged printing 1927)

Von Plato, Jan (1994): *Creating Modern Probability*; Cambridge: Cambridge University Press.

Von Wright, Georg Henrik (1982): “Wittgenstein on probability”; in: G.H. von Wright, *Wittgenstein*, Oxford: Basil Blackwell, ch. 5.