

Field that initiates equation numbering:

DRAFT

Chapter from a book, *The Material Theory of Induction*, now in preparation.

Infinite Lottery Machines

John D. Norton

Department of History and Philosophy of Science

University of Pittsburgh

<http://www.pitt.edu/~jdnorton>

1. Introduction

No single calculus of inductive inference can serve universally. There is even no guarantee that the inductive inferences warranted locally, in some domain, will be regular enough to admit the abstractions that form a calculus. However, in many important cases, when the background facts there warrant it, inductive inferences can be governed by a calculus. By far the most familiar case is the probability calculus.

That many alternative calculi other than the probability calculus are possible is easy to see. Norton (2010) identifies a large class of what are there called “deductively definable” logics of induction. Generating a calculus in the class is easy. It requires little more than picking a function from infinitely many choices.

The harder part is to see whether some specific calculus is warranted in some particular domain. This and the following chapters will provide a few illustrations of unfamiliar cases. In them, the warranted calculus is not the probability calculus. The systems to be investigated are: in this chapter, infinite lottery machines; and, in subsequent chapters, continuum-sized outcome sets, which include nonmeasurable outcomes; indeterministic physical systems; and the quantum spin of electrons.

The focus of this chapter, a fair infinite lottery machine, selects among a countable infinity of outcomes, 1, 2, 3, ... without favor. It allows us to pose a series of inductive problems. In this arrangement, how much support inductively is given to the outcome of some particular number, say 378? Or to some finite set of numbers, say all those between 37 to 256? Or to some infinite set of numbers, such as the even numbers or the prime numbers? The answers to these questions will be supplied by the inductive logic applicable to this domain.

The warranting facts that pick out the logic will be the physical properties of the infinite lottery machine. The inductive logic will be the same for all properly functioning infinite lottery machines. Thus the pertinent warranting facts will be just those that they have in common. That is the fact that they choose a number without favoring any.

The example of the infinite lottery machine has already proven troublesome. We shall see in Section 2 that an unreflective application of the probability calculus to it fails. The literature has explored several ways of modifying the calculus to accommodate the infinite lottery. They include dropping countable additivity and introducing infinitesimal probabilities. In subsequent sections, I will argue that neither of these modifications succeeds. The defining characteristic of the infinite lottery is that it chooses its outcomes without favoring any one. That characteristic is captured formally in the condition of “label independence” of Section 3. It says that the chance of an outcome with some definite number or a set of them is unaffected if we permute the numbers that label the outcomes. This condition, it is argued in Sections 4 and 5, is incompatible with the (finite) additivity of a probability measure. This additivity is the familiar property that, if we have two mutually exclusive outcomes, then we can add their probabilities to find the probability of their disjunction. Thus the chance properties of an infinite lottery machine cannot be represented by a probability measure. Attempts to continue to do so, it is argued in Section 6, amount to altering the background facts presumed. These attempts do not solve the problem but merely exchange it for a different problem that can be solved with a probability measure. Section 7 explores a non-standard calculus that is warranted by specific configurations of an infinite lottery machine. Section 8 outlines how we can give intuitive meaning to the values in the non-standard calculus and use it to make predictions. Section 9 defends the failure of what is there identified as the “containment principle.” Section 10 reports briefly on work elsewhere on the unexpected complications found when we try to determine the extent to which an infinite lottery machine is physically possible. Section 11 concludes.

Finally, the Appendix reviews the so-called “measure problem” of eternal inflation in modern cosmology. It turns out to be essentially the same as the difficulty of fitting an additive probability measure to an infinite lottery machine.

2. The Initial Difficulty

The infinite lottery machine entered the literature because it poses an immediate problem if we wish to use the probability calculus as the applicable inductive logic. That problem arises from a tension between two conditions. First, the machine chooses each number without favor. So each outcome n must have equal probability $P(n)$:

$$\varepsilon = P(1) = P(2) = \dots = P(n) = \dots \quad (1)$$

Second, the outcomes are mutually exclusive and at least one must happen. Hence all these probabilities must sum to unity in the infinite sum:

$$P(1) + P(2) + \dots + P(n) + \dots = 1 \quad (2)$$

No value of ε can satisfy both (1) and (2). For if we choose some $\varepsilon > 0$, no matter how close this ε to zero, then (2) is the summing of infinitely many non-zero ε 's. Summing only finitely many will eventually exceed the unity required in (2). If, instead, we set $\varepsilon = 0$, then (2) is the summing of infinitely many zeroes, which is zero.

Two types of solutions have been proposed in the literature. The most popular, advocated by Bruno de Finetti (1972; §5.17), targets the fact that (2) requires the summing of an infinity of probabilities. This infinite sum operation is qualitatively different from merely summing finitely many probabilities. For the infinite summation is carried out in two steps. First, one sums finitely many terms, up to some large number N , say:

$$S(N) = P(1) + P(2) + \dots + P(N)$$

One then takes the limit of $S(N)$ as N grows infinitely large. De Finetti proposed that we discard this rule of “countable additivity”¹ and employ only the first step, “finite additivity,” in which we are allowed to add only finitely many probabilities. The outcome is that we no longer require

¹ The full condition of countable additivity applies to any infinite set of mutually incompatible outcomes $\{A_1, A_2, \dots, A_n, \dots\}$ and asserts that $P(A_1 \text{ or } A_2 \text{ or } \dots) = P(A_1) + P(A_2) + \dots$, where the ellipses “...” indicate that the formulae continue for all n .

summation condition (2) for the infinite lottery machine; and we can now employ $\varepsilon = 0$ in (1), without running into contradictions. De Finetti's proposal has been subject to extensive critical scrutiny. See, for example, Bartha (2004), Blackwell and Diaconis (1996), Kadane, Schervish, and Seidenfeld (1986), Kadane and O'Hagan (1995) and Williamson (1999).

Setting $\varepsilon = 0$ amounts to setting the probability of each individual number outcome (or any finite set of them) to zero. That seems too severe to some. Might we not manage by assigning a very, very tiny probability—an “infinitesimal” amount—to each outcome? Non-standard analysis provides a mathematically clean way of doing just this. The possibility has been explored by, for example, Benci, Horsten, and Wenmackers (2013) and Wenmackers and Horsten, (2013); and it has been subjected to critical scrutiny by, for example, Pruss (2014), Williamson (2007) and Weintraub (2008).

Neither of these repairs to probabilistic analysis will be pursued further here since, as I will now argue, no such repair is adequate. The infinite lottery requires an even greater departure from normal ideas of probability.

3. Label Independence

To proceed, we must clarify just what is meant by “choosing *without favor*” or, as it is sometime said, having a “fair” lottery. Taking this to mean that each outcome has equal probability is untenable since it presumes that the probabilistic treatment is adequate. We need an analysis that does not make this presumption. In the following, I shall speak of the “chance” of an outcome, where the term will no longer designate a probability. Just what it designates will be determined through the development of the inductive calculus that governs it, in the sections that follow.

What it is to choose without favoring any outcome can be specified through the requirement of “label independence.” The driving intuition is that, when outcomes are chosen *with favor*, then the chances will, in general, differ with different outcomes. Holding a ticket for the outcome labeled “37” may be preferable to, say, “18,” if the outcome labeled “37” is favored over the one labeled “18.” If, however, the choice is made *without favor*, then we should be indifferent to whether we have the outcome labeled “37,” “18” or any other label. Moreover, that indifference should remain no matter how the lottery machine operator switches the labels

around over the various outcomes. We should not care to which outcome our label “37” is attached, for none is favored.

The general requirement is that the chances are unaffected by any permutation of the labels. A permutation moves labels from outcomes to outcomes such that every outcome starts and ends with exactly one a label; no labels are discarded; and no new labels are introduced. More formally, the requirement is:

Label independence

All true statements pertinent to the chances of different outcomes remain true when the labels are arbitrarily permuted.

We can see how it works by taking the case of a finite randomizer, the roulette wheel. Such a wheel has, in the American case, 38 equally sized pockets on its perimeter. It is spun and a ball projected in the opposite direction. The pockets are numbered from 1 to 36, 0 and 00; and the outcome is the pocket in which the ball eventually comes to rest. As long as the wheel is well balanced with equal sized pockets and the croupier spins and projects with vigor, the ball will pass over the wheel many times and arrive with equal chance in each pocket. Under those conditions, the choice of labeling of the pockets is immaterial. We could, without compromising the fairness of the wheel, peel off the labels that mark each pocket and rearrange them in any way we please.

To apply label independence, we start with a statement true of a properly made roulette wheel:

Pockets 11 and 23 are the same size.

Under a permutation that switches label 11 with label 3 and label 23 with label 10, the proposition now asserts a truth expressed in the old labeling as:

Pockets 3 and 10 are the same size.

Proceeding with further permutations, we see that the label independence of the statement amounts to the assertion that any two pockets have the same size. Similarly the following is true of any well functioning roulette wheel:

The ball ends up in pockets 1 to 12,
roughly as often as it does in pockets 13 to 24.

Under label independence, it remains true if we permute the labels of pockets 13 to 24 with those of pockets 25 to 36. It now expresses a truth expressed in the old labeling as

The ball ends up in pockets 1 to 12,
roughly as often as it does in pockets 25 to 36.

Thus the label independence of the second statement reflects the fact that the relative frequency of outcomes in a set of pockets depends merely on the number of pockets in the set.

The qualification “pertinent to the chances” is essential, for there are many statements true of a roulette wheel whose truth is not preserved under arbitrary permutation of the pocket labels. For example, in an American wheel:

Pockets 3 and 4 are diametrically opposite on the wheel.

This statement does not remain true under most permutations of the pocket labels. However, since the statement is not pertinent to the randomizing function of the wheel, the failure does not violate label independence.

4. Abandoning Finite Additivity

There are no surprises when label independence is used to characterize how a finite randomizer, such as a roulette wheel, picks outcomes without favor. Matters change when label independence is applied to an infinite lottery machine. The reason is that labels on infinite sets of outcomes can be permuted in ways that are impossible for finite sets. It is easy to permute them so that the labels for some infinite set of outcomes end up assigned to one of its proper subset. It follows from label independence that the set and its proper subset have the same chance. If chances are probabilities, that means that they have the same probability. Assembling several permutations like this soon contradicts the requirement that the probability of an outcome is the sum of the probabilities of its disjoint parts. That is a striking result that bears being repeated. If outcome A is the disjunction of mutually exclusive outcomes B or C or D, that is,

$$A = (B \text{ or } C \text{ or } D),$$

and B, C and D pairwise contradict, then we can have cases in which

$$\text{Chance (A)} = \text{Chance (B)} = \text{Chance (C)} = \text{Chance (D)} \quad (3)$$

which is incompatible² with finite additivity,³ which requires

² Unless all the probabilities are zero.

$$P(A) = P(B) + P(C) + P(D) \quad (4)$$

That is, the label independence of an infinite lottery machine requires us to abandon finite additivity for a measure of the chance of sets of outcomes. Since finite additivity is essential to the definition of probability, it follows that chances cannot be probabilities for an infinite lottery machine.

5. An Example of the Failure of Finite Additivity

An illustration of the failure of finite additivity in (3) and (4) is provided by an example reported in Bartha (2004, §5) and Norton (2011, pp. 412-15). Assume that the chance function “Ch(.)” measures the chance of the different sets of outcomes of an infinite lottery machine, recalling that the notion of chance employed here, so far, is only loosely defined and need not be a probability measure. For some numbering of the outcomes, the labels on the sets of even numbered outcomes

$$even = \{2, 4, 6, 8, \dots\}$$

and on the sets of odd numbered outcomes

$$odd = \{1, 3, 5, 7, \dots\}$$

can be switched one-one by a permutation:

$$1 \leftrightarrow 2, 3 \leftrightarrow 4, 5 \leftrightarrow 6, 7 \leftrightarrow 8, \dots$$

Hence, by label independence, the two sets must have equal chance:

$$Ch(even) = Ch(odd) \quad (5)$$

Now consider the four sets of every fourth number.

$$one = \{1, 5, 9, 13, \dots\}$$

$$two = \{2, 6, 10, 14, \dots\}$$

$$three = \{3, 7, 11, 15, \dots\}$$

$$four = \{4, 8, 12, 16\}$$

By similar reasoning each of *one*, *two*, *three*, and *four* have equal chance:

³ The full condition of finite additivity applies to any finite set of mutually incompatible outcomes $\{A_1, A_2, \dots, A_n\}$ and asserts that $P(A_1 \text{ or } A_2 \text{ or } \dots \text{ or } A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$.

$$\text{Ch}(\text{one}) = \text{Ch}(\text{two}) = \text{Ch}(\text{three}) = \text{Ch}(\text{four}) \quad (6)$$

So far, nothing untoward has happened. All this is compatible with the $\text{Ch}(\cdot)$ function being a probability measure. This will now change.

Consider two sets of outcomes: *one* and the set whose members are in (*two* or *three* or *four*). Since all the sets are countably infinite, we can have the following two-part permutation of the labels. The first switches one to one the labels on *odd* with those on *one*:

$$1 \leftrightarrow 1, 3 \leftrightarrow 5, 5 \leftrightarrow 9, 7 \leftrightarrow 13, \dots$$

The second part switches one to one the labels on *even* with those of (*two* or *three* or *four*):

$$2 \leftrightarrow 2, 4 \leftrightarrow 3, 6 \leftrightarrow 4, 8 \leftrightarrow 6, 10 \leftrightarrow 7, 12 \leftrightarrow 8, 14 \leftrightarrow 10, 16 \leftrightarrow 11, \dots$$

For convenience, since the set *one* now carries the labels that originated in *odd*, let us also call it *odd**; and similarly (*two* or *three* or *four*) is also called *even**. That is, we have two names for each outcome set:

$$\text{one} = \text{odd}^* \quad (\text{two or three or four}) = \text{even}^*$$

Since the new labels of outcomes in *odd** and *even** can also be switched one-one with each other, analogously to (5), they must also have equal chance. That is:

$$\text{Ch}(\text{even}^*) = \text{Ch}(\text{odd}^*) \quad (7)$$

Combining we have

$$\begin{aligned} \text{Ch}(\text{two}) &= \text{Ch}(\text{three}) = \text{Ch}(\text{four}) && \text{[from (6)]} \\ &= \text{Ch}(\text{one}) && \text{[from (6)]} \\ &= \text{Ch}(\text{odd}^*) && \text{[since one and odd}^* \text{ name the same set]} \\ &= \text{Ch}(\text{even}^*) && \text{[from (7)]} \\ &= \text{Ch}(\text{two or three or four}) && \text{[since (two or three or four) and even}^* \\ &&& \text{name the same set]} \end{aligned}$$

These last equalities violate⁴ finite additivity (4), since a finitely additive probability measure $P(\cdot)$ must satisfy:

$$P(\text{two}) + P(\text{three}) + P(\text{four}) = P(\text{two or three or four})$$

⁴ Unless all the probabilities are zero.

6. Finite Additivity Must Go

The simple example shows that label independence for an infinite lottery is incompatible with the finite additivity of a probability measure. To proceed, at least one of them must be given up. Both Bartha (2005, §5) and Wenmackers and Horsten (2013, p. 41) find giving up finite additivity too great a sacrifice. In my view, we have no choice but to sacrifice finite additivity. For label independence is a defining characteristic of an infinite lottery machine. Without it, we can no longer say that the infinite lottery machine chooses its outcomes without favoring any. There is no comparable necessity for probability measures, other than our comfort and familiarity with them.

To persist in describing the chance properties of an infinite lottery machine by a probability measure is, in effect, to change the problem posed. For no single probability measure can satisfy all the equalities derived above from label independence. We must choose which subset will be satisfied. That choice amounts to adding extra conditions on the operation of the infinite lottery machine. While the augmented problem may be quite well-posed and even interesting, it is a different problem. The extra conditions must breach label independence, so that we no longer describe a device that chooses outcomes without favor. We have not solved the original problem, but merely changed it to a different problem we like better.

To see how this favoring can come about, consider the two equalities (5) and (7). If the chance function is a probability function $P(\cdot)$, then they become

$$P(\text{even}) = P(\text{odd}) = 1/2 \quad (5a)$$

$$P(\text{even}^*) = P(\text{odd}^*) = 1/2 \quad (7a)$$

We cannot uphold both if we note that the probabilistic version of (6) requires

$$P(\text{one}) = P(\text{two}) = P(\text{three}) = P(\text{four}) = 1/4 \quad (6a)$$

For then $P(\text{odd}^*) = P(\text{one}) = 1/4$; while $P(\text{even}^*) = P(\text{two}) + P(\text{three}) + P(\text{four}) = 3/4$, in contradiction with (7a).

To preserve the applicability of a probability measure, we have to block one of (5a) or (7a). A simple strategy is to select a preferred numbering of the outcomes, such as the original labeling, and then define the probability of each set of outcomes in the natural way. That is, we consider the sequence of finite, initial sets

$$\{1\}, \{1, 2\}, \{1, 2, 3\}, \dots, \{1, 2, 3, \dots, n\}, \dots \quad (8)$$

The probability of some nominated outcome set is defined as the limit of the frequency of outcome set members in this sequence. For the outcome *even*, we have

$$\begin{aligned} P(\text{even}) &= \lim_{n \rightarrow \infty} n/2n = 1/2 && n \text{ is even} \\ &= \lim_{n \rightarrow \infty} (n+1)/2n = 1/2 && n \text{ is odd} \end{aligned} \quad (9)$$

Definitions of the form (9) using the sequence (8) gives the expected probabilities (5a) and (6a) for $P(\text{even})$, $P(\text{odd})$, $P(\text{one})$, $P(\text{two})$, $P(\text{three})$ and $P(\text{four})$. However they fail to return (7a), since, as before, we have $P(\text{odd}^*) = P(\text{one}) = 1/4$ and $P(\text{even}^*) = P(\text{two or three or four}) = 3/4$.

There is a second, parallel “starred” analysis that preserves the equality of (7a) while giving up (5a). It proceeds exactly as above, but replaces the sequence (8) with one natural to the starred labeling of outcomes. That is, the starred labels assigned to outcomes after the permutation conform with

$$\begin{aligned} \text{odd}^* &= \{1^*, 3^*, 5^*, 7^*, \dots\} = \{1, 5, 9, 13, \dots\} \\ \text{even}^* &= \{2^*, 4^*, 6^*, 8^*, \dots\} = \{2, 3, 4, 6, 7, 8, 10, 11, 12, \dots\} \end{aligned}$$

In place of (8), it has the sequence:

$$\{1^*\} = \{1\}, \{1^*, 2^*\} = \{1, 2\}, \{1^*, 2^*, 3^*\} = \{1, 2, 5\}, \{1^*, 2^*, 3^*, 4^*\} = \{1, 2, 5, 3\}, \dots \quad (8a)$$

Using the sequence (8a), definitions of probability based on relative frequencies akin to (9), will give starred results that are the reverse of the unstarred results. That is, we shall secure (7a) $P(\text{even}^*) = P(\text{odd}^*) = 1/2$, but not (5a).

In comparing the unstarred and starred analysis, we see how each improperly favors certain outcomes in the judgment of the other. The unstarred analysis gives $P(\text{odd}^*) = 1/4$ and $P(\text{even}^*) = 3/4$, improperly favoring *even*^{*} over *odd*^{*}, according to a starred analysis. However the starred analysis gives $P(\text{odd}) = 1/4$ and $P(\text{even}) = 3/4$, improperly favoring *even* over *odd*, according to an unstarred analysis.

Thus describing an infinite lottery machine with a probability measure replaces the original requirement of selection without favor, by selection under by the added restriction that the selection must respect also a preferred numbering scheme and the limiting ratios native to it.

That some such change in the problem is required if probabilities are to be retained was noted by Edwin Jaynes. He was a leading proponent of objective Bayesianism and a master of the memorable riposte, which he formulated for this case as follows (2003, p.xxii).

Infinite-set paradoxing has become a morbid infection that is today spreading in a way that threatens the very life of probability theory, and it requires immediate

surgical removal. In our system, after this surgery, such paradoxes are avoided automatically; they cannot arise from correct application of our basic rules, because those rules admit only finite sets and infinite sets that arise as well-defined and well-behaved limits of finite sets. The paradoxing was caused by (1) jumping directly into an infinite set without specifying any limiting process to define its properties; and then (2) asking questions whose answers depend on how the limit was approached.

For example, the question: ‘What is the probability that an integer is even?’ can have any answer we please in $(0, 1)$, depending on what limiting process is used to define the ‘set of all integers’ (just as a conditionally convergent series can be made to converge to any number we please, depending on the order in which we arrange the terms).

In our view, an infinite set cannot be said to possess any ‘existence’ and mathematical properties at all – at least, in probability theory – until we have specified the limiting process that is to generate it from a finite set.

The bluster of Jaynes’ riposte cannot cover the fact that he can offer no good reason for eschewing infinite sets that do not come with a preferred ordering or numbering scheme. If we must eschew all such sets, then we are precluding from inductive analysis cases that arise in real science. The problems just rehearsed in Sections 5 and 6 above have played out almost exactly as a foundational problem in recent inflationary cosmology, the “measure problem,” where the lack of a preferred order on an infinite set of pocket universes has precluded introduction of a probability measure over them. The problem is reviewed in the Appendix. This should quell fears that that the problem of fitting a probability measure to an infinite lottery machine is merely the contrarian whimsy of eccentric theorists and idle philosophers. The problem has a connection and application in real science.

7. The Inductive Logic Warranted for an Infinite Lottery Machine

The defining characteristic of an infinite lottery machine is that its choice of outcomes respects label independence. That characteristic rules out an inductive logic whose strengths of support are probability measures. According to the material theory of induction, the background

facts warrant the inductive logic appropriate to the domain. Label independence, the characteristic common to all infinite lottery machines, is the key, warranting fact. It acts powerfully and leads us to the following inductive logic.

7.1 Equal Chance Sets

The logic divides outcomes sets into types such that all sets of the same type must have the same chance. To implement this division, we require that two outcomes sets are of the same type if the members of the two sets can be mapped one-one to one another by a permutation of labels. That means that the outcome sets must have the same size (i.e. cardinality). In addition, the complements of the sets must also be the same size, else the requisite permutation of labels will not be possible. What results are sets of outcomes of the following types:⁵

finite_n: a set with n members, where n is a natural number.

Examples of *finite₃* are {1, 2, 3}, {27, 1026, 5000} and {24, 589, 2001}.

infinite_{co-infinite}: an infinite set whose complement is also infinite.

An example is the infinite set of even numbers {2, 4, 6, ...} since its complement is the infinite set of odd numbers {1, 3, 5, ...}

infinite_{co-finite-n}: an infinite set whose complement is finite of size n.

An example of *infinite_{co-finite-10}* is the set of all numbers greater than 10: {11, 12, 13,...} since its complement is the finite set {1, 2, 3, ..., 10}.

7.2 Chance Values

The requirement of label independence entails that sets of outcomes of the same type must be assigned the same chance. Thus the chance function Ch(.) in this logic can only have the following set of values:

$$\text{Ch}(\textit{finite}_n) = V_n, \text{ where } n = 1, 2, 3, \dots \quad (10a)$$

$$\text{Ch}(\textit{infinite}_{\textit{co-infinite}}) = V_\infty = \text{“as likely as not.”} \quad (10b)$$

$$\text{Ch}(\textit{infinite}_{\textit{co-finite-n}}) = V_{-n}, \text{ where } n = 1, 2, 3, \dots \quad (10c)$$

⁵ Co-infinite means that the complement of the set is infinite. Co-finite means that the complement of the set is finite.

And for completeness we add in the two special cases

$$\text{Ch}(\text{empty-set}) = V_0 = \text{“certain not to happen”} \quad (10d)$$

$$\text{Ch}(\text{all-outcomes}) = V_{-0} = \text{“certain to happen”} \quad (10e)$$

According to (10a), all equal-sized finite sets of outcomes have the same chance: any n membered finite set has the same chance V_n . This is required by label independence since some permutation can always switch the labels between any two finite sets, as long as they are the same size. Similarly, (10b) tells us that all infinite sets that are co-infinite have the same chance. We have already seen an example above in (5) and (7):

$$\text{Ch}(\text{even}) = \text{Ch}(\text{odd}) = \text{Ch}(\text{even}^*) = \text{Ch}(\text{odd}^*) = V_\infty$$

Since each of the four infinite sets are co-infinite, there is a permutation that switches their labels. By label independence, they have the same chance. Since every co-infinite infinite set of outcomes is assigned the same value V_∞ as its complement set, we informally name this value “as likely as not.” Finally, (10c) can be interpreted similarly to (10a).

7.3 Comparing Chance Values

The conditions (10) are powerful restrictions. They preclude the chance function $\text{Ch}(\cdot)$ being an additive probability measure. However they leave the logic underspecified. We do not yet know whether the values V_n, V_∞, V_{-n} are the same or different; and, if they are different, how they compare with one another. To arrive at the conditions (10), we used label invariance only. Further restrictions can enrich the logic.

A qualitative ranking of the strengths of support derives from the idea that the chance of a set of outcomes cannot be diminished if we add further outcomes to the set. This condition induces the relation “ \leq ,” which is read as “is no stronger than.” It obtains between values A and B when the outcomes that realize a value A can be a subset of the outcomes that realize a value B . As a result, the relation inherits the properties of set theoretic inclusion. It is antisymmetric, reflexive and transitive. It is easy to see that:

$$V_0 \leq V_1 \leq V_2 \leq V_3 \leq \dots \leq V_\infty \leq \dots \leq V_{-3} \leq V_{-2} \leq V_{-1} \leq V_{-0} \quad (11)$$

One might think this condition unavoidable. It is not. It is merely familiar and amounts to one construal of the meaning of strength of support. A somewhat similar condition fails in the “specific conditioning logic” of Norton (2010, §11.2).

Further discriminations, if they happen at all, must be warranted by further background facts, whose truth must be recovered from the physical properties of the pertinent chance process. One case that is easy to motivate physically arises if we have an additive measure that is not normalizable. That is, the total measure of its space is infinite. It arises if we have a space in which lengths, areas or volumes are defined, the total space has infinite length, area or volume and the chances of some event occurring in a region of the space are measured by its length, area or volume. This case is developed more fully in the next chapter on “Uncountable Problems” in Section 4. An illustration recounted there derives from steady state cosmology. According to it, the chance of a hydrogen atom being created in some region of our cosmic infinite Euclidean space is proportional to the region’s volume.

To apply the infinite lottery logic this case, we divide the space into infinitely many parts of equal length, area or volume. An outcome $finite_n$ arises when the event is realized in some subset of the space of n of these parts. Its chance is measured by n . Correspondingly, the chance associated with any infinite volume of space will be measured by ∞ . That is, we have:

$$\text{Ch}(finite_n) = V_n = n \quad \text{where } n = 1, 2, 3, \dots \quad (12)$$

$$\text{Ch}(infinite_{co-infinite}) = V_\infty = \text{Ch}(infinite_{co-finite-n}) = V_{-n} = \infty$$

The inequalities relating the various values of V_n in (11) become strict inequalities.

$$V_0 < V_1 < V_2 < V_3 < \dots < V_\infty \quad (11a)$$

If the outcome of the infinite lottery machine lies in some finite set of outcomes, then the chance relations (12) match those of a finite probabilistic randomizer with the same finite set of outcomes. That is, the chances of different outcomes in the finite set will behave like probabilities defined as:

$$P(A|B) = \text{Ch}(A)/\text{Ch}(B) \quad (13)$$

where A is a subset of B and B is a finite set of outcomes.

The conditions (11a) and (13) are not assured. They can fail, depending on the particular physical instantiation of the infinite lottery machine. Such a failure would arise if the randomizer is based on the non-probabilistic, indeterministic systems described in Chapter 15 below. The conditions succeed for the “Spin of a pointer on a dial” device of Norton (2018).

Correspondingly, while label independence does not force it, we may require as an additional assumption in some more specific logic that:⁶

$$V_{\infty} < \dots < V_{-3} < V_{-2} < V_{-1} < V_{-0} \quad (11b)$$

In the following section, we shall see why this additional assumption fits naturally into the formal properties of the chance function.

These inequalities along with relations (10), (11), (12) and (13), all assumed henceforth, characterize an inductive logic native to an infinite lottery machine well enough for us to see that such logics differ significantly from a probabilistic logic.

A curious outcome of the analysis is that this logic is the reverse of the one de Finetti (1972; §5.17) proposed for an infinite lottery. In his logic, additivity was preserved for outcomes comprised of infinite sets; but it was trivialized for outcomes of finite sets, since these latter were all assigned zero probability. In the present logic, non-trivial additivity is maintained for finite sets through (12) and (13), but additivity fails through (10b) for most infinite sets.

8. Interpreting the Inductive Logic

The chance function $Ch(\cdot)$ of Section 7 specifies an inductive logic. Its formal properties are clear. However we may well ask what its quantities mean. What should we think when we learn that some outcome has such and such a chance value? This question is asking less than is usually asked, in the analogous circumstance, when we seek an interpretation of probability. It is not asking for an explicit definition, such as is sought by a relative frequency interpretation of probability or from the subjectivist Bayesian definition of probability in terms of betting quotients. One can have an understanding of a magnitude, adequate for practical applications, without an explicit definition of it. Since the values of the chance function (10) are so unfamiliar, that is all that is sought here.

⁶ Considerations of cardinality make natural the strict inequality $V_{\infty} < V_{-n}$ for all n . However, unlike the case of V_n , I have been unable to conceive possible background facts that would warrant strict inequalities among the individual values of V_{-n} as shown in (11b). Might an inventive reader be able to conceive such facts?

8.1 The Probabilistic Model

The problem of developing some informal understanding of an initially abstruse quantity arises also for ordinary probabilities. We can use its solution as a model for the new chance function. Take the simple case of a coin toss, whose outcomes can be heads H or tails T . How are we to understand the probability assertion that $P(H) = 0.5$? How are we to distinguish that probability assertion from nearby assertions like $P(H) = 0.4$ or $P(H) = 0.6$? To be told that a probability of 0.4 is weaker than a probability of 0.5 is true but merely qualitative and falls well short of the precision we expect.

We gain a better understanding of such assertions, sufficient to discriminate among them, by contriving associated circumstances of either very high or very low probability. For example:

If $P(H) = 0.5$, then, with probability near one, the frequency of H among many, independent coin tosses will be close to 0.5.

If $P(H) = 0.4$, then, with probability near one, the frequency of H among many, independent coin tosses will be close to 0.4.

Sentence like these, by themselves, are not sufficient to give informal meaning to the quantity $P(\cdot)$. All we have is one probability statement, that $P(H) = 0.5$, associated with another statement concerning an outcome with probability near one. Without something further, we will be trapped forever in a self-referential web of statements in which probabilistic assertions are made about other probabilistic assertions, without otherwise clarifying what any probabilistic assertion means. The axioms and definitions used to deduce all these assertions can be modeled in many systems with an extensive quantity whose magnitude is additive. To break out of the self-referring trap, we use a rule that coordinates large and small values of probability with informal judgments of expectation about chancy outcomes:

Rule of coordination for probability.

Very low probability outcomes generally do not happen; and very high probability outcomes generally do.

Thus we come to some understanding of the difference between $P(H) = 0.5$ and $P(H) = 0.4$: we expect each to deliver roughly 50% or 40% H respectively in repeated, independent coin tosses.

This interpretive rule, in various forms, has a long history and has come to be known as “Cournot’s Principle.”⁷ In his canonical treatment of the foundations of probability theory, Kolmogorov (1950, p. 4) has a version of this rule that employs the locution “practically certain”:

- (a) One can be practically certain that if the complex of conditions \mathfrak{S} [Fraktur capital S] is repeated a large number of times, n , then if m be the number of occurrences of event A , the ratio m/n will differ slightly from $P(A)$.
- (b) If $P(A)$ is very small, one can be practically certain that when conditions \mathfrak{S} are realized only once, the event A would not occur at all.

This process of conveying meaning should not be confused with subjective Bayesians’ process of elicitation of probabilities. They determine, for example, that a subject has assigned probability 0.5 to H when the subject accepts even odds on either H or T . The present concern is how the subject, prior to the elicitation, came to judge that 0.5 is the appropriate probability to assign. That in turn requires some prior understanding by the subject of what probability 0.5 means.

8.2 The Analogous Analysis for the Chance Function

This same strategy can be used both to interpret the values of the chance function (10) and, at the same time, to display the predictive powers of the logic. The analogs of very low probability and very high probability outcomes are those with chance V_n and chance V_{-n} . A chance V_n outcome is realized when the number drawn resides in a finite set among the infinitely many possibilities. This is not an outcome we should expect to happen since it is thoroughly swamped by the infinitely many numbers outside the set. A chance V_{-n} happens when the number drawn resides outside some finite set. Since there are infinitely many possibilities

⁷ For a brief survey, see Shafer (2008, §2). One must be careful to treat the rule as nothing more than an informal guide. Otherwise the danger is that one misidentifies very low probability events as strictly impossible and very high probability events as necessary. For de Finetti’s view of the rule, see de Finetti (1974, pp. 180-181). My use of the term “rule of coordination” is intended to recall Reichenbach’s notion of a coordinative principle.

outside the finite set that realize it, this is an outcome we should expect. That is, we have the interpretive rule:

Rule of coordination for chance.

Very low chance outcomes with chance V_n generally do not happen; and very high chance outcomes with chance V_{-n} generally do.

This rule divides outcomes sharply into three sets:

outcomes in one of the *finite*_n, which we do not expect;

outcomes in *infinite*_{co-infinite}, which may or may not happen “as likely as not”; and

outcomes in one of the *finite*_{co-finite-n}, which we do expect.

The application of this rule is simpler than in the probabilistic case for two reasons. First, in the present case, the division of outcomes into unexpected, intermediate and expected is sharp. This sharpness makes it natural to replace the inequalities of (11) by strict inequalities. In the probabilistic case, the division was muddier. Just how low should a probability be before its outcome is not to be expected? If one is pressed, one eventually introduces some arbitrary cutoff, knowing that any cutoff can be challenged if sufficient contrivance is allowed.

Second, the intermediate co-infinite infinite outcomes all are assigned the same chance values of V_∞ . The intermediate outcomes in the probabilistic case, however, are assigned a range of probabilities and further work is needed to distinguish them. For example, we separated the cases of probability 0.5 and 0.4 by considering a large number of independent trials. The comparable analysis is not needed for the chance function. However, as an exercise in applying the chance function, in Section 8.4 below, it is used to determine the chance of various frequencies of outcomes of even and odd numbers in many, independent drawings of an infinite fair lottery.

8.3 Applying the Rule of Coordination

To get a sense of how this rule is used, we can apply it to a simple case. Consider the chance that the number drawn is less than or equal to some large number N . This outcome set has N members and thus has chance V_N . It is an outcome not to be expected. The outcome that the number is greater than N , however, is in the complement set and thus has chance V_{-N} . It is an outcome we do expect. This must appear strange at first. For it tells us that no matter how large

we make N – one million, one quadrillion, one million^{million} – we are sure the number drawn is greater, even though we are certain that some definite, finite number is drawn. There is only strangeness here, but no problem. It is how the chances are in an infinite lottery. All our calculus does is to relate that fact to us.

To explore the application of this rule further and to see how the chance function behaves, consider the case of repeated, independent drawings from the infinite lottery. To be specific, consider 1000 independent drawings of the infinite lottery. The combined outcome of each set of 1000 trials is a 1000-tuple, such as

$$\langle 156, 27, 2398, \dots, 180 \rangle_{1000}$$

Its elements are the numbers drawn successively in independent drawing 1, drawing 2, etc. The subscript 1000 reminds us that there are 1000 elements in the tuple. The set of all such outcomes is Ω_{1000} . The outcome in which number 1 is drawn every time is

$$\mathbf{1}_{1000} = \langle 1, 1, 1, \dots, 1 \rangle_{1000}$$

All the remaining outcomes are just the complement set,⁸ $\Omega_{1000} - \mathbf{1}_{1000}$. Since a permutation of labels maps any tuple in this set to any other, we have from label independence that each tuple of outcomes in Ω_{1000} has an equal chance. Moreover, the set of tuples Ω_{1000} forms a countably infinite set. Thus its outcomes have the chance properties of an infinite lottery.

It follows that we can now apply the chance logic (10). The outcome $\mathbf{1}_{1000}$ has a single member and is of type *finite*₁. The complement set of outcomes that somewhere differ from all-1 tickets, $\Omega_{1000} - \mathbf{1}_{1000}$, is of type *infinite*_{co-finite-1}. Thus:

$$\text{Ch}(\mathbf{1}_{1000}) = V_1 \quad \text{Ch}(\Omega_{1000} - \mathbf{1}_{1000}) = V_{-1}$$

Applying the rule of coordination, we infer that an outcome in which all numbers drawn in 1000 independent repetitions are 1 is not to be expected. Correspondingly an outcome in which at least one of the numbers drawn is not 1 is to be expected.

The example is easy to extend. We might ask whether we should expect all the 1000 drawings to yield the same number, where the same number is found in some finite set, say $\{1, 2,$

⁸ That is, the set of tuples Ω_{1000} excluding the tuple $\mathbf{1}_{1000}$.

3}. That is, the outcome is ($\mathbf{1}_{1000}$ or $\mathbf{2}_{1000}$ or $\mathbf{3}_{1000}$). Proceeding as above, we find this outcome is not to be expected, since

$$\text{Ch}(\mathbf{1}_{1000} \text{ or } \mathbf{2}_{1000} \text{ or } \mathbf{3}_{1000}) = V_3.$$

We get a different result if we ask after the outcome in which all the numbers drawn are the same, but that number can be any in an infinite set, such as the set of all natural numbers. This set of outcomes, ($\mathbf{1}_{1000}$ or $\mathbf{2}_{1000}$ or $\mathbf{3}_{1000}$ or ...), is of type *infinite*_{co-infinite} for which

$$\text{Ch}(\mathbf{1}_{1000} \text{ or } \mathbf{2}_{1000} \text{ or } \mathbf{3}_{1000} \text{ or } \dots) = V_\infty.$$

That is, the outcome of all 1000 numbers drawn being the same is “as likely as not.”

Compare this to another outcome. What is the chance that all the numbers drawn in the 1000 repetitions are less than or equal to some number N , where the numbers need not be the same? This outcome corresponds to N^{1000} tuples in the outcome set Ω_{1000} . Thus we have

$$\text{Ch}(\text{all numbers less than or equal to } N) = V_{N^{1000}}.$$

That is, since N^{1000} is finite, the outcome is one that will generally not happen according to the rule of coordination.

What is the chance that at least one of the numbers in 1000 independent drawings is greater than N ? This outcome set is the complement of the last set considered with N^{1000} members. Thus this outcome set is co-finite infinite and we have

$$\text{Ch}(\text{at least one number greater than } N) = V_{\neg N^{1000}}.$$

That is, the outcome is to be expected according to the rule of coordination.

These last examples are surely surprising to someone whose intuitions about chance have been tutored by the probability calculus. We assign greater chance to all 1000 numbers being the same than we do to all 1000 numbers being any combination of numbers less than or equal to N . This holds no matter how large we make N .

Here are more examples illustrating outcomes to which the “as likely as not” chance of V_∞ is assigned. Consider the numbers drawn in 1000 independent repetitions of the infinite lottery:

all-even: all numbers drawn are even numbers

all-odd: all numbers drawn are odd numbers

all-powers: all numbers drawn are powers of 10,

that is, $10, 10^2, 10^3, 10^4, \dots$

not-all-powers: all numbers drawn are NOT powers of 10,
that is, not and of 10, 10^2 , 10^3 , 10^4 , ...

Each of these outcomes corresponds to sets of tuples in Ω_{1000} of type *infinite*_{co-infinite}. It follows that they have equal chance:

$$\text{Ch}(\textit{all-even}) = \text{Ch}(\textit{all-odd}) = \text{Ch}(\textit{all-powers}) = \text{Ch}(\textit{not-all-powers}) = V_\infty$$

This will seem surprising if we think that there are vastly fewer outcomes in *all-powers* than in *not-all-powers*, since there are vastly fewer powers of ten than numbers that are not powers of ten. Any surprise should be eradicated by recalling that both these sets are countably infinite. The impression that one is bigger than the other is purely an artifact of labeling. Label independence warns us that such artifacts of labeling should be ignored. The two sets are equinumerous and equinumerous in their complements; and thus they have equal chance.

8.4 Relative Frequencies of “as likely as not” Outcomes

In independent repetitions of a finite lottery, we expect with high probability, that roughly half the numbers drawn will be even and half of them odd. That is a consequence of the probabilistic fact that an even number is drawn with probability 1/2. We should not expect similar results in an infinite lottery, for the value V_∞ assigned to both even and odd outcomes is quite removed in its formal properties from a probability 1/2. We cannot expect a high chance to be assigned to outcomes with roughly equal numbers of even and odd numbers in independent repetitions and a low chance to other proportions of even and odd. For otherwise the chance function would be behaving like an additive probability measure.

We can affirm by direct calculation that the chance function of the infinite lottery does not return such favoring of certain relative frequencies. Consider N independent drawings of the lottery. We compute the chances of there being exactly n even outcomes among the N numbers drawn as follows.

The outcome set consist of N -tuples comprising Ω_N . Consider some particular sequence of outcomes, where there are exactly n even numbers amongst all N outcomes and we have fixed the particular order in which *odd* and *even* appears:

$$\langle \textit{odd}, \textit{odd}, \dots, \textit{even}, \textit{odd}, \textit{even}, \textit{even} \rangle_N$$

Since each of *odd* and *even* are realized by infinitely many numbers, the set of N -tuples realizing this particular sequence is infinite. Correspondingly there are infinitely many ways that the complement set of N tuples could be realized. Thus the outcome is co-infinite infinite and it has chance V_∞ .

There are many ways that the n even numbers could be ordered among the N numbers drawn. The number of ways is given by the combinatorial factor $C(N,n) = N!/(n! (N-n)!)$. All that matters for our purposes is that this combinatorial factor is always finite for finite N and n . It follows that there are still infinitely many N -tuples that realize the outcome of n even numbers in any order amongst the N drawings; and also infinitely N -tuples in the complement set. Hence the outcome has the chance value V_∞ . Since n can have any value from 0 to N , we have:

$$\text{Ch}(0 \text{ even in } N) = \text{Ch}(1 \text{ even in } N) = \text{Ch}(2 \text{ even in } N) = \dots = \text{Ch}(N \text{ even in } N) = V_\infty$$

That is, in N independent drawings, no matter how large N , all relative frequencies of even numbers amongst them have equal chance. The chance function does not in any way favor relative frequencies of around one half; or any other frequency.

The chance value described by “as likely as not” would, in the probabilistic context, have to designate a probability of one half; and we would expect roughly one half successes in repeated independent trials. There is nothing as specific in the chance properties of the infinite lottery. An outcome set that is “as likely as not” is merely one whose chance is equal to its complement outcome set.

9. Failure of the Containment Principle

This infinite lottery logic will likely be discomfoting for someone whose intuitions are guided by probability theory. One source of discomfort may be that the removal of elements from an outcome set commonly does not reduce the chances of the outcome. It would seem natural that the set of even numbered outcomes $\{2, 4, 6, 8, \dots\}$ must be assigned greater chance than the set of every fourth numbered outcome $\{4, 8, 12, 16, \dots\}$. This second set is properly contained in the first. However the present logic assigns the same chance to both. We might express the intuition more clearly as:

The containment principle. If a set of outcomes A is properly contained in a set of outcomes B , then the chance of A is strictly less than the chance of B :
 $\text{Ch}(A) < \text{Ch}(B)$.

If the background facts support it, there is no problem with a logic that conforms with this principle. However the principle cannot lay claim to a preferred status. As is always the case, whether a logic has some feature is decided by prevailing background facts. The background fact of label independence entails the failure of the containment principle.

Two further considerations reduce the appeal of the principle:

First, the containment principle has not been uniformly respected in familiar probabilistic applications. There is a probability zero of a dart hitting any particular point on a dartboard of continuum many points. The same zero probability is assigned to the dart hitting any of a countable infinity of points on the dartboard, even if that set contains the single point originally considered. In another example, we follow de Finetti's prescription for the infinite lottery and employ a probability measure that is only finitely additive. Then the probability of drawing a one is the same the probability of drawing any number less one hundred million. Both are zero probability outcomes.

Second, the containment principle by itself is insufficient to induce chances that can compare all sets of outcomes. Since the set of even numbered outcomes is disjoint from the set of odd multiples of three $\{3, 9, 15, 21, 27, \dots\}$, we are left unable to compare their chances. In such cases, we may be inclined to retain the chance assignments of the present logic: if disjoint outcome sets (and their complements) are equinumerous, then they are assigned the same chance. What results, however, is a non-transitive comparison relation for chances. We have from considerations of equinumerosity that:

$$\text{Ch}(\{2, 4, 6, 8, \dots\}) = \text{Ch}(\{3, 9, 15, 21, 27, \dots\})$$

$$\text{Ch}(\{4, 8, 12, 16, \dots\}) = \text{Ch}(\{3, 9, 15, 21, 27, \dots\})$$

If transitivity of the comparison relation for chances is supposed, it follows that:

$$\text{Ch}(\{4, 8, 12, 16, \dots\}) = \text{Ch}(\{2, 4, 6, 8, \dots\}).$$

This equality contradicts the containment principle, which tells us that:

$$\text{Ch}(\{4, 8, 12, 16, \dots\}) < \text{Ch}(\{2, 4, 6, 8, \dots\}).$$

If transitivity is dropped, we will be unable to assign a single value to each chance, but only assign pairwise comparisons of strength. Presumably some accommodation of the two approaches can be found eventually, but it may not be pretty or simple.

In sum, we should use the containment principle when the background facts call for it. When they do not call for it, we should feel no special loss at its failure.

10. Is An Infinite Lottery Machine Physically Possible?

The discussion so far has presumed the physical possibility of an infinite lottery machine. In what sense are they physically possible? Elsewhere (Norton, 2018; Norton and Pruss, 2018, Norton, manuscript a) I have pursued the question in greater detail. The answer proves to be more complicated and much more interesting than one might first imagine.

The natural starting point is to seek some design that employs ordinary probabilistic randomizers, such as coin tosses, die throws and pointers spun on dials. We run into difficulties immediately. We will need infinite powers of discrimination to distinguish among the infinitely many possible pointer outcomes crammed onto the scale etched onto the surface of the dial. If we use coins or dice, we will need to use infinitely many of them to create an outcome space big enough to hold the countable infinity of outcomes of the infinite lottery machine.

If we are undaunted by the task of flipping infinitely many coins or reading pointer positions with infinite precision, the prospects for an infinite lottery machine seem good. Infinitely many coin tosses produce an outcome space of continuum size, that is, an order of infinity higher than that needed for the countably infinite outcomes of the infinite lottery machine. Somewhere in it we would expect to find a countable infinity of outcomes that implement an infinite lottery machine.

However in Norton (2018), as corrected by Norton and Pruss (2018), we found a maddening problem. With some ingenuity, we can use ordinary probabilistic randomizers to form infinite lottery machines. However in every design we could imagine, there was always a probability of zero that the machine would operate successfully. The persistence and recalcitrance of the failure gave the clue that the problem was not merely one of an impoverished imagination for the design of the infinite lottery machines. There was some unidentified matter of principle defeating all attempts.

In Norton (manuscript a) that matter of principle is recovered from what I would otherwise have imagined to be the arcana of measure theory and axiomatic set theory. The probabilistic randomizers will provide us with an outcome space expansive enough to host the infinite lottery outcomes that encode results “1,” “2,” “3,” and so on. If a probability is defined for each of these outcomes, then that probability must be the same for each and can only be zero. For otherwise, if it is greater than zero, we need only sum finitely many of the equal, non-zero probabilities $P(1), P(2), P(3), \dots$ to arrive at a sum greater than one. That sum contradicts the normalization of the probability measure to unity. If, however, we set each of the probabilities $P(1), P(2), P(3), \dots$ to zero, then the probability that any one of the infinite lottery outcomes, 1, 2, 3, ..., arises is zero. For it is given by the sum

$$P(1) + P(2) + P(3) + \dots = 0 + 0 + 0 + \dots = 0$$

That means that the infinite lottery machine operates successfully only with probability zero.

The escape is to use infinite lottery outcomes to encode results “1,” “2,” “3,” ... that are probabilistically nonmeasurable. Norton (manuscript a) describes two designs that do this. The same difficulty besets both. Their designs presume the existence of the nonmeasurable outcome sets, but do not specify which those sets are. That means that, after the randomizers settle into some end state, we cannot know the outcome set to which they belong. The number selected as the infinite lottery outcome is inaccessible to the user, rendering the device useless.

It turns out that, as far as we know, this failure must always happen. For all known examples of nonmeasurable sets are nonconstructive and we have some reason to expect that none can be constructed. That means that we are allowed to assume their existence, commonly by virtue of the axiom of choice of axiomatic set theory, or something equivalent to it.⁹ However there is no explicit description for which they are. We are caught in a dilemma. If an infinite lottery machine based on ordinary probabilistic randomizers is to return a result we can read, it will do so successfully only with probability zero. If we demand a probability of success greater than zero, then we can have it, but the result of the infinite lottery machine will be inaccessible to us.

These results apply only to infinite lottery machines constructed from ordinary probabilistic randomizers. They do not preclude other designs. Norton (2018, manuscript a)

⁹ For more on nonmeasurable sets and the axiom of choice, see Chapter 14.

describes designs based on quantum mechanical systems. In the simplest, one takes a quantum particle in a definite momentum state. It consists of a wave uniformly distributed over space in the direction of the momentum. Divide that space into a countable infinity of intervals of the same size, numbered 1, 2, 3, If we now perform a measurement on the position of the particle, it will manifest with equal chances in each interval. An infinite lottery machine has been implemented.

While the exercise of designing these infinite lottery machines is entertaining, I take a more permissive view of them. For hundreds of years, the paradigm of a probabilistic system in probability theory was the coin toss, die throw and card shuffle. Yet prior to quantum theory, our best science told us that none of these was a true randomizer. Probability theory thrived merely by supposing that these real randomizers were imperfect surrogates for true but unrealizable probabilistic randomizers: idealized ideal coin tosses, die throws and card shuffles. We can, I propose, take the same attitude to infinite lottery machines. They are an idealized case that can be added to our repertoire of idealized randomizers. We can and should ask what inductive logic is adapted them.

Finally, we should separate the issue of the cogency of the design of an infinite lottery machine from the cogency of the infinite lottery logic described in this chapter. We may not be able to specify explicitly which are the infinite lottery outcomes of a probabilistically based machine. But, on the authority of the axiom of choice, they exist. So we can ask what chance each has of being realized; and we should expect a suitable logic of induction to tell us.

11. Conclusion

The infinite lottery remains one of the most popular arguments used to establish that the countable additivity of a probability measure must be reduced to mere finite additivity. What this chapter shows is that the implications of the infinite lottery are still stronger. It requires also that we abandon finite additivity. The existing literature has been reluctant to accept this further conclusion for it requires abandoning probabilities as the gauge of the possibility of the various outcomes. However, as I argued in Section 6, to persist in the use of a finitely additive probability measure for this purpose is to change the problem posed by adding further conditions,

such as a preferred numbering of the outcomes. The original infinite lottery problem is solved by a non-additive logic such as developed in Sections 7 and 8.

The new chance logic of these Sections will seem strange to those already steeped in probabilistic thinking. The strangeness is merely a result of its unfamiliarity. It is easy to lose sight of how abstruse is even the notion of probability. It was once unfamiliar to all of us. Imagine trying to convey to someone new to it that there is a probability of 0.5 that their unborn child will be a girl. We may eventually convey the idea by saying:

“What is the probability of a girl?”

It is the same as getting heads on a fair coin toss.”

This formulation uses a physical randomizer as a benchmarking device.

Now consider the cosmologists described in the appendix. They consider the infinitely many *like* and *unlike* patches spawned by eternal inflation. They find the chance properties of the patches to conform with label independence; and they find themselves confused by the resulting chance behavior. We should be able to use the same benchmarking strategy to clarify these chance properties for them:

“What is the chance of a *like* patch?”

It is the same as the chance of an even number in an infinite, fair lottery.”

Appendix: The “Measure Problem” in Eternal Inflation¹⁰

A1 Inflation and Eternal Inflation

Inflation in cosmology is a brief period of very rapid expansion in the very early universe. It has the same effect as taking a wrinkled rubber sheet and stretching it to an enormous size. The wrinkles are all but eliminated. This smoothing process motivated in large part the introduction of inflation into cosmological theory in the 1980s. The smoothing would explain why the cosmic matter distribution is so uniform on the largest scale and why the geometry of space is so close to flat. It also explains why, contrary to expectations of exotic particle theories, we see no magnetic monopoles. The inflationary stretching of space exiles them to parts of the cosmos we cannot see.

¹⁰ For a fuller discussion of the measure problem and its inductive analysis, see Norton (manuscript).

Under continuing criticism, the status of inflation in modern cosmology remains mixed. It was unclear that there ever was a pressing need to explain these features of the cosmos through further theory. The matter driving inflation was initially supposed to come from novel particle physics: a “GUT,” that is, a grand unified theory. Those efforts failed. The driving matter is now just a novel matter field, the inflaton, posited *ad hoc* with just the right properties. Moreover, the search for a viable form of inflation has led to multiple versions, so that it is not so much a single theory as a program of research.

Nonetheless, the notion has proven quite appealing and it has become a staple, if debated, topic in cosmology. The strongest argument for it comes from its treatment of quantum fluctuations. During inflation, tiny, evanescent quantum fluctuations are amplified to cosmic scales, where they are “frozen in” as classical perturbations in matter density that match the nonuniformities we observe now.

The original idea was that there would be an early period of inflation, driven by the exotic matter of the inflaton field. This rapid expansion would cease and be followed by a more slowly expanding state, driven by familiar forms matter and radiation. Eternal inflation is a variation in which this cessation of inflation never happens universally. Rather it happens in patches, with each patch reverting to a modestly expanding universe with ordinary matter. Each is a pocket universe or little island universe. Outside these patches, inflation continues. Since inflating space grows so much faster than the space of the patches, the universe overall persists eternally in an inflating state, continuously spawning non-inflating pocket universes. One of these pocket universes is our observable universe.

A2 The Measure problem: Should we be here?

The immediate question asked of eternal inflation is whether we should expect a spawned pocket universe to be like our observable universe. It would count against eternal inflation if a universe like ours is exceptional among the non-inflating universes spawned. The measure problem is the problem of finding a way to quantify how much we should expect patches like ours.

The difficulty can be seen in a simplified version of the problem in which we introduce a binary classification: pocket universes *like* ours versus pocket universes *unlike* ours. We gauge the extent to which a universe like ours will come about in eternal inflation by asking after the

distribution of *like* and *unlike* over the pocket universes. It is natural to ask for the probabilities of each. That query leads to trouble.

Alan Guth introduced inflation to cosmology in the early 1980s. Here is his development of the problem (2007, p. 11):

However, as soon as one attempts to define probabilities in an eternally inflating spacetime, one discovers ambiguities. The problem is that the sample space is infinite, in that an eternally inflating universe produces an infinite number of pocket universes. The fraction of universes with any particular property is therefore equal to infinity divided by infinity—a meaningless ratio. To obtain a well-defined answer, one needs to invoke some method of regularization.

Since there is a countable infinity of these pocket universes, we can see the similarity to the infinite lottery problem. It is like asking after the distribution of *even* and *odd* tickets in the lottery. Guth continues the above remarks by making the connection:

To understand the nature of the problem, it is useful to think about the integers as a model system with an infinite number of entities. We can ask, for example, what fraction of the integers are odd. Most people would presumably say that the answer is $1/2$, since the integers alternate between odd and even. That is, if the string of integers is truncated after the N th, then the fraction of odd integers in the string is exactly $1/2$ if N is even, and is $(N + 1)/2N$ if N is odd. In any case, the fraction approaches $1/2$ as N approaches infinity.

However, the ambiguity of the answer can be seen if one imagines other orderings for the integers. One could, if one wished, order the integers as

$$1, 3, 2, 5, 7, 4, 9, 11, 6, \dots, \tag{11}$$

always writing two odd integers followed by one even integer. This series includes each integer exactly once, just like the usual sequence $(1, 2, 3, 4, \dots)$. The integers are just arranged in an unusual order. However, if we truncate the sequence shown in Eq. (11) after the N th entry, and then take the limit $N \rightarrow \infty$, we would conclude that $2/3$ of the integers are odd. Thus, we find that the definition of probability on an infinite set requires some method of truncation, and that the answer can depend nontrivially on the method that is used.

Guth correctly recognizes that recovering a well-defined probability requires us to add something. He calls it “regularization” and it corresponds to imposing an order on the set of outcomes quite analogous to that used in Section 6 above. The difficulty, of course, is that there are multiple choices for the ordering and each typically leads to a different probability measure.

In including regularization in the set up of the problem, Guth presumes more than is needed to arrive at it. The same problem is generated in Section 5 above merely by matching one-to-one infinite sets of the same cardinality. Paul Steinhardt is also one of the founding figures of inflationary cosmology and now one of its sternest critics. He sets up the problem using cardinality considerations alone (2001, p. 42):

In an eternally inflating universe, an infinite number of islands will have properties like the ones we observe, but an infinite number will not. The true outcome of inflation was best summarized by Guth: “In an eternally inflating universe, anything that can happen will happen; in fact, it will happen an infinite number of times.”

So is our universe the exception or the rule? In an infinite collection of islands, it is hard to tell. As an analogy, suppose you have a sack containing a known finite number of quarters and pennies. If you reach in and pick a coin randomly, you can make a firm prediction about which coin you are most likely to choose. If the sack contains an infinite number of quarter and pennies, though, you cannot. To try to assess the probabilities, you sort the coins into piles. You start by putting one quarter into the pile, then one penny, then a second quarter, then a second penny, and so on. This procedure gives you the impression that there is an equal number of each denomination. But then you try a different system, first piling 10 quarters, then one penny, then 10 quarters, then another penny, and so on. Now you have the impression that there are 10 quarters for every penny.

Which method of counting out the coins is right? The answer is neither. For an infinite collection of coins, there are an infinite number of ways of sorting that produce an infinite range of probabilities. So there is no legitimate way to judge which coin is more likely. By the same reasoning, there is no way to judge which kind of island is more likely in an eternally inflating universe.

A3 No Probabilities—No Predictions

Guth seems optimistic that there will be a solution to the measure problem. Steinhardt is pessimistic and uses his pessimism as grounds for criticizing inflationary theory. However they agree that securing probabilities is essential to eternal inflation as a predictive theory. Guth (2007, p. 11) writes: “To extract predictions from the theory, we must therefore learn to distinguish the probable from the improbable.” Steinhardt (2011, p. 42) is more forthright in his concern:

Now you should be disturbed. What does it mean to say that inflation makes certain predictions—that, for example, the universe is uniform or has scale-invariant fluctuations—if anything that can happen will happen an infinite number of times? And if the theory does not make testable predictions, how can cosmologists claim that the theory agrees with observations, as they routinely do? He then reviews with disdain the idea of imposing a measure on the islands (pp. 42-43):
An alternative strategy supposes that islands like our observable universe are the most likely outcome of inflation. Proponents of this approach impose a so-called measure, a specific rule for weighting which kinds of islands are most likely—
analogous to declaring that we must take three quarters for every five pennies when drawing coins from our sack. The notion of a measure, an ad hoc addition, is an open admission that in inflationary theory on its own does not explain or predict anything.

Guth and Steinhardt share an all or nothing view: if probabilities cannot be secured, then the theory has failed as an instrument of prediction. This view is based on a widely accepted but false presumption: that the only precise way to deal with uncertainties is through probabilities. A major goal of this entire work is to show that this presumption is too severe and too narrow. We can still deal formally with uncertainty when probabilities are inapplicable. The background facts may merely warrant an inductive logic that is not probabilistic. In this case, the inductive logic warranted is summarized in the chance function (10).

We should separate the question of whether there is an inductive logic native to the situation from the question of whether we can secure the sorts of prediction we might like. In the case of eternal inflation, there is a well-defined inductive logic applicable. However it turns out not to support the sorts of predictions the cosmologists seek. The difficulty is that the inductive

logic assigns the same chance V_∞ to any universe in which there are infinitely many *like* pocket universes and infinitely many *unlike* pocket universes. Since this combination encompasses virtually all the possibilities that can be realized,¹¹ the logic is unable to discriminate among them usefully, that is, in a way that might privilege *like* universes.

Some prediction is still possible. The chance function (10) has predictive powers, as shown in Section 8 above. They may be weaker than the predictive powers of a full probability measure. But that is all that the specification of the infinite lottery permits.

More generally, we cannot demand that the universe gives us theories of the type that we happen to like. We may prefer theories of indeterministic processes always to be endowed with probabilities, for they enable strong predictions. However the world is under no obligation to provide such theories. Probabilities are not provided by the indeterministic systems described in a later chapter; and the theories are correspondingly weak in predictions. That fact does not make them failures as theories. They just happen to be the best the world will give us.

References

- Bartha, Paul (2004) "Countable Additivity and the de Finetti Lottery," *The British Journal for the Philosophy of Science*, **55**, pp. 301-321.
- Benci, Vieri, Horsten, Leon and Wenmackers, Sylvia (2013) "Non-Archimedean Probability," *Milan Journal of Mathematics*, **81**, pp. 121-51.
- Blackwell, David and Diaconis, Persi (1996) "A Non-Measurable Tail Set," pp. 1-5 in T. S. Ferguson, L. S. Shapley and J. B. McQueen, eds., *Statistics, Probability and Game Theory: Papers in Honor of David Blackwell*. Hayward, CA: Institute of Mathematical Statistics.
- de Finetti, Bruno (1974) *Theory of Probability: A Critical Introductory Treatment*. Vol. 1. Chichester: John Wiley & Sons.
- de Finetti, Bruno (1972) *Probability, Induction and Statistics*. London: John Wiley & Sons.

¹¹ There is an uncountable infinity of possible distributions of *like* and *unlike* over the countable infinity of pocket universes. The case in the main text occupies all of them excepting a countable infinity of exceptions that arise in universes finitely many *like* pocket universes, or in universes with finitely many *unlike* universes.

- Guth, Alan (2007) "Eternal Inflation and its Implications," <https://arxiv.org/abs/hep-th/0702178v1>
- Jaynes, Edwin T. (2003) *Probability Theory: The Logic of Science*. Cambridge: Cambridge University Press.
- Kadane, Joseph B., Schervish, Mark J. and Seidenfeld, Teddy (1986) "Statistical Implications of Finitely Additive Probability" in *Bayesian Inference and Decision Techniques*. Amsterdam: Elsevier Science Publishers, pp. 59-76.
- Kolmogorov, Andrey N. (1950) *Foundations of the Theory of Probability*. Trans. N. Morrison. New York: Chelsea.
- Norton, John D. (2010) "Deductively Definable Logics of Induction." *Journal of Philosophical Logic*, 39 (2010), pp. 617-654.
- Norton, John D. (2011) "Challenges to Bayesian Confirmation Theory," *Philosophy of Statistics, Vol. 7: Handbook of the Philosophy of Science*. Prasanta S. Bandyopadhyay and Malcolm R. Forster (eds.) Elsevier.
- Norton, John D. (2018) "How to Build an Infinite Lottery Machine," *European Journal for Philosophy of Science*. **8**, pp. 71-95.
- Norton, John D. and Pruss, Alexander R. "Correction to John D. Norton 'How to Build an Infinite Lottery Machine,'" *European Journal for Philosophy of Science*. **8**, pp. 143-44.
- Norton, John D. (manuscript) "Eternal Inflation: When Probabilities Fail," Prepared for special edition "Reasoning in Physics," *Synthese*, eds. Ben Eva and Stephan Hartmann. <http://www.pitt.edu/~jdnorton/homepage/cv.html> or <http://philsci-archive.pitt.edu/14401/>
- Norton, John D. (manuscript a) "How NOT to Build an Infinite Lottery Machine."
- Pruss, Alexander R. (2014) "Infinitesimals are Too Small for Countably Infinite Fair Lotteries," *Synthese*, **19**, pp. 1051-57.
- Shafer, Glenn (2008) "The Game-Theoretic Framework for Probability," Ch. 1 (pp. 3-15) in B. Bouchon-Meunier, C. Marsala, M. Rifqi and R. R. Yager, *Uncertainty and Intelligent Information Systems*. Singapore: World Scientific.
- Steinhardt, Paul J. (2011) "The Inflation Debate: Is the Theory as the Heart of Modern Cosmology Deeply Flawed?" *Scientific American*, April 2011, pp. 36-43.
- Weintraub, Ruth (2008) "How Probable Is an Infinite Sequence of Heads? A Reply to Williamson," *Analysis*, **68.3**, pp. 247-50.

Wenmackers, Sylvia and Horsten, Leon and (2013) "Fair Infinite Lotteries," *Synthese*, **190**, pp. 37-61.

Williamson, Timothy (2007) "How Probable is an Infinite Sequence of Heads," *Analysis*, **67.3**, pp.173-80