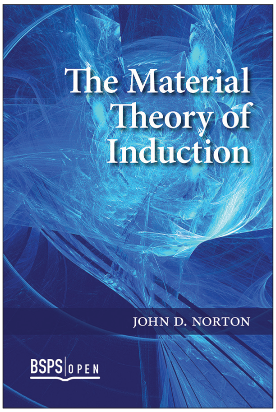

THE MATERIAL THEORY OF INDUCTION

by John D. Norton

ISBN 978-1-77385-254-6

Acknowledgement: *We acknowledge the wording around open access used by Australian publisher,* **re.press**, *and thank them for giving us permission to adapt their wording to our policy* http://www.re-press.org

13

Infinite Lottery Machines

13.1. Introduction

No single calculus of inductive inference can serve universally. There is not even any 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 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 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 chapter and the next three will provide a few illustrations of unfamiliar cases. In these cases, 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 infinite lottery machine that is the focus of this chapter 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 these domains.

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 in the existing literature. We shall see in Section 13.2 that an unreflective application of the probability calculus to it fails. The literature has explored several ways of modifying the calculus to accommodate an 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 an infinite lottery is that it chooses its outcomes without favoring any. This characteristic is captured formally in the condition of "label independence" described in Section 13.3. 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 13.4 and 13.5, 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 do so, I argue in Section 13.6, amount to altering the background facts presumed. These attempts do not solve the problem but merely exchange the problem for a different one that can be solved with a probability measure. Section 13.7 explores a non-standard calculus that is warranted by specific configurations of an infinite lottery machine. Section 13.8 outlines how we can give intuitive meaning to the values in the non-standard calculus and use it to make predictions. Section 13.9 extends the logic to repeated independent drawings of the lottery. Section 13.10 uses the extension to show that the chances of frequencies of outcomes in these repeated drawings do not conform with probabilistic expectations so that frequencies cannot be used to reintroduce probabilities. Section 13.11 defends the failure of what is identified as the "containment principle." Section 13.12

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 13.13 offers some concluding discussion.

Finally, Appendix 13.A 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.

13.2. The Initial Difficulty

An 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. This problem arises from a tension between two conditions. First, the machine chooses each number without favor. So each outcome n must have equal probability $P(n)$:

$$\varepsilon = P(1) = P(2) = \ldots = P(n) = \ldots. \tag{1}$$

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

$$P(1) + P(2) + \ldots + P(n) + \ldots = 1. \tag{2}$$

No value of ε can satisfy both (1) and (2). For if we choose some $\varepsilon > 0$, no matter how close this ε is to zero, then (26) 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 zeros, 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

$$S(N) = P(1) + P(2) + \ldots + P(N).$$

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 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.[2]

Setting $\varepsilon = 0$ amounts to setting the probability of each individual number outcome (or any finite set of them) to zero. This seems too severe to some. Might we not manage by assigning a 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, for example, by 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 the repairs to probabilistic analysis will be pursued further here, since, as I will below argue, no such repair is adequate. The infinite lottery requires an even greater departure from normal ideas of probability.

13.3. Label Independence

To proceed, we must clarify just what is meant by "choosing *without favor*," or, as it is sometimes said, having a "fair" lottery. Taking this to mean that each outcome has equal probability is untenable, since this 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. What it designates will be determined through the development of the inductive calculus that governs it.

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

2 See, for example, Bartha (2004), Blackwell and Diaconis (1996), Kadane, Schervish, and Seidenfeld (1986), Kadane and O'Hagan (1995) and Williamson (1999).

What it is to choose without favor can be specified through the requirement of "label independence." The driving intuition is that when outcomes are chosen *with* favor, then the chances will generally 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, this 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 label; no labels are discarded; and no new labels are introduced. More formally, the requirement is the following:

> *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, thirty-eight 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 with pass over the wheel many times and arrive with equal chance in each pocket. Under those conditions, the choice of labeling 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.

13.4. 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 subsets. It follows from label independence that the set and its proper subset have the same chance. If chances are probabilities, it 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. This is a striking result that bears repeating. 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

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

which is incompatible[3] with finite additivity,[4] which requires

$$P(A) = P(B) + P(C) + P(D). \tag{4}$$

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.

13.5. 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 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

3 Unless all the probabilities are zero.

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

probability measure. For some numbering of the outcomes, the labels on the sets of even-numbered outcomes[5]

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

and the labels on the sets of odd-numbered outcomes

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

can be switched one-to-one by a permutation:

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

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

$$\mathrm{Ch}(even) = \mathrm{Ch}(odd) \tag{5}$$

Now, consider the four sets of every fourth number:

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

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

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

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

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

$$\mathrm{Ch}(one) = \mathrm{Ch}(two) = \mathrm{Ch}(three) = \mathrm{Ch}(four). \tag{6}$$

So far, nothing untoward has happened. All of this is compatible with the Ch(.) function being a probability measure. This will now change.

Consider two sets of outcomes: set *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 the labels one-to-one on *odd* with those on *one*:

$$1 \leftrightarrow 1, 3 \leftrightarrow 5, 5 \leftrightarrow 9, 7 \leftrightarrow 13, \ldots.$$

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

5 Here and henceforth I move without warning between a set representation of an outcome, *even* = {2, 4, 6, …} and an equivalent propositional representation, *even* = 2 or 4 or 6 or ….

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

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:

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

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

$$\mathrm{Ch}(even^*) = \mathrm{Ch}(odd^*) \tag{7}$$

Combining this, we have

$$\begin{aligned}
\mathrm{Ch}(two) = \mathrm{Ch}(three) = \mathrm{Ch}(four) \quad & \text{[from (30)]} \\
= \mathrm{Ch}(one) \quad & \text{[from (30)]} \\
= \mathrm{Ch}(odd^*) \quad & \text{[since } one \text{ and } odd^* \text{ name the same set]} \\
= \mathrm{Ch}(even^*) \quad & \text{[from (31)]} \\
= \mathrm{Ch}(two \text{ or } three \text{ or } four) \quad & \text{[since (} two \text{ or } three \text{ or } four \text{) and } even^* \\
& \text{name the same set]}
\end{aligned}$$

These last equalities violate[6] finite additivity (4), since a finitely additive probability measure *P*(.) must satisfy,

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

13.6. Finite Additivity Must Go

The simple example above 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

6 Unless all the probabilities are zero.

lottery machine. Without it, we can no longer say that the infinite lottery machine chooses its outcomes without favor. 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. This 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; we have merely changed the problem to one 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 $P(.)$, then they become

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

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

We cannot uphold both if we note that the probabilistic version of (30) requires

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

For then $P(odd^*) = P(one) = 1/4$; while $P(even^*) = P(two) + P(three) + P(four) = 3/4$, 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

$$\{1\}, \{1, 2\}, \{1, 2, 3\}, \ \ldots, \{1, 2, 3, \ldots, n\} \ , \ldots. \tag{8}$$

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

$$P(even) = \mathrm{Lim}_{n\to\infty}\, n/2n = 1/2 \qquad n \text{ is even}$$

$$= \mathrm{Lim}_{n\to\infty}\, (n+1)/2n = 1/2 \qquad n \text{ is odd.} \qquad (9)$$

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

$$odd^* = \{1^*, 3^*, 5^*, 7^*, \ldots\} = \{1, 5, 9, 13, \ldots\}$$

$$even^* = \{2^*, 4^*, 6^*, 8^*, \ldots\} = \{2, 3, 4, 6, 7, 8, 10, 11, 12, \ldots.\}.$$

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

$$\{1^*\} = \{1\}, \{1^*, 2^*\} = \{1, 2\}, \{1^*, 2^*, 3^*\} = \{1, 2, 5\},\ \{1^*, 2^*, 3^*, 4^*\} = \{1, 2, 5, 3\}, \ldots. \qquad (8a)$$

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*(*even**) = *P*(*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*(*odd**) = 1/4 and *P*(*even**) = 3/4, improperly favoring *even** over *odd**, according to a starred analysis. However, the starred analysis gives *P*(*odd*) = 1/4 and *P*(*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 with 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 (2003). He was a leading proponent of objective Bayesianism and a master of the memorable riposte, which he formulated for this case as follows:

> 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. (p. xxii)

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 rehearsed in Sections 13.5 and 13.6 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. This problem is reviewed in Appendix 13.A. This should quell fears 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.

13.7. 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. This 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.

13.7.1. Equal Chance Sets

The logic divides outcome sets into types such that all sets of the same type must have the same chance. To implement this division, we require that two outcome sets are of the same type if the members of the two sets can be mapped one-to-one by a permutation of labels. This 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:[7]

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

Examples of $finite_3$ are {1, 2, 3}, {27, 1026, 5000}, and {24, 589, 2001}.

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

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

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$_{\text{co-finite-}n}$: an infinite set whose complement is finite of size n.

An example of *infinite*$_{\text{co-finite-10}}$ is the set of all numbers greater than 10: {11, 12, 13,…) since its complement is the finite set {1, 2, 3, …, 10}.

13.7.2. 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:

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

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

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

And for completeness we add in the two special cases

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

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

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, (34b) 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):

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

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).

13.7.3. Comparing Chance Values

The conditions (10) are powerful restrictions. They preclude the chance function Ch(.) being an additive probability measure. However, they leave the logic underspecified. We do not yet know whether the values V_n, V_∞, and 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 "≤," 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

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

One might think that this condition is 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 any, 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 in Section 14.4. An illustration presented there derives from steady-state cosmology. Accordingly, 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 to this case, we divide the space into an infinite number of 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

$$\mathrm{Ch}(\mathit{finite}_n) = V_n = n \qquad \text{where } n = 1, 2, 3, \ldots \tag{12}$$

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

The inequalities relating the various values of V_n in (35) become strict inequalities.

$$V_0 < V_1 < V_2 < V_3 < \ldots < V_\infty \tag{11a}$$

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

$$P(A \mid B) = \mathrm{Ch}(A)/\mathrm{Ch}(B), \tag{13}$$

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. 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[8]

$$V_\infty < \ldots < V_{-3} < V_{-2} < V_{-1} < V_{-0}. \tag{11b}$$

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

8 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 of possible background facts that would warrant strict inequalities among the individual values of V_{-n} as shown in (35b). Might an inventive reader be able to conceive of such facts?

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 the 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.

13.8. Interpreting the Inductive Logic

The chance function Ch(.) of Section 13.7 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 asks less than is usually asked in the analogous circumstance when we seek an interpretation of probability. It does not ask for an explicit definition, such as would be 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.

13.8.1. 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 outcome 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.

Sentences like these, by themselves, are not sufficient to give informal meaning to the quantity $P(.)$. All we have is one probability statement, that $P(H) = 0.5$, associated with another statement concerning an outcome with a 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 of these assertions can be modeled in many systems with an extensive quantity whose magnitude is additive. To break out of the self-referential trap, we use a rule that coordinates large and small values of probability with informal judgments of expectation about chance 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."[9] In Andrey Kolmogorov's (1950, p. 4) canonical treatment of the foundations of probability theory, he has a version of this rule that employs the locution "practically certain":

(a) One can be practically certain that if the complex of conditions S is repeated a large number of times, n, then if

9 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. 180–81). My use of the term "rule of coordination" is intended to recall Reichenbach's notion of a coordinative principle.

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 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 was the appropriate probability to assign. This in turn requires some prior understanding by the subject of what probability 0.5 means.

13.8.2. 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 infinite possibilities. This is not an outcome we should expect to happen, since it is thoroughly swamped by the infinite numbers outside the set. A chance V_{-n} happens when the number drawn lies outside some finite set. Since there are infinite possibilities outside the finite set that realize it, this is an outcome we should expect. That is, we have the following 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.

The rule divides outcomes sharply into three sets:

- outcomes in one of the $\textit{finite}_n$ sets, which we do not expect;
- outcomes in $\textit{infinite}_{\text{co-infinite}}$ sets, which may or may not happen "as likely as not"; and
- outcomes in one of the $\textit{infinite}_{\text{co-finite-}n}$ sets, 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 13.8.4 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.

13.8.3. Applying the Rule of Coordination

To illustrate how the rule of coordination is used, we 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$^{\text{million}}$—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 the fact to us.

13.9. Repeated, Independent, Infinite Lottery Drawings[10]

13.9.1. Applying Label Independence

To explore the application of the rule of coordination further and to see how the chance function behaves, consider the case of repeated independent drawings from a sequence of identical infinite lottery machines. We will consider the case of *N* independent drawings from *N* machines: machine$_1$, machine$_2$, …, machine$_N$. The combined outcome of *N* drawings will form an *N*-tuple such as

$$<156, 27, 2398, \ldots, 180>_N,$$

where the subscript *N* reminds us that there are *N* elements in the tuple. The set of all such outcomes is Ω_N. It is countably infinite, since it is formed as a finite tuple of elements of a countably infinite set.

Label independence can be implemented once again. We consider permutations of the labels on the outcomes of each lottery machine individually. Under such permutations, any *N*-tuple can be mapped to any other *N*-tuple. Thus, label independence requires that the outcome represented by each *N*-tuple each has a chance.

Label independence allows us to form equal chance sets of outcome sets, analogous to the equal chance sets of Section 13.7.1. Consider, for example, the set of all *N*-tuples such that every element in each of the member *N*-tuples is an even number. We will write this as[11]

$$\textit{all-even} = [\textit{even, even, \ldots, even}]_N = \{<n_1, n_2, n_3, \ldots, n_N>_N\text{: all } n_i \textit{ even}\}.$$

Analogously we have

$$\textit{all-odd} = [\textit{odd, odd, \ldots, odd}]_N = \{<n_1, n_2, n_3, \ldots, n_N>_N\text{: all } n_i \textit{ odd}\}.$$

10 The analysis of Sections 13.8 and 13.9 was decisively advanced by ideas that emerged in an energetic email exchange with Matthew W. Parker. I thank him for this and also for helpful remarks on the present text.

11 The square bracket notation […] is used to preclude the misreading that *all-even* is an *N*-tuple of sets, whose first, second, third, … members are each the sets of even drawings on machine$_1$, machine$_2$, machine$_3$, ….

When it happens that two sets of outcomes can be mapped onto each other by a label permutation, then label independence requires that the two sets have the same chance. Since they can be so mapped, *all-even* and *all-odd* have the same chance. They belong to the same equal chance set of outcome sets.

This shows that the inductive logic induced by label independence on repeated, independent drawings is similar in structure to that induced on single drawings. We shall see below that the full structure induced for the repeated case is more complicated. However, there are simple sectors in the logic that are formally the same as the logic that applies to single drawings.

13.9.2. A Simple Sector

A simple sector consists of a set of equal chance sets, where those equal chance sets can be totally ordered by set inclusion. That is, the equal chance sets form a chain such that the outcomes of each equal chance set is a subset of those higher in the chain. Since the set of all outcomes Ω_N is countably infinite, the equal chance sets will be of the type familiar from Section 13.7.1, namely finite_{n}, $\text{infinite}_{\text{co-infinite}}$, and $\text{infinite}_{\text{co-finite-n}}$. Because they are also totally ordered, we can assign the chance values V_0, V_1, …, V_∞, … V_{-1}, V_{-0} of (34). If all the cardinalities are not realized by the equal chance sets, then the sector will only have a subset of these values. Thus the equal chance sets of a simple sector follow the same logic as that governing equal chance sets of single drawings.

A note of caution is in order: there are many simple sectors in the outcome space of repeated drawings. The chance values only have a meaning within the sector in which they are defined relative to the chance of the other outcomes in the sector. Without further justification, we cannot assume that the chance of $V_{\text{something}}$ in the outcome space of a single drawing has the same meaning chance of $V_{\text{something}}$ in a simple sector of the outcome space of repeated drawings.

An example of a simple sector is the set of all outcomes in which all drawings return the same number. The outcome in which number 1 is drawn every time is

$$\mathbf{1}_N = \langle 1, 1, 1, \ldots, 1 \rangle_N$$

with an obvious extension of the notation to all 2, all 3, … outcomes. Set complementation with the simple sector gives a notion of negation. For example[12]

$$\textit{not}\ \mathbf{1}_N = \mathbf{2}_N \text{ or } \mathbf{3}_N \text{ or } \mathbf{4}_N \text{ or } \ldots$$

$$\textit{not}\ \mathbf{2}_N = \mathbf{1}_N \text{ or } \mathbf{3}_N \text{ or } \mathbf{4}_N \text{ or } \ldots.$$

The outcome $\mathbf{1}_N$ has a single member and is of type $\textit{finite}_1$. The complement *not* $\mathbf{1}_N$ is of type $\textit{infinite}_{\text{co-finite-1}}$. Thus,

$$\text{Ch}(\mathbf{1}_N) = V_1 \qquad \text{Ch}(\textit{not}\ \mathbf{1}_N) = V_{-1}.$$

Applying the rule of coordination, we infer that an outcome in which all numbers drawn in *N* independent repetitions are 1 is not to be expected in relation to other outcomes in the sector. Correspondingly, an outcome in which none of the numbers drawn is 1 is to be expected.

To identify further members in the sector, we ask whether we should expect all the *N* drawings to yield the same number, where the same number is found in some finite set, say {1, 2, 3}. That is, the outcome is ($\mathbf{1}_N$ or $\mathbf{2}_N$ or $\mathbf{3}_N$). Proceeding as above, we find this outcome is not to be expected, since

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

We get a different result if we ask about the outcome in which all the numbers drawn are the same, but that the number can be any in an infinite set of type $\textit{infinite}_{\text{co-infinite}}$, such as the set of all even numbers or the set of all odd numbers. These two outcomes are ($\mathbf{2}_N$ or $\mathbf{4}_N$ or $\mathbf{6}_N$ or …) and ($\mathbf{1}_N$ or $\mathbf{2}_N$ or $\mathbf{3}_N$ or …). Since these two outcomes can be mapped onto each other by a permutation of labels and because they are of type $\textit{infinite}_{\text{co-infinite}}$, we assign the same value

$$\text{Ch}(\mathbf{2}_N \text{ or } \mathbf{4}_N \text{ or } \mathbf{6}_N \text{ or } \ldots) = V_\infty$$

$$\text{Ch}(\mathbf{1}_N \text{ or } \mathbf{3}_N \text{ or } \mathbf{5}_N \text{ or } \ldots) = V_\infty.$$

These outcomes are "as likely as not" in this sector.

12 As before, I move without warning between the set representation of the outcome *not* $\mathbf{1}_N = \{\mathbf{2}_N, \mathbf{3}_N, \mathbf{4}_N, \ldots\}$ and its equivalent propositional representation *not* $\mathbf{1}_N = \mathbf{2}_N$ or $\mathbf{3}_N$ or $\mathbf{4}_N$ or … .

13.9.3. A Finite Simple Sector

All the finite outcome sets in the last simple sector above are subsets of another simple sector. Consider the outcome in which all the numbers drawn in N repetitions are less than or equal to some big, finite number *Big*, where the numbers drawn need not be the same. This outcome corresponds to a set of Big^N tuples in the outcome set Ω_N. Thus, we have

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

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

This is a new sector since a permutation of labels cannot map the set of tuples here assigned the value V_{Big^N} onto the set assigned the value V_{Big^N} in the simple sector of Section 13.9.2. For example, consider the $finite_2$ equal chance sets in each sector. The sector in this section will have outcomes like

$$<2, 1, 1, \ldots, 1>_N \text{ or } <3, 1, 1, \ldots, 1>_N.$$

No permutation of labels can map these onto the tuples, such as

$$<4, 4, 4, \ldots, 4>_N \text{ or } <5, 5, 5, \ldots, 5>_N$$

in the corresponding $finite_2$ equal chance sets of the simple sector of Section 13.9.2.

We cannot directly compare chance values across different sectors. However, our rule of coordination enables us to make some coarser judgments. What of the outcome that at least one of the numbers in N independent drawings is greater than *Big*? This outcome set is the complement of the last set considered with Big^N members. Thus, this outcome set is co-finite infinite so that the outcome is to be expected according to the rule of coordination. That is, no matter how big we make *Big*, we must always expect that at least one of the numbers drawn in N drawings will be greater.

Similarly, we cannot directly compare the chance values across the different sectors of Sections 13.9.2 and 13.9.3. However, our rule of coordination, applied to tuples of drawings, tells us that outcomes realized by finitely many tuples of drawings generally do not happen. If we now assume that outcomes realized by infinitely many tuples of drawings are more likely than the finite case, we arrive at a result that is surely surprising to someone whose intuitions about chance have been tutored by the probability calculus. It is more likely that all N numbers drawn are the same than it is that all N numbers drawn are less than or equal to some number *Big*, no matter how big we make it. This holds no matter how large we make N.

13.9.4. A "Likely As Not" Sector

Here are examples that illustrate outcomes to which the "as likely as not" chance of V_∞ is assigned. Consider the numbers drawn in N 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, …
> *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_N of type *infinite*$_{\text{co-infinite}}$. They can each be mapped onto any other by a permutation of the labels on the individual lottery machines. 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 dispelled by recalling that both of 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 in these examples are equinumerous and equinumerous in their complements, and they can be mapped onto each other by a label permutation.

13.9.5. Further Sectors

The chance logic of repeated independent infinite lottery drawings includes further sectors with more complicated properties. An indication of the nature of these sectors follows from consideration of two independent drawings. Consider the outcome that the first number drawn is 1 and that the second number drawn is even—that is, [1, *even*]—and then another outcome [1 or 2, *even*]. Both can be mapped one-to-one by label permutations onto infinite-co-infinite sets of pairs. However, no permutation of labels can map [1, *even*] to [1 or 2, *even*]. Thus, they cannot be required by label independence to have the same chance value. We would need to assign them different chance values. In an obvious notation, they might be $V_{1,\infty}$ and $V_{2,\infty}$. In this notation, the outcome [*even*, *even*] would be assigned the value $V_{\infty,\infty}$. The applicable chance logic would then reside in relations analogous to those of (35), such as $V_{1,\infty} \leq V_{2,\infty} \leq \ldots \leq V_{\infty,\infty}$; and $V_{1,\infty} = V_{\infty,1}$; $V_{2,\infty} = V_{\infty,2}$; etc.

13.10. Relative Frequencies of "As Likely As Not" Outcomes

13.10.1. Can Frequencies Reintroduce Probabilities?

The inductive logic induced by label independence precludes an ordinary probabilistic logic. We might wonder, however, whether probabilities can be reintroduced indirectly by an empirical approach. We carry out many independent drawings and let the limiting behavior of the frequencies reintroduce probabilities. This approach would succeed with a finite lottery. In independent repetitions, we expect with high probability that roughly half of 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 shall see in this section by direct calculation that the chance function of the infinite lottery does not return the favoring of relative frequencies of odd and even outcomes such as would be needed to reintroduce a probability of one half for each.

13.10.2. Odd and Even Outcomes

Consider $N > 1$ independent drawings of the lottery as in Section 13.9. The outcome sets that interest us are sets of N-tuples of the form

$$[odd, odd, \ldots, even, odd, even, even]_N$$
$$= \{<n_1, n_2, n_3, \ldots, n_N>_N :$$
$$n_i \text{ is an odd number in the positions marked "}odd\text{"}$$
$$\text{and an even number in the positions marked "}even\text{"}\}.$$

Since each *odd* and *even* are realized by infinitely many numbers, the set of N-tuples realizing any particular outcome set of the form [*odd*, *odd*, …, *even*, *odd*, *even*, *even*]$_N$ is infinite. Correspondingly, there are infinitely many ways that the complement set could be realized. Thus, the outcome is co-infinite infinite, and it has chance V_∞ of the simple sector of Section 13.9.3.

Permuting the labels on the individual lottery machine outcomes, we find that each of these outcome sets can be mapped onto any other. For example, the outcome set

$$[odd, odd, \ldots, even, odd, even, even]_N$$

can be mapped onto the outcome set

$$all\text{-}odd = [odd, odd, \ldots, odd, odd, odd, odd]_N.$$

We take the lottery machines in the positions marked "*even*" in the first outcome set and apply a permutation of labels that switches odd and even numbers. It follows that all the outcome sets of odd and even outcomes in this subsection have equal chances.

13.10.3. Frequencies of Even Outcomes

Our concern is not just the outcome sets of Section 13.10.2. We want to know the chances of n even numbers in N independent draws. These chances are assigned to larger outcome sets. The case of $n = 0$ is the *all-odd* tuple above. The case of $n = 1$ is realized as the union of N outcome sets

$$
\begin{aligned}
1\ even = {} & [even, odd, \ldots, odd, odd, odd, odd]_N \\
& \cup [odd, even, \ldots, odd, odd, odd, odd]_N \cup \\
& \ldots \\
& \cup [odd, odd, \ldots, odd, odd, odd, even]_N.
\end{aligned}
$$

In general, the number of these outcome sets to be joined to form the set of *n even* outcomes is given by the combinatorial factor $C(N, n) = N!/(n!(N - n)!)$. This combinatorial factor is always finite for finite *N* and *n*. It follows that there are still infinitely many *N*-tuples of individual outcome numbers that realize the outcome of exactly *n* even numbers in any order among the *N* drawings; and also infinitely *N*-tuples in the complement set.

As a result, it is natural to assign the chance value V_∞ to each outcome of *n* even numbers among *N* draws for any *n*. We might then continue with the natural supposition that each outcome of *n* even numbers among *N* draws has the same chance for any *n*. This was a conclusion I drew in an earlier version of this chapter and reported in a paper (Norton 2018a, §9). Unfortunately, the inference to this conclusion is a fallacy, and I retract it. That the outcomes have the same chance requires that they be in the same sector of the infinite logic. The values V_∞ reported might be drawn from different sectors. Then they would have an immediate meaning only within each sector. To conclude that they represent equal chances requires further argumentation. Ideally, we would need to show that permuting the labels takes us from one outcome of *n* even numbers to any other, which would show that they are within the same sector after all. This has not been shown and cannot be shown.

For it is easy to show that the outcome set of $n = 0$ even numbers drawn cannot be mapped by a label permutation onto the outcome set of *n* even numbers drawn, where $0 < n < N$. To see this, for the purpose of a *reductio*, assume otherwise: that there is such a mapping for some particular value of $0 < n < N$. Then a permutation of labels must include mappings of *N*-tuples of the form

$$<o_{1,1}, o_{1,2}, o_{1,3}, \ldots o_{1,N}> \rightarrow <e_{1,1}, ?, ?, \ldots, ?>$$

$$<o_{2,1}, o_{2,2}, o_{2,3}, \ldots o_{2,N}> \rightarrow <?, e_{2,2}, ?, \ldots, ?>$$

$$\ldots$$

$$<o_{N,1}, o_{N,2}, o_{N,3}, \ldots o_{N,N}> \rightarrow <?, ?, ?, \ldots, e_{N,N}>.$$

Here, $o_{1,1}, o_{1,2}, \ldots, o_{N,N}$ are odd numbers that enter into N-tuples that map to N-tuples with even numbers $e_{1,1}, e_{2,2}, \ldots, e_{N,N}$ in the positions shown. The "?, ?, ?, …" represent further numbers that may be odd or even, but have at least one odd number in each N-tuple.

Since the label permutations are carried out independently on each machine, it now follows that the label permutation on the set of machines must also include the map

$$<o_{1,1}, o_{2,2}, o_{3,3}, \ldots, o_{N,N}> \rightarrow <e_{1,1}, e_{2,2}, e_{3,3}, \ldots, e_{N,N}>.$$

However, this mapping is not included in the mapping supposed, for an N-tuple drawn from $n = 0$ *even* outcome set is mapped to an N-tuple drawn from the $n = N$ *even* outcome set. This contradiction completes the *reductio.*

While not all outcome sets with n even numbers can be mapped onto each other. There are a few mappings that succeed. We can map the outcome set with n even numbers among N draws onto the outcome set with $N - n$ even outcomes merely by a permutation that switches everywhere odd and even numbers in each lottery machine. Thus we have

$$\text{Ch}(n\ even) = \text{Ch}(N - n\ even) \text{ for all } 0 \leq n \leq N.$$

In Appendix 13.B, it is shown that this last possibility exhausts all the possibilities for equivalences under label permutation in the case of *n even* outcomes. That is, it is shown that a label permutation cannot map the outcome set *n even* to the outcome set *m even* unless $n = N - m$.

In the following two sections, we shall see that we can infer enough equivalences under label permutation to show that the essential point reported is correct: the chances of n even outcomes do not make likely a stabilization of frequencies that accord with probabilistic expectations.

13.10.4. The Chances of *N* Odd Versus *N* Even in *N* Drawings

The simplest case arises with the two extremes *all-even* and *all-odd*. They are in the same sector, since a permutation of the individual lottery labels can map one onto the other. To probe their chance behavior, consider another property:

$$div\ m = \text{set of numbers divisible by } m$$

and its complement *not div m*. The outcomes *even* and *odd* are the special case of $m = 2$. We have from earlier that a permutation of labels can map each of *even, odd, div m, not div m* onto each other. So they individually have the same chance. It now follows immediately that the same is true of the *N* tuples:

$$all\text{-}even = [even, even, \ldots, even, even, even, even]_N$$

$$all\text{-}odd = [odd, odd, \ldots, odd, odd, odd, odd]_N$$

$$all\text{-}div\ m = [div\ m, div\ m, \ldots, div\ m, div\ m, div\ m, div\ m]_N$$

$$all\text{-}not\ div\ m = [\ not\ div\ m, not\ div\ m, \ldots, not\ div\ m, not\ div\ m, not\ div\ m, not\ div\ m]_N$$

They have equal chance, so we may write:

$$\text{Ch}(N\ div\ m \text{ in } N) = \text{Ch}(N\ even \text{ in } N) = \text{Ch}(N\ odd \text{ in } N) = \text{Ch}(N\ not\text{-}div\ m \text{ in } N).$$

These equalities differ markedly from probabilistic expectations. Since we have P(*div m*) = $1/m$ and *P*(*not-div m*) = $(m - 1)/m$, we expect

$$P(N\ div\ m \text{ in } N) = [\ 1/m\ ]^N << P(N \text{ not-}div\ m \text{ in } N) = [(m-1)/m\ ]^N.$$

That is, the outcome (*N not-div m* in *N*) is $(m - 1)^N$ times as probable as outcome (*N div m* in *N*). It is the basis of the probabilistic expectation that *not-div m* outcomes are likely to occur much more frequently than *div m* outcomes (for $m > 2$). The equalities of the chance function do not reflect this probabilistic favoring or the associated expectations concerning frequencies.

13.10.5. Chances of Intermediate *N* Even Drawings in *N* Drawings

The preceding section has shown that the chance of frequencies of *div m* in *N* drawings is independent of *m* for the extreme $n = N$ case of *all-div m*. This independence of the chances from *m* holds for all values of *n*. That is, the chance of 0, 1, 2, … occurrences of a *div m* number in *N* drawings is independent of the value of *m*. Below, I sketch a diagrammatic proof for the simple case of $N = 2$. The proof will then be generalized to all *N*.

In two independent drawings, we will represent the four possible outcomes sets as

$$OO = [odd, odd] \quad OE = [odd, even] \quad EO = [even, odd] \quad EE = [even, even].$$

The frequency $n = 0$ corresponds to *OO*; $n = 1$ to (*OE* or *EO*); and $n = 2$ to *EE*. Figure 13.1 lays out the pairs of individual number outcomes in a grid. (It only shows a finite corner of the infinite grid.) The first number drawn is on the horizontal axis, and the second number drawn is on the vertical axis. The set of pairs that comprise *OO* is shown by the distribution of the labels "*OO*," and so on for the remaining outcomes.

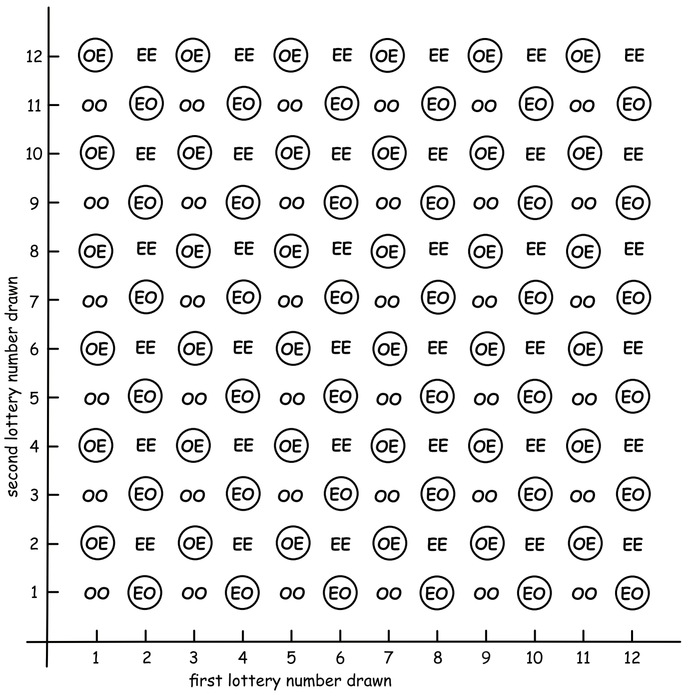

Figure 13.1. Distribution of outcomes *OO*, *OE*, *EO*, and *EE* in a two-lottery outcome space.

We will permute the labels so that the outcome sets for $n = 0$, $n = 1$, and $n = 2$ *even* outcomes coincide with the outcome sets for $n = 0$, $n = 1$, and $n = 2$ *div* 6 outcomes.

A permutation of the labels of the first lottery can be represented in the figure by leaving the labels in their positions on the axes and permuting the columns associated with the first lottery's numbers. The requisite permutation shifts the first five odd-numbered columns—1, 3, 5, 7, 9—to the left; and then places the first even-numbered column, 2, after it; and so on for the all the column numbers: five odd-numbered columns, then an even-numbered column, repeatedly. The result is shown in Figure 13.2.

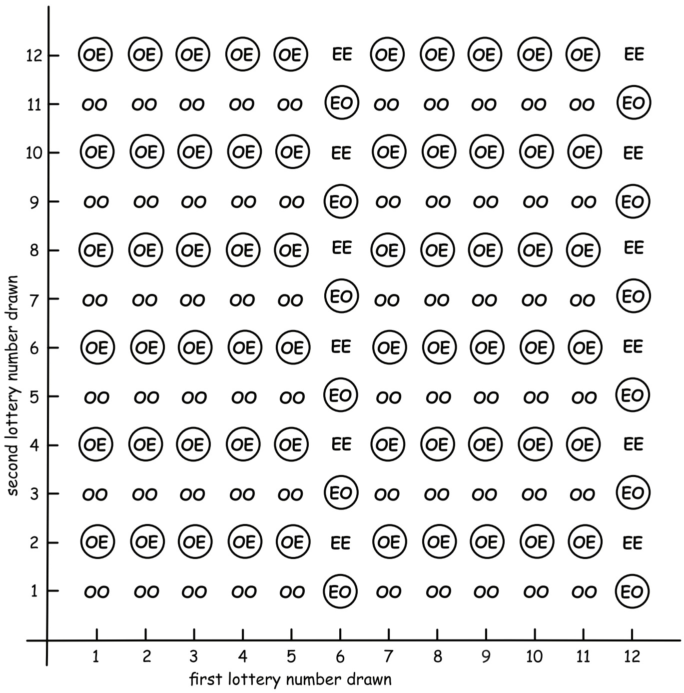

Figure 13.2. Result of permuting the columns.

To complete the manipulation, we perform the same permutation on the labels of the second lottery. That is, we perform the corresponding permutation of the rows to which the second lottery's numbers are associated. The result is shown in Figure 13.3.

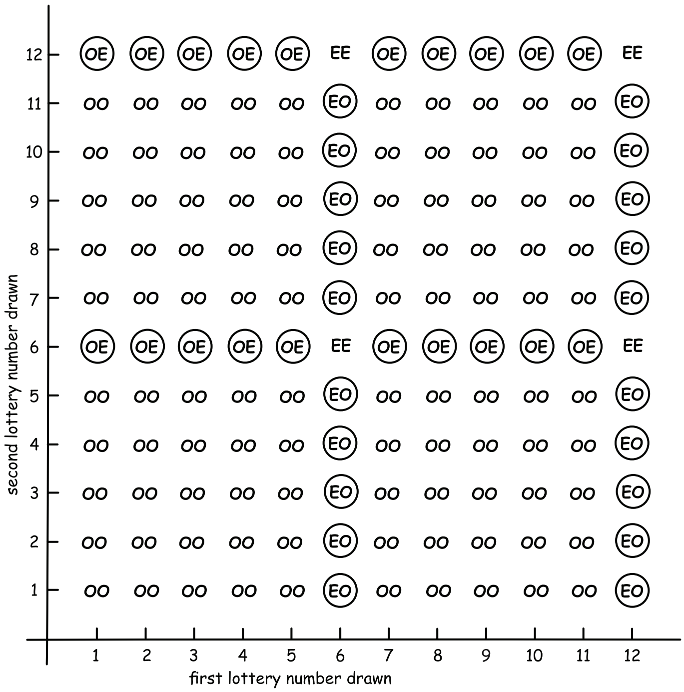

Figure 13.3. Result of permuting the columns and rows.

We read from Figure 13.3 that the outcomes have been relocated as follows:

$n = 0$ *even* outcomes (*OO*) coincides with $n = 0$ *div* 6 outcomes

$n = 1$ *even* outcomes (*OE* or *EO*) coincides with $n = 1$ *div* 6 outcomes

$n = 2$ *even* outcomes (*EE*) coincides with $n = 2$ *div* 6 outcomes.

Thus, the chances of *n even* outcomes equals the chances of *n div* 6 outcomes for all *n*.

The figure shows the manipulation for the case of $m = 6$. It is clear that it will succeed for any value of $m > 2$. It follows that the chances of the frequencies are independent of whether we are asking about even numbers or numbers divisible by 6 or 10 or 100 or 1,000. That is, the chances of these frequencies do not conform with the probabilistic expectations that even numbers appear in repeated trials roughly half of the time and that those divisible by 6 or 10 or 100 or 1,000 appear roughly 1/6 or 1/10 or 1/100 or 1/1,000th of the time, respectively.

13.10.6. The General Case[13]

The general result is that the chances of *n div m* outcomes in *N* drawings is independent of the value of *m* for all $0 \leq n \leq N$.

To see it, first note that there is a permutation of the label numbers of one lottery machine such that the set *div m* is mapped exactly onto the set *div k* for any *m*, $k > 1$. That is, under the permutation, all number labels divisible by *m* are switched with all number labels divisible by *k*. The construction of the $N = 2$ case displays the permutation for the case of $m = 2$ and $k = 6$.

Consider any *N*-tuple of outcomes that has exactly *n* outcomes divisible by *m*—that is, drawn from the set *div m*. Under the permutation, this *N*-tuple is mapped to one that has exactly *n* outcomes divisible by *k*—that is, drawn from the set *div k*. Now consider the set of all *N*-tuples with exactly *n* outcomes divisible by *m*. The same permutation will map it to the set of all *N*-tuples with exactly *n* outcomes divisible by *k*. Thus, label independence entails that the two sets have the same chance, and we can write

$$\mathrm{Ch}(n\ div\ m \text{ in } N) = \mathrm{Ch}(n\ div\ k \text{ in } N) = \mathrm{Ch}(n\ even \text{ in } N)$$

for all $0 \leq n \leq N$ *and* any *m*, $k > 1$. Since the outcomes of *n even* and *N-n even* may be mapped onto each other, we can extend these equalities of chances:

$$\mathrm{Ch}(n\ div\ m \text{ in } N) = \mathrm{Ch}(n\ even \text{ in } N) = \mathrm{Ch}(N\text{-}n\ even \text{ in } N) = \mathrm{Ch}(N\text{-}n\ even\ m \text{ in } N)$$

for all $0 \leq n \leq N$.

13.10.7. Frequencies Do Not Give Us Probabilities

What these results show is that the tempting strategy for reintroducing probabilities fails. The temptation is to say “Do the experiment. Run many

13 I thank Matthew W. Parker for this proof.

independent drawings from lottery machines. Read the limiting frequencies in many drawings. They will reveal to you the probabilities hidden in the lottery machines!"

The strategy fails since the chances of different frequencies do not mass in a way that would reveal probabilities. Probabilistic intuitions would lead us to expect that drawing all *N* numbers divisible by 100 in *N* draws would be much less likely that drawing all *N* numbers *not* divisible by 100 in *N* draws. Yet they have the same chance, so we have no reason to expect the second over the first.

The same probabilistic intuitions would lead us to expect that the most likely numbers of even drawings in *N* drawings would cluster around *N*/2. Numbers of even drawings far from *N*/2 would be unlikely. From this clustering, we could recover a probability of one half for an even number. The trouble is that this same clustering around *N*/2 is likely for outcomes divisible by 10, 100, or 1,000. We would then have to infer that numbers divisible by 10, 100, or 1,000, or any other number greater than 2, also have a probability of one half. No ordinary probability distribution can realize these probabilities.[14]

The calculations reviewed in this section and in Appendix 13.B show that the chances of securing *n* or *m even* numbers in *N* repeated independent draws from infinite lottery machines are incomparable for most *n* and *m*. Thus, this section leaves open whether imposition of further background facts will lead to further relations that will lead to chances favoring certain frequencies of outcomes. However, what has been shown is that if there is any favoring, it is not of a type that can be used to reveal underlying probabilities as long as the fair character of the infinite lottery is preserved.

14 Assume otherwise. Then the probability of drawing a number divisible by 2^r is one half, for any $r > 1$. Since the probability of drawing a number divisible by 2 is also one half, it follows the probability of drawing numbers divisible only by $2^1, 2^2, \ldots, 2^{r-1}$, is zero. But since *r* can be set as large as we like, we infer that the chance of a number divisible by any power of two is zero, which contradicts the probability of one half for even numbers.

13.11. Failure of the Containment Principle

The infinite lottery logic will likely be discomforting 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, …} must be assigned greater chance than the set of every fourth numbered outcome {4, 8, 12, 16, …}. 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*: Ch(*A*) < 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 that consists of a continuum of 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 as the probability of drawing any number less than 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, …}, we are 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

Ch({2, 4, 6, 8, …}) = Ch({3, 9, 15, 21, 27, …})

Ch({4, 8, 12, 16, …}) = Ch({3, 9, 15, 21, 27, …}).

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

Ch({4, 8, 12, 16, …}) = Ch({2, 4, 6, 8, …}).

This equality contradicts the containment principle, which tells us that

Ch({4, 8, 12, 16, …}) < Ch({2, 4, 6, 8, …}).

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.

13.12. Is an Infinite Lottery Machine Physically Possible?

The discussion so far has presumed the physical possibility of an infinite lottery machine. But in what sense are they physically possible? Elsewhere (Norton, 2018; Norton and Pruss, 2018, Norton, 2020) I have pursued the question is 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 this much bigger space 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. But 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 suggested 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 (2020) the 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 the probability is greater than zero, we need only sum finitely many of the equal, non-zero probabilities $P(1)$, $P(2)$, $P(3)$, … 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 $P(1)$, $P(2)$, $P(3)$, … 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

$$P(1) + P(2) + P(3) + \ldots = 0 + 0 + 0 + \ldots = 0.$$

This 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 (2020) 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. This 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 non-constructive, and we have some reason to expect that none can be constructed. This 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.[15] 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, 2020) describes designs based on quantum mechanical systems. In the simplest such design, 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. We 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

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

thrived merely by supposing that these real randomizers were imperfect surrogates for true but unrealizable probabilistic randomizers: idealized 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.

13.13. 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 13.6, 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 13.7 and 13.8.

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 the notion of probability even is. 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 Appendix 13.A. 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 13.A: The "Measure Problem" in Eternal Inflation[16]

13.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.

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 Grand Unified Theory ("GUT"). These efforts failed. The driving matter is now just a novel matter field, the inflation, posited ad hoc with just the right properties. Moreover, the search for a viable form of inflation

16 For a fuller discussion of the measure problem and its inductive analysis, see Norton (2018a).

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 non-uniformities 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 of 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.

13.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 were 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 (2007) introduced inflation to cosmology in the early 1980s. Here is his development of the problem:

> 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. (p. 11)

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 following 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 Nth, 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
>
> $$1, 3, 2, 5, 7, 4, 9, 11, 6, \ldots, \qquad (14)$$
>
> 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, …). The integers are just arranged in an unusual order. However, if we truncate the sequence shown in Eq. (14) after the Nth entry, and then take the limit $N \to \infty$, 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 non-trivially 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 13.6 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 13.5 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:

> 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 differ-

> ent 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. (2001, p. 42)

13.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 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? (2011, p. 42)

He then reviews with disdain the idea of imposing a measure on the islands:

> 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. (pp. 42–43)

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 book 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,[17] 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 Sections 13.9 and 13.10 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

17 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 except a countable infinity of exceptions that arise in universes finitely many *like* pocket universes, or in universes with finitely many *unlike* universes.

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.

Appendix 13.B: Inequivalences under Label Permutation of Outcomes of Many Independent Drawings

The numbers drawn independently from N infinite lottery machines form an N-tuple $<n_1, n_2, n_3, \ldots, n_N>_N$. These N-tuples can be grouped into "ordered parity sets" such as $[odd, odd, \ldots, even, odd, even, even]_N$ defined in the main text in Section 13.10.2. The outcome sets of primary interest are those with n *even* numbers in any order. They are the "unordered parity sets," written "(n, N)":

$$(n, N) = \text{Union of all ordered parity sets } [parity, \ldots, parity]_N \text{ with exactly } n \text{ } even.$$

where *parity* is either *even* or *odd*. The following is to be shown:

Theorem

No label permutation can map the unordered parity set (n, N) onto (m, N), for all $0 \le n \le N$, excepting the trivial case of $n = m$, implemented by an identity map on labels, and the case of $n = N - m$, implemented by a label permutation that switches all odd with all even numbers.

Proof

The case of $n = 0$ and $0 < m < N$ has been shown in Section 13.10.3. Switching "even" for "odd" in that demonstration shows the case of $n = N$ and $0 < m < N$. Here we need only consider $0 < n, m < N$ in the theorem.

Assume for purposes of a *reductio* that there exists a label permutation f that maps the N-tuple $<n_1, n_2, n_3, \ldots, n_N>_N$ to $<f(n_1), f(n_2), f(n_3), \ldots, f(n_N)>_N$ such that unordered parity set (n, N) is mapped onto (m, N), where n does not equal $N - m$.

It may be the case that a label permutation maps every member of some *ordered* parity set of (n, N) onto elements of the same ordered parity set of (m, N). The mapping is "onto" so that the image of the ordered parity set of (n, N) coincides with the ordered parity set of (m, N). We shall say that the label permutation respects ordered parity sets just if this last property is true for every ordered parity set of (n, N).

There are $N!/(n!(N-n)!)$ ordered parity sets that are subsets of (n, N); and $N!/(m!(N-m)!)$ ordered parity sets that are subsets of (m, N). Unless we have the cases excepted in the theorem, $n = m$ or $n = N - m$, these two combinatorial factors are unequal. It follows that there can be no one-to-one label permutation that respects ordered parity sets for the cases considered in the theorem.

For example, there are four ordered parity sets for (1,4): *EOOO*, *OEOO*, *OOEO*, *OOOE*, written here in compact notation with "*E*" = *even* and "*O*" = *odd*. There are six ordered parity sets for (2, 4): *EEOO*, *EOEO*, *EOOE*, *OEEO*, *OOEE*. A label permutation that respects ordered parity sets would have to map the members of each of the *EEOO*, *EOEO*, … of (2, 4) onto distinct ordered parity sets *EOOO*, *OEOO*, … of (1, 4). Since there are six of the former and four of the latter, this is impossible.

Set n as the number of evens for which $N!/(n!(N-n)!) > N!/(m!(N-m)!)$. (There will always be an inequality since the case of equality $n = N - m$ is excluded.) Since the label permutation cannot respect ordered parity sets, it follows that the permutation must "cross over" the boundaries somewhere of the ordered parity sets. That is, there must be two N-tuples that map as

$$\mathbf{R} = \langle r_1, r_2, r_3, \ldots, r_N\rangle_N \text{ maps to } f(\mathbf{R}) = \langle f(r_1), f(r_2), f(r_3), \ldots, f(r_N)\rangle_N$$

$$\mathbf{S} = \langle s_1, s_2, s_3, \ldots, s_N\rangle_N \text{ maps to } f(\mathbf{S}) = \langle f(s_1), f(s_2), f(s_3), \ldots, f(s_N)\rangle_N,$$

where $f(\mathbf{R})$ and $f(\mathbf{S})$ belong to the same ordered parity set of (m, N), but **R** and **S** belong to different ordered parity sets of (n, N).

To proceed, we form a new N-tuple $\mathbf{T} = \langle t_1, t_2, t_3, \ldots, t_N\rangle_N$ by the rule

$$t_i = r_i \text{ if } r_i \text{ is even; or if both } r_i \text{ and } s_i \text{ are odd.}$$

$$= s_i \text{ if } r_i \text{ is odd and } s_i \text{ is even.}$$

Each of **R** and **S** have n even numbers in their tuples. However, the positioning of the even numbers in their N-tuples must be different somewhere, since **R** and **S** come from different ordered parity sets. The definition of **T** is designed to collect all the even numbers from **R** and **S** such that **T** has at least one more even number than **R** and **S**. For example, if **R** = <1, 1, 2, 2> and **S** = <1, 2, 1, 2>, then **T** = <1, 2, 2, 2>. That is, **T** belongs to an unordered parity set, (n', N), where $n' > n$.

The label permutation f maps **T** as

$$\mathbf{T} = <t_1, t_2, t_3, \ldots, t_N>_N \text{ maps to } f(\mathbf{T}) = <f(t_1), f(t_2), f(t_3), \ldots, f(n_N)>_N.$$

Each $f(t_i)$ is either $f(r_i)$ or $f(s_i)$. Since f(**R**) and f(**S**) are both members of the same ordered parity set (m, N), it follows that f(**T**) is a member of the same ordered parity set (m, N). That is, the label permutation f maps an N-tuple **T** in (n', N), where $n' > n$, to an N-tuple f(**T**) in (m, N). Since a label permutation is invertible, it follows that there is no N-tuple in (n, N) that the label permutation maps to f(**T**). This mapping of **T** contradicts the initial assumption that the label permutation maps (n, N) to (m, N) and completes the *reductio* needed to establish the theorem.

REFERENCES

Bartha, Paul. 2004. “Countable Additivity and the de Finetti Lottery.” *The British Journal for the Philosophy of Science* 55: pp. 301–21.

Benci, Vieri, Leon Horsten, and Sylvia Wenmackers. 2013. “Non-Archimedean Probability.” *Milan Journal of Mathematics* 81: pp. 121–51.

Blackwell, David and Persi Diaconis. 1996. “A Non-Measurable Tail Set.” In *Statistics, Probability and Game Theory: Papers in Honor of David Blackwell*, edited by T. S Ferguson, L. S. Shapley, and J. B. McQueen, pp. 1–5. Hayward, CA: Institute of Mathematical Statistics.

de Finetti, Bruno. 1972. *Probability, Induction and Statistics*. London: John Wiley & Sons.

———. 1974. *Theory of Probability: A Critical Introductory Treatment*. Vol. 1. Chichester: John Wiley & Sons.

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., Mark J. Schervish, and Teddy Seidenfeld. 1986. “Statistical Implications of Finitely Additive Probability.” In *Bayesian Inference and Decision Techniques*, pp. 59–76. Amsterdam: Elsevier Science Publishers.

Kolmogorov, Andrey N. 1950. *Foundations of the Theory of Probability*, translated by N. Morrison. New York: Chelsea Publishing Company.

Norton, John D. 2010. "Deductively Definable Logics of Induction." *Journal of Philosophical Logic* 39: pp. 617–54.

———. 2011. "Challenges to Bayesian Confirmation Theory." In *Handbook of the Philosophy of Science.* Vol. 7, *Philosophy of Statistics*, edited by Prasanta S. Bandyopadhyay and Malcolm R. Forster, pp. 391–439. Amsterdam: Elsevier.

———. 2018. "How to Build an Infinite Lottery Machine." *European Journal for Philosophy of Science* 8, pp. 71–95.

———. 2018a. "Eternal Inflation: When Probabilities Fail." *Synthese*, https://doi.org/10.1007/s11229-018-1734-7 Preprint, http://philsci-archive.pitt.edu/14401/

———. 2020. "How NOT to Build an Infinite Lottery Machine." *Studies in History and Philosophy of Scienc*e 82: pp. 1–8

Norton, John D. and Alexander R. Pruss. 2018. "Correction to John D. Norton 'How to Build an Infinite Lottery Machine'." *European Journal for Philosophy of Science* 8: pp. 143–44.

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." In *Uncertainty and Intelligent Information Systems*, edited by B. Bouchon-Meunier, C. Marsala, M. Rifqi and R. R. Yager, pp. 3–15. 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 Leon Horsten. 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.