The Paradox of Deterministic Probabilities

Valia Allori
Philosophy, Northern Illinois University
vallori@niu.edu

To appear in: *Inquiry: An Interdisciplinary Journal of Philosophy*

Abstract

This paper aims to investigate the so-called paradox of deterministic probabilities: in a deterministic world, all probabilities should be subjective; however, they also seem to play important explanatory and predictive roles which suggest they are objective. The problem is then to understand what these deterministic probabilities are. Recent proposed solutions of this paradox are the Mentaculus vision, the range account of probability, and a version of frequentism based on typicality. All these approaches aim at defining deterministic objective probabilities as to make them compatible with determinism. In this paper I argue that one can think of the equivalent of subjective and objective deterministic probabilities in terms of typicality. Also, I show that only what I identify with objective probabilities play the necessary explanatory and predictive roles, while the subjective components essentially guide beliefs. In this way, the paradox is solved.

Keywords: typicality, probability, determinism, Mentaculus vision, range account, Statistical Mechanics

1. Setting up the Paradox: Probability and Determinism

Assume that classical mechanics is the true theory of the world. It has been argued that statistical mechanics, namely classical mechanics applied to certain systems composed by a vast number of Newtonian particles, is able to explain macroscopic laws, such as the laws of thermodynamics. Since classical mechanics is deterministic, the future behavior of any physical system is fixed given the initial conditions and the laws. So, a Laplacian demon could exactly calculate everything that happens, at any level of description. There is no room for uncertainty of outcomes or probability that a state of affairs will realize other than zero and one: either something certainly happens, or it never happens. Nonetheless, as I summarize in the next section, in statistical mechanics we routinely use probability talk when we describe macroscopic objects: if an ice cube is taken out of the refrigerator, we talk about its probability of melting within two minutes; if a gas in a container is allowed to expand freely, we talk about the probability of the gas spreading in the whole room. These are probabilities arising in a deterministic framework, so let's call them *deterministic probabilities*.

But what are these deterministic probabilities? On the one hand, if the world is deterministic, the straightforward way of interpreting such probabilities seeems to think of them as subjective rather than objective: they are a measure of our ignorance of initial conditions, or of our limited capabilities in solving the necessary equations. We are not a Laplacian demon, so we cannot exactly compute what will happen to a macroscopic physical system such as an ice cube or a gas: we do not possess all the information and, even if we did, we could not exactly compute the particles' motion. It seems that it is because of this ignorance that we use probability talk.[1]

Nonetheless, deterministic probabilities seem to have important roles which would be inexplicable, if they were subjective. That is, deterministic probabilities need to be *objective*. In fact, probabilities seem explanatory: the fact that the probability that an ice cube melting in the next 5 minutes is high explains why the ice cube is gone in an hour. Or the fact that the probability that the gas expanding freely is high explains why the gas will eventually occupy the whole room. Or the fact that the probability of getting a tail tossing a fair coin is ½ explains why if we toss a fair coin one hundred times, we roughly get 50 tails and 50 heads. In addition, probabilities guide belief. For example, consider Bayesian inference. As evidence accumulates, the degree of belief in a given hypothesis becomes very high or very low. Bayesian inference calculates a numerical estimate of the degree of belief in the hypothesis after evidence has been observed. So following Bayesian inference yields knowledge of rational belief concerning probabilities.

From these considerations, namely the observations that deterministic probabilities seem subjective when looking at the underlying deterministic dynamics, but need to be objective when looking at their role in our scientific discourse and practice, arises the so-called *paradox of deterministic probabilities* (Loewer 2001): *How can we account for the objective roles of deterministic probabilities in laws, explanations and predictions, if they are subjective?*

After summarizing in the next section how deterministic probabilities come about within a prototypical deterministic theory such as statistical mechanics, in Section 3 I discuss how David Albert and Barry Loewer use what their *Mentaculus vision*, in the framework of the Humean account of laws, to solve this paradox. They argue that deterministic probabilities are actually objective, even if they seem subjective. Instead in this paper I argue for a solution of the paradox in terms of the notion of typicality. In Section 4 I discuss the idea of explanation based on typicality, and then I show in Section 5 how deterministic probabilities can be understood in this framework. My approach is similar to typicality frequentism, according to which deterministic probabilities are objective across-world typical frequencies. (Hubert 2021). However, in

[1] For instance, Popper (1985) famously rejected the idea that objective probabilities are compatible with determinism.

contrast with this view, I argue that one could define what we think of as objective deterministic probabilities as typical *single-world* relative frequencies, and I show that they can play the relevant explanatory and predictive roles. In addition, one could use typicality to make sense of what we think of subjective deterministic probabilities. These are ignorance probabilities which guide our credences and define single event probabilities. Similarly to the range account of probabilities, according to which probabilities are ranges of values in a space of possible states, in my approach what we regard as subjective deterministic probabilities describing our ignorance are defined in terms of volume in phase space, and therefore are across worlds. They are not objective features of the world, and so they are neither explanatory nor predictive of any physical fact. I conclude in Section 6 explaining how the paradox is solved. The paradox arises from the assumption that there is only one type of probability which plays two incompatible roles: it is both subjective (because of determinism) and objective (because of its explanatory, predictive and nomological roles). Instead my view can account for both of these features: typical single world frequencies play the role of objective deterministic probabilities in that they are explanatory and predictive; instead typical across-worlds phase space volume ratios capture the essence of subjective deterministic probabilities, expressing our ignorance on the details of the situation and guiding our beliefs and expectations.

2. Prototypical Deterministic Probabilities: Statistical Mechanics

Where does probability talk enter in a deterministic theory such as Newtonian mechanics? In most cases, in classical mechanics one explains a phenomenon by solving the Newton equation, either exactly or using some suitable approximation. This is compatible with Hempel's deductive-nomological model (1965): a phenomenon is explained if it can be deduced from a law of nature. Think for instance of the derivation of the acceleration of bodies on the Moon with this simple formula: $a = \frac{GM}{R^2}$, where G is the gravitational constant, and M and R are respectively the mass and the radius of the Moon. Think also to rigid body mechanics, where Newton's law holds for extended solid bodies when applied to the center of mass and when substituting, in the case of rotations, the moment of inertia instead of the mass. One can also derive from the Newtonian dynamics less fundamental laws, like for instance Kepler's laws of planetary motion, using suitable approximations. Things instead get more complicated when one deals with non-rigid bodies, such as fluids and gases. In the case of gases, macroscopic phenomena are described by the laws of thermodynamics, in particular the second law according to which a quantity called entropy never decreases. Statistical mechanics uses statistical methods because of the large number of the objects (the particles in the gas) involved in the calculations, and provides a framework to account for the laws of thermodynamics in terms of the microscopic Newtonian dynamics.[2] Since the instantaneous description of a particle is given by its instantaneous position and velocity, the complete dynamical description of a macroscopic body at a fixed time is the *microstate* X, given by all the instantaneous positions and velocities of the N particles composing the macroscopic body. The set of all the microstates constitutes phase space, Ω. Any macroscopic description is determined by the microscopic one, but not *vice versa*: there are many microstates corresponding to the same macroscopic description, the *macrostate* Γ. Microstates that belong to the same macrostate are all macroscopically similar: any microstate has the same values of the macroscopic (i.e. thermodynamic) variables, such as, e.g., pressure or temperature. Macrostates provide a partition of phase space into disjoint regions. Among all macrostates, the *equilibrium state* Γ_{eq} is such that a system in that state will not change: a gas completely spread out in a room will not change temperature, volume or pressure spontaneously. In contrast, a system not in equilibrium will tend to reach the equilibrium state: a gas concentrated in a corner of a room will tend to spread out. This is because the equilibrium microstate happens to be incredibly larger than any other macrostate,[3] so that the vast majority of microstates in a given microstate will eventually 'fall into' the equilibrium state, and once there it will stay there (figure 1).

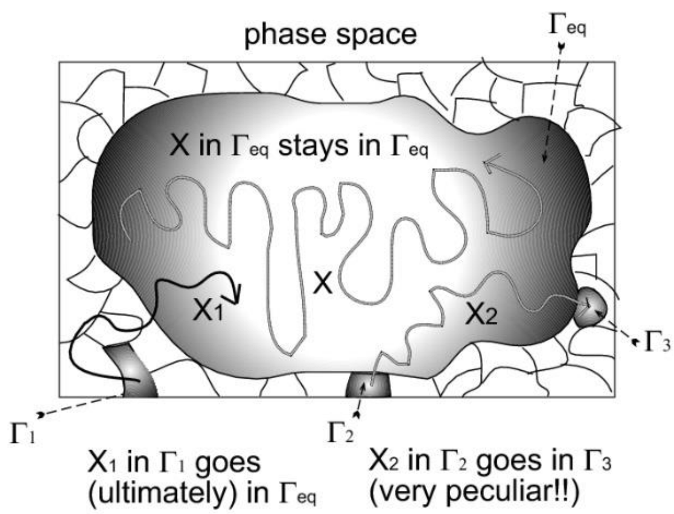

Figure 1: The expected behavior: most microstates will fall into the equilibrium microstate.

---

[2] See Albert (2000) for a clear exposition. See also see e.g. Bricmont (1995). Penrose (1999). Goldstein (2001) and references therein.

[3] To give an idea of how large the equilibrium state is, consider a gas in a corner of a box divided in two regions of volume V. When the separation wall is removed, the gas expands to reach its new equilibrium state where it occupies the whole volume $2V$. The phase space volume increases of the same amount for any particle of the gas; therefore there is a transition from a region of volume V^N in phase space to a region of volume $2^N V^N$ and $V_{eq}/V_{initial} = 2^N \sim 2^{10^{23}} \sim 10^{10^{23}}$.

This is the expected behavior of the microstate. Only very few microstates will not enter the equilibrium state or will get out of it fairly quickly. They constitute an unexpected or exceptional, behavior. The fact that most microstates evolve towards equilibrium explains the second law of thermodynamics. In fact, Boltzmann's entropy is defined as a measure of the volume of the phase space occupied by the macrostate to which a given microstate belongs to. Therefore, most microstates move from a macrostate of a smaller volume to one of a bigger volume, entropy will accordingly increase.

It sems natural at this point to say that it is *likely or probable* that entropy will not decrease, and this is indeed the way in which probability talk enters the framework of a deterministic theory. This is technically done by introducing a probability distribution over the initial microstate of the universe, sometimes called the Statistical Postulate. However, because of the time reversal symmetry of the Newtonian law, if it's likely for entropy to increase toward the future, it is also likely for it to increase toward the past, and the latter is not what happens. A popular way of breaking the symmetry is the so-called Past Hypothesis, namely the assumption that the universe begun in an microstate whose entropy was incredibly small.[4] In this way, Albert (2000, 2015) and Loewer (2001, 2004, 2012, 2020) have proposed the so-called *Mentaculus vision*, according to which only three ingredients are needed to explain the macroscopic phenomena and their laws: 1) the Newtonian dynamics; 2) the Past Hypothesis; and 3) the Statistical Postulate.

3. The Mentaculus Vision: Deterministic Chances

But what exactly are the statistical mechanical probabilities? This is an instance of the paradox of deterministic probabilities presented in Section 1. In a deterministic theory there seems to be no room of the objective probability needed to account for their roles in in the theory. As Loewer (2001) puts it:

1. these probabilities are explanatory: "So, for example, that it is a law that the half-life of some substance is one second explains why the substance is almost all gone by the end of an hour" [*ibid.*];
2. they should satisfy the probability axioms;
3. they should guide beliefs, and they should apply to both types of events and to particular events: "For example, statistical mechanics provides the chance that ice cubes (of a certain size and temperature) placed in a bowl of water (of a certain size and temperature) melt within 5 minutes and the chance that *this* ice cube now placed in *this* bowl of water will melt within 5 minutes" [*ibid.*];
4. they may change in time;

[4] The names 'Statistical Postulate' and 'Past Hypothesis' are due to Albert (2000).

5. it should make sense to have a probability distribution over the initial conditions of the universe; and
6. it has to be possible for chances different from 0 and 1 to appear in deterministic theories.

Loewer argues that epistemic accounts of probability fail capture these features, as well as other approaches to objective probabilities such as propensities and frequentist interpretation. Instead he proposes that the Mentaculus vision can solve this paradox. The idea is that laws are contingent generalizations implied by the theory that best combine simplicity and informativeness; and probabilities come into the picture by allowing some of the axioms of the system to specify probabilities. That is, the contingent generalizations are the laws and the probabilistic statements are the probabilities. In the framework of statistical mechanics adding a probability distribution to the initial conditions to the laws increases informativeness with little cost in simplicity. As such, it earns its status as a law. The probability measure on initial state of the whole universe is a probability distribution in that if a proposition with high probability matches the actual facts, it has a better fit than a proposition with low probability. This distribution assigns probabilities to everything, including single events. Moreover, an agent needs to constrain her belief according to these probabilities according to another axiom, the *Principal Principle*, which roughly states that an agent ought to adjust her credence according to the probability of the proposition given by the best system. This is a way in which one can define objectively deterministic probabilities, and therefore solve the paradox.[5]

I argue instead for a solution of the paradox which does not reduce all probability talk in a deterministic framework to talk about suitable objective probabilities. Rather, I argue that the notion of typicality can make sense of both types of deterministic probability talk. I explain what typicality is in the next section, and I connect it to deterministic probabilities in Section 5.

4. Explanation based on Typicality: 'Expect Typical Events'

There is another influential reading of the statistical mechanical explanation of the macroscopic laws which does not rely on the notion of probability but rather uses the weaker notion of typicality.[6] According to the proponents of the typicality approach, to say that most microstates belonging to the same microstate behave thermodynamically (i.e., they are entropy-increasing) is to say that the typical microstate in that microstate is entropy-increasing. In turn, that means that the extension of the typical set, namely

[5] For a closely related but distinct Humean account, see Hoefer (2007, 2011, 2019).
[6] See, e.g., Lebowitz (1981, 1993a, 1993b, 1999), Goldstein (2001).

the set of entropy-increasing microstates, is very big or, alternatively, the set of exceptions E to the typical behavior is very small. This idea can be mathematically implemented introducing a typicality measure μ which allows measuring the size of the sets in phase space. The proposed typicality measure is the Lebesgue-Liouville measure, the uniform measure, which is the same as the probability measure used in the Mentaculus vision. This measure is such that the set of exceptions is really small: (putting $\mu\ (\Omega)$, the measure of the whole phase space, as equal to 1) $\mu(E) << 1$. More precisely, the measure of the set of exceptions (with any tolerance ε) gets smaller and smaller as N, the number of particles, gets bigger and bigger. Consider for example a series of tossing of a fair coin. Write 1 for heads, 0 otherwise. A possible sequence of eleven flips is given by $X = (1,1,1,0,1,0,0,1,0,0,1)$. The corresponding phase space Ω in this context is the space of all possible combinations of 0s and 1s in N places, if N is the number of tosses in a sequence. In this example $N = 11$ and the possible sequences are $2N = 1024$, i.e. Ω has 1024 elements. The frequency $n(X)$ of 1s in a sequence X is given by counting the times 1 comes up in X: $n(X) = \frac{1}{N}\sum_{i=1}^{N} N_i$ where N_i is 1 when the i-th place of X is 1, and 0 otherwise. Here $n(X) = 6/11$. The more the number N of flips in a sequence grows, the more the number of sequences that do *not* have a frequency of 1s that is equal (or very close) to ½ decreases.

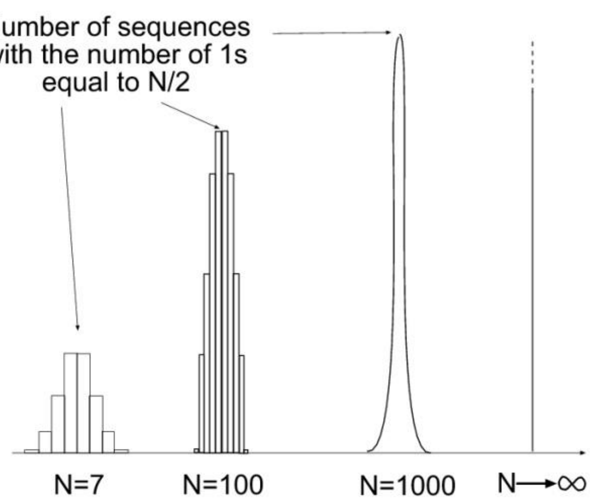

Figure 2: As *N* increases, the set of exceptions (atypical sequences) gets smaller and smaller.

Call 'expected behavior' having a frequency equal (or close) to ½, and call E the set of those 'exceptions' to it. This example with $N = 11$ is not really good to see this since N is too small, but it is possible to actually count the expected and exceptional sequences. An 'exceptional' sequence is for instance $X = (1,0,0,0,0,0,0,0,0,0,0)$, while an `expected'

sequence is $X = (1,1,1,0,1,0,0,1,0,0,1)$. In detail, in this case,[7] defining E with a tolerance $\varepsilon = 1/22$ such that $E = \{X \in \Omega: n(X) \in (5/11, 6/11)\}$, we have that, while all the possible sequences are 1024, those in E (the exceptions) are 100. Therefore, $\mu(E) = 100/1024 = 0.0976$ that is sufficiently less than 1 to get the idea. In general, for large N, we have $\mu(E) << 1$. That is, the number of sequences such that $n(X)$ is not ½ (and not a number close to ½) are few (figure 2).
In the language of thermodynamics, take X as the microstate, Ω as the phase space and μ as typicality measure. The frequency n is analog to the thermodynamic variables and defines the macrostate: any microstate with the same amount of 1s, regardless the ordering, belongs to the same macrostate. The typical behavior in this case is the entropy-increasing behavior.

As we have seen, thus, in the typicality approach in statistical mechanics the typicality measure is used to count the number of states and it is such that the set of exceptions to the entropy-increasing, expected, behavior is really small. More precisely, if we want to explain why a macroscopic system remains in equilibrium, with $\mu(E) \ll 1$ we mean that the measure of the exceptions (namely the extension of the set of microstates whose temporal evolution brings the system out of the equilibrium state) is very small with respect of the measure of the equilibrium state: $\mu_{\Gamma_{eq}}(E) = \frac{\mu(E)}{\mu(\Gamma_{eq})} \ll 1$.
Instead, when the system is not in equilibrium but it belongs to a macrostate Γ, and we want to explain why it will eventually reach equilibrium, then with $\mu(E) \ll 1$ we mean that the measure of the exceptions (namely the extension of microstates whose temporal evolution brings the system in the equilibrium state) is very small with respect to the measure of t Γ: $\mu_\Gamma(E) = \mu(E)/\mu(\Gamma)$<<1. Once we know μ, in order to take into account the facts of our world, we build μ_Γ, conditionalizing μ on the realization on events Γ that have happened in our world.

A side note: why the Lebesgue-Liouville measure and not another measure? A complete discussion is outside of the scope of this paper, but let me say that some proponents of the typicality approach argue that the choice is due to the stationarity of the measure, as in this way there is no privileged initial time.[8]

A remark on explanation based on typicality. As we have seen earlier, in a deterministic framework, one could maintain that, similarly to what advocated by Hempel's deductive-nomological model, a phenomenon is explained if it results as an exact or approximate solution of the relevant equation. To extend this idea to statistical mechanics, one would have to show that all solutions of the microscopic equations lead to an increasing entropy. However, we have seen, this does not happen, as there are

---

[7] Remember in fact that the number of sequences containing N elements with k times 2 symbols (that is, the number of sequences of N elements that have frequency of 1s equal to k/N) is given by $g(k,N) = N(N-1)\ldots(N-k+1)/k!$.

[8] See Goldstein (2001). Others instead wish to justify it on Bayesian terms (Bricmont 2001, 2020).

always entropy-decreasing microstates. Nonetheless, and this is what typicality does, one can show that these microstates are very few, and that most microstates are entropy-increasing. In other words, it seems satisfactory to say that we have explained a phenomenon if we have shown that it holds for typical initial conditions, since we cannot do it for all of them. As Bricmont (2001, 2020) has pointed out, if something is typical, no further explanation seems to be required because we can expect this phenomenon to happen. For instance, the fact that a gas expands freely is not surprising: this is what we expect to happen, given what we have discussed so far. It would be surprising if it did not expand.[9] A less satisfactory explanation would be, on the contrary, one which is true only for very special initial conditions. In fact, one can always find an initial condition that will account for a phenomenon ("a gas completely distributed in a room which goes back into a corner of the room"), but this is not what one would expect.[10]

Also notice that in typicality explanations, details do not matter. Having proven that gases typically expand is enough: we do not need to show that gases expand with a 98% probability. To explain the second law of thermodynamics, we needed to show that most microstates behave suitably, without specifying exactly how many. We do not need to show that "entropy will increase with 0.98 probability;" rather it is enough to show that "entropy will typically increase."[11]

5. Typicality Talk and Probability Talk

In the previous sections I have shown how typicality has been introduced in a deterministic theory such as statistical mechanics. It is a measure on phase space that allows to suitably count the microstates in a macrostate, and in particular it is a measure that can be placed on initial conditions. In this section I wish to discuss how typicality can help understating deterministic probability talk. First, I argue that the probabilities which play a role in predictions and explanation, which therefore can be dubbed 'objective,' can be understood in terms of typicality as single world typical frequencies (Section 5.1). They guide belief in the sense that we are justified in expecting what is typical. Moreover, the probability talk dealing with ignorance, namely the 'subjective'

---

[9] This is connected to Hempel's notion of expectability (Hempel 1964). This is unsurprising, as the explanation based on typicality seems to be an extension of the deductive-nomological model (see also Wilhelm 2019, Allori 2020). Moreover, this idea is compatible with Cournot's principle: if an event has a very low probability, "then one can be practically certain that the event will not occur" (Kolmogorov 1933).

[10] For more on typicality and explanation, see Lazarovici and Reichert (2015), Wilhelm (2019), Crane and Wilhelm (2020), Allori (2020).

[11] For more on the relation between typicality and probability, see Volchan (2007), Goldstein (2011), Pitowsky (2012), Hemmo and Shenker (2012), Lazarovici and Reichert (2015), Wilhelm (2019), Allori (2020), Maudlin (2007, 2020), Hubert (2021).

or epistemic component of deterministic probabilities, are understood as being about what happens in different worlds compatible with the actual macrostate and can be used to make sense of single event probabilities.

5.1 Objective Probabilities: Predictions, Explanations, and Relative Frequencies

In this section I argue that the typicality measure can be used to define *single world typical relative frequencies* which capture the essence of what we have called 'objective deterministic probabilities' by playing the necessary explanatory and predictive roles. They are also prescriptive, in the sense that they are a guide to belief as they explain why we are confident that typical things will happen.

In the philosophical literature, objective probabilities are often connected with empirical distributions, as proposed by frequentist (actual or hypothetical) accounts according to which probabilities are fundamentally (actual or hypothetical) relative frequencies. These approaches have been criticized so much[12] that frequentism can be barely regarded as a living view. However, a new version of frequentism has been recently proposed according to which probabilities are long term typical frequencies.[13] This account shares some features with the so-called range account, which identifies probabilities with ranges of values in a suitable space of possible states.[14] Both the range account and typicality frequentism develop the idea that the probability of an event is the set of initial conditions leading to the event over the total initial conditions (von Kries 1886). However, while von Kries' account is epistemic (depending on the agent knowledge, the relevant sets may change), the range account uses the work of Poincare (1912) and & Hopf (1934) on the method of arbitrary functions to define probabilities in terms of suitable probability distributions.[15] The main problem with the range account is to justify the choice of the probability measure, which is solved in typicality frequentism because, in contrast with the range account, it uses a typicality measure dictated by the relevant physical situation.[16]

While I am sympathetic with typicality frequentism, I think it has some difficulties. As anticipated above, typicality frequentism takes the probabilities in deterministic theories to be relative frequencies understood in terms of typicality. I disagree with the idea that all deterministic probabilities are of this form: I think that there are subjective deterministic probabilities that can be interpreted in terms of

---

[12] Jeffries (1992), Hajek (1997, 2007, 2009).
[13] Hubert (2021).
[14] Abrams (2000), Rosenthal (2010, 2016), Strevens (2003, 2008, 2011, 2013).
[15] See also Myrvold's account of 'almost objective chances' (2016, 2020, 2021), which is based on the method of arbitrary functions, but which is however not objective.
[16] See Hubert (2021) for a discussion.

typicality, more in agreement with the spirit of the range account, as I discuss in the following section.

More importantly, in typicality frequentism each microstate in a given macrostate represents a possible universe, so that by saying that it is typical for a free gas to expand, one means that *in most universes* compatible with the current macrostate the free gas expands. This is also what happens in the Mentaculus vision, with respect to the probability measure. However, I do not think this is correct: by saying that it is typical for a free gas to expand, one should mean with this that *in most repeated experiments in this world* the free gas expands. In fact, to say that a theory explains a phenomenon should be to say that it accounts for why the phenomenon happens in our world, not in other worlds. If typical frequencies are to play the role of objective probabilities, they have to be relative frequencies *in a single world*, generated by repeating the same experiments with similar initial conditions. The relative frequency which we want to explain and predict with our theory is a well-defined regularity in a given class of repeated experiments with initial macroscopic preparations which are substantially identical. That is, when an experimenter prepares a set of repeated experiments in this way, she will obtain empirical regularities, or empirical distributions, which are the relative frequencies of the various outcomes. If we repeat *M* times *N* coin tosses (with *N* and *M* sufficiently big) they tend to display half heads and half tails (or half 1s and half 0s). That is, suppose we toss 100 coins 1000 times: the first series of 100 coin tosses X_1 is taken in San Diego at 5 pm today, the second X_2 is taken in Chicago tomorrow at 8 am, the third X_3 in Venice on December 8th at 4 pm and so on. The (actual) frequency of heads of the generic j-th series is $n(X_j) = \sum_i^{1000} N_i/100$, where N_i is the number of 1s in the j-th series ($j = 1,\ldots,1000$), and the set $\rho_{emp} = \{N(X_1), N(X_2), N(X_3), \ldots., N(X_{1000})\}$ is the relevant empirical distribution, which shows the pattern that, for most coin tosses, half will be 1s and half will be 0s. Alternatively, take a set of N gases concentrated in one corner of N similar boxes; let them evolve freely; check what has happened after 2 hours, say; record the N results: the first gas spreads out in the first box; the second spreads out in the second box; and so on. In general, the empirical distributions present patterns: for instance, most gases spread across the entire container. These are *single world* frequencies.

Instead *across-worlds* frequencies, which are the ones considered by typicality frequentism, are neither explanatory nor can have predictive power (however, see the next section). In fact, if we repeat the same experiment (for example, the free expansion of a gas initially concentrated in a corner of a box) in different universes (whatever this would mean), we observe that entropy increases in most of them. Nevertheless, this fact does not help understanding why in *this* universe, if we prepare the gas in a corner of a box, in most repetitions of this experiment it will spread all over, and it does not allow us to predict that in this universe most gases will expand. In other words, the relative frequencies that have an explanatory and predictive role are the distributions in the actual world: the (typical) frequencies in *this* universe.

These single world typical frequencies are explanatory in the sense that the distribution predicted by the theory matches the observed distribution. In particular, in statistical mechanics that means that the observed distribution ρ_{emp} and the theoretical distribution ρ_{theo} are close: $|\ \rho_{emp} - \ \rho_{theo}| < \varepsilon$., with some positive ε. In the example of the coin tosses, typicality explains the pattern that for the typical sequence half results will be 1s and half 0s: $|\frac{1}{N}\sum_{i=1}^{N} N_i - \frac{1}{2}| \ll \varepsilon$. That is, there exist atypical sequences X_k such that frequencies $n(X_k)$ is very different from ½ which therefore belong to the set of exceptions E, but they are such that they are very few (figure 3).

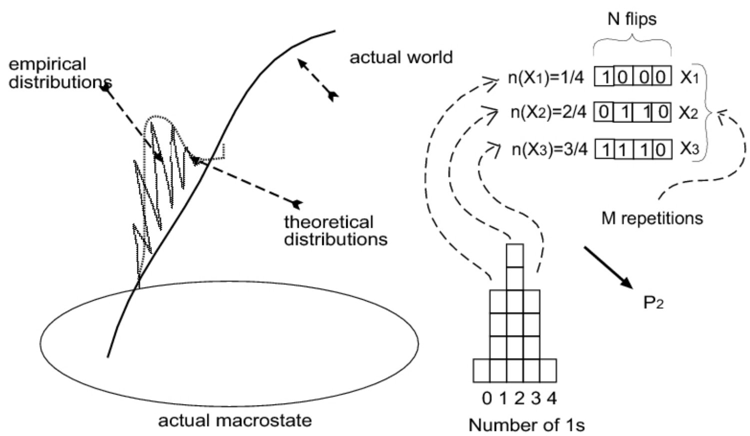

Figure 3: Empirical distributions are in one world, not across worlds

Therefore, the empirical distributions are explained if they are typical in this world as just explained: it's typical for free gases to expand, for coin tosses to display half heads and half tails, for ice cubes to melt, and so on. So, to conclude, I think that *single world typical frequencies* play the role of *objective deterministic probabilities* in our discourse are, and that, in virtue of their connection with typicality, they play the suitable explanatory and predictive roles that Loewer required of deterministic probabilities.

5.2 *Subjective Probabilities: Ignorance, Belief, Updating, and Single Events*

I have argued in the previous section that single world typical frequencies capture the essence of what we have dubbed objective deterministic probabilities: we talk about objective probabilities in a deterministic framework, we are talking about single world typical frequencies.

However, typicality can also be used to understand the probability talk which has to do with our lack of information. Consider a single event probability. Some have argued that single event probabilities are meaningful,[17] while others instead thought that 'the probability of winning this battle' has no place in the theory of probability.[18] Nonetheless, I think we can give a definition of probability of a single event as follows. What is the meaning of the sentence: "The probability of tossing a fair coin and getting heads is 50%"? We have seen it in the previous section: repeated tossing of fair coins will typically show half tails and half heads. However, what about the sentence: "The probability that tossing *this* fair coin and getting heads is 50%"? There is no repetition here, so we cannot describe it like before. I think however that this can be understood in terms of *across-world* frequencies: 50 microstates out of the 100 which are compatible with the actual macrostate of this coin will evolve into a macrostate where the coin has landed heads. In other words, there is a fact of the matter of what the actual microstate of the coin is, which determines how the coin will eventually land. However, we do not know what the actual microstate is, therefore we cannot compute where the coin will end up, so we have to consider all the microstates. Since we are spanning over the macrostate, these microstates represent possible worlds macroscopically similar to the actual world. If the coin is typical, most microstates compatible with the coin's macrostate will display a sequence of half heads and half tails. Therefore, we can think of the probability of getting heads in terms of volume in phase space. That is, the probability of getting heads is given by the number of possible microstates in the initial macrostate which lead to the final macrostate corresponding to heads, over the total number of microstates in the initial microstate: $P = V_{head}/(V_{head} + V_{tail})$, where V_{head} and V_{tail} are respectively the volumes in phase space of the microstate corresponding to the coin landing heads and landing tails. We could use the indifference principle because the ignorance about the actual microstate is real: we do not know which microstate is actual, and we have no reason to think something is privileged over the other (figure 4). This is essentially how to understand the subjective probability of a fair coin landing heads. Notice how this probability is *across-worlds*, so that it does not specify a feature of our world. It describes our ignorance about which world we are in. This approach is similar to the range account, which, unlike typicality frequentism, prescribes that probabilities are fundamentally connected to volumes in a relevant phase space.

This ignorance probability can be interpreted as the *degree of belief* of an external observer making an inference on the probability this coin will land heads. That is, the observer does not know which microstate they are in, so they do not know what will happen exactly, but they do know the statistics of the typical coin. This plays a role in the justification of why they have a given degree of belief about coins landing heads. The observer's confidence in the probability of this coin coming up head being ½ comes

[17] Hajek (2009) and La Caze (2016).

[18] von Mises quoted in Gilles (2000).

from the fact that we assume the world is typical and fair. If repeated coin tosses display an atypical distribution (such as all tails) one would expect the coin to be unfair and they will reshape the volumes in phase space accordingly. That is, given that probability is used here to comply with ignorance of the actual microstate, if we earn new information we can modify and update the set of possible accessible states.

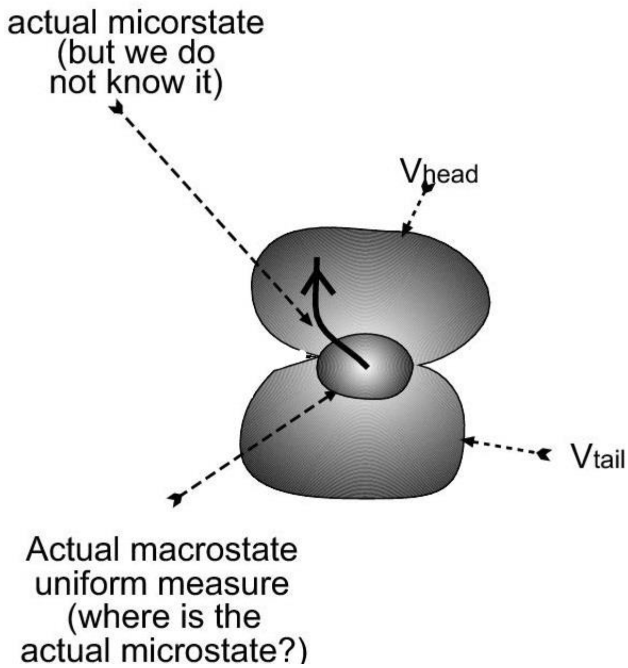

Figure 4: The phase space accessible to the coin is divided into two volumes corresponding to the coin landing heads and landing tail.

Given that the measure of typicality is temporally stationary, one can make sense of the fact that probabilities can be time indexed. Indeed, the subjective probability which guides degrees of belief changes once one updates their information about the actual microstate as discussed above.

Notice again that this subjective probability guiding beliefs does not reflect an objective feature of the world. Rather it is about our ignorance of the actual macrostate. Using the typicality measure to guide one's credence does not play any role in the explanation or the prediction of any physical fact. The fact that I do not know the actual microstate does not explain that, say, this coin lands heads. This is cashed out by thinking of single event probabilities as being about what happens to different worlds compatible with the actual macrostate, while objective deterministic probabilities are single world frequencies.

6. The Paradox is Solved

The paradox of deterministic probability is solved in my approach as follows:

1) There is a typicality measure over initial conditions.
   My approach differs from the Mentaculus vision in that it has a typicality measure rather than a probability measure, like typicality frequentism and (some versions of) the range account.
2) This typicality measure allows to define both objective and subjective deterministic probabilities.
   The role of objective probabilities is fulfilled by typical relative frequencies in our world, and this is distinct from all the previous approaches (Mentaculus, range account, typicality frequentism), which all think of objective probabilities as across worlds. The empirical distributions are explained by the theory if they are typical. In this approach, subjective probabilities are ignorance probabilities, which allow us to define single event probabilities. Unlike typicality frequentism, which remains silent on them, in my approach subjective probabilities can be accounted for: ignorance probabilities use the uniform distribution *as a practical rule* in order to predict some macroscopic behavior, and are defined in terms of ratios of volumes in phase space.
3) Only objective deterministic probabilities play the necessary predictive and explanatory roles we want probability to have.
   Objective probabilities explain by looking at single world frequencies and sowing they are typical. In contrast, subjective probabilities run across worlds. And this is compatible with the fact that in my approach subjective probabilities are neither explanatory nor predictive, in contrast with what happens within the Menatculus vision.

Acknowledgements

I am grateful to an anonymous reviewer for their insightful comments, and to the participants of the 2014 Indiana Philosophical Association for their helpful remarks on a previous version of this paper. In addition, I would like to thank Jean Bricmont for useful and enlightening discussions on topics connected with this paper.

References

- M. Abrams. "Mechanistic Probability." *Synthese* 187(2): 343–75 (2012).
- D. Z. Albert. *Time and Chance*. Harvard University Press (2000).
- V. Allori. "Some Reflections on the Statistical Postulate: Typicality, Probability and Explanation between Deterministic and Indeterministic Theories." In: V. Allori (ed.), *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism, and Laws of Nature*: 65-111. World Scientific (2020).
- J. Bricmont. "Science of Chaos or Chaos in Science?" *Annals of the New York Academy of Sciences* 775 (1): 131–175 (1995).
- J. Bricmont. "Bayes, Boltzmann and Bohm: Probabilities in Physics." In: J. Bricmont, D. Dürr, M.C. Galavotti, G.C. Ghirardi, F. Petruccione, and N. Zanghi (eds.) *Chance in Physics: Foundations and Perspectives*, Lecture Notes in Physics 574, Springer-Verlag (2001).
- J. Bricmont. "Probabilistic Explanations and the Derivation of Macroscopic Laws". In V. Allori (ed.) *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*: 36-61. World Scientific (2020).
- C. Callender. "There is No Puzzle about the Low Entropy Past". In: C. Hitchcock (ed.). *Contemporary Debates in Philosophy of Science*: 240-255. London: Blackwell (2004).
- A. A. Cournot. *Exposition de la Theorie des Chances et des Probabilites*. Paris: Librairie de L. Hachette (1843).
- R. Frigg. "Typicality and the Approach to Equilibrium in Boltzmannian Statistical Mechanics." *Philosophy of Science* 76: 997–1008 (2009).
- R. Frigg. "Why Typicality Does Not Explain the Approach to Equilibrium." In: M. Suarez (ed.) *Probabilities, Causes and Propensities in Physics*: 77-93. Synthese Library. Springer (2010).
- D. Gillies. *Philosophical Theories of Probabilities.* London: Routledge (2000).
- S. Goldstein. "Boltzmann's Approach to Statistical Mechanics." In: J. Bricmont, D. Dürr, M.C. Galavotti, G.C. Ghirardi, F. Petruccione, and N. Zanghi (eds.) *Chance in Physics: Foundations and Perspectives*, Lecture Notes in Physics 574, Springer-Verlag (2001).
- S. Goldstein. "Typicality and Notions of Probability in Physics." In: Y. Ben-Menahem and M. Hemmo (eds.) *Probability in Physics*: 59–71. Heidelberg: Springer (2012).
- S. Goldstein, and J. Lebowitz. "On the (Boltzmann) Entropy of Nonequilibrium Systems." *Physica D* 193: 53–66 (2004).
- A. Hájek. "'Mises Redux' – Redux: Fifteen Arguments Against Finite Frequentism." *Erkenntnis* 45 (2/3): 209–27 (1996).
- A. Hájek. "The Reference Class Problem is Your Problem Too." *Synthese* 156(3): 563–85 (2007).
- A. Hájek. "Fifteen Arguments Against Hypothetical Frequentism." *Erkenntnis* 70(2): 211–35 (2009).

- M. Hemmo and O. Shenker. *The Road to Maxwell's Demon: Conceptual Foundations of Statistical Mechanics*. Cambridge University Press (2012).
- C. G. Hempel. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. The Free Press, New York (1965).
- C. Hoefer. "The Third Way on Objective Probability: A Sceptic's Guide to Objective Chance." Mind 116 (463): 549–96 (2007).
- C. Hoefer. "Physics and the Humean Approach to Probability." In: C. Beisbart and S. Hartmann (eds.) *Probabilities in Physics*: 321–37. New York: Oxford University Press (2011).
- C. Hoefer. *Chance in the World: A Humean Guide to Objective Chance*. New York: Oxford University Press (2019).
- E. Hopf. "On Causality, Statistics, and Probability." *Journal of Mathematics and Physics* 13: 51-102 (1934).
- M. Hubert. "Reviving Frequentism." *Synthese*:1-30 (2021).
- R. Jeffrey. "Mises Redux." In: *Probability and the Art of Judgment*: 192-202. Cambridge: Cambridge University Press (1992).
- A. Kolomogoroff. "Grundbegriffe Der Wahrscheinlichkeitsrechnung." *Ergebnisse Der Mathematik Und Ihrer Grenzgebiete.* 1. Folge. Springer-Verlag, Berlin Heidelberg (1933).
- A. La Caze. "Frequentism." In: A. Hájek and C. Hitchcock (eds.) *The Oxford Handbook of Probability and Philosophy*: 341–59. Oxford: Oxford University Press (2016).
- D. Lazarovici and P. Reichert. "Typicality, Irreversibility and the Status of Macroscopic Laws." *Erkenntnis* 80 (4): 689–716 (2015).
- J. Lebowitz. "Microscopic Dynamics and Macroscopic Laws." *Annals New York Academy of Sciences*: 220-233 (1981).
- J. Lebowitz. "Boltzmann's Entropy and Time's Arrow.: *Physics Today* (September): 32–38 (1993a).
- J. Lebowitz. "Macroscopic Laws, Microscopic Dynamics, Time's Arrow and Boltzmann's Entropy." *Physica A* 194: 1–27 (1993b).
- J. Lebowitz. "Statistical Mechanics: A Selective Review of Two Central Issues. *Reviews of Modern Physics* 71: 346–357 (1999).
- B. Loewer. "Determinism and Chance." *Studies in the History of Modern Physics* 32 (4): 609–620 (2001).
- B. Loewer, "David Lewis's Humean Theory of Objective Chance." *Philosophy of Science* **71**(5): 1115–1125 (2004).
- B. Loewer. "The Mentaculus Vision." In: V. Allori (ed.) *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism, and Laws of Nature.* World scientific (2020).
- T. Maudlin. "What Could be Objective about Probabilities?" *Studies in History and Philosophy of Modern Physics*, 38(2):275–91 (2007).

- T. Maudlin. "The Grammar of Typicality." In: V. Allori (ed.) *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism, and Laws of Nature*: 231–251. World scientific (2020).
- W. C. Myrvold. "Probabilities in Statistical Mechanics." In: A. Hájek and C. Hitchcock (eds.) *The Oxford Handbook of Probability and Philosophy*. Oxford: Oxford University Press, (2016).
- W. C. Myrvold. "Explaining Thermodynamics: What Remains to be Done?" In: V. Allori, (ed.) *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*. World Scientific (2020).
- W. Myrvold *Beyond Chance and Credence: A Theory of Hybrid Probabilities*. Xford University Press (2021).
- H. Crane and I. Wilhelm. "The Logic of Typicality." In: V Allori (ed.) *Statistical Mechanics and Scientific Explanation*: *Determinism, Indeterminism and Laws of Nature*: 173-229. Worlds Scientific (2020).
- D. Ruelle. *Chance and Chaos*. Princeton Science Library (1991).
- R. Penrose. *The Emperor's New Mind.* Oxford: Oxford University Press (1999).
- I. Pitowsky. "Typicality and the Role of the Lebesgue Measure in Statistical Mechanics." In: J. Ben-Menahem and M. Hemmo (eds.): *Probability in Physics* 51-58. Berlin and New York: Springer (2012).
- H. Poincare'. *Calcul des probabilites*. Paris: Gauthier-Villars (1912).
- K. Popper, *Quantum Theory and the Schism in Physics*. New York: Rowman & Littlefield (1985).
- J. Rosenthal. "The Natural-Range Conception of Probability." In G. Ernst and A. Hüttemann (eds.) *Time, Chance and Reduction: Philosophical Aspects of Statistical Mechanics*: 71–91. Cambridge, UK: Cambridge University Press (2010).
- J. Rosenthal. "Johannes von Kries's Range Conception, the Method of Arbitrary Functions, and Related Modern Approaches to Probability. *Journal for General Philosophy of Science* 47(1): 151–70 (2016).
- M. Strevens. "Do Large Probabilities Explain Better?" *Philosophy of Science* 67: 366–90 (2000).
- M. Strevens. *Bigger than Chaos: Understanding Complexity through Probability*. Cambridge, MA: Harvard University Press (2003).
- M. Strevens. *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press (2008).
- M. Strevens. "Probability out of Determinism." In: C. Beisbart and S. Hartmann (eds.) *Probabilities in Physics*: 339–64. New York: Oxford University Press (2011).
- M. Strevens. *Tychomancy: Inferring Probability from Causal Structure*. Cambridge, MA: HarvardUniversity Press (2013).

- S. B. Volchan. "Probability as typicality." *Studies in History and Philosophy of Modern Physics* 38 (4):801–14 (2007).
- J. von Kries. *Die Principien der Wahrscheinlichkeitsrechnung*. J. C. B. Mohr (1886).
- R. von Mises. *Probability, Statistics and Truth*. New York: Macmillan, 1928/1957.
- I. Wilhelm. "Typical: A Theory of Typicality and Typicality Explanation," *The British Journal for the Philosophy of Science*. (2019).