Essay Review

Probability in Classical Statistical Mechanics

Janneke van Lith[*]

Y.M. Guttmann, *The Concept of Probability in Statistical Physics* (Cambridge: Cambridge University Press, 1999), xi + 267 pp; ISBN 0-521-62128-3 (hardback); price = £35.00, $54.95.

1 Introduction

How can probabilistic statements be understood in the deterministic theory of classical statistical mechanics? Are they epistemic, i.e. do they reflect our ignorance of the exact state of a system, or objective, in the sense that they represent relative frequencies of some sort? Are there perhaps other options? These questions are central in Guttmann's book.

The core of the book is formed by the discussion of four positions, spread over epistemic and objective interpretations of probability, which all get a 'charitable exposition' (to use Guttmann's words) but are in the end dismissed. In the fifth chapter, emerging from the foregoing discussion, three different solution schemes come to the fore. The book ends with three appendices containing varied background information.

The author is undeniably original. Especially the chapters on the Haar measure (Ch. 3) and on topological dynamics (Ch. 4) constitute a fresh line of approach, and contain stimulating new ideas. In these chapters the striking connections between measure theoretical and topological notions are explored, and the implications for the philosophical foundations of statistical mechanics are investigated. Guttmann gives a comprehensive account of these two frameworks for statistical mechanics, which each in their own way provide a method of dealing with incomplete information. Within measure theory lack of information about the exact phase point of a system is translated into a statement about the probability that the system is in a certain region in phase space. With the tools of topological dynamics one can describe the approximate behaviour of systems in terms of nearby trajectories. Guttmann also discusses surprising connections such as the 'duality theorem' (p. 158) which allows one to translate theorems directly from one framework into the other.

In many respects the book falls somewhat outside the traditional literature on the foundations of statistical mechanics. This could have made it a highly valuable contribution to this area, but unfortunately it doesn't live up to its promise. Inaccuracies and errors, both in argumentation and in technical detail, are present in too large numbers. Therefore, the book is not suited as a first introduction to the subject (as it is meant to be). Only someone who is already familiar with the subject, and prepared to check everything, will be able to extract valuable ideas.

That these errors may have far-reaching consequences is illustrated by the following quotation:

> 'The canonical distribution, when restricted to any subensemble, remains essentially the same. This fact, in and of itself, implies that we cannot improve our predictions by taking into account additional information.' (p. 210)

[*]Department of Philosophy, Heidelberglaan 8, 3584 CS Utrecht, The Netherlands. E-mail: Janneke.vanLith@phil.uu.nl.

Guttmann repeatedly stresses the importance of this observation, and to a large extent bases his judgement about the ergodic approach and the works of Jaynes and Gibbs on it (see pp. 53, 59–60, 197–198, 210, 214). However, it is blatantly wrong. When restricted to a sub*system* the canonical distribution retains its exponential form. But suppose for instance that new information reveals that the system's energy lies in the tail of the canonical distribution. Certainly restriction to the sub*ensemble* of systems with this energy value will improve our predictions.

An example of mathematical inaccuracy is Guttmann's discussion of the problem of "updating" in the theory of subjective probabilities. Here, the question is how to revise a prior probability assignment when new information is received. He discusses a theorem by Seidenfeld (1986, pp. 471–473), which presents conditions under which two rival methods of updating, namely the Maximum Entropy formalism and Bayesian conditionalisation, agree. Guttmann reports that there is no conflict between these two methods when 'new information leads us to a new probability distribution p^* that concentrates on a set of events with [prior] probability zero or [prior] probability one' (p. 59). In fact, however, Seidenfeld's result proves that there is no conflict if all constraints under which entropy is maximised are of the form $E[I_e] = 0$ or $E[I_e] = 1$ (i.e. the expectation value of the indicator function for some event e is either zero or one). Thus, he formulates particular conditions on the constraints, not on the posterior distribution p^*. Another example is that Guttmann incorrectly limits the validity of the ergodic theorem (p. 77) to ergodic evolutions.

There are innumerable typos and sloppy notation, which make the text difficult to read. Pages 47–48 are a pain to the eye, where the symbol x has two different meanings within a single expression, where symbols are missing, and where ∞ is denoted as 00. Another example is pp. 137–139; here B denotes both an algebra and a specific element of that algebra, and some errors are made in copying from (Banach, 1937): $\{b_n\} \subset S_n$ in the fifth axiom of congruence should read $\{b_n\} \subset G_n$, and the order of limits in P4 is wrong. These make it impossible to understand the text without checking the reference oneself.

My general opinion is that due to its sloppiness this book has missed the opportunity to become the standard reference on probability in statistical physics. Still it offers enough interesting material to be of value to the specialist, and I hope especially that the parts on the connection between measure theory and topological dynamics (Ch. 4, Sections 5.2 and 5.3) will encourage further research. In the following I will single out two issues: ergodic theory (Ch. 2) and the Haar measure (Ch. 3).

2 Ergodic theory

The branch of mathematics called ergodic theory was developed in the early twentieth century in connection with problems in the foundations of statistical mechanics. Ergodic theory is clearly relevant for foundational issues, but still opinions diverge on the role it can play, the particular questions that it addresses and the answers it can give (see Van Lith, 2001 for a survey of the foundational roles of ergodic theory). An example is the very interesting outline of a subjectivist approach based on ergodic theory which Guttmann presents in his second chapter. His description of the standard ergodic approach is also surprising and unorthodox, as I will indicate below. Familiar theorems appear in a different perspective when used in an unexpected new foundational role.

The orthodox use of ergodic theory in the foundations of statistical mechanics proceeds via Birkhoff's ergodic theorem and its corollary. The theorem states that for a measure-preserving dynamical system infinite time averages exist for almost all initial conditions. The corollary states that if the system obeys a certain dynamical condition, called metrical transitivity or ergodicity, those infinite time averages are equal to the microcanonical phase average, again for almost all initial conditions. Physics textbook wisdom has it that thermodynamic measurements take such a long time compared to microscopic relaxation times that they can be taken to yield infinite time averages. The above ergodic theorems then provide the justification for calculating microcanonical phase averages for predicting measurement results.

In his discussion of the ergodic approach Guttmann gives a non-orthodox twist to this story by *identifying* probabilities with infinite time averages. That is, Guttmann considers a frequency interpretation of probability where the repetitions occur in a single system in the course of time, rather than (as is standard) across an ensemble of identically prepared systems. With probabilities identified with time averages, the ergodic theorem plays a different role than outlined above: It demonstrates under what conditions probabilities are equal to the microcanonical measure. But probabilities so understood cannot change in the course of time, making an embedding of this interpretation of equilibrium statistical mechanics into the general non-equilibrium theory awkward if not impossible. The appendix III is evidence of this; here Guttmann nevertheless tries to extend his version of frequentism to non-equilibrium statistical mechanics, resulting in such unfortunate claims as that a probability density that evolves in the infinite time limit into a stationary density is merely a "notational variant" of the latter.

Very interesting is the subjectivist version of the ergodic approach, presented in Sections 2.7 and 2.8. Here Guttmann presents a generalisation of the well-known representation theorem of De Finetti, and the opportunities this theorem offers for a subjective interpretation of probability. The theorem (the ergodic decomposition theorem) says that every stationary probability measure on a bounded space S can be decomposed uniquely into ergodic components (i.e. into probability measures that are ergodic on a subspace of S). The ergodic components in turn can be connected to relative times of sojourn in their respective subspaces, and so with objective features of the system. This means that every personal probability ascription, as long as it is stationary, can be represented as a subjective expectation of objective frequencies.

This representation theorem holds for stationary measures only. Guttmann makes an effort to justify stationarity, and discusses four possible justifications that according to him all fail. This is because to a subjectivist only coherence requirements are compelling, but stationarity cannot be enforced on grounds of coherence. But the same can be said of exchangeability in De Finetti's original theorem and still the representation theorem is seen as very important in that context.

3 Haar measure

According to the corollary of the ergodic theorem that holds for metrically transitive systems, there may be a measure zero set of exceptions to the equality of time and phase averages. In the third chapter, Guttmann makes a heroic attempt to get rid of these exceptions. He does so by considering the Haar measure on the group of Hamiltonian time evolution operators, which is supposed to form a bridge between the Lebesgue measure on the one hand, and the time measure (i.e. the measure that attributes to each set the relative time of sojourn), and thus infinite time averages on the other. The goal of the chapter is to investigate the dynamical conditions needed to establish that this bridge is one of strict identity. A brave goal indeed, since counterexamples can easily be imagined in the presence of metrical transitivity, thus demonstrating that the sought-for dynamical conditions must be even stronger. Such counterexamples of course form a set of measure zero, but nevertheless concrete cases can be pointed out; for instance periodic phase trajectories generally yield time averages very different from phase averages. It is therefore not a surprise that in the end Guttmann dismisses the Haar construction because of limited applicability. Nevertheless it contains many interesting and valuable ideas.

Let me outline the way in which Guttmann connects the Haar measure to both phase and time measures. In any book on topological groups one can find a treatment of the Haar measure, being the measure on a topological group which is invariant under the group action. The existence and uniqueness of such a measure were shown by Haar in the 1920s by an explicit construction. In the present application, the group consists of the Hamiltonian time evolution operators, which is a one-parameter transformation group. If this group is represented by the time parameter, its Haar measure is given by the Lebesgue measure on the real line. However, it is not the Lebesgue measure on the real line, but on the phase space that is of interest for the present application. Guttmann manages to make a connection with the latter by modifying Haar's original construction. Also, he argues that the measure thus constructed on phase space is equal without exception to the time measure. Let me explain this in some detail.

The key notion is that of congruence. Let $\langle S, \mathcal{B} \rangle$ be a measurable space. Following Banach's treatment of the Haar measure, Guttmann presents an axiomatic definition of a congruence relation between elements of $\mathcal{B}$. Next, a function is defined in terms of this notion of congruence which is shown to be a measure; basically, it assigns equal values to congruent sets. The important point is that the axioms of congruence can be satisfied, in certain cases, by use of a group: Two sets A and B are congruent iff there is an element g of the group connecting them, $gA = B$. The connection with the time measure is guaranteed by the fact that the group at issue is the group of time evolutions. However, in order for this construction to work, the group has to be *transitive*, i.e. for all points $a, b \in S$ there must be a transformation g such that $ga = b$.

In the original Haar construction the space S is the topological group itself. Therefore, transitivity isn't a problem: It follows immediately from the group structure. But in Guttmann's version of it, S is the phase space. Transitivity now presents itself as a huge obstacle. In fact, it demands that all points in phase space lie on a single trajectory; this is nothing less than the original Ergodic Hypothesis which is known to be false if the dimension of phase space is larger than one!

The rest of the chapter is a struggle to surmount the requirement of transitivity. Guttmann discusses several possibilities to weaken transitivity, and investigates in those cases whether a congruence relation can still be defined. It is a pity that he loses sight of the aim of connecting phase and time averages in this discussion; that is, he does not check whether the modified congruence relations still enable one to define a measure that (like the Haar measure) is connected to both phase and time averages.

So what does this leave us with? Does Guttmann make headway with respect to the ergodic approach, or does the Haar construction only work for systems obeying the Ergodic Hypothesis? In fact there is one small victory here, but Guttmann forgets to stress it: For so-called minimal systems (p. 143), i.e. systems for which every phase trajectory is dense, the group of time evolution operators defines a congruence relation that indeed leads to the standard measure on phase space.

4 Three solutions?

The book ends with three proposed solutions that emerge from the foregoing discussion. Here is how Guttmann formulates the problem:

> 'If probabilities are not physical parameters that we discover by methods of observation and measurement only, how can we justify our willingness to be guided by them? How can we explain the utility of probabilities?' (p. 190)

4.1 The extension of the subjectivist framework

The first solution is subjectivist in spirit, and consists in a subjective interpretation of both probabilities (as degrees of belief) and distances between phase points (as judgements of similarity). In order to find the "subjective distance" between two points s and s' that represent the system M in phase space, one has to imagine two hypothetical "epistemic situations" E and E' formed by combining one's current body of beliefs supplemented with the knowledge that M is in state s or s' respectively. Next, one has to make a "similarity judgement" of E and E'. Thus, s and s' are close iff E and E' are similar. Guttmann makes it plausible that this concept of subjective distance indeed obeys the metric axioms. Just like subjective probabilities, subjective distances may vary from person to person. Also, subjective distances need not coincide with objective notions of distance such as distance along a phase trajectory.

Guttmann himself isn't entirely satisfied with this solution, but I think he pitches his demands high. First, he sees problems that are based on the confusion between subensembles and subsystems mentioned earlier (see p. 198). Second, he jumps too quickly from giving an interpretation of probability, and explaining its utility, to justifying a particular expression for the probability distribution. This appears from his discussion of regularity (pp. 196–197), which is a property of the microcanonical measure that connects measure theoretical and topological notions. According to Guttmann the fact that the microcanonical measure is regular imposes extra conditions for subjective probabilities and distances, but that is only the case if the goal is to justify the microcanonical measure, rather than the utility of probabilities in general. A consistent subjective interpretation of both probabilities and distances seems perfectly possible.

4.2 Reformulation of the aims of the ergodic approach

As his second solution Guttmann presents a framework which encompasses both measure theory and topological dynamics. The central notion is an abstract system $\langle S, \mathcal{B}, W, T \rangle$, which generalises both dynamical and topological systems. As usual S is the phase space and T the group of time evolution operators. $\mathcal{B}$ is a set of subsets of S, and it generalises the algebra (measure theory) and the collection of open sets (topological dynamics). W is the subset of $\mathcal{B}$ of sets that count as "small"; this generalises measure zero sets (measure theory) and sets of first category (topological dynamics).

Guttmann gives a convincing presentation of the abstract theory that so emerges. Various quantities that have played important roles in earlier chapters (ergodicity, relative frequencies, recurrence, boundedness) are now defined in terms of abstract systems. And indeed, by choosing either the measure theoretical or the topological dynamical interpretation of $\mathcal{B}$ and W, the abstract notions reduce to the familiar ones. The framework gives a beautiful insight in the structural similarity of the two modes of description of statistical mechanical systems.

But does it deliver what Guttmann claims, namely that 'within this very elementary framework, we can arrive at satisfactory explanations of the emergence of equilibrium in particular and the phenomenon of deterministic randomness in general' (p. 200)? I do not think so. In order to explain the mentioned phenomena, some connection between the abstract formalism and the empirical world has to be established. Presenting a synthetic framework does not relieve one of the task of giving an interpretation of the concepts of the theory, be it measure theoretic, or topological dynamical, or some alternative. Thus, what Guttmann presents here falls short as an answer to the interpretation problem. It is more like refraining from interpretation.

4.3 Pragmatist foundations of SM

In a sense, pragmatism is an easy and attractive standpoint. Instead of explaining the utility of probabilities one could just accept this utility and take advantage of it. Indeed, as Guttmann argues, it is not necessary either to reduce probabilities to objective physical properties or to construe them as degrees of belief. But even a pragmatist has to give some interpretation to probabilities, if only to be able to judge whether the theory is in agreement with empirical reality. Here, the pragmatist solution that Guttmann presents does not come into its own.

The pragmatist approach that Guttmann presents as his third solution to the interpretation question is the "Gibbs-Khinchin formulation" of statistical mechanics. He claims that Khinchin's contributions to the foundations of statistical mechanics add to Gibbs's pragmatist viewpoint the necessary justification for using the (micro-)canonical distribution. I think this is based on a misconception of Khinchin's theory (1949). According to Guttmann, Khinchin *derived* the microcanonical distribution by using the central limit theorem (p. 213). But in fact Khinchin *presupposed* the microcanonical distribution, and just applied the central limit theorem to it in order to investigate its asymptotic behaviour. So the question where the microcanonical distribution comes from still stands. After all, even Khinchin himself did not seek the answer to this question in his central limit theorem but in ergodic theory.

None of these solutions is worked out in the necessary detail, and Guttmann acknowledges this fact. In broad lines I agree with the author on the possible directions where solutions can be found, though one may wonder whether pragmatism doesn't simply boil down to evading the question. A subjective interpretation of probability seems unproblematic, though one then has to accept that probabilities may differ from person to person, and thus that it is impossible to justify the particular probability distributions that are used in statistical mechanics on coherence grounds.

The combination of measure theory and topological dynamics is especially promising, although I have a somewhat different approach in mind, namely the one offered by Malament and Zabell (1980). They argue that probabilities in statistical mechanics should obey a certain topological notion (continuity under translations), which implies absolute continuity of the measure (a measure theoretical notion). Together with some other assumptions (to wit, ergodicity and stationarity) this singles out the standard measure uniquely, and thus gives a justification for the standard measure (see also Vranas, 1998). This however raises further questions about the justification of the topological notions at issue. For a detailed study of such questions Guttmann's book offers plenty of points of departure.

Acknowledgements – It is a pleasure to thank Jos Uffink and Dennis Dieks for helpful discussion.

References

[1] Banach, S. (1937) 'On Haar's Measure', in: S. Saks, *Theory of the Integral*, Monografie Matematyczne T. VII (Warszawa: Z Subwencji Funduszu Kultury Narodowej), pp. 314–319.

[2] Khinchin, A.I. (1949), *Mathematical Foundations of Statistical Mechanics*, (New York: Dover).

[3] Malament, D.B. and Zabell, S.L. (1980) 'Why Gibbs Phase Averages work – the Role of Ergodic Theory', *Philosophy of Science* **47**, pp. 339–349.

[4] Seidenfeld, T. (1986) 'Entropy and Uncertainty', *Philosophy of Science* **53**, 467–491.

[5] Van Lith, J.H. (2001) 'Ergodic Theory, Interpretations of Probability and the Foundations of Statistical Mechanics', *Studies in History and Philosophy of Modern Physics* **32**(4), pp. ?? – ??.

[6] Vranas, P.B.M. (1998) 'Epsilon-Ergodicity and the Success of Equilibrium Statistical Mechanics', *Philosophy of Science* **65**, pp. 688–708.