Unit 1: Central limit theorem and diffusion

Introductory remarks: What is Statistical Physics?

31 Jan

Mathematical sciences such as physics aim to determine the laws of nature and understand how these govern experimental observations—both in everyday circumstances and under extreme conditions. This mathematical understanding is typically guided by reproducing a set of observations, with the resulting framework then used to make predictions for other "observables".

Over the past few centuries this process has been tremendously successful, with theoretical physics accurately predicting experimental and observational results from sub-atomic through to supra-galactic scales. Modern physics labs can create a vacuum better than in outer space and the coldest temperatures in the known universe, as well as going to the other extreme to reach temperatures of millions of degrees and pressures millions of times atmospheric pressure at sea level. Amazingly, many aspects of these realms of physics can be theoretically described by mathematics developed centuries ago.[1]

Statistical physics is one domain in which simple mathematical principles enable amazing predictive capabilities. Initially developed in the nineteenth century, statistical physics remains a pillar of modern physics, and will retain this position in years to come. The foundations of statistical physics lie in the use of probability theory to mathematically describe experimental observations and corresponding laws of nature that involve stochastic randomness rather than being perfectly predictable.

The lack of perfect predictability in statistical physics is a matter of practicality rather than one of principle. It results from working with a large number of degrees of freedom—that is, a large number of independent objects such as atoms. For illustration, Avogadro's number $N_A \approx 6.022 \times 10^{23}$ is the large number of molecules in everyday amounts of familiar substances—about 18 grams of water or about 22 litres of air at sea-level atmospheric pressure ($\approx$101 kPa). Specifying the positions and velocities of $\sim 10^{23}$ objects would require far more information than could be stored even in the memory of the biggest existing supercomputers. Statistical physics instead produces simple mathematical descriptions of large-scale properties such as temperature, pressure and diffusion, which are generally of such outstanding quality that the underlying 'randomness' is effectively invisible.

Historically, the difficulty of detecting the stochastic processes underlying such *thermodynamic* properties made it challenging to convince skeptics that atoms and molecules really exist. Ludwig Boltzmann, a prominent early developer of statistical physics, endured a constant struggle to defend his ideas, which likely contributed to his deteriorating mental health and eventual suicide in 1906. A crucial advance to convincingly establish the existence of atoms was Albert

[1] Eugene Wigner's famous article, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" (1960), and subsequent work in the philosophy of physics, elaborates on why this may be considered 'amazing'. These lecture notes will not comment extensively on philosophy.

Einstein's use of statistical physics to explain the observed "Brownian motion" of particles suspended in fluids—this work was part of Einstein's "miracle year" in 1905, along with special relativity and early contributions to quantum physics. More recent applications of statistical physics include explaining why stars don't collapse under the 'weight' of their own gravity, and identifying effects of dark matter in temperature fluctuations observable in the *cosmic microwave background* lingering from the early years of the universe.

In this unit we will focus on some of the foundational mathematics that will underlie our later development and application of statistical ensembles. Looking back to Boltzmann's times, we can consider the following question one of his opponents might have asked: *If the pressure of a gas in a container results from molecules stochastically colliding with the walls of that container and pushing them out, then how can the pressure be so stable and reproducible, rather than itself fluctuating stochastically?* The mathematical answer lies in the **law of large numbers** and the **central limit theorem**, which we will learn and apply to the physics of diffusion in one dimension.

1.1 Probability foundations

We begin by building a more formal mathematical framework around the concept of probability, through a sequence of definitions.

First, a random **experiment** $\mathcal{E}$ involves setting up, manipulating and/or observing some (physical or hypothetical) system with some element of randomness. Flipping a coin is a simple random experiment. In the context of the statistical ensembles that will be the focus of this module, a typical experiment will be to allow a collection of particles to evolve in time, subject to certain constraints.

Each time an experiment is performed, the world comes out in some **state** ω. The definition of the experiment and the state must include all objects of interest, and may include more besides. When flipping a coin, for example, the full state could contain information not only about the final orientation of the coin, but also about its position—did it land on the floor or on a cat?

The **set of all states** Ω collects all possible states ω that the given experiment $\mathcal{E}$ can produce, and is therefore intricately tied to $\mathcal{E}$ itself.

We are generally not interested in all aspects of the full state ω. For example, we won't care where a flipped coin lands. Instead we're typically only interested in whether it lands heads up or tails up—and we may want to set aside any state that doesn't cleanly map on to those options. The **measurement** $X(\omega)$ extracts and quantifies this information, acting as a function that maps the state ω to a number that we can mathematically manipulate. If we repeat a fixed experiment $\mathcal{E}$ many times and carry out the measurement X on each resulting state ω, we will obtain a sequence of numbers $X(\omega)$ that behave as a *random variable*.

Acting with the measurement X on all of the possible states in the set Ω

defines the **set of all outcomes** (or **outcome space**) A:

$$X : \Omega \to A.$$

That is, A collects all possible measurement results that the given experiment $\mathcal{E}$ and measurement X can produce. A can be finite, countably infinite, or uncountably infinite (i.e., continuous).

Let's consider some examples to clarify these definitions. With an experiment of rolling a six-sided die and measuring the number (1–6) that comes out on top, what is the set of all outcomes A? What additional information could be present in a given states ω?

$A = \{1, 2, 3, 4, 5, 6\}$

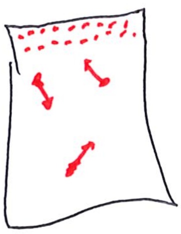
$\omega \supset$ position, orientation, temperature, colour, phase of moon

What is the outcome space A if we toss a coin four times and each time measure whether it lands heads up (H) or tails up (T)?

$A = \{$ HHHH, HHHT, HHTH, HHTT, ... THTT, TTHT, TTTH, TTTT $\}$

$\#A = 2^4 = 16$

What information could characterize a state ω for a gas of 10^{23} argon atoms in a container? What might be interesting to measure?



$\omega \supset \{$ 10^{23} positions, 10^{23} velocity, isotope (mass), electronic state & spin orientation etc... $\}$

Measure

Temperature

Volume → density (gradients)

Pressure

We can generalize the concept of measurement by introducing a unique number as a *label* to characterize each state ω in the set Ω. This would provide a label function $L(\omega)$ as a random variable. Our condition of uniqueness makes $L(\omega)$ isomorphic, so that the label can be used interchangeably with the full state,

$$\omega \longleftrightarrow L(\omega).$$

While the measurements $X(\omega)$ we consider will generally not produce a unique number for each ω, we will design them precisely to remove irrelevant information that doesn't interest us. Ignoring that irrelevant information leaves us free to interchange the set of outcomes A for the set of states Ω. (Some textbooks may never distinguish between A vs Ω in the first place, though this can be a source of confusion.)

Only a couple of definitions remain. The next is to define an **event** to be any subset of the set of all outcomes A. For example, events resulting from rolling a die could include (i) rolling a 6, (ii) rolling anything but a 6, (iii) rolling any even number, and many more. Collecting all events of interest defines the **set of events** (or **event space**) $\mathcal{F}$.

We are now prepared for the final foundational definition in this section, the **probability** P of an event in the set $\mathcal{F}$. Mathematically, P is a *measure function*, 31 Jan

$$P : \mathcal{F} \to [0, 1],$$

which must satisfy the following two requirements:

1. The probability of a countable union of mutually exclusive events must equal the sum of the probabilities of each of these events.

2. The probability of the outcome space ($\mathcal{F} = A$) must equal 1 (even if A is uncountable). This simply means that the experiment $\mathcal{E}$ must produce an outcome. If no outcome were produced, it would not make sense to say that the experiment had occurred.

Combining the outcome space, event space and probability measure gives us a *probability space* $(A, \mathcal{F}, P)$.

For example, consider an experiment that can only produce N possible states, so that

$$\Omega = \{\omega_1, \omega_2, \cdots, \omega_N\}.$$

If two states are identical, $\omega_i = \omega_j$, they must produce the same measurement outcomes $X(\omega_i) = X(\omega_j)$, which implies the contrapositive

$$X(\omega_i) \neq X(\omega_j) \qquad \Longrightarrow \qquad \omega_i \neq \omega_j.$$

On the other hand, as described above, it is possible to have $X(\omega_i) = X(\omega_j)$ even when $\omega_i \neq \omega_j$. This means that the size n of the outcome space A may be smaller than the size of Ω, $n \leq N$. We can write

$$A = \{X_1, X_2, \cdots, X_n\},$$