Time-Dependent Statistical Mechanics
3B. Elementary Probability Theory (continued)

© Hans C. Andersen

September 24, 2009

Here we continue our discussion of elementary probability. We will not discuss this in class.

1 Introduction

We often make probabilistic statements of the form: "The probability that such-and-such is true is equal to ...." This is done in mathematical derivations that involve other types of statistical reasoning, such as the use of probability distribution functions of the type discussed in previous notes. It is also often done in discussing the properties of microscopic but probabilistic models of dynamics. Here we discuss some of the ground rules for using this type of language. (This is by no means intended as a complete introduction to these aspects of probability theory.)

2 The probability of a statement

Consider some rather well defined situation that can be reproduced many times (in principle) and in which the observable features of the situation can be different each time the system is created.

- The situation could be an experiment of the type we discussed in class, where a large system is prepared and then observed.

- The situation could be a hypothetical one in which we pick a particular molecule in some sort of system and then observe what happens to it.

Let S be some statement that can be made about the situation. Examples might be:

- At time t, the number of atoms within 7.0 Angstroms of the origin is less than 5. (This is an example of a statement that might be made in the first situation above, if the experimental system were an atomic liquid.)

- The molecule does not suffer a collision between time t_1 and time t_2. (This is an example of a statement that might be made in the second situation above, in which we picked a molecule at random from a dilute gas of molecules.)

If S is such a statement, then the probability of S (or the probability that S is true) is usually denoted $P(S)$. It can defined to be the fraction of the situations in which S is true if the situation is repeated many times in completely independent ways.

Often we have one statement that is the negative of the other. Thus, if S_1 is the statement "The molecule does not suffer a collision between time t_1 and time t_2" and S_2 is the statement "The molecule does suffer a collision between time t_1 and time t_2", then each statement is the negative of the other. In any particular realization of the situation, one of them must be true but they can't both be true. So

$$P(S_1) + P(S_2) = 1$$

3 The joint probability of two statements

If S_1 and S_2 are two statements, then the joint probability of S_1 and S_2 (or the probability that both S_1 and S_2 are true) is often denoted $P(S_1, S_2)$. It can be defined to be the fraction of the situations in which both S_1 and S_2 are true if the situation is repeated many times in completely independent ways.

Note the following examples:

- If S_1 and S_2 are the negatives of one another, then $P(S_1, S_2) = 0$.

- If S_1 and S_2 are the same statement, then $P(S_1, S_1) = P(S_1)$.

4 The statistical independence of statements

Two statements S_1 and S_2 are statistically independent[1] if

$$P(S_1, S_2) = P(S_1)P(S_2)$$

Note the following examples:

- If S_1 and S_2 are the negatives of one another, then they are not statistically independent, as can be seen from the equations above.
- What happens if S_1 and S_2 are the same statement? Is S_1 statistically independent of S_1? (Work it out. The answer may surprise you, but it will make sense if you think about it.)

5 Conditional probability of two statements

If S_1 and S_2 are two statements, the probability of S_1 given S_2 (or the probability that S_1 is true given that S_2 is true) is often denoted $P(S_1|S_2)$. It is defined as

$$P(S_1|S_2) = \frac{P(S_1, S_2)}{P(S_2)}$$

It is equal to the fraction of those situations in which S_2 is true that also have S_1 true.[2]

A probability of this form is called a conditional probability.

This equation leads to the following equation that can be used to calculate joint probabilities. The first is

$$P(S_1, S_2) = P(S_1|S_2)P(S_2)$$

Interchanging the two statements, we also have

$$P(S_1, S_2) = P(S_2|S_1)P(S_1)$$

[1] I'm not sure this is standard terminology.

[2] In other words, suppose we create the situation many times and ignore those situations in which S_2 is false. Then considering only the remaining situations in which S_2 is true, what fraction of them have S_1 true?

6 The exponential distribution

This section concerns the types of statistical models that are usually used for first order chemical kinetics and radioactive decay. But we shall discuss it from a slightly different point of view.

Suppose we consider a situation in which we set up some initial conditions, we set the clock to $t = 0$, and then we wait for some specific event to happen. Let T be the time at which the event happens. (This is sometimes called a 'waiting time', since this is the time you wait before the event occurs.) T is a random variable, and we assume that the random process that causes the event has the following characteristic: If, at time t, the event has not yet happened, the probability that it will happen between time t and $t + dt$ is simply λdt, where λ is a constant that is independent of t.

We want to know the probability distribution function for the time T. Call this $P_T(t)$. It is defined so that $P_T(t)dt$ is the probability that $t \leq T \leq t + dt$.

There are a variety of ways of solving this problem. Here is one of the more straightforward.[3]

The only potentially confusing thing is the translation of the information we are given into a form in which we can use it. A proper translation is: The probability that the event happens between t and $t + dt$, given that it did not happen between time 0 and time t, is equal to λdt.

Then by the following little trick, we can get the appropriate equation. Note that the statement "the event happens between t and $t + dt$" is true only if the statement "the event does not happen before t" is also true. Thus $P_T(t)dt =$ the probability that (the event happens between t and $t + dt$) $=$ the probability that (the event happens between t and $t+dt$) and (the event does not happen before t) $=$[the probability that the event does not happen before t]$\times$[the probability that the event happens between t and $t+dt$ given that it does not happen before t]

$$= \left(\int_t^\infty dt' \, P_T(t') \right) (\lambda dt)$$

Hence

$$P_T(t) = \lambda \int_t^\infty dt' P_T(t')$$

[3] If you can come up with a more straightforward derivation, please let me know.

This is an integral equation for $P_T(t)$, but it can easily be converted into a differential equation by differentiating.

$$\dot{P}_T(t) = -\lambda P_T(t)$$

The solution is

$$P_T(t) = P_T(0)e^{-\lambda t}$$

The initial value can be gotten either from the normalization condition or from the statement that $P_T(0)dt = \lambda dt$, which follows from the information given.

$$P_T(t) = \lambda e^{-\lambda t}$$

Thus the time at which the event happens has an exponential distribution function. The mean of the distribution is

$$\langle T \rangle = \int_0^\infty dt\, t P(t) = 1/\lambda$$

Thus the mean waiting time is the reciprocal of λ, which was introduced in the statement of the model.