Physics 114 Statistical Mechanics Spring 2018
Seminar 4

Overview:

This week we study the fundamentals of *probability theory*. We go over the definition of a probability distribution, both for *discrete* and *continuous random variables*. We practice counting states for systems of interest, so that we can derive probability distributions - like the ubiquitous *binomial distribution*. We learn to calculate the *moments* of a probability distribution, with especial emphasis on the first two: the mean and variance. For discrete distributions involving a large number N, we will find it useful to use *Stirling's approximation*. We (finally!) encounter a definition of entropy based on the probability of observing a macrostate. Section 3.4.1 of G&T talks about entropy in terms of *uncertainty*. A recommended reading by Pratt talks about *ignorance*. With these in mind, we are meant to believe that $S(\Omega) \propto ln\Omega$ is an excellent definition of how uncertain/ignorant we are about the results of a measurement, if all measurement results, Ω, are equally probable. The more general case, if a measurement result has a probability of P_i, would be $S = -k\ P_i\ lnP_i$.

(Next week, we will continue to think about in entropy in a statistical way. Section 14.8 of B&B will call $S = -k\ P_i\ lnP_i$ the *Gibbs entropy* of a thermodynamic system. B&B Ch. 15 defines the *Shannon entropy*, which applies beyond thermodynamics, to fields like cryptography, data compression and quantum mechanics.)

Required Reading:

G&T Sections

- Ch. 3
  We will return to Bayes theorem next week, so you can omit section 3.2.4 on Bayesian inference. You can also omit section 3.11.2, a proof of the Central Limit Theorem.
- Section 4.1

B&B sections

- Section 1.4
- Ch. 3
- Appendices C.1 , C.2, C.3 and C. 13

Recommended Reading:

- There is a short reading on "ignorance" and Lagrange's undetermined multipliers "Electronic readings and Resources" area of our Moodle site.
- Calculating Lagrange's undetermined multipliers: a Wolfram Widget: `http://www.wolframalpha.com/widgets/view.jsp?id=1451afdfe5a25b2a316377c1cd488883`
  There is also this demo, a Wolfram Demonstrations Project: `http://demonstrations.wolfram.com/LagrangeMultipliersInTwoDimensions/`

Concept checklist from Readings:

- Prof. Park of Williams College told us "*Probabilty* doesn't tell you whether you will win, but it tells you how to play the game." It also tells us how gas molecules, electrons, photons, ... play their games.
- We imagine a space of *exclusive outcomes* of experiments $\{x_i\}$, called the *sample space*. The function $P(x_i)$ represents the likelihood of seeing result x_i upon *one trial* of the experiment. If we do many trials, the number of times outcome x_i occurs is proportional to $P(x_i)$.
- $P(x_i)$ is for outcomes with *discrete* results, like numbers show on a pair of dice. If experiments have *continuously distributed* outcomes, like the location x of a particle, we adopt the notation $p(x)$. Now x is a real number and $p(x)$ is the *probability density*. The likelihood of seeing any outcome x is zero! However, $p(x)dx$ is the probability of seeing x fall somewhere within x and $x + dx$. It is nonzero, in general, for finite dx.
- Some rules for probability are $P(x_i) \geq 0$ and $\Sigma_i P(x_i) = 1$ ... probabilities are positive and normalized. Normalization for $p(x)$ would be $\int p(x)dx = 1$.
- The *addition rule* of probability applies when we do an experiment with an exclusive outcome, but ask a less-exclusive question. For example, we can ask about the likelihood of an outcome being x_i *OR* x_j when the experiment is done once. The answer is:
  $Prob(x_i \; OR \; x_j) = P(x_i) + P(x_j)$.
- Another rule has to do with doing more than one trial, or doing trials of two different kinds of experiments as in B&B section 3.6 which talks about "independent random variables". The most basic question is: What is the likelihood of seeing x_i *AND* x_j as results of two trials. If trials are independent, the *multiplication rule* applies: $Prob(x_i \; AND \; x_j) = P(x_i)P(x_j)$.

- *Conditional probabilities* are needed where experimental outcomes are *not independent*, as they relate to the question we ask. For example, suppose we toss two dice, look at the first but not the second, and ask about the sum of the values shown. The first die has value x_{seen}. Call this outcome Event B. We could ask for the probability that the total value shown on both die, $x_{seen} + x_{unseen}$, is greater than 7. Call this outcome Event A. What is $Prob(A\ occurs,\ given\ that\ B\ occurs)$? This is written $P(A|B)$. In general, $P(A|B) \neq P(A)$. This is because the information we received from Event B is meaningful. For example, it is more likely for the total to exceed 7 if Event B was $x_{seen} = 6$, as opposed the $x_{seen} = 2$.

- The completeness of our sample space implies

  - $P(A) = P(A|B) + P(A| \dashv B)$
    where $\dashv B$ is the situation where outcome B does *not* occur.
  - $P(A\ AND\ B) = P(A|B)P(B) = P(B|A)P(A)$

- The *mean* or *average* of a probability distribution is an important concept: $\bar{x} = \Sigma_i x_i P(x_i)$. For continuous distributions, it is $\bar{x} = \int xp(x)dx$.

- A related idea, is the expected value of a function, $f(x)$. If outcome x has a probability density $p(x)$, then
  $\bar{f} = \int f(x)p(x)dx$.

- The *moments* of a probability distribution are $\mu_j \equiv \overline{x^j}$. The first moment, μ_1 is the mean. Often, people calculate "central moments" instead: $\Delta\mu_j \equiv \overline{(x - \bar{x})^j}$. The first central moment is thus zero. The second central moment, $\Delta\mu_2$ is also known as the *variance*:
  $\sigma^2 = \overline{(x - \bar{x})^2}$. Its square root is, for many distributions we will encounter, a measure of the "width" of the distribution, in that most outcomes fall within $\pm\sigma$ of the mean, $\bar{x}$.

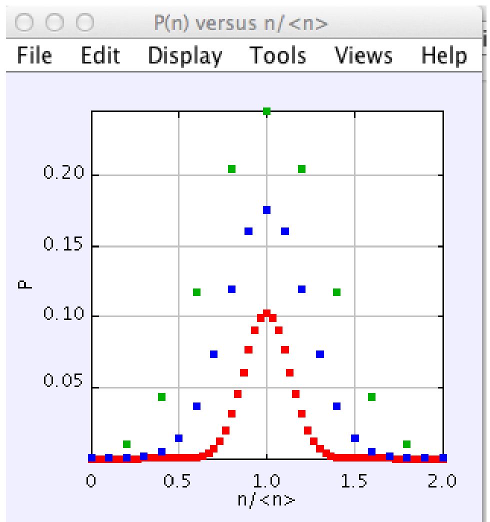
- Having more than one independent random variable, finding mean or variance of a product or sum, is treated in B&B sections 3.5 and 3.6.

- A *Bernoulli process* concerns an object which can have only two states, for example, a coin showing heads or tails, but N independent objects are considered ... either by repeating the experiment with one object N times, or by examining N objects all at once. An important parameter of the distribution is p, the probability of the first outcome (heads, say) Thus $q = 1 - p$ is the probability of tails. A "fair" coin has $p = q$. The distribution $P_N(n)$ with n the number of heads among N tosses is the *Binomial distribution*.

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

Please know how to calculate its mean, Np, and variance, Npq, and note that the entries of *Pascal's triangle* correspond to the values $P_N(n)$.

- Here is a graph of $P_N(n)$ from the G&T Binomial applet for $N = 10, 20$ and 60. If you look at the page from which the applet loads,

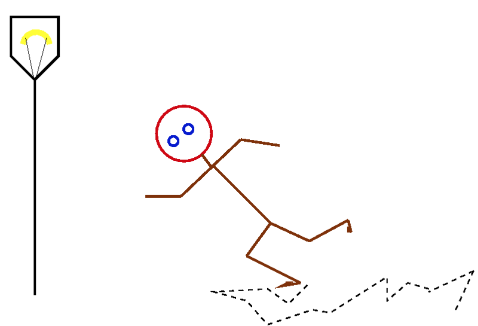


http://stp.clarku.edu/simulations/binomial/index.html, you will see that for $N > 20$, they use *Stirling's approximation* for $N!$:

$$lnN! \approx NlnN - N + \frac{1}{2}ln(2\pi N)$$

We often need $N!$ for very large N values, so please get comfortable with Stirling's approximation, derived in Appendix C.3 of B&B.

- Another Bernoulli process is the *random walk*. Random walks in space (shape of polymers) and time (paths of photons in the stellar interior) show up a lot in physics!



- Mathematical functions and integrals everyone should know:
  - The factorial integral (in Appendix C1 of B&B)
  - The gaussian integrals (in Appendix C.2 of B&B)
  - The combination (or binomial coefficient) $C(n,r) \equiv \begin{pmatrix} n \\ r \end{pmatrix} \equiv \frac{n!}{(n-r)!\ r!}$

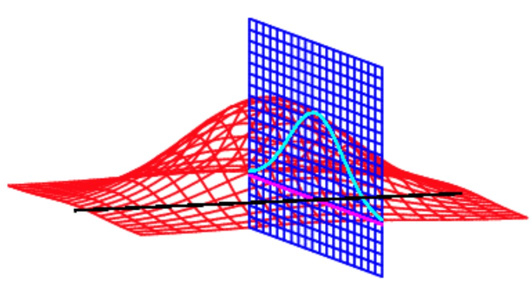
- Trials of an experiment yield *samples* for a histogram. You can estimate the moments of the true, underlying probability distribution from the histogram. The *law of large numbers* says that the more trials you do, the closer the mean, variance, ... will come to the one predicted by the true distribution.

- Suppose you find the *sum* of the results of a set of trials: $S = \Sigma_i^N s_i$. Even if $p(s_i)$ is not a Gaussian distribution, $p(S)$ approaches a *Gaussian* as you sum an arbitrarily large number, N, of results. This is the *Central Limit Theorem.*

$$p(S) = \frac{1}{\sqrt{2\pi\sigma_S^2}}\ exp[-(S-\bar{S})^2/2\sigma_S^2] \quad ; \quad \bar{S} = N\bar{s}\ ; \quad \sigma_S^2 = N\sigma^2$$

- How can we quantify ignorance? G&T section 3.4.1 says that *uncertainty* in a system with probabilities $\{P_i\}$ can be characterized by a function $S(\{P_i\})$. In order for uncertainties for subsystems to be additive, $S = -\Sigma_i P_i ln(P_i)$. In the special case that $P_i = 1/\Omega$ for all i, one has $S = ln\Omega$.

- Normalization is a *constraint* on the function $p(x)$. The method of *Lagrange's undetermined multipliers* is a clever way to maximize a function subject to a constraint like this. (It also works for multiple constraints ... we'll exploit this next week.)



- The normalization constraint on $p(x)$ is a simple idea with huge physical consequences. S subject to this constraint implies that all states are equally probable. This is discussed in the Pratt reading. (Next week, a second constraint is applied: the mean of the energy is known. This will lead us to the *Boltzmann distribution* of states: $P_j \propto e^{-\beta E_J}$.)

- G&T Ch. 4.1 begins to outlines the methods of stat mech:
  - Specify macrostates and the microstates that contribute to each macrostate.
  - Choose the *ensemble*. This is a collection of identically-prepared systems, like the different trials from probability theory. For example, if energy is the same for all members, this is called the *microcanonical ensemble*
  - Calculate statistical properties.
- Section 4.1 of G&T has a classic example: distinguishable particles with *two spin states* (identical statistics to the atoms of Section 1.4 in B&B). (Next week we'll read Section 4.2 and consider another classic example: the *Einstein solid*.)

Warmup Problem: *Due before seminar, Monday eve or, more realistically, mid-day on Tuesday*

1: Independent spins G&T Problem 4.1

Presentation: Monte Carlo Integration You may base this presentation on G&T Problem 3.60 if you wish. Please explain the concept of doing integration by Monte Carlo. Show us the solution to this problem. Use their software if you wish, or write your own. If you happen to know about "importance sampling" or would like a reference from me, to make your integral that much more efficient, just ask :-)

Regular problems:

1: I heart Random numbers ... a numerical problem

This is a problem asking you to do some amount of numerical computation. Using whatever computing environment you feel good about ...

a) Generate a set of 200 random numbers $\{x_i\}$,where $i = 1, ..., 200$. These numbers should be uniformly distributed between 0 and 1. Please give the lines of code you used to generate these, as your answer to this problem.

b) For a uniform distribution of random numbers, $p(x)dx = dx$. In other words, $p(x) = 1$ for all x. Plot a histogram of these numbers, to see if indeed all numbers seem equally likely.

c) Find the mean and variance of your numbers, $\bar{x}$ and σ^2. Comment on whether they are close to what'd you expect from a uniform probability distribution between 0 and 1.

d) Calculate a new random number X which is the sum of all 200 numbers: $X = \Sigma_{i=1}^{200} x_i$. Do this lots of times, until you have a set of 500 numbers $\{X_j\}$ where $j = 1, 500$. Plot a histogram of these numbers. How close is your histogram to the shape of a Gaussian? Does the probability distribution of these numbers, $p(X)$, have a mean and variance that is close to what is predicted by the central limit theorem?

2: Random walks and the binomial distribution G&T Problem 3.35

3: Basic probability:
i) **What are the rules?** G&T Problem 3.7
Note: The person introducing this problem could do i) ...
ii) **Likelihood of various outcomes?** G&T Problem 3.18
Note: ii) assumes gender is binary :-{

iii) **Expectation values:** B&B Problem 3.5 parts (a), (c), and (e)

4: How big is that pond, and how many fish are in it?
i) G&T Problem 3.58
ii) G&T Problem 3.59

5: Monte Carlo Integration G&T Problem 3.60

6: Poisson distribution B&B Problem 3.3

7: Binomial distribution for magnets and gasses
i) G&T Problem 3.27
ii) G&T Problem 3.34

8: Lagrange's undetermined multipliers G&T Problem 3.51

9: Stirling's approximation G&T Problem 3.33