Probability Density

Lecture 4

Physics 342
Quantum Mechanics I

Monday, February 4th, 2008

We review the basic notions of probability, in particular the role of probability density in determining the fundamental quantities relevant to a statistical interpretation: expectation values and variance (for example). First, we will think about the description from a discrete point of view, then generalize to the continuum.

4.1 Discrete Random Variables

Consider a measurement that returns an integer. Griffiths covers this with an example involving age, where the measurement is "what is your age?", and the answers are restricted to integers. But the quantum mechanical form of averaging refers to experiments performed over and over on "identically prepared systems", so consider an example in this spirit: A hopper of ping-pong balls is released by pulling on a lever. Each time we pull the lever, we get a different number of ping-pong balls, with the variation coming from our lever-pull-speed, and simple geometry of balls falling through an opening. Suppose we pull the lever a hundred times, and count the number of balls that come out after each pull. Our system has a maximum (the number of balls in the hopper), and we expect careful lever-pulling to result in roughly the same number of balls during each experimental run.

For the results of the hundred experimental observations, we denote $N(j)$ to be the number of times we got j balls. So we can make a histogram of the results, shown in Figure 4.1.

The total number of times we actually counted is given by $N = \sum_{j=0}^{\infty} N(j)$, for our setup, $N = 100$, and the range of j can be restricted to the non-zero values from our histogram. The probability that we get j balls, based on

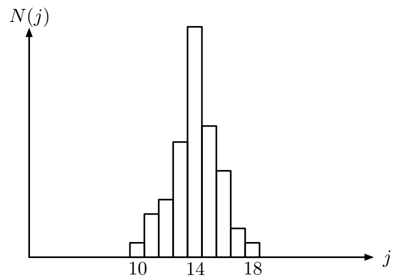


Figure 4.1: The number $N(j)$ of times we got j balls in the ball hopper experiment.

our experimental data, is:

$$P(j) = \frac{N(j)}{N}. \tag{4.1}$$

The average value returned to us, denoted $\langle j \rangle$ is

$$\boxed{\langle j \rangle = \frac{1}{N} \sum_{j=0}^{\infty} j\, N(j) = \sum_{j=0}^{\infty} j\, P(j),} \tag{4.2}$$

so we take each j, and multiply by the probability we get it. Suppose we have a function of j: $f(j)$ – the average of the function is again the functional value for a particular j times the probability we got j:

$$\langle f(j) \rangle = \sum_{j=0}^{\infty} f(j)\, P(j). \tag{4.3}$$

As a specific example, take $f(j) = j^2$, then

$$\boxed{\langle j^2 \rangle = \sum_{j=0}^{\infty} j^2\, P(j).} \tag{4.4}$$

Finally, we want a measure of the "width" of the distribution – the deviation from the average. We could take the average of: $j - \langle j \rangle$, but this gives:

$$\langle j - \langle j \rangle \rangle = \sum_{j=0}^{\infty} (j - \langle j \rangle)\, P(j) = \langle j \rangle - \langle j \rangle \underbrace{\sum_{j=0}^{\infty} P(j)}_{=1} = 0, \tag{4.5}$$

and that is not a good measure – the issue is the sign, the deviation from the average occurs on both the left and right of the average. We can square the summand, and average the squares (all positive) – this is called the variance:

$$\begin{aligned}\sigma^2 = \langle\langle j - \langle j \rangle\rangle^2\rangle &= \sum_{j=0}^{\infty}(j^2 - 2\langle j \rangle j + \langle j \rangle^2)\, P(j) \\ &= \sum_{j=0}^{\infty} j^2\, P(j) - 2\,\langle j \rangle \sum_{j=0}^{\infty} j\, P(j) + \langle j \rangle^2 \sum_{j=0}^{\infty} P(j) \\ &= \langle j^2 \rangle - \langle j \rangle^2,\end{aligned} \tag{4.6}$$

and the square root of this gives us the "standard deviation", the quantity we are interested in (although it is equivalent to the variance).

For the data shown in Figure 4.1, we can build $N(j)$ and $P(j)$:

j	$N(j)$	$P(j)$
10	2	.02
11	6	.06
12	8	.08
13	16	.16
14	32	.32
15	18	.18
16	12	.12
17	4	.04
18	2	.02
total	100	1

with all other values of j having $N(j) = 0$. From these, we can calculate the mean, variance and standard deviation:

$$\langle j \rangle = 14.04 \qquad \sigma^2 = 2.68 \qquad \sigma = 1.64 \tag{4.7}$$

Notice that none of the measurements returned the expected value (which is not even an integer). We can also compute, as an additional piece of information equivalent to σ, the analytically intractable $\langle |j - \langle j \rangle| \rangle$:

$$\langle |j - \langle j \rangle| \rangle = \sum_{j=0}^{\infty} |j - \langle j \rangle|\, P(j) = 1.1. \tag{4.8}$$

The standard deviation, and indeed, the absolute deviation average above are both "smaller" than a half-width type of eyeball estimate. The fact is that the smaller deviations are much more heavily weighted since $P(j)$ is sharply peaked about them, so even though the larger deviations have greater numerical value, the probability with which they occur is vanishingly small.

These expressions require us to have the experimental data in hand – we must have constructed the histogram and be able to form $P(j)$ numerically as in the table above. That's fine as far as actual experiments go, but the program of quantum mechanics is a theoretical description of $P(j)$ (in essence) – so we expect idealized distributions. Indeed, the underlying structure of non-relativistic quantum mechanics is determined by a PDE that tells us, given a particular physical setup, what the continuous analogue of $P(j)$ is.

Discrete Example

On the discrete side, we can develop idealized probability distributions, $P(j)$ for various scenarios. For example, we know that the probability of rolling a 3 on a six sided die is $1/6$ – although if we actually roll the die six times, we may or may not achieve this. The probability that we will get a 3 on a six-sided die by rolling it 10 times is given by the binomial distribution:

$$P(n, \ell, p) = \begin{pmatrix} n \\ \ell \end{pmatrix} p^{\ell} (1-p)^{(n-\ell)} \qquad \begin{pmatrix} n \\ \ell \end{pmatrix} = \frac{n!}{\ell!\,(n-\ell)!} \tag{4.9}$$

which gives the probability of getting ℓ occurences of an event, that has a single-event probability of p, in n tries. To get a three on a single roll has probability $1/6$, and if we roll $n = 10$ times, and count the number of threes, we will get, for example, 2 occurences with probability $P(10, 2, 1/6) \sim .29$. Let's see how we would construct this particular distribution for $n = 4$ – each throw is independent, so we have probability $1/6$ for each roll of the die – to get $\ell = 1$ counts of 3 in our four rolls, we must have three non-three rolls, and this occurs with probability $\left(\frac{5}{6}\right)^3$, and then we need one roll where we get a three, with probability $\frac{1}{6}$ – there are four different places for the 3 to go, so we expect:

$$P(4, 1, 1/6) = 4 \left(\frac{5}{6}\right)^3 \frac{1}{6} = .39. \tag{4.10}$$

Two counts of three requires two threes and two other numbers: $\left(\frac{5}{6}\right)^2 \left(\frac{1}{6}\right)^2$ and can occur in six different ways, so we have:

$$P(4,2,1/6) = 6\left(\frac{5}{6}\right)^2 \left(\frac{1}{6}\right)^2 = .12, \tag{4.11}$$

and we can see how this "logic" is worked out in the formula itself:

$$P(n,\ell,p) = \underbrace{\binom{n}{\ell}}_{\text{number of ways}} \underbrace{p^\ell}_{\text{count}} \underbrace{(1-p)^{(n-\ell)}}_{\text{other}}. \tag{4.12}$$

It is pretty clear that, for our example, we have

$$\sum_{\ell=0}^{4} P(4,\ell,1/6) = 1, \tag{4.13}$$

or in words: "The probability that in four rolls, we get either 0, 1, 2, 3, or 4 rolls of three is 1". In addition, we can calculate the mean and standard deviation associated with $P(4,\ell,1/6)$:

$$\langle j \rangle = \frac{2}{3} \qquad \sigma = \frac{\sqrt{5}}{3}. \tag{4.14}$$

Indeed, because of its nice analytical form, these have general expressions for $P(n,\ell,p)$:

$$\langle j \rangle = n\,p \qquad \sigma^2 = n\,p\,(1-p). \tag{4.15}$$

With an expression like the binomial distribution, we can compute any number of statistically relevant/interesting quantities. But there is no dynamical rule (other than common sense) we are using to determine the probabilities – each new type of experiment, whether it is ping pong ball counting or die rolling requires some analysis. The basic idea is the same in quantum mechanics – if one can determine the probability density, all manner of predictive and statistical information is available – the big difference is that Schrödinger's equation gives us a systematic (if at times difficult) way to calculate the probability distribution of physical situations specified[1], in that case, by a classical Hamiltonian.

That life should be this way is not our current issue, we will begin by positing Schrödinger's equation, with motivation to follow. Note that motivation is *all* we can do, one cannot deduce the equation any more than you can deduce Maxwell's equations. But before we begin, we have to pass from the discrete probabilities discussed above to continuous *probability density*.

4.2 Probability Density

We are used to continuous density descriptions of sources for various theories. In E&M, for example, we have charge density $\rho(x, y, z, t)$ that tells us the charge per unit volume in a specific location at a specific time. For taut strings, we have a mass density μ, mass per unit length (a degenerate three dimensional distribution) that governs the propagation of waves on the string. Well, just as these densities tell us about quantities of interest per unit volume, we can define a *probability density*, that tells us the "probability per unit volume" of some process. In terms of use (calculating average, variance, etc.), the probability density in a continuous setting is very much like the probability $P(j)$ for the ping pong balls, just smeared out appropriately over a continuum of values.

From a density point of view, we know how to find the total "stuff" (charge, mass, probability) enclosed in a volume – suppose $\rho(\mathbf{r})$ represents the probability density for a particle to exist at a particular location – then the probability that the particle is within a small volume centered at $\mathbf{r}_0$ is: $dP(\mathbf{r}_0) = \rho(\mathbf{r}_0)\, d\tau$. In one dimension, this reduces to $dP(x_0) = \rho(x_0)\, dx$, and we have the usual sorts of degenerate cases.

By full integration, we can calculate interesting probabilities – for a ball of radius R centered at the origin, the probability of finding a particle inside the ball (if we are interpreting our probability density in terms of particle

[1]Note that quantum mechanics is not alone in this regard – statistical mechanics often proceeds from a natural Newton's second law with stochastic noise (possibly representing unresolved degrees of freedom in the problem) – in that case there is a natural way to develop the Fokker-Planck equation for various potentials, and this governs the probability density in that field.

existence at a location) is:

$$P(R) = \int_{B(R,0)} \rho(\mathbf{r}\,')\,d\tau' = \int_0^{2\pi}\int_0^{\pi}\int_0^{R} \rho(\mathbf{r}\,')\,r'^2\,\sin\theta'\,dr'\,d\theta'\,d\phi' \quad (4.16)$$

and if we were talking about charge density, this would represent the total charge enclosed in a sphere of radius R.

We have a side requirement on the probability density that does not exist for other types of density – the total probability, over all space, now, must not only be finite, it must be $= 1$. So there is a normalization in the above:

$$P(R \longrightarrow \infty) = 1. \quad (4.17)$$

For example, suppose we have $\rho(\mathbf{r}\,') = \frac{A}{r}\,e^{-r/a}$ for r the distance from the origin and a some constant with the units of length – then the constant A is fixed:

$$\begin{aligned} P(R \longrightarrow \infty) &= \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \frac{A}{r'}\,e^{-r'/a}\,r'^2\,\sin\theta'\,dr'\,d\theta'\,d\phi' \\ &= 4\,\pi\,A\int_0^{\infty} r'\,e^{-r'/a}\,dr' \\ &= 4\,\pi\,A\,a^2 \end{aligned} \quad (4.18)$$

Now, the requirement is that this should be $= 1$, so we set $A = \frac{1}{4\pi\,a^2}$. Our density is $\rho(\mathbf{r}) = \frac{e^{-r/a}}{4\pi a^2\,r}$ (notice that the units here are, correctly, 1/volume).

4.2.1 Moments

We saw that the analogue of "total charge" for a probability density must be one, that's necessary in order to give integrals of the density the interpretation of a probability. We can also ask for the analogue of an electric dipole moment:

$$\mathbf{p} = \int \rho(\mathbf{r}\,')\,\mathbf{r}\,'\,d\tau' \quad (4.19)$$

from electrostatics has a familiar form on the probabilistic side – it's what we would call the expectation value of $\mathbf{r}$ (compare with (4.2)). We would write this:

$$\langle\mathbf{r}\rangle = \int \rho(\mathbf{r}\,')\,\mathbf{r}\,'\,d\tau'. \quad (4.20)$$

Similarly, we can construct average values of functions:

$$\langle f(\mathbf{r}) \rangle = \int \rho(\mathbf{r}\,') \, f(\mathbf{r}\,') \, d\tau', \tag{4.21}$$

which are, in theory decomposable in terms of higher "moments" of the distribution (the probabilistic version of multipole moments).

In addition to the average, we can also define the variance in a manner similar to the discrete case – in one dimension, for example, we have:

$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2. \tag{4.22}$$

For our density from above, we can construct the average "radius" via (integrating over all space, now, written in spherical coordinates):

$$\begin{aligned}
\langle r \rangle &= \int r' \, \rho(r') \, d\tau' = \int_0^{2\pi} \int_0^{\pi} \int_0^{\infty} r' \, \rho(r') \, r'^2 \, \sin\theta' \, dr' \, d\theta' \, d\phi' \\
&= \frac{4\,\pi}{4\,\pi\,a^2} \underbrace{\int_0^{\infty} e^{-r'/a} \, r'^2 \, dr'}_{=2\,a^3} \\
&= 2\,a.
\end{aligned} \tag{4.23}$$

4.3 Quantum Mechanics

So we can calculate stuff, but what are these probability densities probability densities *of*, and what equation governs them? We do not have a physical interpretation for a probability density without knowing how it is generated. From the above, $\rho(\mathbf{r})$ could be the probability density that an apple exists at the point $\mathbf{r}$. Quantum mechanics gives a very specific interpretation to probability densities governed by the Schrödinger equation, and this is what allows us to sensibly discuss the physics of these (particular) densities.

As a final note, we are used to dealing with densities as sources for fields – charge density provides the source for electric fields, energy density is the source for gravitational fields. Quantum mechanics is different – here, the goal will be the probability density, so we are *finding* the distribution of probability in space given a physical system, typified by its potential.

Homework

Reading: Griffiths, pp. 5–11.

Problem 4.1

Griffiths 1.11. You are implicitly told the distribution $\rho(\theta)$ in the problem set-up – all you need to do is normalize it, then calculate a bunch of expectation values.

Problem 4.2

For the probability density

$$\rho(r) = \begin{cases} A\,r & r \le R \\ 0 & r > R \end{cases} \tag{4.24}$$

where $dP = 4\,\pi\,\rho(r)\,r^2\,dr$ describes the probability of finding a particle in a spherical shell of thickness dr located at r (in spherical coordinates):

a. Find A by requiring that the probability of finding the particle somewhere is 1 (remember, we are in three dimensions, so you need to integrate over all space).

b. What is the probability of finding the particle in a sphere with radius between $\frac{1}{2}\,R$ and R?

c. Compute $\langle r \rangle = \int r\,\rho(r)\,d\tau$ – this represents the "most likely" radius at which to find the particle.

Problem 4.3

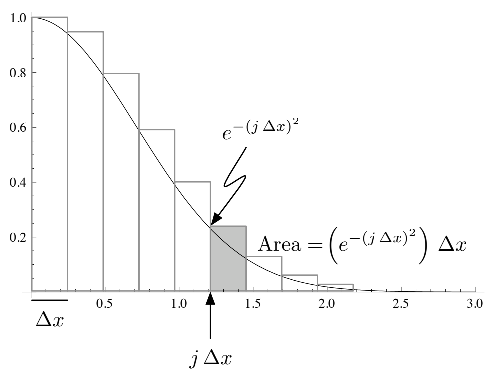
Griffiths 1.3. Here we study a famous continuous distribution – the Gaussian – look at the tables in the back of the book to help with the integral.

Optional: Suppose we did not know that $\int_0^\infty e^{-x^2}\,dx = \frac{1}{2}\,\sqrt{\pi}$, how could we get an estimate of the numerical value of this integral?

Referring to the figure below – we can make boxes whose area approximates the area under the curve. Formally, we have

$$\int_0^a f(x)\,dx \sim \sum_{j=0}^{N} f(j\,\Delta x)\,\Delta x \tag{4.25}$$

where $(N+1)\,\Delta x = a$ serves to define Δx.



From the value of the function at $x = 5.0$: $e^{-25} \sim 10^{-11}$, we set $a = 5$ as a reasonable cutoff. In Mathematica, we can use the following commands to perform the sum in (4.25) using `NN` boxes:

```
In[1]:= NN = 100;
        a = 5.0;
        DX = a / (1 + NN);
        Sum[Exp[-(j DX)^2] DX, {j, 0, NN}]

Out[4]= 0.910979
```

For what value of `NN` (roughly) do you get an accuracy of $.01\,\%$?