Physics 200-05

Probabilities and expectation values

Probabilities

Probabilities mathematically are "measures" defined on a list of events. They have certain properties. If $p(A)$ represeentes the probability that event of type A will occur, and $p(B)$ represents the probability that event B could occur, and if A and B are mutually exclusive (Ie, it is impossible that any event could be both of type A abd type B) then the probability that an event of either type A or type B could occur is the sum of the two probabilities $p(A\ or\ B) = p(A) + p(B)$. For example, when one throws a single die, the type of event where a single pip occurs on the uppermost side of the die and a type where three pips occur on the upper surface are mutually exclusive. You cannot have both occuring in one event. Then the probability that either one or three pips show on the top is just the sum of probabilities that one or three occur.

If there are M possibilities of type A_m, and they are all mutually exclusive, and if in each event at least one of those types MUST occur, then

$$\sum_m p(A_m) = 1 \tag{1}$$

These probabilities are defined separately from frequencies, but, probabilities have their link with the real world through frequencies of occurance. If one has N independent trials, independent meaning that there is nothing, except the physical situation which links those N different trials, and in which we do not believe that there is any other link between them, then if the probability of event of type A in any single trial is $p(A)$, then the frequency with which the A type event would be expected to occur in those N trials is $r(A) = p(A)N$. Ie, in those N different independent trials, $p(A)$ is roughly the fraction of events in which type A events would occur. I say roughly because there is believed to be no causal influence which ensures that frequency. Ie, the very statement of independence of the trials means that there is nothing in any trial which keeps track of, or can influence, any other trial.

This fraction $\frac{r(A)}{N}$ is not exact. In various attempts one would expect a spread of values for this fraction of order $\sqrt{\frac{p(A)(1-p(A))}{N}}$ I.e., over a very large number of trials N the difference between this fraction of times A would occur and the probability of A occuring would decrease with a rate proportional to $\frac{1}{\sqrt{N}}$.

Much philosophical literature has been devoted to try to understand what these probabilities really mean. Are they simply a shorthand for frequencies [Frequentists]? But then what do they mean before any trials have occured? Are they simply a measure of your own degree of belief in what will happen [Bayesians]? But then why does the dice behave in a way such that the frequencies are in accord with your beliefs?. Are they a measure of your ignorance [Maxwell, Boltzman,..]– since you have no good reason to prefer one outcome over another because of your ignorance, all outcomes are equal, and thus have equal probability? I will not persue these issues here.

Expectation value Let us say that the events are divided into types by the value of some parameter. Ie, the type A above means that some parameter has some specific value say a. Now in N trials, that value will be different in the various trials. The number of trials in which the value is a will be $r(a)$. Let there be M possible values for that parameter, $a_1,\ a_2,\ ...a_M$. and the number of trials in which some value a_m occurs is $r(a^m)$. Then since in each trial one of those values must occur, and since the total number of trials is N then the $r(a_m)$ must sum up to N.

If we collect all N values, a quantity that is often of interest is the average value. This is the sum of the values occuring in all of the trials over the total number of tials.

$$Average(a) = \frac{\sum_m r(a_m)a_m}{N} = \sum \frac{r(a_m)}{N} a_m \tag{2}$$

Above we found that if we have the notion of probabilities then we would expect that the ratio, $\frac{r(a_m)}{N}$ should just be an approximation to $p(a_m)$. Thus the average value approximates the *expectation value*

$$\bar{a} = \sum_m a_m p(a_m) \tag{3}$$

Note that this is NOT the value to be expected in any trial. For example if there are two possibilities +1 and -1, and the probablilites of each are 1/2, then the expectation value is 0, even though that value could never occur in any trial. If you want you could just think of it as the expected average value. (Note that because the number of occurances $r(a_m)$ will change from trial to trial, the average value will also change.)

Standard Deviation There is another value which is often used with probability distributions, and that is the standard deviation. This is a measure of by how much one expects the value in any one trial to vary.

In N trials, it is the square root of the average square of difference between the value actually occuring an any trial, and the average value.

$$StD = \sqrt{\frac{\sum_m r(a_m)(a_m - Average(a))^2}{N}} \tag{4}$$

Again using the approximation of probabilities, this leads to the notion of the uncertainty, written as

$$\Delta a = \sqrt{\sum_m p(a_m)(a_m - \bar{a})^2} \tag{5}$$

This is a measure of how much the value varies from time to time. Again, it is not the difference between any value and the average value. For the example above where the probabilities of ± 1 were $1/2$, the uncertainty would be

$$\Delta = \sqrt{\frac{1}{2}(1-0)^2 + \frac{1}{2}(-1-0)^2} = \frac{1}{\sqrt{2}} \tag{6}$$

Quantum Expectation values

Let us say that we have a system with a state $|\psi\rangle$, and we have the matrix A representing some physical attribute of the system. Let us assume that A is and MxM matrix, so it has M eigenvalues and eigenvectors. Assume that these eigenvalues are a_m and the eigenvectors are $|a_m\rangle$. Then we can write the state $|\psi\rangle$ as

$$|\psi\rangle = \sum_m \alpha_m |a_m\rangle \tag{7}$$

where α_m are called the amplitudes for a_m in the state ψ. Because eigenvectors are always normalised (have unit value $\langle a_m||a_m\rangle = 1$) we find that

$$\langle\psi|A|\psi\rangle = (\sum_{m'} \alpha^*_{m'}\langle a'_m|)A(\sum_m \alpha_m|a_m\rangle) \tag{8}$$

$$= (\sum_{m'} \alpha^*_{m'}\langle a'_m|)(\sum_m \alpha_m A|a_m\rangle) \tag{9}$$

$$= (\sum_{m'} \alpha^*_{m'}\langle a'_m|)(\sum_m \alpha_m a_m|a_m\rangle \tag{10}$$

$$= \sum_m{}' \sum_m \alpha^*_{m'}\alpha_m a_m \langle a_{m'}||a_m\rangle \tag{11}$$

Now, since the eigenvectors for different eigenvalues are orthogonal, this last line becomes

$$\sum_{m}' \sum_{m} \alpha_{m'}^* \alpha_m a_m \langle a_{m'} || a_m \rangle \tag{12}$$

$$= \sum_m \alpha_m^* \alpha_m a_m \tag{13}$$

Now, the Born taught us that the quantity $\alpha_m^* \alpha_m$ should be regarded as the probability that on a measurement, the value a_m should be obtained. Thus, we find

$$\langle \psi | A | \psi \rangle = \sum_m p(a_m) a_m = \bar{a} \tag{14}$$

Thus the expectation value (average value) of the eigenvalues is just the multiplication of the bra-state with A and then with the ket state.

By an exactly similar procedure we find

$$\Delta a \equiv \Delta A = \langle \psi | \left(A - \langle \psi | A | \psi \rangle I \right)^2 | \psi \rangle \tag{15}$$

Since $\langle \psi | A | \psi \rangle$is just a number, and since $AI = A$ this can also be written as

$$\Delta A = \langle \psi | \left(A^2 - \langle \psi | A | \psi \rangle^2 I \right) | \psi \rangle \tag{16}$$

Two level systems Since there are only two possible values for any attribute, and since the total probabilities add up to 1, the expectation value of any of the three Pauli operators completely determines the probabilities.

$$p(+1) + p(-1) = 1 \tag{17}$$

$$< \Sigma >= (+1) p(+1, \Sigma) + (-1) p(-1, \Sigma) \tag{18}$$

from which we find that

$$p(+1, \Sigma) = \frac{1}{2}(1+ < \Sigma >) \tag{19}$$

$$p(-1, \Sigma) = \frac{1}{2}(1- < \Sigma >) \tag{20}$$

where Σ stands for one of the Pauli matrices. Note that the probabilities are specific to the Pauli matrix, which is why I wrote $p(+1\Sigma)$ to remind us of that. (eg, the probability of the values ± 1 for say Σ_3 in a state are different from those for Σ_1.

There is another interesting fact about these matricies. Given any state vector $|\psi\rangle$, there always exists a combination of the Pauli matrices such that this state is the $+1$ eigenvector for that matrix. The proof is easy. Let me write the state as

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \tag{21}$$

where we have $|a|^2 + |b|^2 = 1$ since we want a normalized state vector. Note here $|a|^2$ just means a^*a. Then, it is easy to see that the matrix

$$A = \begin{pmatrix} |a|^2 - |b|^2 & ab^* \\ a^*b & |b|^2 - |a|^2 \end{pmatrix} \tag{22}$$

has the vector $|\psi\rangle$ as eigenvector

$$A|\psi\rangle = \begin{pmatrix} |a|^2 - |b|^2 & ab^* \\ a^*b & |b|^2 - |a|^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \tag{23}$$

$$= \begin{pmatrix} (|a|^2 - |b|^2)a + 2ab^*b \\ 2a^*ba + (|b|^2 - |a|^2)b \end{pmatrix} \tag{24}$$

$$= \begin{pmatrix} (|a|^2 + |b|^2)a \\ (|a|^2 + |b|^2)b \end{pmatrix} \tag{25}$$

$$= \begin{pmatrix} a \\ b \end{pmatrix} = |\psi\rangle \tag{26}$$

However, writing this in terms of the Pauli matrices we find that

$$A = (ab^* + a^*b)\Sigma_1 + i(ab^* - a^*b)\Sigma_2 + (|a|^2 - |b|^2)\Sigma_3 \tag{27}$$

But we also have

$$\langle\psi|\Sigma_1|\psi\rangle = (ab^* + a^*b) \tag{28}$$

$$\langle\psi|\Sigma_2|\psi\rangle = i(ab^* - a^*b) \tag{29}$$

$$\langle\psi|\Sigma_3|\psi\rangle = |a|^2 - |b|^2 \tag{30}$$

Thus we finally have

$$A =< \Sigma_1 > \Sigma_1 + < \Sigma_2 > \Sigma_2 + < \Sigma_3 > \Sigma_3 \tag{31}$$

Ie, the combination of the Pauli sigma matrices for which $|\psi\rangle$ is the +1 eigenvector is just determined by teh expectation value of the Sigma matrices for that eigenvector.

Also, since $|\psi\rangle$ is the plus 1 eigenvector of A we have

$$\langle\psi|A|\psi\rangle = \langle\psi||\psi\rangle = 1 \tag{32}$$

but

$$\langle\psi|A|\psi\rangle = \langle\psi|(<\Sigma_1>\Sigma_1+<\Sigma_2>\Sigma_2+<\Sigma_3>\Sigma_3)|\psi\rangle \tag{33}$$

$$=<\Sigma_1>^2+<\Sigma_2>^2+<\Sigma_3>^2 \tag{34}$$

Since the state $|\psi\rangle$ in the above was arbitrary, we find that

$$<\Sigma_1>^2+<\Sigma_2>^2+<\Sigma_3>^2=1 \tag{35}$$

for all states $|\psi\rangle$. Thus if the expectation value of any one of the Pauli sigma matrices is one, then both of the others **must** be zero expectation value .

This also shows that there is a one to one relation between the ket vectors of a two level system, and a physical three dimensional real vector in space.

Finally, if $|\psi\rangle$ is the +1 eigenvector of a matrix

$$B = \vec{\beta}\vec{\Sigma} \tag{36}$$

(where $\vec{\beta}\cdot\vec{\beta}=1$) then the expectation value of some other matrix

$$G = \vec{\gamma}\cdot\vec{\Sigma} \tag{37}$$

is

$$\langle\psi|G|\psi\rangle = \vec{\gamma}\cdot\vec{\beta} \tag{38}$$

since

$$\langle\psi|\vec{\Sigma}|\psi\rangle = \vec{\beta} \tag{39}$$