Review of basic probability concepts

Both measurements of observable quantities and the wave function that defines a state of a quantum system have statistical interpretation. For this reason it is useful to review some basic concepts of the probability theory. For the most part we will only deal with elementary probability theory in this quantum mechanics course. Hence the review is short.

Some quantities in quatum mechanics may take discrete values. Examples include the energy levels of a particle in a potential well of the projection of the orbital momentum on a designated axis. For simplicity let us assume that we deal with a range of integer values. We can then introduce the probability that k-th value occurs upon a random pick/measurement

$$P(k) \qquad\qquad 0 \leq P(k) \leq 1$$

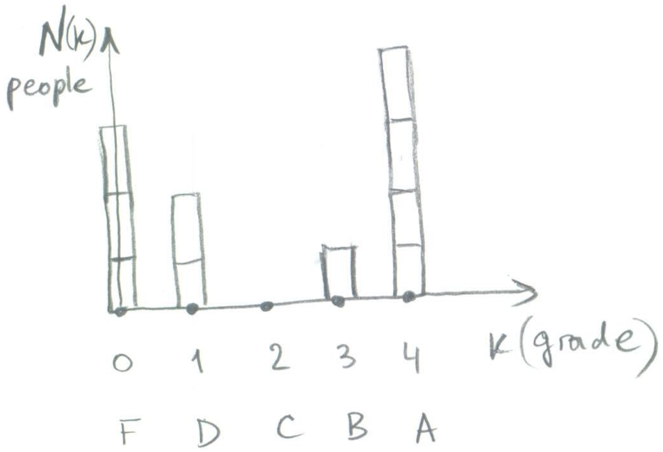
The probability of measuring/picking any value out of the allowed range of values must be equal to 1 (or 100%). Hence

$$\sum_{k}{}' P(k) = 1$$

Note that the range of possible values k can be either finite or infinite. It depends on the specifics of a particular quantity. For example, if we throw a dice the possible outcomes are 1 to 6. and each probability is $\frac{1}{6}$:

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$$

We could think of a slightly more sophisticated example. Suppose we have a group of 10 students who took a quantum mechanics class. It is known that four of them obtained A's, one obtained B, nobody obtained a C, two people got Ds and three people failed the class. Then the histogram of their performance would look as follows



$N(0) = 3$
$N(1) = 2$
$N(2) = 0$
$N(3) = 1$
$N(4) = 4$

$k = 0..4$

The total number of students is

$$N = \sum_{k=0}^{4} N(k) = 10$$

Now if we randomly select a person out of that group of students then the probability that this selected person's grade was k is

$$P(k) = \frac{N(k)}{N}$$

$P(0) = \frac{3}{10}$ $P(1) = \frac{2}{10}$ $P(2) = 0$

$P(3) = \frac{1}{10}$ $P(4) = \frac{4}{10}$

and

$$\sum_{k=0}^{4} P(k) = 1$$

Looking at the histogram we can easily see that the most probable grade was 4 (A). This is the grade that was obtained by the largest number of students.

We can introduce the median — the value of k such that the the number of students who obtained a grade smaller than the median is equal to the number of students who obtained a larger grade. In our case the median grade is 2. In case of discrete distributions the median may not always be uniquely defined.

Now, a more commonly used quantity is the average (or mean, or expected) value. It is defined as

$$\langle k \rangle = \sum_k k P(k)$$

In our case

$$\langle k \rangle = \frac{0 \cdot 3 + 1 \cdot 2 + 2 \cdot 0 + 3 \cdot 1 + 4 \cdot 4}{10} = 2.1$$

We can see that the average (or mean) value may not necessarily be the same as the most probable value.

We can go further and introduce the so-called moments of our distribution — average square, average cube, and so on. The average square is particularly useful

$$\langle k^2 \rangle = \sum_k k^2 P(k)$$

$$\langle k^2 \rangle = \frac{0^2 \cdot 3 + 1^2 \cdot 2 + 2^2 \cdot 0 + 3^2 \cdot 1 + 4^2 \cdot 4}{10} = 7.5$$

Finally, we can go even further and introduce the average value of any function/quantity with our distribution:

$$\langle f(k) \rangle = \sum_k f(k) P(k)$$

Note that in general $\langle k^2 \rangle \neq \langle k \rangle^2$

In order to characterize how widespread or peaked / localized a distribution is we can calculate how far on average each value deviates from the mean. We could try to define $\delta k = k - \langle k \rangle$. However we would see that

$$\langle \delta k \rangle = \sum_k \left(k - \langle k \rangle \right) P(k) = \sum_k k P(k) - \langle k \rangle \underbrace{\sum P(k)}_{1} =$$

$$= \langle k \rangle - \langle k \rangle = 0$$

This is hardy a surprise as we would expect that on average the negative values of $k - \langle k \rangle$ are cancelled by the positive values of $k - \langle k \rangle$.
Using the modulus $|k - \langle k \rangle|$ instead of $k - \langle k \rangle$ is a possibility. However, the modulus function is not smooth at zero. That creates some potential issues. To avoid this issue we can square before averaging, i.e. let us do $\langle (\delta k)^2 \rangle$. The resulting quantity is known as the variance

$$(\Delta k)^2 = \langle (\delta k)^2 \rangle = \sum_k (\delta k)^2 P(k) = \sum_k (k - \langle k \rangle)^2 P(k)$$

$$= \sum_k \left(k^2 - 2k\langle k \rangle + \langle k \rangle^2 \right) P(k) = \sum_k k^2 P(k) - 2\langle k \rangle \sum_k k P(k)$$

$$+ \langle k \rangle^2 \sum_k P(k) = \langle k^2 \rangle - 2\langle k \rangle \langle k \rangle + \langle k \rangle^2 = \langle k^2 \rangle - \langle k \rangle^2$$

The square root of the variance (which has the same physical units as k) is called the standard deviation

$$\Delta k = \sqrt{\langle k^2 \rangle - \langle k \rangle^2}$$

Now let us turn our attention to the case when the quantity of interest is not discrete, but continuous. In this case it is customary to talk about the probability density rather than just probability. If x is a discrete variable (e.g. particle position) and $\rho(x)$ is the probability density then $\rho(x)\,dx$ gives the probability that the given quantity lies in the range $[x, x+dx]$. In the language of mathematicians $\rho(x)$ is called PDF – probability density function.

The integral over a certain interval values of x gives the probability that the quantity of interest lies in this interval

$$P(a \leq x \leq b) = \int_a^b \rho(x)\,dx$$

The above expression is related to the so called cumulative distribution function (CDF). The term CDF is commonly used by mathematicians. The CDF is defined as

$$F(t) = \int_{-\infty}^{t} \rho(x)\,dx$$

The characteristics of a distribution that we introduced for discrete variables can be extended to the continuous case in the obvious way

$$\langle x \rangle = \int_{-\infty}^{+\infty} x\,\rho(x)\,dx \qquad \langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2 \rho(x)\,dx \qquad \langle f(x) \rangle = \int_{-\infty}^{+\infty} f(x)\,\rho(x)\,dx$$

$$\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle} \qquad \text{while} \qquad \int_{-\infty}^{+\infty} \rho(x)\,dx = 1$$