Tuesday 10^{th} May 2005

Quantum Physics Handout

Probabilities and Probability Densities

Probabilities

Suppose you throw N darts at a dart board and record the scores. The results of the first 21 throws are shown in Fig. 1.

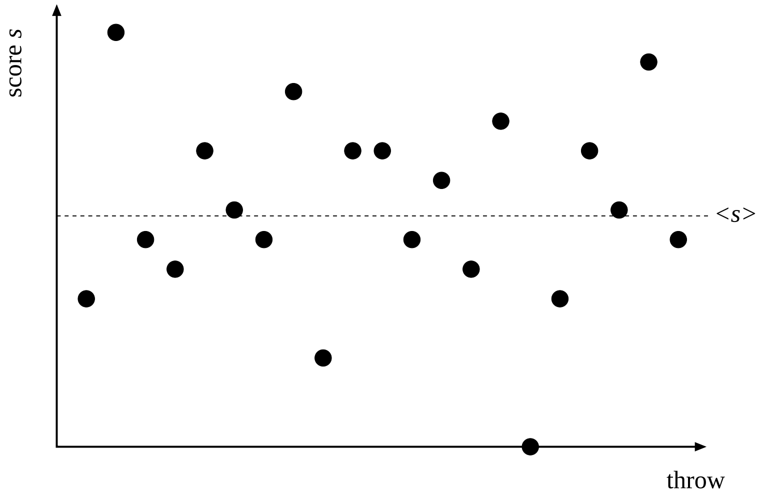

Figure 1: The first 21 scores recorded in the dart game

Another way to summarise the data is to plot the number n_s of darts obtaining each possible score s. An example is shown in Fig. 2. Since no single dart can score more than 60, $n_s = 0$ for all $s > 60$.

The total number of darts thrown is equal to the number n_0 that scored 0 plus the number n_1 that scored 1 plus the number n_2 that scored 2 plus ...

$$n_0 + n_1 + n_2 + \ldots \; = \; N \; ,$$

or, equivalently,

$$\sum_{\text{score } s=0}^{\infty} n_s \; = \; N \; .$$

Dividing both sides of this equation by N gives

$$\sum_{s=0}^{\infty} p_s \; = \; 1 \; ,$$

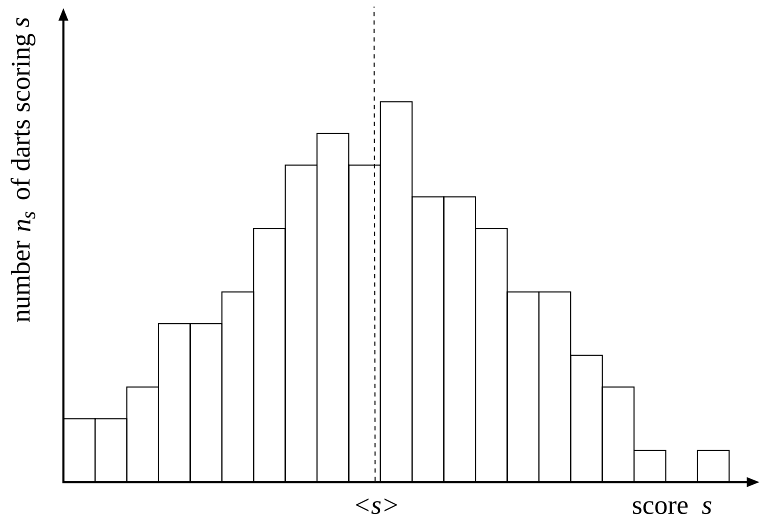

Figure 2: The number of darts n_s scoring each possible value of s

where $p_s \equiv n_s/N$. In the limit of large N (a very long game), p_s normally becomes independent of N (assuming that the player does not get tired!). This limiting value is called the *probability that a dart scores s points.*

Expected Values

The mean or expected score per dart, denoted $\langle s \rangle$, is given by

$$\langle s \rangle \;=\; \lim_{N\to\infty} \frac{\text{sum of scores of all } N \text{ darts}}{N} \; .$$

Since n_0 darts scored 0, n_1 scored 1, n_2 scored 2, and so on, we have

$$\text{sum of scores of all } N \text{ darts} \;=\; \sum_{s=0}^{\infty} s n_s \; .$$

This enables us to re-express $\langle s \rangle$ in terms of probabilities:

$$\langle s \rangle \;=\; \lim_{N\to\infty} \frac{1}{N} \sum_{s=0}^{\infty} s n_s \;=\; \sum_{s=0}^{\infty} s p_s \; .$$

In other words,

$$\begin{aligned} \langle s \rangle \;=\; & (0 \times \text{Prob. of scoring } 0) + (1 \times \text{Prob. of scoring } 1) + \ldots \\ & + (s \times \text{Prob. of scoring } s) + \ldots \; . \end{aligned}$$

More generally, the mean or expected value of any function $g(s)$ of the score s is given by:

$$\langle g(s) \rangle = \lim_{N\to\infty} \frac{\sum_{s=0}^{\infty} g(s) n_s}{N} = \sum_{s=0}^{\infty} g(s) p_s \,.$$

For example, the expected value of s^2,

$$\langle s^2 \rangle = \lim_{N\to\infty} \frac{\sum_{s=0}^{\infty} s^2 n_s}{N} = \sum_{s=0}^{\infty} s^2 p_s \,.$$

Variance and Standard Deviation

Evaluating $\langle s \rangle$ tells us the mean score per dart but provides no information about the spread of scores. We would also like to know the typical distance of a point in Fig. 1 from the horizontal line $s = \langle s \rangle$ and the typical width of the distribution in Fig. 2.

The simplest possible measure of spread or uncertainty, the expected value of $s - \langle s \rangle$, is no good:

$$\begin{aligned} \langle s - \langle s \rangle \rangle &= \sum_{s=0}^{\infty} (s - \langle s \rangle) p_s \\ &= \sum_{s=0}^{\infty} s p_s - \langle s \rangle \sum_{s=0}^{\infty} p_s \\ &= \langle s \rangle - \langle s \rangle \quad \text{(remember that } \textstyle\sum_{s=0}^{\infty} p_s = 1) \\ &= 0 \,. \end{aligned}$$

Because s is equally likely to lie above or below $\langle s \rangle$, the positive and negative contributions to the average cancel and $\langle s - \langle s \rangle \rangle$ is zero.

One way to avoid this cancellation would be to work out the expected value of $|s - \langle s \rangle|$, but the modulus function is mathematically awkward because the slope of $|x|$ changes discontinuously as x passes through the origin. A mathematically more convenient measure of the width of a probability distribution is the *standard deviation* σ, defined to be the (positive) square root of the *variance*

$$\sigma^2 = \left\langle (s - \langle s \rangle)^2 \right\rangle \,.$$

In words: the variance σ^2 is the expected value of the square of the distance of the score s from the mean score $\langle s \rangle$; and the standard deviation σ is the square root of the variance. The standard deviation is also called the root mean square (rms) width of the probability distribution. Unlike the variance, the standard deviation always has the same physical dimensions as the random variable s. (Both σ and σ^2 are dimensionless in our darts example.)

Another useful formula for the variance is

$$\sigma^2 = \langle s^2 \rangle - \langle s \rangle^2 \, .$$

This can be derived from the original definition as follows:

$$\begin{aligned}
\left\langle (s - \langle s \rangle)^2 \right\rangle &= \sum_{s=0}^{\infty} (s - \langle s \rangle)^2 \, p_s \\
&= \sum_{s=0}^{\infty} \left(s^2 - 2\langle s \rangle s + \langle s \rangle^2 \right) p_s \\
&= \sum_{s=0}^{\infty} s^2 p_s - 2\langle s \rangle \sum_{s=0}^{\infty} s p_s + \langle s \rangle^2 \sum_{s=0}^{\infty} p_s \\
&= \langle s^2 \rangle - 2\langle s \rangle \langle s \rangle + \langle s \rangle^2 \\
&= \langle s^2 \rangle - \langle s \rangle^2 \, .
\end{aligned}$$

Probability Densities

Until now we have been considering quantities such as scores in a dart game that can only take discrete (separate, quantised) values. A few small adjustments are required to apply the ideas of probability theory to continuous variables such as the heights of people.

Instead of throwing N darts, imagine that you measure the heights of N people. The probability that anybody in your sample is exactly 1.8 m tall is zero. Some people may be roughly 1.8 m tall, and a few may be very close to 1.8 m tall, but there is no chance of finding someone who is *exactly* 1.8 m tall (plus or minus nothing).

We can, however, ask about the number of people $n(h, h + \Delta h)$ with heights between h and $h + \Delta h$, where Δh is finite. Given a large enough sample, the ratio $n(h, h+\Delta h)/N$ is independent of N and we can define the corresponding probability

$$p(h, h + \Delta h) \;=\; \lim_{N \to \infty} \frac{n(h, h + \Delta h)}{N} \, ,$$

just as in the discrete case.

If Δh is small enough, the number of people with heights between h and $h + \Delta h$ ought to be proportional to Δh. (For example, one would expect the number of people with heights between 1.8000 m and 1.8002 m to be roughly double the number with heights between 1.8000 m and 1.8001 m.) This suggests defining a *probability density function* (pdf) $f(h)$ via:

$$p(h, h + \Delta h) \;=\; f(h)\,\Delta h\;.$$

As long as Δh is small enough, $f(h)$ should be independent of the value of Δh. The pdf is the central quantity in all applications of probability theory to continuous random variables.

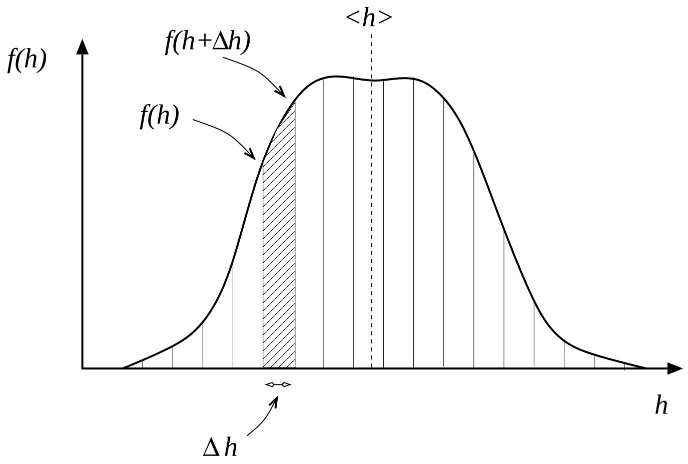

Figure 3: A probability density function

Figure 3 shows an example pdf. For small Δh, the probability $f(h)\Delta h$ that the height of a randomly chosen person lies between h and $h + \Delta h$ is approximately equal to the shaded area. The probability that the height lies between $h_{\rm low}$ and $h_{\rm high}$ (where $h_{\rm low}$ and $h_{\rm high}$ need not be close) is the area under the pdf from $h_{\rm low}$ to $h_{\rm high}$:

$$p(h_{\rm low}, h_{\rm high}) \;=\; \int_{h_{\rm low}}^{h_{\rm high}} f(h)\,dh\;.$$

Since the probability that h lies somewhere between zero and infinity is equal to 1, it follows that

$$\int_0^\infty f(h)\,dh \;=\; 1\;.$$

Expected Values

Following the darts example, the expected height $\langle h\rangle$ is defined via:

$$\langle h\rangle \;\approx\; (0 \times \text{Prob. height is between 0 and } \Delta h)$$

$$
\begin{aligned}
&+ (\Delta h \times \text{Prob. height is between } \Delta h \text{ and } 2\Delta h) \\
&+ \ldots \\
&+ (n\Delta h \times \text{Prob. height is between } n\Delta h \text{ and } (n+1)\Delta h) \\
&+ \ldots \\
\approx\ & \sum_{n=0}^{\infty} n\Delta h\, f(n\Delta h)\, \Delta h \\
=\ & \sum_{n=0}^{\infty} h_n f(h_n)\, \Delta h \ ,
\end{aligned}
$$

where $h_n \equiv n\Delta h$ is the value of the height h at the left-hand edge of the n^{th} strip of width Δh. In the limit as $\Delta h \to 0$, the summation turn into an integral and the $\approx$ signs become $=$ signs:

$$\langle h \rangle \ = \ \int_0^\infty h f(h)\, dh \ .$$

Just as in the discrete case, the expected value of any function $g(h)$ of the height h is defined via:

$$\langle g(h) \rangle \ = \ \int_0^\infty g(h) f(h)\, dh \ .$$

Variance and Standard Deviation

The variance σ^2 is defined exactly as in the discrete case:

$$\sigma^2 \ = \ \left\langle (h - \langle h \rangle)^2 \right\rangle \ = \ \langle h^2 \rangle - \langle h \rangle^2 \ ,$$

but the expectation values are now given by integrals,

$$\sigma^2 \ = \ \int_0^\infty (h - \langle h \rangle)^2\, f(h)\, dh \ = \ \int_0^\infty h^2\, f(h)\, dh - \left(\int_0^\infty h\, f(h)\, dh \right)^2 \ ,$$

instead of summations. The standard deviation is still the square root of the variance.