Avoiding Confusion in Probability

A tale of a Year 12 class

A few years ago, I taught a Year 12 Maths A class. At the beginning of the first probability unit I decided to try to get a feel for the students' understanding of probability. I placed a cup in one corner of the room, took a piece of chalk to the opposite corner, faced the corner and announced that I was about to throw the chalk over my shoulder and try to get it into the cup. Then I asked what they thought was the probability that I would succeed.

Without exception, the students answered 'half', '50-50' or words to that effect. This rather flattering estimate definitely didn't result from the fact that they had seen me play basketball. Their reasoning was that probability is the number of favourable outcomes over the total number of outcomes, $n(\mathrm{E})/n(\mathrm{S})$, and there were two possible outcomes here – in the cup and out of it. They all agreed on this. This was something they had obviously learned well in Year 11.

We then spent quite a while in debate, neither I nor them managing to shift the opinion of the other. It boiled down to different definitions of probability: mine was the fraction of times the event would occur in the long run; theirs was $n(\mathrm{E})/n(\mathrm{S})$. Of course, on a matter of definition, logical debate is pointless. It came down to the relative authority of their Year 11 teacher and me – and I lost.

A resolution came only when we invited their Year 11 teacher to come in and talk about it. As I expected, this teacher was under no illusion that probability was defined as $n(\mathrm{E})/n(\mathrm{S})$ and after a few minutes, the class was happy to accept a new definition.

This is a long story to make a point that students are often confused about very basic concepts in probability, even after several years of teaching on the subject. Many textbooks present these basic concepts in ways that are misleading or just plain wrong.

In this article I am going to suggest a way of looking at elementary probability (Year 8 to 10 stuff) that I have found allows less chance of confusion.

There are basically four things to learn about quantifying probability:

- the definition of probability
- approximating probabilities using data
- determining probabilities of simple events by symmetry
- deriving probabilities of compound events from known probabilities of the component simple events

The definition of probability

For school mathematics, probability is best defined as the fraction of times something would happen in the long run (i.e. if tried a very large number of times).

Approximating probabilities using data

This definition leads to collection of data as the obvious way to determine probability. Of course, data will only ever give an approximation to the true probability. Students should have plenty of hands-on practice at finding probabilities from data, exploring issues like reliability of results from different numbers of trials and fallacies like memory and dependence of the outcome of tossing a coin on which way up it was before it was tossed etc.

Data used to determine a probability can be procured in two ways. In some cases the data might be generated in an experiment conducted for the purpose of finding the probability. In other cases, the data may already be in existence – pre-existing data. Medical probabilities are usually determined from pre-existing data rather than by deliberately infecting healthy people for the sake of an experiment.

Determining probabilities of simple events by symmetry

In approximating probabilities using data, students might notice that certain probabilities seem to come to simple fractions, while other don’t. For example the probability of getting a head when tossing a coin seems to come to one half, while the probability of a matchbox landing on its side might come to about 23%. Many students will have an explanation for the nice fractions. The explanation lies in the symmetry of the experiment. In the case of tossing a coin, there is no difference between the two sides of the coin that will make one side any more likely to land upwards than the other. When rolling a die, there is no such difference between the six faces. When throwing a matchbox, however, there is a material difference between the different faces that will make it more likely that the matchbox will land on some faces than on other. Probabilities for the matchbox cannot be determined by symmetry.

Symmetry is the only way to determine the probability of simple events other than using data. The concept of symmetry is not mentioned in most textbooks. Yet it is relied upon in most questions.

Deriving probabilities of compound events

This is the calculation of probabilities of compound events from known probabilities of the component simple events.

Probabilities of multi-outcome events are calculated by adding the probabilities of the component events. To calculate of the probability that a matchbox lands on its side or on its end, we need to know the probabilities of these simple events. Let's say they are 0.23 and 0.06 respectively. The probability of landing on its side or its end is then 0.23 + 0.06, i.e. 0.29. Another example is the probability of drawing a king or queen from a pack of cards. There are four kings and four queens, each with a probability of $^{1}/_{52}$, giving a total of $^{8}/_{52}$.

Now the rule of $p = n(\mathrm{E})/n(\mathrm{S})$ can be applied to the card situation because all the simple probabilities are the same, but this rule cannot be applied to the matchbox – nor to the chalk! There is symmetry in the card situation, but not in the matchbox or chalk situations. It is important to ensure that students gain the concept of addition of probabilities and that, if the $p = n(\mathrm{E})/n(\mathrm{S})$ rule is introduced, it is done as a shortcut to the addition procedure that can be applied when all the probabilities being added are the same.

Probabilities of multi-stage experiments, i.e. the probability that one thing happens AND that another thing happens, are calculated by multiplying the component probabilities. An example is calculating the probability that two matchboxes will both land on their sides. This is 0.23×0.23, i.e. 0.053.

These techniques are crystallised into the addition and multiplications rules, though other techniques like tree diagrams and tables are often used as an alternative to using the multiplication. Later, the more general versions of these rules can be introduced for situations where events are not mutually exclusive or not independent.

Experimental and theoretical determination of probability

So probabilities can be approximated using data. Or they can be determined using symmetry. Or they can be determined from other known probabilities using the addition and multiplication rules. The latter two methods are often grouped under the heading of determining probability theoretically. This grouping, however, tends to obscure the important distinction between the two methods. It may be better to talk about approximating probabilities using data, determining probabilities by symmetry, and deriving probabilities from other known probabilities, rather than using the two-way classification into experimental and theoretical methods.

As well as this, use of the terms 'experimental probability' and 'theoretical probability' is problematic. These terms convey the impression that there are two types of probability. As if somehow the probability of getting a head on a coin is ½, but if you do a one-trial experiment and get a tail, then the probability is 0. This is not only misleading; it is not true. There is only one probability for an event – the probability of getting a head is ½. There are just different ways of finding it.