



## Understanding true probability, model estimates and experimental estimates

**True probability** is the (almost always) unknown actual probability that an event will occur in a given situation. The actual probability of a coin landing heads up is affected by the position from which it is tossed, the asymmetry of the two faces of the coin etc, so is not exactly 0.5, though the probability of a fair coin landing heads will be very close to 0.5. We can find out about the unknown true probability by observation (experiment) or by trying to understand the situation and modelling it.

In probability an **experiment** is one or more trials of a probability situation. An **experimental estimate** is calculated from observation as the number of successful trials divided by the total number of trials. In the long run (over many trials), the experimental estimate may approach the true probability.

An experimental estimate that a coin will land heads if it is tossed 20 times and lands heads up 14 times is  $14/20 = 0.7$ .

A **probability model** is a representation of a situation involving probability. Probability models can incorporate experimental estimates and assumptions about the situation (e.g., independence). These assumptions may be based on an idealised view of the world and an understanding of the mathematics of probability.

A **model estimate** is an estimate of the probability that an event will occur, based on a probability model. The model estimate of a fair coin landing heads is 0.5.

If a model is a good representation of the situation, the experimental estimate over many trials will be close to the model estimate.

A model must always be considered in context. A good model is one which is fit for the purpose for which it is being used. When tossing an approximately fair coin, the model estimate of  $P(\text{heads}) = 0.5$  is a good model for most purposes. A transport system modelling the timing of traffic lights to get a smooth flow of traffic will require a more complex model, tested against experimental observations to ensure that it is fit for the purpose.

In some situations there is no obvious theoretical model, so we can only estimate the probabilities and probability distributions via experiment. These estimates can be used as a basis for building a theoretical model. For instance, to develop a model of the probability of getting a basketball through the hoop, an initial model might assume a constant probability of 0.5. As data is gathered, there could be successive refinements of the model so that it becomes a better estimate of the true probability. The data might indicate that the probability of getting the ball in the hoop is closer to 0.2 and that it changes over time.

Sometimes we might think that an obvious theoretical model applies, but experimental estimates demonstrate that our model is a poor one. There is now a need to find a better model using the estimates from the experiments. We might initially model the result of spinning a weighted coin as  $P(\text{heads}) = 0.5$  but realise that that estimate is a poor one and use data to improve it.