

SIMULATION AND UNCERTAINTY

Tony O'Hagan

Department of Probability and Statistics
The University of Sheffield
Western Bank
Sheffield, S10 2TN, UK

ABSTRACT

In the talk, I will address the following two themes:

Elicitation. How do we quantify expert judgments in probabilistic form? The probabilities are necessarily subjective. How do we get inside the expert's head to measure them? What are the pitfalls in doing so? How do we minimize the subjectivity?

Parameter uncertainty in simulation. Given that parameter uncertainty is epistemic, how do we quantify it in practice? Does it have to involve elicitation? How do we efficiently propagate parameter uncertainty through a model, particularly when that model is itself stochastic because it is dealing with aleatory uncertainties?

I will illustrate both themes by looking at Bayesian clinical trial simulation.

1 TWO KINDS OF UNCERTAINTY

Simulation is an almost ubiquitous tool of modern research. Uncertainty is an almost ubiquitous feature of modern life. The two come together in a number of different ways. In many areas of science and technology, people build sophisticated mathematical models to simulate real-world phenomena. Such simulators can be extremely complex, and their computer implementations may take hours to run. Generally, however, they are deterministic; if run again with the same settings they will produce exactly the same outputs. Yet there is uncertainty here – uncertainty in the adequacy of the model, in its correct implementation, in the values of the various input parameters that must be set to run it, and in other parameters that are hard-coded in the equations. In the I-Sim setting, however, simulation usually means something stochastic. The results of such simulations are not deterministic, and when we run them again with the same settings we can expect different outputs. Here, simulation and uncertainty are firmly linked, yet uncertainty can enter in subtle ways, as well as through the obvious stochastic nature of the simulation.

It is often useful to distinguish two kinds of uncertainty. The first is aleatory uncertainty (after the Latin *alea* meaning a die, as in the quotation 'alea jacta est' – the die

is cast), which is associated with randomness or inherent variability in phenomena. This is the familiar kind of statistical uncertainty and it is quantified by the familiar idea of probability as the long-run relative frequency with which some event happens.

The second kind is epistemic uncertainty (after the Greek *episteme* meaning knowledge), which is associated with imperfect knowledge. Some things that I am uncertain about are: whether the next US president will be a Republican; the atomic weight of Ruthenium; how many people died at the battle of Agincourt. My uncertainty about all of these is epistemic. If I want to quantify that uncertainty, I cannot use the familiar definition of probability because none of these events can be considered random and repeatable. (If you doubt this statement in respect of the event that the next US president is Republican, simply ask yourself whether you would be willing to estimate its probability just by the historical relative frequency with which the Republican Party has won the presidency.)

All of these are unique, one-off things, and we cannot apply the usual definition of probability, yet people are generally quite happy to use the language of probability when referring to them. For instance, we talk of the probability of rain this afternoon, although the weather conditions today are unique and we would not think that the historic frequency of rainy afternoons would make a good estimate.

It is in the nature of epistemic uncertainty that it is subjective and impermanent. Your imperfect knowledge is not the same as mine, and knowledge changes over time. Reading this, you are much more uncertain about my height than I am; after seeing my talk you will be much less uncertain than you are now, but still more uncertain than I am.

2 TWO KINDS OF PROBABILITY

In order to address epistemic uncertainties, we need a new definition of probability. The conventional one, based on long-run frequency of occurrence, is called frequency probability. The new one is called subjective probability, or personal probability, and it is defined simply as a measure of the degree of belief that you have in the truth of some uncertain proposition.

Note that subjective probability can deal with both kinds of uncertainty, but frequency probability only exists for aleatory uncertainties. Notice also that it is by no means easy to specify a precise probability for anything, using either definition! If we want to specify a frequency probability, then in principle we must repeat the event in question an infinite number of times, in order to find the exact limiting relative frequency. Nobody has *ever* done this, so nobody has *ever* known the probability of *anything* according to that definition. You may be used to saying that the probability that a coin falls 'heads' when tossed is one-half, or the probability that a die lands as a '6' when rolled is one-sixth, but these are actually just beliefs – nobody has done the experiment.

Subjective probabilities are no easier to specify with absolute precision, because the beliefs are in the person's head. If you ask me to specify my probability for the next US president being a Republican, I might say 0.4, but I don't mean 0.4000000000. I don't even necessarily mean 0.4 to the nearest 0.1, because I'm not sure that I can assess my own knowledge and beliefs that accurately. The process of quantifying subjective probabilities is one of the themes of this talk.

One thing is clear, though. If I ever were to find out a frequency probability, then my subjective probability for that event would be the same. That is, if I knew that in a huge (effectively infinite) run of repetitions of some event, it actually occurred on 86.2357% of occasions, then my probability for it occurring on the next repetition will be 0.862357. (This is not only kind of obvious but can be proved mathematically.)

So subjective probability encompasses frequency probability as a special case.

3 TWO KINDS OF STATISTICS

The different notions of probability lie at the heart of the debate between Bayesian and frequentist statisticians. Statistical inference is all about trying to learn about the values of unknown parameters using the data. The parameters are unknown because of our lack of knowledge, and we hope to reduce our uncertainty through observing the data. Almost invariably, the uncertainty about these parameters is wholly (or at least partly) epistemic. Conventional frequentist statistical methods rely on the frequency definition of probability, and as a result they do not regard parameters as having probabilities. In contrast, Bayesian statistics is generally seen as founded on subjective probability, and it can and does provide probability statements about parameters.

These distinctions can be difficult and surprising, because for instance a 90% confidence interval for some unknown parameter is almost universally interpreted as saying that there is a 90% probability that the parameter lies in the given interval. Yet this is precisely what it cannot mean,

since that would be making a probability statement about the parameter. All the probabilities in frequentist theory are statements about the data, which is regarded as random and repeatable. All the statements apply in the context of long-run repetition of the experiment that gave rise to the observed data. The correct interpretation of a 90% confidence interval is that if the experiment were repeated many times, and the same rule for calculating a confidence interval were applied to each such dataset, then 90% of those intervals would contain the true value of the parameter.

This is not the place to delve into these issues more deeply, but the point is that there are fundamental differences between the two theories of statistics, and the statements of frequentist analyses are almost always misinterpreted. This talk will explore some implications of the two kinds of uncertainty and probability in the context of simulation.

AUTHOR BIOGRAPHY

TONY O'HAGAN is Professor of Statistics at the University of Sheffield. His research field is Bayesian Statistics, in which he has made many contributions to the theory, methodology and applications. Relevant current research interests are the elicitation of expert knowledge and quantifying uncertainty in complex mechanistic models.

