

In praise of Bayes

Sep 30th 2000



Bayesianism is a controversial but increasingly popular approach to statistics that offers many benefits-although not everyone is persuaded of its validity

IT IS not often that a man born 300 years ago suddenly springs back to life. But that is what has happened to the Reverend Thomas Bayes, an 18th-century Presbyterian minister and mathematician-in spirit, at least, if not in body. Over the past decade the value of a statistical method outlined by Bayes in a paper first published in 1763 has become increasingly apparent and has resulted in a blossoming of "Bayesian" methods in scientific fields ranging from archaeology to computing. Bayes's fans have restored his tomb and posted pictures of it on the Internet, and a celebratory bash is planned for next year to mark the 300th anniversary of his birth. There is even a Bayes songbook-though, since Bayesians are an academic bunch, it is available only in the obscure file formats that are used for scientific papers.

Proponents of the Bayesian approach argue that it has many advantages over traditional, "frequentist" statistical methods. Expressing scientific results in Bayesian terms, they suggest, makes them easier to understand and makes borderline or inconclusive results less prone to misinterpretation. Bayesians claim that their methods could make clinical trials of drugs faster and fairer, and computers easier to use. There are even suggestions that Bayes's ideas could prompt a re-evaluation of fundamental scientific concepts of evidence and causality. Not bad for an old dead white male.

Previous convictions

The essence of the Bayesian approach is to provide a mathematical rule explaining how you should change your existing beliefs in the light of new evidence. In other words, it allows scientists to combine new data with their existing knowledge or expertise.

The canonical example is to imagine that a precocious newborn observes his first sunset, and wonders whether the sun will rise again or not. He assigns equal prior probabilities to both possible outcomes, and represents this by placing one white and one black marble into a bag. The following day, when the sun rises, the child places another white marble in the bag. The probability that a marble plucked randomly from the bag will be white (ie, the child's degree of belief in future sunrises) has thus gone from a half to two-thirds. After sunrise the next day, the child adds another white marble, and the probability (and thus the degree of belief) goes from two-thirds to three-quarters. And so on. Gradually, the initial belief that the sun is just as likely as not to rise each morning is modified to become a near-certainty that the sun will always rise.

In a Bayesian analysis, in other words, a set of observations should be seen as something that changes opinion, rather than as a means of determining ultimate truth. In the case of a drug trial, for example, it is possible to evaluate and compare the degree to which a sceptic and an enthusiast would be convinced by a particular set of results. Only if the sceptic can be convinced should a drug be licensed for use.

This is far more subtle than the traditional way of presenting results, in which an outcome is deemed statistically significant only if there is a better than 95% chance that it could not have occurred by chance. The problem, according to Robert Matthews, a mathematician at Aston University in Birmingham, is that medical researchers have failed to understand that subtlety. In a paper to be published shortly in the *Journal of Statistical Planning and Inference*, he sets out to demystify the Bayesian approach, and explains how to apply it after the event to existing data.

Patients in clinical trials will soon benefit. Bayesian methods offer the possibility of modifying a trial while it is being conducted, something that is impossible with traditional statistics. Andy Grieve and his colleagues at Pfizer, a drug firm, are intending to do just that.

Traditionally, dose-allocation trials-in which the aim is to establish the most effective dose of a new drug-involve giving different groups of patients different doses and evaluating the results once the trial has finished. This is fine from a statistical point of view, but unfair on those patients who turn out to have been given non-optimal doses. Rather than analysing the results at the end of a trial, Dr Grieve's method will evaluate patients' responses during it, and adjust the doses accordingly. The advantage of this over the traditional approach that it maximises the medical benefit to all participants. It also means that fewer people are needed to conduct a trial, because participants on non-optimal doses can have those doses changed to increase the amount of data collected near the optimal dose.

Pfizer is intending to conduct a trial using this new method, and the plan is to re-analyse the data once it is completed in ways that will satisfy both Bayesians and non-Bayesians. This kind of parallel approach is likely to become increasingly common, at least until Bayesianism has been more widely accepted.

Bayesian methods can also be used to decide between several competing hypotheses, by seeing which is most consistent with the available data. This idea was recently used to determine the date of construction of "Seahenge", an ancient timber circle found off the coast of Norfolk, in eastern England. Results from tree-ring dating were inconclusive, suggesting such divergent dates as 2019BC, 2050BC and 2454BC. So six samples from the monument's central stump were radiocarbon-dated, and the results were used to evaluate the three tree-ring possibilities via Bayesian analysis. The evidence was overwhelmingly in favour of 2050BC, and was inconsistent with either of the other two tree-ring dates.

Decision-making using Bayesian methods has many applications in software, as well. Perhaps the best-known example is Microsoft's Office Assistant, which appears as a somewhat irritating anthropomorphic paper-clip that tries to help the user. When a user calls up the assistant, Bayesian methods are used to analyse recent actions in order to try to work out what the user is attempting to do, with this calculation constantly being modified in the light of new actions. (Unfortunately, a non-Bayesian approach is used to decide when to make the paper-clip pop up on its own, adding to the annoyance of many users.) According to Eric Horvitz, a statistician in Microsoft's research division, future products will try to determine users' intentions more broadly, so as to speed things up. Having worked out which link on a web article a user is most likely to click on, for example, the computer could fetch the corresponding article in advance, so that it appears more quickly.

Bayes is still, however, the focus of much controversy. Larry Wasserman, a statistician at Carnegie Mellon University, in Pittsburgh, says that although Bayesianism is becoming more acceptable, it is no panacea, and when used indiscriminately it becomes "more a religion than a science". Perhaps the grandest claims made for Bayesian methods are those of Judea Pearl, a computer scientist at the University of California, Los Angeles. Dr Pearl has suggested that by analysing scientific data using a Bayesian approach it may be possible to distinguish between correlation (in which two phenomena, such as smoking and lung cancer, occur together) and causation (in which one actually causes the other). This kind of claim makes many scientists, including many Bayesians, throw up their hands in horror. Evidently there is life in the old reverend yet.