STATISTICS

Forecasting with finesse

Paul Ormerod assesses a Bayesian take on predicting everything from poker games to climate change.

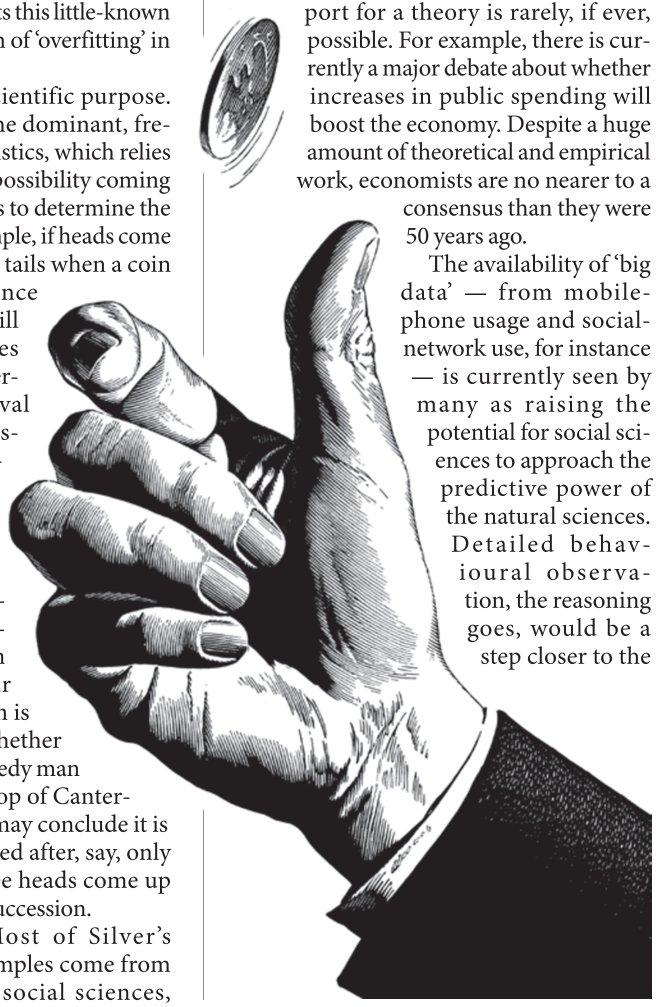
Nate Silver is an intriguing man. A statistician by training, he developed a system for forecasting baseball performance that has had real influence on how the top teams evaluate potential players. His blog made accurate predictions about the 2008 US Presidential election. He has a regular slot in the *New York Times*. And now, in his thoughtful, engaging *The Signal and the Noise*, he offers an array of fascinating examples of forecasting, from baseball and elections to poker, chess, stock markets, terrorist attacks, earthquakes and climate change.

The 'signal' in the title refers to genuine information, which can be used for prediction. 'Noise' is the purely random component of data, which cannot. A serious problem with many forecasting models is that they try to explain too much, and end up 'explaining' the noise. Silver documents this little-known but fundamental problem of 'overfitting' in clear terms.

Silver has a serious scientific purpose. He is highly critical of the dominant, frequentist approach in statistics, which relies on the frequency of one possibility coming up over a number of trials to determine the true probability. For example, if heads come up more frequently than tails when a coin is tossed, your confidence that the coin is biased will increase the more tosses you observe. Silver is a fervent advocate of the rival approach: Bayesian statistics. This is just as mathematically rigorous, but is a more heuristic approach, in which you form a view about the chances of an event happening and revise this estimate as new information comes to light. So, your prior estimate that a coin is biased will depend on whether the person flipping is a seedy man in a bar or the Archbishop of Canterbury. If the former, you may conclude it is biased after, say, only three heads come up in succession.

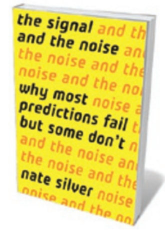
Most of Silver's examples come from the social sciences, which encompass the really difficult problems, such as forecasting the economy. In the natural sciences, a theory that explains the past will also, in general, be usable for making predictions about the future. Not so the social sciences, in which a clear distinction usually needs to be made between explanation and prediction. For example, there are models that describe well why price changes in stock markets in the past exhibit the features they do; these are based on how traders behave. Forecasting is much harder; and 'predictive' models are essentially pure extrapolations, over very short time periods, of past data, with no behavioural content.

In the natural sciences, theories can be developed and verified — or, to be more accurate, not falsified — by the evidence obtained from replicable experiments. In the social sciences, such strong support for a theory is rarely, if ever, possible. For example, there is currently a major debate about whether increases in public spending will boost the economy. Despite a huge amount of theoretical and empirical work, economists are no nearer to a consensus than they were 50 years ago.



GETTY

The availability of 'big data' — from mobile-phone usage and social-network use, for instance — is currently seen by many as raising the potential for social sciences to approach the predictive power of the natural sciences. Detailed behavioural observation, the reasoning goes, would be a step closer to the realm of replicable experiments. Silver is profoundly sceptical about this possibility, and argues that a massive increase in data will make predictions more prone to failure, not less. As he puts it, "the number of *meaningful* relationships ... is [tiny]... there isn't any more truth in the world than there was before the Internet".



The Signal and the Noise: Why So Many Predictions Fail — but Some Don't
NATE SILVER
Penguin: 2012. 352 pp. $27.95

Silver points out a key critique of Bayesian analysis, which is that it introduces the 'unscientific' concept of personal judgement. But in the social sciences, this is required all the time, even when the frequentist approach is used.

A revealing case emerged in 2007, when the Bank of England produced its 'fan charts', showing the potential range of outcomes for economic growth over the next five years. This indicated that the probability of a UK recession occurring in 2008–12 was essentially zero. But the economists had used data from only 1993–2006 to calibrate the probability distribution of growth rates; they had convinced themselves that their 1990s theories had solved problems of the economy once and for all. Using the frequentist approach, they had judged data from before 1993 as irrelevant — which we now know to be profoundly wrong.

Silver discusses the poor record of economic forecasting at length, and is correct in stressing the importance of understanding the data, rather than just pouring it into a statistical package and pressing the button. He does not, however, make clear how a Bayesian approach could have predicted the financial crisis.

Nor does he discuss the extent to which a given data set contains any signal at all, rather than being completely dominated by noise. In data sets in the social sciences, there are often many factors that can influence outcomes, and behaviour varies over time; so the data produced by a system is often indistinguishable from a random series. In such cases, neither the frequentist nor the Bayesian approach will produce successful forecasts.

But, overall, Silver does make a good case for a Bayesian approach. His book is a lively, well-argued corrective to the prevailing view that we need large amounts of data before we can make intelligent and accurate forecasts. ■

Paul Ormerod *is an economist and complex-systems theorist, and author of* Positive Linking: How Networks Can Revolutionise the World.
e-mail: pormerod@volterra.co.uk

➔ NATURE.COM
For more on Bayes' theorum of probability, see:
go.nature.com/2b3ecn