<u>"The Signal & the Noise" - A Paean to Intellectual Honesty</u>

Best known today for his success as a forecaster in the recent US presidential contest (he forecast the result correctly in every individual state), Nate Silver's fierce commitment to objectivity stems from a rather unexpected source – his experience as an online poker player. He learnt that, for anybody numerate enough to understand the odds involved, the most important cause of failure in this activity is overconfidence: believing that you have seen a pattern (the Signal) when all you have really seen is random fluctuations (the Noise). In this book he applies this and a variety of related lessons to a wide range of different activities where forecasts are involved, from baseball player statistics to climatology and from global economics to earthquakes. He explains why some fields – such as earthquakes – have proved highly resistant to forecasting progress, whereas others – such as hurricane prediction - have seen valuable advances. In every case he shows in detail how self-deception, whether due to emotional bias or to external pressures, can lead forecasters astray. And he demonstrates the lesson by painstakingly identifying the limitations of his own critiques.

A key central theme turns out to be the contrast between the "Fisherian" and "Bayesian" traditions of statistics, an apparently arcane dispute which actually appears to be of great importance to an understanding of the growth of objective knowledge. The essential difference is that Bayesian statistics (invented in the 1760's) actually requires you to start out with a preconception of what kind of relationship you expect to see, whereas Fisherian statistics only takes the "facts" as input. As a result, R. A. Fisher claimed that his approach was far sounder, essentially free from bias, and so for most of the twentieth century it was only Fisher's approach that was taught. But in the present century Bayesian statistics has made a big comeback, and Silver convincingly argues that, paradoxically, it is the more "honest" approach for many applications. The reason for this is that the objective outcome of conventional statistics means nothing without an understanding of the context. A "97% confidence level" in, say, a correlation between some sporting result and a political election result tells us that such a correlation has only a 3% probability of occurring by chance, but begs the question of what else could have caused it; whereas a similar correlation between, say, some property of a material and its exact composition does indeed give grounds for confidence in a causal relationship. In Bayesian statistics, consideration of the context is built into the method. Given an initial estimate of confidence in a relationship, every new piece of data improves that estimate.

Silver even suggests that Bayesian statistics can be thought of as a model for the "scientific method". He doesn't really spell out how, but I think I can see the point. To ultra-summarise, the traditional theory of scientific knowledge stressed the process of "induction", whereby repeated observation of similar facts is taken as grounds for propounding a new scientific hypothesis. Against this, Karl Popper famously asked how many white swans you need to observe to prove the proposition that all swans are white. (They aren't, of course). Popper's account suggested that the hypothesis has to come first, from which predictions can be derived (the process of "deduction"), and the role of observation is to check the predictions so as to eliminate false hypotheses. Popper is now rather unfashionable among philosophers – the actual work of scientists too often looks like induction for one thing – but the workings of Bayes' theorem suggests a kind of hybrid process, which perhaps can be applied to qualitative as well as to quantitative knowledge. Yes, the starting point has to be a hypothesis, as Popper claimed, but observations don't just serve to falsify it or otherwise, they can also refine it, build it, extend it maybe in unexpected directions. Induction and deduction, in short, have to work hand in hand. It seems a good model to keep in mind when considering scientific debates.

Roger Haines