Paradigms for Statistical Inference

Murray Aitkin
University of Melbourne

April 16, 2012

1 References

Kuhn, T.S. (1962) *The Structure of Scientific Revolutions* (3rd edn. 1996). University of Chicago Press.

Kuhn, T.S. (2000) *The Road since Structure*, eds. J. Conant and J. Haugeland. University of Chicago Press.

Examples of recent paradigm challenge

Jaynes, J. (1976) *The Origin of Consciousness in the Breakdown of the Bicameral Mind*. Boston: Houghton Mifflin.

Gammage, B. (2011) *The Biggest Estate on Earth: How Aborigines Made Australia*. Sydney: Allen and Unwin.

2 Kuhn's concerns

Thomas Kuhn (1922-1996) was professor of linguistics and philosophy at MIT. He was trained as a theoretical physicist, and became interested in the history of science as a graduate student when asked to jointly teach an experimental course in physical science for non-science students. This course contained descriptions of out-of-date scientific theory and practice which made a deep impact on him, and this developed into a concern for what he saw as a *mismatch* between the progress of science as described in standard textbooks for students, and the reality of its progress as described historically by scientists at the time. A particularly serious aspect of this mis-representation was the idea that the progress of science was a smooth process of cumulative discovery which has led to the present state of our understanding in an *evoutionary* way.

Kuhn's great contribution in his book and other writings was to challenge this representation fundamentally – to stand it on its head – by arguing that on the contrary, science progressed by *revolutionary* changes, in what he termed *paradigms for research*.

Paradigms are

> ... universally recognized scientific achievements that, for a time, provide model problems and solutions for a community of researchers." (1962 p. viii.)

A scientific paradigm provides a *framework* for the progress of the science, what Kuhn calls "normal science": it determines

what is to be observed and scrutinized;

the kind of questions that are supposed to be asked and probed for answers in relation to this subject;

how these questions are to be structured;

how the results of scientific investigations should be interpreted.

(Wikipedia)

3 The evolution of paradigms

> ... [O]ne of the things a scientific community acquires with a paradigm is a criterion for choosing problems that, while the paradigm can be taken for granted, can be assumed to have solutions. To a great extent these are the only problems that the community will admit as scientific or encourage its members to undertake. Other problems, including many that had previously been standard, are rejected as metaphysical, as the concern of another discipline, or sometimes just too problematic to be worth the time.

Once adopted, a new paradigm provides a stimulus to the development of research – some previously intractable problems are now seen to have new solutions, while other problems which were formerly seen as important are now abandoned because the new paradigm specifies other fruitful directions which are now to be followed. Much of the contribution of the old paradigm however remains in the new paradigm, with the new paradigm giving essentially the same solutions as the old.

"Normal science" now becomes the formulation and solution of the problems that are described by the new paradigm, and the success and importance of the paradigm are determined by the speed with which these new problems are solved, and in turn suggest further extensions of the problems which can be solved within the paradigm.

> To scientists, at least, the results gained in normal research are significant because they add to the scope and precision with which the paradigm can be applied. ... Bringing a normal research problem to a conclusion is achieving the anticipated in a new way, and it requires the solution of all sorts of complex instrumental, conceptual, and mathematical puzzles. The man who succeeds proves himself an

> expert puzzle-solver, and the challenge of the puzzle is an important part of what usually drives him on. (pp. 35-6)
>
> The scientific enterprise as a whole does from time to time prove useful, open up new territory, display order, and test long-accepted belief. Nevertheless, *the individual* engaged on a normal research problem *is almost never doing any of these things*. Once engaged, his motivation is of a rather different sort. What then challenges him is the conviction that, if only he is skilful enough, he will succeed in solving a puzzle that no one before has solved or solved so well. Many of the greatest scientific minds have devoted all of their professional attention to demanding puzzles of this sort. On most occasions any particular field of specialization offers nothing else to do, a fact that makes it no less fascinating to the proper sort of addict. (p. 38, author's emphasis)

As the paradigm is articulated further, difficulties begin to arise – the problems become more difficult to solve, and for some problems it is not clear that a solution can be found:

> The early attacks on the resistant problem will have followed the paradigm rules quite closely. But with continuing resistance, more and more of the attacks upon it will have involved some minor or not so minor articulations of the paradigm ... Through this proliferation of divergent articulations, (...more and more frequently described as *ad hoc* adjustments), the rules of normal science become increasingly blurred. Though there is still a paradigm, few practitioners prove to be entirely agreed about what it is. (p. 83)

In most cases, the problem is eventually solved, by a very difficult process, *within the paradigm*, which then continues until a further problem becomes even more difficult, and an *anomaly* – an intractable difficulty – appears. The paradigm is stuck, and is in *crisis*.

4 Revolutionary changes in paradigm

What happens then? This is Kuhn's fundamental argument. the crisis can be resolved only by *a change in the paradigm*; the difficulty arises from the structure of the present paradigm:

> Anomaly appears only against the background provided by the paradigm. The more precise and far-reaching that paradigm is, the more sensitive an indicator it provides of anomaly and hence for an occasion of paradigm change. (p. 65)
>
> For what is it that differentiates normal science from science in a crisis state? ... What we previously called the puzzles that constitute

normal science exist only because no paradigm that provides a basis for scientific research ever completely resolves all its problems... [E]very problem that normal science sees as a puzzle can be seen, from another viewpoint, as a counterinstance and thus a source of crisis. (p. 79)

5 Reactions to paradigm crisis

X-rays ... were greeted not only with surprise but with shock. Lord Kelvin at first pronounced them an elaborate hoax. Others, though they could not doubt the evidence, were clearly staggered by it. Though X-rays were not prohibited by established theory, they violated deeply entrenched expectations ... Previously completed work on normal projects would now have to be done again because earlier scientists had failed to recognize and control a relevant variable. X-rays, to be sure, opened up a new field and thus added to the potential domain of normal science. But they also ... changed fields that had already existed. In the process they denied previously paradigmatic types of instrumentation their right to that title. (p. 59)

... what scientists never do when confronted by even severe and prolonged anomalies. ...[T]hey do not renounce the paradigm that has led them into crisis. They do not, that is, treat the anomalies as counter-instances... They will devise numerous articulations and *ad hoc* modifications of their theory in order to eliminate any apparent conflict. ...[E]ven a discrepancy unaccountably larger than that experienced in other applications of the theory need not draw any profound response. There are always some discrepancies. Even the most stubborn ones usually respond at last to normal practice. Very often scientists are willing to wait, particularly if there are many problems available in other parts of the field. (pp. 77 and 81)

The invention of ... new theories regularly, and appropriately, evokes the same response [of resistance] from some of the specialists on whose areas of special competence they impinge. For these men the new theory implies a change in the rules governing the prior practice of normal science. Inevitably, therefore, it reflects upon much scientific work they have already successfully completed. That is why a new theory, however special its range of applications, is seldom or never just an increment to what is already known. Its assimilation requires the reconstruction of prior theory and the re-evaluation of prior fact, an intrinsically revolutionary process that is seldom completed by a single man and never overnight. (p. 7)

6 How does the conversion from the old paradigm to the new come about?

> [W]hen ...a new paradigm has been evoked, scientists will be reluctant to embrace it unless convinced that two all-important conditions are being met. First, the new paradigm must seem to resolve some outstanding and generally recognized problem that can be met in no other way. Second, the new paradigm must promise to preserve a relatively large part of the concrete problem-solving ability that has accrued to science through its predecessors. (p. 169)

> Fortunately, there is also another sort of consideration that can lead scientists to reject an old paradigm in favor of a new. These are the arguments, rarely made entirely explicit, that appeal to the individual's sense of the appropriate or the aesthetic – the new theory is said to be "neater", "more suitable," or "simpler" than the old. (p. 155)

> The transfer of allegiance from paradigm to paradigm is a conversion experience that cannot be forced. Lifelong resistance, particularly from those whose productive careers have committed them to an older tradition of normal science, is not a violation of scientific standards but an index of the nature of scientific research itself. The source of resistance is the assurance that the older paradigm will ultimately solve all its problems, that nature can be shoved into the box the paradigm provides. ... [This] assurance is what makes normal or puzzle-solving science possible. (p. 151)

> A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it. (p. 151 quoting Max Planck)

7 The relevance of Kuhn to statistics

Statistics is not one of the physical or social sciences: we do not build theories based on experimental or observational data – that is the role of the scientist with whom the statistician may be collaborating. Statistics is an *enabling technology*, to use a current buzz-word: we know how to *design* experiments and surveys and *use* experimental or observational data to bear on scientific questions, through probability models and their statistical analysis.

So it is not obvious that Kuhn's paradigm views are relevant to the inferential disputes among statisticians. However, I quoted very liberally from Kuhn to be able to show that his description of paradigm change and conflict *fits very well* with the disputes in statistics.

The term paradigm has come to be commonly used in statistics in *the Bayesian paradigm*, and less frequently, in *the frequentist paradigm*. These multiple usages of the term imply, in the Kuhnian sense, that there are separate communities of researchers following these paradigms, since the paradigms are inconsistent in their principles (axioms) for the analysis of data. We will call these separate communities *schools*. This might not matter if these schools attended to different problems in the applications of statistics, but it *does* matter since the applications of the schools overlap very substantially, and the Bayesian and frequentist schools at least claim *universality* in the breadth of their applications.

There are other paradigms (the likelihood paradigm for example), and other approaches to data analysis, which do not have, or do not claim, the generality of the Bayesian and frequentist schools. The survey sampling design-based theory is not usually called a paradigm, though there are exceptions – the "total survey error paradigm" for example (Groves and Lyberg 2010). The survey sampling school had in the past a much narrower focus, and sometimes seemed to be obsessively fixed on the estimation of the population mean alone. This has changed considerably in recent years, and the claims made for the generality of its theory have widened; we will call this approach the *survey sampling paradigm*.

There are many practising statisticians who would not regard themselves as members of only one, or any, school. Many statistical scientists are agnostic about theories/paradigms: "I'm not a Bayesian, but I use what works. I'm an empiricist – if MCMC works I use it." (Without a paradigm, how can one tell if anything works?)

Statements like these make plain that statistics does not have a *single* paradigm, and the conflict among the major schools is a clear example of what Kuhn calls the *pre-paradigm* state of theories:

> There are schools in the sciences, communities, that is, which approach the same subject from incompatible viewpoints. But they are far rarer there than in other [non-science] fields; they are always in competition, and their competition is usually quickly ended. (p. 177)
>
> Probably the most striking [issue of scientific community structure] ... is the transition from the pre- to the post-paradigm period in the development of the scientific field. Before it occurs, a number of schools compete for the domination of a given field. Afterward, in the wake of some notable scientific achievement, the number of schools is greatly reduced, ordinarily to one, and a more efficient mode of scientific practice begins. (p. 178)
>
> Normal science can proceed without rules only so long as the relevant scientific community accepts without question the particular problem-solutions already achieved. Rules should therefore become important and the characteristic unconcern about them should vanish whenever paradigms or models are felt to be insecure.... The pre-paradigm period, in particular, is regularly marked by frequent and deep debates over legitimate methods, problems, and standards

> of solution, though these serve rather to define schools than to produce agreement. (p. 47)
>
> When paradigms enter, as they must, into a debate about paradigm choice, their role is necessarily circular. Each group uses its own paradigm to argue in that paradigm's defense .... To the extent ... that two scientific schools disgree about what is a problem and what is a solution, they will invariably talk through each other when debating the relative merits of their respective paradigms. In the partially circular arguments that regularly result, each paradigm will be shown to satisfy more or less the criteria that it dictates for itself and to fall short of a few of those dictated by its opponent. (p. 94)
>
> [Wh]atever its force, the status of the circular argument is only that of persuasion. It cannot be made logically or even probabilistically compelling for those who refuse to step into the circle. The premises and values shared by the two parties to a debate over paradigms are not sufficiently extensive for that. ... [In] paradigm choice there is no standard higher than the assent of the relevant community. (p. 109)

8 The frequentist paradigm

Bayesian theory began as the theory of inverse probability, proposed by Thomas Bayes (posthumously) in 1763. Remarkably, Bayes's theorem remains the centrepiece of the current Bayesian paradigm – its extreme simplicity, as the simple inversion of conditional probability, is unshakeable. What became shakeable was Bayes's postulate of the uniform prior for a parameter as a general principle, in the absence of information to the contrary. By itself this was not a sufficient criterion for opposition to the theory – there was no other.

In the pre-Fisherian period, many statisticians were agnostic, in the sense of using "what works". Karl Pearson was one: he published papers using Bayes's theorem to obtain (normal) posterior distributions with a flat prior and normal model, and other papers using normal repeated-sampling distributions, without any sense of conflict.

Fisher's definition of the likelihood (though not by that name) in his first 1912 paper gave a firm basis for a new paradigm, though it was not till his 1922 paper that an explicit theory was put forward, with some of its properties, extended in his 1925 paper. The 1928-33 work on hypothesis testing by Neyman and Pearson extended the Fisher paradigm in a different direction, inconsistent with Fisher's own views, and the current version of the paradigm is sometimes called the Fisher-Neyman-Pearson, or generally the frequentist paradigm.

This was a truly revolutionary paradigm, in the Kuhnian sense. It provided a theoretical foundation for statistical analysis that did not depend on a prior specification for the model parameter, yet obtained exact or asymptotically exact results (in the sense of random variables with exact sampling distributions) for a very wide range of important probability and statistical models. The benefits of this paradigm have been enormous, and it is still taught as the standard paradigm in most statistics departments, though the Bayesian paradigm is taught increasingly as an alternative, or (much less commonly) as *the* paradigm.

9 Revival of the Bayesian paradigm

An important paper in psychology stimulated interest in Bayesian theory (Edwards W, Lindman H & Savage L J. Bayesian statistical inference for psychological research. Psychol. Rev. 70:193-242, 1963), and this developed slowly with the increasing power of computers.

From about 1990 onwards, Bayesian methods began a major revival. The key to this revival was the development of Markov chain Monte Carlo methods, initially through the Data Augmentation algorithm of Tanner and Wong (1987), which was a Bayesian extension of the immensely important maximum likelihood EM algorithm of Dempster, Laird and Rubin (1977 – this paper had a number of Bayesian extensions and suggested extensions). This development and its rapidly increasing success illustrates another aspect of Kuhn's argument (pp. 5-6), that the need for a new paradigm develops when the current paradigm seems unable to deal with the increasing complexity of problems, which its own success has brought to the forefront of analysis.

10 The survey sampling paradigm

In his 1934 sampling paper, Neyman overthrew the then-current theory of purposive sampling and established the theory of random stratified cluster sampling, the foundation of the survey sampling paradigm. This led to a very substantial development of the paradigm.

Its principles were different from those of the frequentist school: though repeated sampling gave the probabilistic argument, as in the frequentist school, its basis was the repeated sampling of the *indicator variables* which define whether a population member was included in the sample or not. As there was no population model for the variable of interest, there was no likelihood in the parameters of the population. The principal question of interest was how to estimate the population mean or total of a variable taking into account the design of the survey.

In his review of the foundations of survey sampling, Smith (1976) concluded (p. 193):

> The basic question to ask is why should finite population inference be different from inferences made in the rest of statistics? I have yet to find a satisfactory answer. My view is that survey statisticians should accept their responsibility for providing stochastic models for

> finite populations in the same way as statisticians in the experimental sciences. These models can then be treated within the framework of conventional theories of inference. The problems with the Neyman approach then disappear to be replaced by disputes between frequentists, Bayesians, empirical Bayesians, fiducialists and so on. But at least these disputes are common to all braches of statistics and sample surveys are no longer seen as an outlier.

To repeat Kuhn, "... [the paradigms,] for a time, provide model problems and solutions for a community of researchers." I argue that both the frequentist and the survey sampling paradigms have stalled. They no longer provide the solutions to their problems – they have exhausted their scientific contributions – and need to be replaced by a Bayesian paradigm which does provide the solutions to its, and their, problems. This paradigm however has a major unsolved puzzle, or anomaly, which endangers its scientific contribution, and its practitioners have not been able to provide the solution to this critical problem.

I discuss these claims below, and give a resolution of the puzzle/anomaly which provides an extended Bayesian paradigm which includes the survey sampling field and is free of the anomaly.