Stat Science & Phil of Science: Where Do/Should They Meet in 2010 (& beyond)?

The question: Where should we meet?

Is one I have wrestled with for months in preparation for this conference…

I mean, literally, I agonized over the meeting place, as several of you know…over a period of months I sent at least 8 of you round to come here, to smell if the paint fumes had gone? How hot? How noisy?

All that data did finally severely test, and pass, the claim that room H103 in Connaught House was fine…But the conclusion was recently falsified.

It was a clear success story to the error statistical philosophy: That the sandblasting is on the other side of building H on Monday and Tuesday (I realized on Wednesday) doesn’t do a thing to rule out that they’d be shifting their machines to the 103 side by Thursday…and into next week!
Skeptic that I am, I returned Thursday to discover my fears were correct…!

In this little intro, I will give some very general points

- focusing mainly on today’s topic of statistical inference
- to try and distill core themes, concepts and problems, and
- lead directly into my own presentation.
Central job of philosophers of science

- to help resolve the conceptual, logical, and methodological discomforts of scientists

- especially in fields that deal with issues of scientific knowledge, evidence, inference, learning despite uncertainties and errors?

The risk of error enters because we want to move beyond the evidence or data

The conclusion is "evidence transcending": The premises can be true while the conclusion inferred may be false without a logical contradiction.

This frees us to talk about “induction” without presupposing certain forms of ampliative inference
...in particular, without presupposing there’s just one role for probability (however it’s interpreted)

While probability naturally arises in capturing induction there are 2 key traditions as to its role: to measure:

1. How much confidence, belief, support to assign a hypothesis

2. How reliably a hypothesis was tested, how well-tested, probed, or corroborated it is

This contrast is at the heart of a philosophical scrutiny of statistical accounts.
In scrutinizing a statistical account at a philosophical or foundational level...we ask: Does it provide an adequate characterization of scientific reasoning, evidence, inference, testing?

- it should be **ascertainable** (must be able to apply it)
- it should be **relevant** to the tasks required of the inference tools
- it **should not be in conflict** with intuitions about inference (science, evidence)—it must have a **principled** way to avoid these conflicts (not merely a reconstruction game)

Right away we are confronted with issues that *appeal to contrasting “philosophical theories”*
Moreover, these questions require asking: *what can various methods be used for?*

- This is distinct from what a method’s founders may have had in mind, or textbook accounts
- Demands standing “one level removed” from common interpretations/applications of methods
We do not want to rehash the “statistics wars” of the 60’s, 70’s, 80’s, 90’s, — up to the present:

• even though the significance test controversy is still hotly debated among practitioners (in psychology, epidemiology, ecology, economics)

• even though it seems each generation fights the ‘statistics wars’ anew, with journalistic reforms, and task forces set up to stem automatic, recipe-like uses of statistics that have long been deplored.

If we are to make progress in these decades old controversies, as well as tackle current ones, we need to get beyond the shallow to a much deeper scrutiny

• revisit some very old examples
• given the familiarity of these examples to this group, my emphasis will be on how they fit into the bigger story I am trying to tell
Relevance for Statistical Science Practice?

I would never be so bold as to suggest that unclarity about philosophical foundations in any way hampers progress in statistical practice…

Only in certain moments is there a need for a philosophical or self-reflective standpoint in practice

Increasingly, over the past few years, say, some issues (mostly within Bayesian statistical circles) seem to cry out for philosophical ministrations

It’s almost as if some of the generals from the older statistics (Bayesian-Frequentist) battles are wondering—with good reason—just who (if anyone) won the war?
Some statisticians suggest that throwing different and competing methods at a problem is all to the good —It increases the chance at least one will be right

This may be so, but one needs to understand how to interpret and relate competing answers…...which goes back to philosophical underpinnings...

Others think that these conflicts are bad for the profession and seek some kind of [Bayesian-Frequentist] “unification” or “reconciliations”

“We [statisticians ] are not blameless….we have not made a concerted professional effort to provide the scientific world with unified testing methodology…and so are tacit accomplices in the unfortunate statistical situation (J. Berger, 2003)
Not waiting for philosophers….“we focus less on what is correct philosophically than on ‘what is correct methodologically logically…professional agreement on statistical philosophy is not on the immediate horizon, but this should not stop us from agreeing on methodology.”

But what is correct methodologically depends on what is correct philosophically

If we look deeply (philosophically, foundationally) at
  (a) what’s put forward as a unification that everyone should love

  (b) and reactions to them

we unearth the gems we need looking for to both understand and get beyond current impasses
The unificationists (e.g., “default” Bayesians) think Bayesians and frequentists should be content with their offerings, but neither seem to, …why?

Finding ways to equate posterior probabilities with frequentist error probabilities (however clever) masks underlying conflicts: we get agreements on numbers that fail both as degrees of belief and as relevant error probabilities….

(I.J. Good spoke of a “Bayes-Non-Bayes compromise” but as with today’s unifiers*, it is always the non-Bayesian who is required to do (nearly all) the compromising)

*generally some form of O-Bayesianism or “default” Bayes
But subjective Bayesians don’t seem to like the unifications much either: Dennis Lindley: They focus too much on technique at the expense of the “Bayesian standpoint” (i.e., updating degrees of belief) (1997) (commenting on Bernardo)

➢ But there is growing confusion as to what the Bayesian or frequentist standpoints are …

➢ Andrew Gelman, a practicing Bayesian statistician (co-wrote Bayesian Data Analysis), recently said:

“The main point where we disagree with many Bayesians is that we do not think that Bayesian methods are useful for giving the posterior probability that a hypothesis is true... Bayesian inference is good for deductive inference within a model, .. for evaluating a model, we prefer …what Cox and Hinkley call ‘pure significance testing’”)
Jay Kadane, venerable subjective Bayesian: “The growth in use and popularity of Bayesian methods has stunned many of us who were involved in exploring their implications decades ago. The result, …is that there are users of these methods who do not understand the philosophical basis of the methods they are using, and hence may misinterpret or badly use the results. ….. No doubt helping people to use Bayesian methods more appropriately is an important task of our time”

I agree, and frequentist statistical philosophers should be included (enough years of exile).
The philosophical doctor is in…

My diagnosis (to whittle it down):

The current situation is a result of never having been clear on contrasting views on
(a) the roles of probability in ampliative inference and

(b) the nature and goals of inductive/statistical inference in relation to scientific inquiry

➤ What is correct methodologically turns on philosophy
➤ Methodology without philosophy is shallow

To cure this malady we need to dig (or drill?) at deep rather than a shallow levels…

Drilling analogy seems appropriate, given the obsession (in the U.S.) with the oil spill