Statistical Methods, part 1
Module 1: Science in practice

Dan Hedlin
Stockholm University
November 2012

Why philosophy of science?

- Science is extremely important and extremely influential
- Everybody has heard about science but nobody knows what it is
- Non-scientists need often evaluate scientific results (or take some kind of stand)

“The Art of Doing Science”

- By Sven Ove Hansson, professor of philosophy at Royal College of Technology in Stockholm
- Bridges the gap between practical science and philosophy of science
- A good introductory book on philosophy of science is “What is this thing called science?” by A.F. Chalmers

Two types of knowledge

1. Episteme (greek), knowledge of facts. Includes knowledge on how things are. Communicated with statements (“earth’s orbit around the sun is nearly elliptical”)
2. Techne, action knowledge. Includes knowledge on how you do things and why and when they give the desired results. Communicated with actions (in practice usually statements describing the actions).

- “pure” science and “applied” science is fairly close to episteme and techne, respectively. Compare *pure mathematics* and *applied mathematics*. In many branches of science pure and applied seem to approach each other
- Is there “pure statistics”?
- Statistical theory (e.g. theory of inference), is it pure or applied?

Not all knowledge is scientific

1. Scientific knowledge (in a broad sense, including humanities)
2. Other types of knowledge. E.g. Carolina's phone number is 0705641654 (fictitious obviously) or how to make nice custard.
3. (Pseudoscience = false science)

- Many sciences have gone from a focus on special knowledge to a focus on universal knowledge. Perhaps seen most clearly in biology (Hansson, page 8).
- “Special knowledge”, e.g. on species or a particular habitat. “Universal knowledge”, e.g. on evolution, ecological systems in general or genetics (there is a subfield called mathematical biology)
- What about statistics? Special or universal?

- The two levels (special and universal) support and depend on each other. Probably true for all sciences.
- Special knowledge gives inspiration and provides and environment for fact checks to the universal side of knowledge.
- Universal -> special, through providing tools, theorems, knowledge of general facts
- Other ‘causes’ and effects?

Science in action knowledge

- Science is used in two ways to gain/enhance action knowledge:

1. To solve problems
2. To study how solutions work in practice

- 1 and 2 address different questions. The combination of 1 and 2 is powerful. Alone each of them is not worth much.
- Examples in statistics?

Spectrum from objective to subjective

- (from Donald Gillies, Philosophical theories of probability, page 175 and following)
- *Fully objective*. Completely independent from us humans. Examples: the death of a distant star. Pace of disintegration of a radioactive atom.
- *Artefactual*. Ex, a pair of scissors. Objective, exists in the real world. Another example: The Big Dipper

continued

The scissors and the star constellation are there, objectively, but they are seen as a pair of scissors or as the Big Dipper only by humans.

- *Intersubjective*. Subjective (not there in the real world), but consensus in a group (in which it is intersubjective)
- *Subjective*. An individual's view. May be shared by others but little contact with one another

- Consider: “What is the ML estimate of the proportion of zeros in an a class of events that either takes values 0 or 1 based on n iid observations of those events. The events follow the distribution Bin(p,n) and the MLE is the sample proportion of zeros”.
- Is this, “the MLE is the sample proportion of zeros” a statement that is fully objective, artefactual, intersubjective, subjective?
- Can you say that it is a true statement?

Probabilities

- Fully objective. The disintegration of atoms in radioactive material follows a Poisson process. Hence the process is fully objective.
- Artefactual. E.g. probability that the child of a mother with cleft lip and a father without will have a cleft lip. (Why not fully objective?)

- Intersubjective. “There is *very high confidence* that the net effect of human activities since 1750 has been one of warming.” (IPCC AR4, Summary for policymakers, page 5). Why not artefactual?
- Subjective. I believe that my research paper on editing will be accepted by JOS or similar journal with a probability larger than 0.9 even though the average is about 0.2.

- Clearly (or is it?), objectivity is an aim in science to strive for, even in cases where you cannot reach it.

1. We may have intersubjective notions that are not objective, e.g. prejudices.
2. Objective notions must be intersubjective. The demand for objectivity is the same for all people since it refers to a single reality that we

try to describe as accurately as possible (Hansson page 12).

Note that artefactual and intersubjective statements are open to criticism since they are not fully objective.

Can statistics ever be fully objective?

- There is a continuum from intersubjective to artefactual

Science must be intersubjective

- Science must be *at least* intersubjective.
- Why? Conclusions in science should resist critical assessment by others (Hansson p. 12).
- If not, consensus fades away. If a conclusion is deserted after having been criticised, then there is no consensus, hence no intersubjectivity.
- So is Bayesian inference not scientific then?

What is science?

- *Science is the systematic search for such knowledge that is independent from any single individual, but that anyone could rediscover or verify.* (Hansson, source http://www.vof.se/visa-english)
- Note the words **systematic**, **independent**, **anyone**. Without any of them the definition would be very different (and flawed).

- Knowledge that science has found is subject to change
- Does this mean that scientific knowledge = don’t know (because it is going to change anyway)
- Principle: one should always use the best available knowledge

- Science is a human activity
- Rational discussion is one crucial part of science
- Science is anti-authoritarian (the scientific community, however, may be authoritarian)

Observations are theory-dependent

- Observations depend in general on theory, they are ‘theory-laden’ or ‘theory-dependent’
- Sometimes there is a lot of theory between what you see and the interpretation of what you see. E.g. advanced equipment at a hospital.
- Sometimes a little, eg mechanical scales

Observations independent of theory

- A correlation coefficient, is that heavily theory-laden or just a little?
- Are observations independent of theory at all possible?

Four types of observation

1. Experiment. Treatments are manipulated and the response is observed for each treatment.
2. Controlled observation. Like an experiment but you cannot or do not wish to influence the events. Pre-determined rules for what to note and how. E.g. observations on what pupils in a classroom say.

3. Uncontrolled observation. Like above but not planned ahead of time but still well documented.
4. Rumours, hearsay, undocumented observations, documentation lost, etc

- Note of course that a sociologist or social anthropologist may record rumours rigorously if the rumours are the object of research

No ideal observations

- Many reasons why you cannot make the ideal observations, including
  - Ethical reasons
  - Expensive, awkward, impractical
  - It takes 25 years , you can’t wait that long
  - The characteristic is latent (e.g. stress)

Two ways to get around it

- Methodological adaption (adapt your methods). E.g. what happens in the body when you freeze to death? Instead you put well-informed volunteers in cold water and pick them up before they are in danger, and extrapolate.

- Or you ask people on stress-related things ("Does it happen more often than once a week that you can't sleep..." etc) rather than taking blood-samples to examine stress levels.
- Vignettes
- Many other methodological adaptations

Second way around it

- Observe a substitute ("object adaptation"). In statistics you may decide to observe a proxy or let a proxy fill out the questionnaire. Hansson's stone age example is particularly nice (p. 41)
- The blood sample on previous page may actually be a mere substitute for "stress"
- Not having ideal observations makes observations weaker and open to disbelief.

Hypotheses

- A hypothesis in research is what you as a researcher tries to prove. Often there are strong prior reasons to believe that the hypothesis is true. (Statisticians may mix it up with the null hypothesis which you usually do not believe in)
- Sometimes there are really two alternatives, more or less equally plausible

Possible to prove a hypothesis?

- Can you prove that your research hypothesis is true?
- Hardly. Lend strong support though. Best by making a good prediction (?)
- In statistics, how can you make and *check* a prediction, using only one single dataset?

Successful prediction

- Successful prediction is generally considered as strong support of a hypothesis
- The support is stronger, the more exact the confirmed prediction is
- If there are several confirmations, their joint force increases the more independent they are of each other

The value of simplicity

- Occam’s razor. No theoretical elements should be introduced unless empirical evidence shows that they are needed (Hansson p. 65)
- Why? This is a major reason:
- Theoretical assumptions that are not forced upon us by the facts can be chosen differently by different people
- So in order to achieve intersubjectivity we should refrain from making assumptions that have no empirical justification

- "... we are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances". Isaac Newton
- Everything should be made as simple as possible, but no simpler (Einstein)
- The crucial issue is always if and in that case how observations can be made that adjudicate between the rival hypotheses (i.e. not which one of them is the simplest) Hansson p. 66
- Note: simplicity is a research strategy, not “science” or “truth”
- In statistics there is a loose concept called overfitting. We will shortly come back to this.

Models

- Three types of models:

1. Iconic models. E.g. Maps or model aeroplane
2. Analogue models. E.g. watch
3. Symbolic models, including mathematical models

- What kind of model is a scatter plot? The graph of a time series? A jittered plot?

- One interesting feature of many models used in practice is that they are false, and known to be false
- “All models are wrong but some of them are useful” George Box
- What is meant by “overfitting”?

Benefits of models

- Explicit mathematical models in statistics are useful because
- Often you have implicit models instead, if you do not acknowledge models explicitly
- Explicit models are testable against data (though not always)
- Allows for simulation, prediction and sensitivity studies – all helps strengthen model

- The flexibility of models makes validation more persuasive. E.g. Exploring data with different, or a series of more refined models allows for deep insight into the real-world issue that the models represent

Practical gains

- Also more practical benefits including
- Facilitates communication
- Thus comparisons of rivalling hypotheses
- Makes it easier to identify ingredients of models as more important than others. E.g. which parameters that are important
- One major purpose of models is to simplify; explicit models make it visible what has been simplified

Model misspecification

- How do we know that models are useful (recall that models are ‘wrong’)
- ‘Misspecified’ means that the model is “too wrong” to be useful

Two hard problems with models

- Relate them to the real world
- Communicate them to other people

Real-world applications

- On p. 72 Hansson reasons that in technology there is little scope for idealisation (the bridge will fall down if you don't build it for the real world)
- However, in technology you don't worry about mathematical exactness and optimality
- Similar to statistics?

- Actually there may be two ”levels” in statistics:

1. A lot of idealisation, e.g. ”under this model we find that”
2. As close to reality as possible, as in medical research