Teaching Statistics

Statistical literacy explained?

Posted on February 1, 2013 by Dr Paul Hewson

I've been trying to set out a working definition of statistical literacy (a phrase I don't like – I much prefer "Quantitative Reasoning" as it seems rather harsh to accuse someone of being illiterate). I do like Milo Schield (2011) "Statistical Literacy" (Fifth Edition) but want to move beyond that into formal inferential statistical methods. However, as I want to be very general, I don't want to use the GAISE definitions (brilliant as they are) as they are too focussed on formal education. Similarly, I find Joel Best (2001) "Damned Lies and Statistics" University of California Press a vital read, but a little too focused on sociology (however vital) for what I need. Daniel Kahnemann (2011) "Thinking Fast and Slow" Penguin also deals rather excellently with many of the problems we have understanding the world. The following represents an attempt to synthesize ideas in one place from all these sources. Hopefully it doesn't destroy the excellence that exists in those publications, but does capture something of what we are trying to achieve in our "statistical literacy" courses.

Read all about it. If we can critique news stories and journal articles in an informed way we are statistically literate. Believing nothing and accepting everything are the two unthinking extremes – all points in-between require skill, judgement and knowledge. Who created this statistic, why did they create that statistic, how did they created that statistic, and what difference could the answers to the previous three questions make to my interpretation?

Data beat anecdotes. We may need to learn from a bigger picture than relying entirely on our own experience.

Troublesome heuristics. Did you know if you had a room with 23 strangers you have better than 50% chance that two people share the same birthday? Our information processing is prone to many biases that lead us to misunderstand the world around us.

Not everything that counts can be counted. We should never confuse simplistic measurements with the phenomenon we are trying to measure. Ticks on boxes on a survey are not the phenomena, they represent one perception of that phenomena.

Interpreting results in context. Any study is subject to a variety of biases. We need to identify common sources of bias in studies and consider their impact. What kind of subjects were selected. Did they all respond. What effect does the data collection mechanism have?

How to make things count. We need to be able to formulate a question that in such a way that it can be addressed with data. Having done that we either need to obtain or collect appropriate data ourselves to answer that question.

Exploratory data analysis. We need to be able to select and interpret appropriate graphs and tables as well as summary statistics. The difference between column and row percentages, and whether we are talking about percentage difference or percentage point differences are subtle and important skills.

Variability is natural, predictable and quantifiable. We do not anticipate data with zero variability. Statistical science is the science of dealing with variability – either by modelling it or designing studies to take it into account in a way that leads us to useful conclusions.

Sampling methods. Random sampling allows the results of studies to be extended to the population from which the sample was taken. Being able to judge a set of results on a sample and determine the target population to which these results can be generalised is a key skill for users of statistics.

Experimental design. Random allocation in comparative experiments allows cause and effect conclusions to be drawn. One of the most powerful ideas is statistics is one of the simplest.

Association is not causation. Knowing how to interpret results when we see an apparent association is a subtle skill. In randomised experiments we might assume evidence of causation. In observational studies we might consider possible confounding variables that could explain the association.

Appropriate use of statistical inference. Even if you haven't yet started to use Bayesian inference you need to appreciate that not all statistical investigations have to end with a p-value. If you are still using frequentist inference remember that practical significance does not equate to statistical significance, that failure to detect a significant result may reflect low sample size and that finding a significant result may reflect large sample size.

Communicating results in a clear and helpful way. There are pitfalls even in the production of simple graphs and tables. Being able to communicate all aspects of a statistical investigation in a positive, constructively critical way allowing others to check assumptions and processes is vital. There is also growing interest in synthesising results – meaning as a producer of statistical information you have to provide appropriate information to allow this to happen.

If the only tool you have is a hammer, every problem looks like a nail. One key statistical skill is knowing when further knowledge is needed. Statistical science is a vast and growing field. It may be that a specialist, a course, a textbook or other learning resource might provide better methods to solve a particular problem.

About Dr Paul Hewson
Started out in life as a Biologist. Did MSc Applied Statistics at Sheffield Hallam, then PhD at Exeter University (both part time). I have been lecturing at Plymouth since 2004.
View all posts by Dr Paul Hewson →

This entry was posted in statistical literacy and tagged GAISE, statistical literacy. Bookmark the permalink.