

Criteria for a persuasive statistical argument: MAGIC

Adapted from Abelson, Robert P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum, pp. 12-14.

Magnitude: How big is the effect? It's a small thing to say C++ is faster than Python. How much faster? Does the ratio vary with the application? (See <http://shootout.alioth.debian.org/> for one way to compare.) How much will BC's carbon dioxide emissions be reduced by the proposed carbon tax? How does that compare with the level of reduction the province wants to achieve, and how much in turn does that compare with the amount of reduction that may be required globally?

Articulation: How readily can the details be summarized into memorable principles? "C++ is eighteen times as fast as Python" is simple enough. But that number is an average of 18 benchmarks (the recurrence of 18 is coincidental, by the way). If we only consider 3 benchmarks for bioinformatics applications (fasta, k-nucleotide, regex-dna), it slips to "C++ is seven times faster than Python". And if we consider only token-passing amongst threads (thread-ring benchmark), Python can do the task using its standard libraries, while C++ has no standard libraries for threading. Should we conclude "Python is infinitely faster than C++ for threaded applications"? (You could do it using the C subset of C++ and the pthread library, but is that really C++?)

Generality: How widely does this conclusion apply? Does it cover lots of cases or only a few? This is specific to the audience. An audience of general computer scientists might consider the full 20 benchmarks the most general. An audience of molecular biologists might only be interested in the 3 benchmarks related to DNA sequences.

Interestingness: How important is the issue addressed? How surprising is the conclusion? How much does the conclusion require a change of behaviour? Or how much does the conclusion remove a worry that might stall the readers from doing something they are doing now or want to do in the future?

Credibility: Given the methods used to gather the and analyze the data, how much should we trust the results? And how much do the results contradict other well-accepted understandings?

Using these criteria when reading

When reading someone else's statistical argument, evaluate their argument using these criteria.

Using these criteria when writing

When writing, consider how you could meet each criterion most strongly, *while remaining true to your data*. If your magnitude is small, don't inflate it. Instead, consider how you might strengthen your case according to other criteria. Perhaps the small effect will surprise your readers because they don't expect it to be there at all, or because they expect it to be larger. Or maybe that small effect will have a big outcome in user preferences. The force of your argument depends upon how well you meet these criteria as a group, not individually.