What do we need to communicate to students, so that they understand confidence intervals?

H_WATT JUL 11, 2022 08:46AM

I've almost entirely given up on trying to explain the intricacy of different stats and devote almost all my time with students trying to indoctrinate them into scientific thinking generally - primarily focussing on humility, parsimony, scientific disinterest and celebrating failure. The statistical problems all seem to fall out of more basic philosophical misunderstandings.

There are two problems here that I think are being addressed within this discussion. (1) How do we teach students the proper definition of a confidence interval (through repeated sampling etc.) so they understand the theory. (2) How do we teach students to seek meaning from their data through confidence intervals and arraive at appropriate conclusions? For the latter, I challenge anyone to derive real meaning from a CI without resorting to an incorrect definition of it, if not out loud, at least in your head.

I don't think I get the point? I'd actually believe it's the other way around - you can easily teach them to make sense of it in practice (correctly) without fully or completely correctly understand their definition — H_WATT

I think it really pays off to spend enough time toying with the concepts of sample/population, even though it seems so basic at first.

The compatibility approach is like a more "likelihood" based way of explaining things. Likelihood is of course a famously obscure topic, although I have to say I quite like that way of phrasing things…

I like compatibility for p-values but I find it quite difficult to grasp for CI — H_WATT

I agree, and I've used the concept of meta-analysis as a proxy to understand this - while pointing out it's not the same thing and a sampling distribution is theoretical

While it is good to understand what we are teaching, how much of what we understand is it worth trying to teach students?

The other thing that helps is to explain that we aren't interested in the sample itself, just for what it can tell us about the population.

I found students got a lot less confused when I spelled out that sampling distributions are almost always a hypothetical construct - you don't actually go out and sample lots of times, this is just what you'd expect if you did.

I think people confuse repeated samples and repeated studies

Although they are practically the same if we think about identical repeated studies. — H_WATT

yes, but usually they aren't — H_WATT

I think there is a bit of an issue learning about CI in relation to other samples (e.g. that 95% of the results fall within the CI) - the relevance of the CI is mostly in relation to the population mean - coming from the central limit theorem

I was taught that if sampled 100 times you'd expect 95% of results to fall within the 95% CI

To teach what confidence intervals mean is much the same as teaching what a p value means -both are a great deal harder than I realised for many years.

Someone just shared an awesome modelling tool with me outside the chat. I'm definitely going to start using this to discuss 95CI and sampling distributions: https://www.esci-dances.thenewstatistics.com/

Get them to read Bad Science by Ben Goldacre

I'm in the existentially difficult position that I've been teaching frequentist stats for 11 years, but I no longer really believe in them. The philosophical problems with them just run too deep. I'm not convinced that Bayesian stats are the easy answer to this problem, but I'm keen to learn. Starting over like this is really hard though.

There is, what I think, a nice definition of a confidence intervals in a book called, "The New Statistics with R - An introduction for biologists", which is as follows: "A common mistake is to state that we are 95% confident that the 'true' parameter value is in the interval. But, either the parameter is in the interval (100%) or it is not (0%) – we just don't know. So what does the 95% refer to? Confidence intervals come from the school of frequentist statistics, which is named for the idea of a hypothetical repeated series of samples. Confidence intervals are designed so that, for an imaginary long run of repeated samples, the interval will capture the 'true' value in 95% of cases. So, we can say we are 95% confident that our interval includes the 'true' value (in this long-term sense).

Population values and population means are different things.

The difference between repeated samples and means of repeated studies

Lately I had to argue for NOT including 95CI in a paper specifically because I didn't think the results were generalisable and didn't want to imply otherwise.

Practically, even if it's wrong, we all interpret them in our head as 'there's a 95% chance the true population parameter is inside this interval' whether we like it or not.

Difference between standard deviation and confidence intervals

The way I was taught it was that 95% CI describes the area where you'd expect 95% of hypothetical study results to fall, but lately I've been doubting that this is actually the case. So much to do with sampling distributions is taken on faith and rarely tested.

Relevance to effect sizes

They need to know that it is all about population parameter but we only have one sample in our hand

we aren't just interested in study participants, but more than

at undergraduate level, they only need to learn that they are measures of precision and understand that larger intervals mean less precise results

95% chance that the mean for a given parameter for a given sample will fall within the 95% CI of that parameter for another sample from the same population.

Frequentists' view of statistics needs to be established before explaining any statistical tools

CIs could be called compatibility intervals

I don't find that students understand confidence intervals until they understand the concept of sampling/inference

To understand research articles

We want students to be able to communicate about confidence limits in their analysis of data sets.

they need to know what corresponds to "no difference" for the estimate they are using!

Bayesian credible interval is more meaningful

understand estimates and standard errors

the distinction between CI and 'ranges of the data'

likelihoods and how they are defined

Clinical implications-what does it mean in clinical practice

have people heard about calling them intervals of uncertainty instead of confidence?

What it's NOT (i.e. 95% of your readings are not going to lie within the CI)

Yes !! — ANONYMOUS

Sample size

What happens to confidence intervals as the population size goes up

they should learn central limit theorem

That repeated sampling is very importants

That 95% CIs are measures of precision and not dispersion.

great — H_WATT

Measure of precision

******

******