Twitter Thread by Sean McClure

Sean McClure
@sean_a_mcclure

@nntaleb 1/10

Science is assumed to be “evidence-based” but that term alone doesn’t mean much. What constitutes good evidence? How is evidence being used? Is it supporting or refuting a hypothesis? Was the hypothesis and experimental design predetermined or found ex post facto?

@engexplain @nntaleb 2/10
The reality is you can find “evidence” for almost any narrative. Limit the sample size, cherry-pick studies, etc. Systematic reviews, meta analyses, and randomized controlled trials are all susceptible to selective interpretation/narrative fallacy.

@engexplain @nntaleb 3/10
At the heart of the problem is the over-reliance on simplistic statistical techniques that do little more than quantify 2 things moving together.

@engexplain @nntaleb 4/10
Take Pearson’s correlation, based on covariance. Variation can increase simultaneously across 2 variables for countless reasons, most of which are spurious. Yet this simple notion of “causality” undergirds much of scientific literature.

@engexplain @nntaleb 5/10
Information-theoretic (entropy based) approaches on the other hand can assess *general* measures of dependence. Rather than some specialized (linear) view based on concurrent variation, entropy encompasses the amount of information contained in and between variables.

@engexplain @nntaleb 6/10
If you were genuinely interested in giving the term “evidence” an authentic and reliable meaning then the methods used to underpin an assertion would be rigorous.

@engexplain @nntaleb 7/10
We wouldn’t look to conveniently simplistic methods to denote something as evidential, rather we would look for a measure capable of assessing the expected amount of information held in a random variable; there is nothing more fundamental than information.

@engexplain @nntaleb 8/10
Consider Mutual Information (MI), which quantifies the amount of information obtained about one random variable through observing another random variable. This observing of the relationship between variables is what measurement and evidence is all about.

@engexplain @nntaleb 9/10
MI determines how different joint entropy is from marginal entropies. If there is a genuine dependence between variables we would expect information gathered from all variables at once (joint) to be less than the sum of information from independent variables (marginals).

@engexplain @nntaleb 10/10
If “evidence-based” science was genuinely invested in authentic measurement it would leverage *general* measures of dependence; that demands an approach rooted in information-theory. Without entropy you’re just picking data, choosing a narrative, and calling it “evidence.”