

Commentary

Spurious associations?

I am sure you have heard, or read, or maybe even used the phrase “spurious association”, and you probably know what it means: the presence (or magnitude) of an association between E and D that does not correspond to the presence (or magnitude) of the effect of E on D . Let’s recall, however, the meaning of the two words that make up the phrase:

Here are explanations that may be found on the internet.

Spurious: false; fake; not being what it purports to be; apparently but not actually valid; the opposite of genuine

Association: any relationship between two measured quantities that renders them statistically dependent

Assembling the two explanations together, we get something like the following:

Spurious association: a false (or fake or not actually valid) relationship (or statistical dependency) between two variables.

Ah?

How does statistical dependency cross the line between true dependency and false dependency?

Between genuine dependency and fake dependency? Between valid dependency and invalid dependency? Fabricated data, perhaps. But that’s not the usual meaning of “spurious association”.

“It is called ‘spurious’ because it is not part of what we seek to estimate – the causal effect...”, someone tried to argue.

Well, the argument appears to be valid, but it is not actually valid (that is, it is spurious): “spurious” – a synonym of “not actually valid” – is not an adjective that may be attached to “statistical dependency”. Statistical dependency between E and D does not purport to reflect the effect of E on D . It is an innocent mathematical idea which makes no causal claims whatsoever. People occasionally make spurious claims (causal and others); associations do not.

All associations arise from some causal structure (and sometimes from conditioning). None is spurious. Since we like associations that arise from causal paths alone – and don’t like those that arise from other paths (confounding paths and induced paths) – let’s call the latter “disliked associations” rather than “spurious associations”. Fair enough?