Some Surprising Facts about (the problem of) Surprising Facts

Deborah Mayo

A common intuition about evidence is that if data x have been used to construct a hypothesis H(x), then x should not be used again in support of H(x). It is no surprise that x fits H(x), if H(x) was deliberately constructed to accord with x. The question as to when and why we should avoid such "double-counting" continues to be the subject of debate in philosophy and statistics. It arises as a prohibition against data mining, hunting for significance, tuning on the signal, and ad hoc hypotheses, and in favor of use-novel and predesignated hypotheses. I have argued that it is the severity or probativeness of the test—or lack of it—that should determine if a double-use of data is admissible. I examine a number of surprising ambiguities and unexpected facts that continue to bedevil this debate.