THE PARADOX BETWEEN TRUTH AND LIES

Boris Aberšek
University of Maribor, Slovenia
E-mail: boris.abersek@um.si

Research work and the interpretation of results require a great deal of caution. One needs to look at more than just numbers. Simply averaging and calculating percentages is not only useless but can distort conclusions. Simpson's paradox is merely the most obvious and most instructional pitfall awaiting those who artfully juggle with numbers and percentages.

There is a saying in the field of artificial intelligence: "*Hard things are easy; easy things are hard*", called Moravec's paradox, after Hans Moravec. Moravec wrote in 1988:

> *"It is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility"* (Moravec, 1988, 15).

Similarly, Minsky emphasized that the most difficult human skills to reverse engineer are those that are *unconscious*, as he wrote:

> *"In general, we're least aware of what our minds do best. We're more aware of simple processes that don't work well than of complex ones that work flawlessly"* (Minsky, 1986, 29).

In today's age of fast and multifarious publications, we often come across such paradoxes, for example, that research and statistical analysis may indeed have been carried out correctly (properly), but the interpretation of the results is inadequate, or even incorrect, or misleading. To relate this hypothesis to Moravec's paradox, one could reformulate the latter by saying that *'to do research is easy; to discuss results is difficult'.* But why is this so? Let us consider this issue from the perspective of another paradox, Simpson's paradox.

Simpson (1951) and before him, Karl Pearson (1895), have observed that certain correlations disappear if we average and observe the characteristics on an entire population, instead of analyzing individual subgroups separately. However, Simpson's idea was only deemed as a paradox as late as 1972, when Canadian statistician Colin Ross Blyth found that *sometimes correlations do not disappear but become reverse.* In such cases, by averaging an entire population, we come to a conclusion that is not true. Blyth referred to this phenomenon as *Simpson's paradox,* although there is in fact nothing paradoxical about it, it is just so unintuitive that it appears unfathomable to the human brain. Therefore, it would be more aptly called *Simpson's reversal.* By the same token, Simpson's paradox is not a malfunction of statistics, but a simple fact: *in order to perform a proper statistical treatment of some phenomenon, we need to understand the phenomenon we are analyzing. Simply averaging and calculating percentages is not only useless, but can distort conclusions, especially in the field of education.*

https://doi.org/10.33225/pec/21.79.834
ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 79, No. 6, 2021

Simpson's paradox is important for three critical reasons. First, people often expect statistical relationships to be immutable. They often are not. The relationship between two variables might increase, decrease, or even change direction depending on the set of variables being controlled. Second, Simpson's paradox is not simply an obscure phenomenon of interest only to a small group of statisticians. Simpson's paradox is actually one of a large class of association paradoxes. Third, Simpson's paradox reminds researchers that causal inferences, particularly in nonexperimental studies, can be hazardous. Uncontrolled and even unobserved variables that would eliminate or reverse the association observed between two variables might exist.

Definition

Consider three random variables X, Y, and Z. Define a $2 \times 2 \times K$ cross-classification table by assuming that X and Y can be coded either 0 or 1, and Z can be assigned values from 1 to K. The marginal association between X and Y is assessed by collapsing across or aggregating over the levels of Z. The partial association between X and Y controlling for Z is the association between X and Y at each level of Z or after adjusting for the levels of Z. Simpson's paradox is said to have occurred when the pattern of marginal association and the pattern of partial association differ.

Various indices exist for assessing the association between two variables. For categorical variables, the odds ratio and the relative risk ratio are the two most common measures of association. Simpson's paradox is the name applied to differences in the association between two categorical variables, regardless of how that association is measured.

Illustration

As a starting point, let us consider an old example, which dates back to the year 1973. At the University of California, Berkeley, only 35 percent of the female applicants were admitted to graduate school, and 44 percent of the male applicants. Such a difference cannot be the result of chance. Assuming that men and women are equally capable, the only possible conclusion, staring us right in the face, is that the university acted in a discriminatory manner. That's why it got sued. But let us now argue as to why simply averaging and calculating percentages is not only useless but also distorts conclusions [1],[2]!

We have to consider not only the numbers, but other data as well. At Berkeley, the admission of candidates is the responsibility of individual departments, so the university took a closer look at who was to blame for such gender discrimination. It turned out that there were no wrongdoers in any of the departments. Some departments accepted more female candidates and others more male candidates, but there were no major deviations. *What happened?* Through careful analysis, researchers found that some departments have very popular study programs for which there is a lot of interest, so the percentage of admissions is low. Such programs include, for example, the social sciences. Other study programs, such as science and engineering, are less popular and desirable among the candidates, and the applications are so few that the acceptance rates are very high. At Berkeley it happened that women mostly applied to highly competitive social sciences' study programs, and because these programs were so competitive, the acceptance rates were low, whereas men disproportionately applied to less popular departments to study science and engineering programs, and these departments had high acceptance rates. Even though the departments accepted both genders in a balanced way – in fact, they even slightly favored

---

1 https://www.brookings.edu/blog/social-mobility-memos/2015/07/29/when-average-isnt-good-enough-simpsons-paradox-in-education-and-earnings/

2 https://www.refsmmat.com/posts/2016-05-08-simpsons-paradox-berkeley.html

accepting female students – there were more and more unaccepted female students among all registered female student applications, in all departments. Therefore, we can observe that *simply averaging and calculating percentages is not only useless but distorts the conclusions.* It follows from the above that *for proper statistical treatment we need to understand the phenomenon we are analyzing.*

Analyses showed that the university did not discriminate students in the admission process. This does not mean, however, that there was no discrimination. Research has clearly shown that the observed discrimination occurs beforehand, at all levels of education and the society. This paradox, however, is not a malfunction of statistics, but a simple fact: *in order to perform a proper statistical treatment of some phenomenon, we need to fully comprehend the phenomenon we are analyzing.*

Avoiding Simpson’s Paradox

Although it might be easy to explain why Simpson’s paradox occurs when presented with an example, determining when Simpson’s paradox will occur is more challenging. In experimental research, in which individuals are randomly assigned to treatment conditions, Simpson’s paradox should not occur, no matter what additional variables are included in the analysis. This assumes, of course, that the randomization is effective and that assignment to treatment condition is independent of possible covariates. If so, regardless of whether these covariates are related to the outcome, Simpson’s paradox cannot occur. In nonexperimental, or nonrandomized, research, such as a cross-sectional study in which a sample is selected and then the members of the sample are simultaneously classified with respect to all of the study variables, Simpson’s paradox can be avoided if certain conditions are satisfied. The problem with nonexperimental research is that these conditions will rarely be known to be satisfied a priori.

Conclusion

In summing up both paradoxes, Moravec's and Simpson’s, we can conclude that it is not enough simply to “conduct research” and produce statistics using all its high-flown statistical methods *without really knowing and understanding the problem we are dealing with.* We can quickly see that what should by definition be hard (planning and conducting research) becomes easy, and what should be easy (interpreting the results obtained) becomes hard, and that without understanding the problem itself, this can quickly lead to wrong conclusions.

References

Minsky, M. (1986). *The society of mind*. Simon and Schuster.
Moravec, H. (*1988). Mind children.* Harvard University Press
Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *13*(2), 238-241. https://www.jstor.org/stable/2984065
Pearson, K. (1895). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, *58*, 240–242. https://www.jstor.org/stable/115794

Received: *September 10, 2021* Accepted: *November 30, 2021*

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 79, No. 6, 2021

Cite as: Aberšek, B. (2021). The paradox between truth and lies. *Problems of Education in the 21st Century*, *79*(6), 834-837. https://doi.org/10.33225/pec/21.79.834

Boris Aberšek	PhD, Professor, University of Maribor, Faculty of Natural Sciences and Mathematics, Koroška cesta 160, 2000 Maribor, Slovenia. E-mail: boris.abersek@um.si Website: https://scholar.google.com/citations?user=aRid0w4AAAAJ&hl=en ORCID: https://orcid.org/0000-0002-4198-4240