

When Numbers Deceive: The Suppositions Latent in Simpson Paradox.

Mathematics is said to be the language of certainty. Numbers are definite and trustworthy, and statistics seem to have definite answers. But mathematics can be unkind, at times, and can produce misleading results even when the statistics is right. A notable example of such is the Simpson Paradox, a statistical phenomenon where trends seen in disaggregated groups are no longer seen or even are reversed when the groups are aggregated.

It seems a hard task on the surface, so we should dissect it. When one alternative is superior in each and every respect, you would expect it to emerge the winner on the balance sheet. But that's not always the case. Thing is, averages do not only rely on the percentages, also on how the data is distributed among various groups. Sometimes this results in a scenario where the general outcome would appear contrary to what you would anticipate considering each category separately. It is something of a paradox but this is how averages operate.

To explain this, consider the case of two hospitals that carry out surgeries based on the success rate. Suppose Hospital A is more successful than Hospital B in minor and major surgeries. Logically, we may find that Hospital A also has to be better in general. But when the outcomes of the two kinds of surgeries are added together, Hospital B might seem like it is more successful as a whole.

This occurs due to weighted averages. Hospital A may have a lower overall success rate than Hospital B since Hospital A does a lot more difficult work and Hospital B does a lot of easier work. This is because the surgeries are difficult to perform, something that is not easily visible to affect the results.

Putting all of the data in a group together can actually alter the influence that each group has on the total average. This is because larger groups make a greater impact on the ultimate number and that can even reverse the trend in the opposite way.

The Simpson Paradox is not just some strange mathematical phenomenon but one that has practical significance. One famous example of this was found in university admissions

numbers. At a rough glance it seemed that the figures were inclined against women when you considered the total number of those getting in. However, when you separated it out by department, you found that in fact, many departments were admitting women at a higher percentage than men. So what was going on? As it happened women were making applications to the department that were indeed very competitive and did not take as many applicants as they could and it seemed like they were discriminated against once you simply looked at the overall numbers.

The same may happen with medical research. The same treatment may seem ineffective with combined data, yet when patients are divided into groups that are relevant during the treatment (i.e., by age or disease severity), the given treatment may seem to be working in each of these groups significantly better. No matter what the aggregate statistics say, it may be misleading to make conclusions without taking into account these hidden variables.

You see, the Simpson paradox is a big thing since it demonstrates to us that our intuition may fail when analyzing data. We believe that when percentages and averages are used we can see the exact picture, but math tells us that it is not always the case. Weird things can occur when we add dissimilar sets of data. Then, it is not only a matter of doing the math but of the way the data is organized. This is quite crucial, since when we fail to do so, we may find ourselves making mistakes or failing to capture crucial information. When we make time to actually examine the data and its assembly we can get a clearer understanding of what is actually happening.

The Paradox of Simpson is an excellent demonstration of the deceitfulness of stats. Numbers can be misleading, you see. To get a better idea of what is happening, you must consider the factors that are influencing the data. It is also possible that something that appears as an obvious trend is not the same when you go to the bottom of the list and look at the way the information is organized. It is as though it is impossible to view the surface of the matter, you need to take into account all of the various factors that could be affecting the outcomes. That is the trick to making a real sense of what the data is actually informing you.

The Simpson Paradox is a terrific example of how math can be counter-intuitive. It teaches us that because something appears to be true, it is not necessarily true. When we glimpse at data, we may believe we are seeing a pattern but math can unveil some hidden relationships that alter everything. Doing calculations is not the point, it is about knowing what the numbers are saying. Math assists us in determining what questions to ask in the first place

and this is what makes it so powerful. Through math, we are able to discover things that we have never known about, and have a more profound insight into the world around us.

-END-