Simpson's Paradox – Sex Descrimination, Smoking, and Altruism

Elliott Sober

Sex Discrimination at Berkeley

The University of Berkeley some years back found evidence that there was sex discrimination in admission to graduate school (Cartwright 1979). It was observed that

(1) A smaller percentage of women were admitted to graduate school than men.

But when the University looked more closely at the evidence, they also found that:

(2) In each department, the percentage of women admitted is the same as the percentage of men.

How can (1) and (2) both be true? The answer is that this is an instance of Simpson's Paradox. But what does that mean?

Let's look at an example. Suppose 1000 women and 1000 men apply to graduate school and that there are two departments, D1 and D2. In accordance with (1) suppose that 220 women and 380 men are admitted. But how can (2) be true? Suppose that the admissions rates for men and women in the two departments are:

Admission Rates		
	D1	D2
Male	20%	40%
Female	20%	40%

But suppose that most men apply to D2 and most women apply to D1:

Numbers Applying		
	D1	D2
Male	100	900
Female	900	100

Given the admission rates and the numbers applying, the result is:

Numbers Admitted		
	D1	D2
Male	20	360
Female	180	40

Notice from this last table that 380 men out of the 1000 applying are admitted, while 220 women out of the 1000 applying are. In this example, propositions (1) and (2) are both true.

Here is a fallacious argument: "Since each department admits the same percentage of women as men, the University as a whole must admit the same percentage of women as men. This seems clear -- what is true in each part must also be true in the whole." This line of reasoning will be correct if there is no correlation between an applicant's sex and the department applied to. But, in this example, these two characteristics are not independent. There is a correlation between being female and applying to department D1; equivalently, there is a correlation between being male and applying to department D2. This is because:

$$\Pr(\text{applying to D1} \mid \text{female}) = 0.9 > \Pr(\text{applying to D1} \mid \text{male}) = 0.1.$$

Dominance Arguments in Decision Theory

A related point: In decision theory, dominance arguments are fallacious when acts and states of the world are non-independent, as the following example from Hacking (p. 120) shows:

.

	Die young	Die old
Don't Smoke	−5	95
Smoke	0	100

The four entries in the above table indicate that if you die young, you'd rather smoke than not (since 0 > -5) and if you die old, you'd rather smoke than not (since 100 > 95). Does this mean that smoking has higher expected utility than not smoking? No. This would follow if smoking and longevity were independent, but they are not. Suppose the correlation is so strong that if you smoke you'll die young, and if you don't smoke, you'll die old. This means that the expected utility of smoking is 0 and the expected utility of not smoking is 95.

The Evolution of Altruism

Here are two definitional truths in evolutionary theory:

(3) Altruists are less fit than selfish individuals who live in the same group.
(4) When a population evolves under the control of natural selection, fitter traits increase in frequency and less fit traits decline.

It may seem to follow from these two truisms that altruism cannot evolve when the evolutionary process is controlled by natural selection. However, this does not follow; to see why we need to think about Simpson's Paradox.

Suppose there are two groups that each contain altruists and selfish individuals. Each group contains 100 individuals:

	Group 1	Group 2
Altriusts	90	10
Selfish Individuals	10	90

Suppose that the individuals reproduce asexually, with an individual's fitness (its number of offspring) depending on its own phenotype (altruistic or selfish) and also on the type of group it inhabits:

	Group 1	Group 2
Altruists	3	1
Selfish Individuals	4	2

Notice that in each group, altruists are less fit than selfish individuals. However the average fitnesses of the two traits are

$w_A = 0.9(3) + 0.1(1) = 2.8$ $\qquad$ $w_S = 0.1(4) + 0.9(2) = 2.2$

The average fitness of altruism is greater than the average fitness of selfishness. Simpson's Paradox strikes again!

Let's suppose that individuals in the first generation reproduce and then die. This is what the next generation will look like in the two groups (assuming that offspring have the same phenotypes as their parents):

	Group 1	Group 2
Altruists	270	10
Selfish Individuals	40	180

:

The first generation of the two-group metapopulation has 50% altruism. In the next generation there are 280 altruists and 220 selfish individuals in total, so the frequency of altruism has gone up. However, the frequency of altruism declines in each group. It drops from 90% to 270/310 = 87% in Group 1 and from 10% to 10/190 = 5% in Group 2.

References

Cartwright, Nancy (1979): "Causal Laws and Effective Strategies," *Nous* 13: 419-437.

Hacking, Ian (2001): *An Introduction to Probability and Inductive Logic*. Cambridge University Press.

Simpson, E.H. (1951): "The Interpretation of Interaction in Contingency Tables." *Journal of the Royal Statistical Society B* 13: 238-241.

Sober, E. (1988):. "What Is Evolutionary Altruism?" In B. Linsky and M. Matthen (eds.), *New Essays on Philosophy and Biology* (Canadian Journal of Philosophy Supp. Vol XIV), University of Calgary Press, pp. 75-99.