Simpson's Paradox: Covid Death Rates UK

Milo Schield, University of New Mexico

Confounding is arguably the #1 problem in dealing with observational data. Yet, it is "the elephant in the room": it is all but absent from our intro statistics courses.

Simpson's paradox is the most extreme form of confounding. But simply showing the problem does not help students understand how or why it occurs. Focusing on the disparity of the elderly helps students understand how or why it occurs.

But our minds look for solutions to problems. Is there any way to take into account (control for) this confounding disparity – without using a computer? Yes!

Introduce standardizing: a simple weighted average calculation. Consider this data on UK Covid Delta deaths between February and August 2021.

Crude association: *Those vaccinated are 2.4 times as likely to die as [are] those unvaccinated.* But, this crude association doesn't take into account (control for) age.

Vaccinated are more than twice as likely to die as are the un-vaccinated				Unvaccinated are more then 3 times as likely to die as are the vaccinated after controlling for age.						
			Crude	Number of Cases			----Weights ----- Stdrdized		
Death rates	<50	50+	All	<50	50+	All	<50	50+	All
Un-vac	0.03%	5.96%	0.17%	147,612	3,440	151,054	0.977	0.023	0.71%
Vaccinated	0.02%	1.68%	0.41%	89,807	27,307	117,115	0.767	0.233	0.21%
0.17% = 0.977*0.03%+0.023*5.96%				237,419	30,747	268,169	0.885	0.115	
0.41% = 0.767*0.02%+0.233*1.68%						0.71% = 0.885*0.03%+0.115*5.96%			
50+ are 10 times as prevalent among the vaccinated (23%) as among the unvaccinated (2.3%).									
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1009243/Technical_Briefing_20.pdf									

Standardizing gives both groups the same mixture of young and old as shown on the right side above. *After controlling for age, those unvaccinated are 3 times as likely to die as [are] those vaccinated.*

Using simple weighted averages, students can calculate the standardized death rates after taking into account age. They can see how a crude comparison (a mixed fruit comparison) can be converted into an adjusted comparison (an apples and apples comparison). Now, teachers can put these problems on a test and expect students to work out the standardized results.

For more on confounding and standardizing, see Schield (2021): "Teaching Confounding" at www.StatLit.org/pdf/2021-Schield-USCOTS.pdf

The U. of New Mexico is offering MATH1300: a new confounder-based statistical literacy class. See the forthcoming paper at www.StatLit.org/pdf/2021-Schield-ASA.pdf