CSE 250A Quiz 1

Tuesday October 9, 2012

Instructions. You should do this quiz in partnership with exactly one other student. Write both your names at the top of this page. Discuss the answer to the question with each other, and then write your joint answer below the question. It is ok if you overhear other students' discussions; you still need to decide if they are right or wrong. You have ten minutes.

Question. Let A, E, and M be binary random variables. Is it possible for the following inequalities all to be true simultaneously?

$$\begin{aligned} p(A=1|M=1) &> p(A=1|M=0) \\ p(A=1|M=1,E=1) &< p(A=1|M=0,E=1) \\ p(A=1|M=1,E=0) &< p(A=1|M=0,E=0) \end{aligned}$$

If yes, use your imagination to give sensible real-world meanings for the three random variables, and explain, briefly but clearly, how the inequalities can be all true. If no, give an argument using the laws of probability theory.

Answer: The inequalities can be all true. This situation is called Simpson's paradox. A famous context where the paradox was encountered is admissions to Berkeley in 1973; for details see Wikipedia. Here $A = 1$ means admitted, $M = 1$ means male, and $E = 1$ means applied to engineering. Overall, men were more likely than women to be admitted: $p(A = 1|M = 1) > p(A = 1|M = 0)$. However, among engineering applicants, women were more likely to be admitted: $p(A = 1|M = 1, E = 1) < p(A = 1|M = 0, E = 1)$. And the same was true among non-engineering applicants: $p(A = 1|M = 1, E = 0) < p(A = 1|M = 0, E = 0)$.

The three inequalities can be all true if the overall admission rate for engineering is higher, and men are more likely to apply to engineering: $p(A = 1|E = 1) > p(A = 1|E = 0)$ and $p(M = 1|E = 1) > p(M = 1|E = 0)$.

Grading guide. Correct answers will not necessarily be very clear. It is difficult to give a watertight answer without going into more detail than is possible in ten minutes.

The answer "no" is wrong. Students can get partial credit if they have an argument that is clear and correct except for a specific mistake, such as a wrongly quoted rule of reasoning with probabilities. If this is the case, try to find the error. If you cannot locate a specific mistake easily, then take off points for lack of clarity. A clear argument with a clear mistake is easier to fix than a generally confusing argument, so the former is preferable and hence worth more points.

Detailed answer. A fully detailed answer (not expected for the quiz) has to work out a complete scenario where all three inequalities are satisfied, because it would be inordinately difficult to prove abstractly that the three inequalities do not lead to a contradiction. A complete scenario has to give numbers (counts of instances) for each of the eight possible combinations of values of the three random variables. Here is an example of a full scenario:

E	M	A	number
0	0	0	50
0	0	1	20
0	1	0	10
0	1	1	2
1	0	0	5
1	0	1	10
1	1	0	40
1	1	1	40

With the numbers in the table above,

$$\begin{aligned}
p(A=1|M=1,E=1) &= 40/80 = 0.5 \\
p(A=1|M=0,E=1) &= 10/15 = 0.67 \\
p(A=1|M=1,E=0) &= 2/12 = 0.17 \\
p(A=1|M=0,E=0) &= 20/70 = 0.29 \\
p(A=1|M=1) &= (2+40)/(10+2+40+40) = 0.46 \\
p(A=1|M=0) &= (20+10)/(50+20+5+10) = 0.35
\end{aligned}$$

so the three inequalities of Simpson's paradox are satisfied.

The example above shows that drawing real-world conclusions based on observed probabilities can be fraught with complexity. Do the data show that women are discriminated against, or not? The obvious conclusion is different depending on whether or not the data are disaggregated according to major. One might say that it is always appropriate to disaggregate, but then sample sizes may be too small to yield statistical significance.

A reasonable real-world conclusion, if the imaginary data above were true, might be that there is no discrimination against women in admission decisions, but there is a discrepancy in the provision of services: more slots are provided in the subjects preferred by men than in those preferred by women. Whether this discrepancy is defensible, or perhaps even desirable, is a separate issue.

Simpson's paradox is an important concern in real applications of machine learning. For example, in what is called A/B testing, a website tries two different versions of an advertisement, with the goal of maximizing the probability that a viewer clicks on the advertisement. Let $A = 1$ mean "the viewer clicks on the advertisement," let $M = 1$ mean "the advertisement is version one," and let $E = 1$ mean "the time of day is the evening." Simpson's paradox arises if version one seems to perform better overall, but version two performs better in the evening, and also separately during the rest of the day.

In the advertising case, part of the issue is that maximizing the probability of clicking

$$\max_m p(A = 1|M = m)$$

is not the correct objective. Instead, the website should maximize the total number of clicks, since it is paid per click. From this point of view, version two is indubitably better.

The advertising case illustrates the importance of distinguishing between variables that an agent controls, and variables that are controlled by the environment. Here, the website controls M but does not control E or A. If the website runs the experiment so that M is independent of E, that is $p(M = 1|E = 0) = p(M = 1|E = 1)$, then it will find that $p(A = 1|M = 1) < p(A = 1|M = 0)$ regardless of the value of $p(E = 1)$. In the special case where $p(M = 1|E = 0) = p(M = 1|E = 1) = 0.5$, version two will get a higher total number of clicks than version one, again regardless of the value of $p(E = 1)$.