Discussion Section 3

Case Study: Monte Hall

AMS 131 3 Jul 18

①

Case study: Death Penalty

Z_1 is a confounding factor (CF) if

① plausible that Z_1, Y are associated

② ———— Z_1, X ————

U, V associated $\leftrightarrow$ as $U \uparrow$, $V \uparrow \downarrow$ on average, & vice versa

case study: Monte Hall problem

Y_i = {you initially choose door i}

M_j = {Monte Hall then opens door j}

C_k = {car actually behind door k}

$i, j, k = 1, 2, 3$

without loss of generality

you pick door 1 ($Y1$) & Monte ($M2$) opens door 2 to reveal a goat

PLAN AHEA

we want to compare $P(C1 \mid M2, Y1)$ and with $P(C3 \mid M2, Y1)$. (11)

This is like ELISA:

	~~monte~~ Monte	ELISA
unknown:	location of car	true HIV status
data:	monte showing you a goat behind door 2	what ELISA said

we want $P(\text{unknown} \mid \text{data})$ but problem setup gave us $P(\text{data} \mid \text{unknown})$

So let's use Bayes's Theorem to reverse order of conditioning in odds form

$$\underbrace{\frac{P(C1 \mid M2, Y1)}{P(C3 \mid M2, Y1)}}_{\text{posterior odds}} = \underbrace{\left[\frac{P(C1)}{P(C3)}\right]}_{\text{prior odds}} \cdot \underbrace{\left[\frac{P(M2, Y1 \mid C1)}{P(M2, Y1 \mid C3)}\right]}_{\text{Bayes factor}}$$

$P(C2 \mid M2, Y1) = 0$

Now by the rules $P(c1) = P(c3) = \frac{1}{3}$,

so the prior odds are $\frac{P(c1)}{P(c3)} = \frac{1/3}{1/3} = 1$

---

to evaluate probabilities like

$P(m2, y1 \mid c1)$, let's use the

and

general form of product rule

for and:

$$\frac{P(m2, y1 \mid c1)}{P(m2, y1 \mid c3)} = \frac{\cancelto{1}{P(y1 \mid c1)} \cdot P(m2 \mid y1, c1)}{\cancel{P(y1 \mid c3)} \cdot P(m2 \mid y1, c3)}$$

but y_i and c_j are independent so

$P(y1 \mid c1) = P(y1) = \frac{1}{3}$ and

$P(y1 \mid c3) = P(y1) = \frac{1}{3}$ so

$$\frac{P(c1 \mid m2, y1)}{P(c3 \mid m2, y1)} = \frac{P(m2 \mid y1, c1)}{P(m2 \mid y1, c3)} = \frac{1/2}{1} = \frac{1}{2}$$

so: after A2 (given Y1, C1),

the posterior odds in favor of

car behind door 3 are 2 : 1,

so $P(C3 \mid A2, Y1) = \frac{2}{3}$ &

you should switch.

case study: Cromwell's Rule

for any D such that $P(D) > 0$

& any A,

Dennis Lindley

(a) if $P(A) = 0$ then $P(A \mid D) = 0$

and

(b) if $P(A) = 1$ then $P(A \mid D) = 1$

D = data	P(A) prior information about A
A = unknown	P(A | D) posterior info about A

Disc. Sec. 3

case study: death penalty

AMS 131
8 Aug 17

(1)

outcome variable (effect) (Y): $\begin{cases} 1 & \text{if death penalty (DP)} \\ 0 & \text{not} \end{cases}$

predictor variable (cause) (X): $\begin{cases} 1 & \text{if defendant white (DW)} \\ 0 & \text{not (DB)} \end{cases}$ defendant black

basic design: observational study

threat to validity: bias from potential confounding factors (PCFs) $Z_1, \dots, Z_k$

one possible PCF (Z): $\begin{cases} 1 & \text{if victim white (VW)} \\ 0 & \text{not (VB)} \end{cases}$ ← victim black

as Z changes from VB to VW, quite possible that $P(DP) \uparrow$

as Z changes from VB to VW, quite possible that $P(DW) \uparrow$

so Z (ethnicity of victim) is a PCF; control for it by <u>holding it constant</u>.

study relationship between DP imposition (2)
and ethnicity of defendant separately for
VB and VW

(bottom table) (middle table)

naive analysis based only on top (aggregate) table:

if a murder is chosen at random

$$P(DP) = \frac{36}{326} \doteq 11.0\%$$

$$P(DP \mid DW) = \frac{19}{160} \doteq 11.9\%$$

$$P(DP \mid DB) = \frac{17}{166} \doteq 10.2\%$$

it appears that white defendants receive the death penalty more often than black defendants, which is a surprise

analysis of middle table (VW)

$$P(DP \mid VW) = \frac{30}{214} \doteq 14.0\%$$

$$P(DP \mid VW, DW) = \frac{19}{151} \doteq 12.6\%$$

$$P(DP \mid VW, DB) = \frac{11}{63} \doteq 17.5\%$$

holding ethnicity of victim constant at White, the rate of imposition of the death penalty rises (!), from 11.0% (top table) to 14.0%, and now Black defendants get the DP more often than white defendants.

analysis of bottom table (VB)

$P(DP \mid VB) =$	$\frac{6}{112} \doteq 5.4\%$	
$P(DP \mid VB, DW) =$	$\frac{0}{9} = 0\%$	
$P(DP \mid VB, DB) =$	$\frac{6}{103} \doteq 5.8\%$	

holding ethnicity of victim constant at Black, the rate of imposition of the death penalty falls (!), from 11.0% (top table) to 5.4%, and (again) now Black defendants get the DP more often than White defendants.

so: overall, in the aggregate (top table), as ethnicity of defendant moves from Black to White, $P(DP)$ goes <u>up</u>, but separately for each of VW (middle table) and VB (bottom table), as ethnicity of defendant moves from Black to White, $P(DP)$ goes <u>down</u> (i.e., the relationship reverses direction): this is a called <u>Simpson's Paradox</u>, and there's nothing paradoxical going on.

why did this happen?

① Murder victims typically ④ know their murderer.

② In the U.S., white people tend to hang out with white people, and Black with Black.

③ Therefore white defendants are mostly murdering white victims.

④ Judges & juries in the U.S. impose the death penalty more often when the victim is White than when the victim is Black.

Homework for you, not to turn in: show that ethnicity of victim is indeed a PCF here, by computing and comparing

P(DP)	P(DW)
P(DP | VB)	P(DW | VB)
P(DP | VW)	P(DW | VW)