this time: conditional probability; discrete distributions

next time: continuous distributions; expectation

$Y_i = \{$ you initially choose door $i\}$

$M_j = \{$ Monte Hall then opens door $j\}$

$C_k = \{$ car actually behind door $k\}$

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

you pick door 1 (Y_1) & Monte opens door 2 to reveal a goat (M_2)

without loss of generality

read DS ch. 3,4

office hours:

M noon-1pm CB BE 312 C/D

M 1-2pm DD BE 357C

Tu 10-11 am CB BE 312 C/D

W noon-1pm DD BE 357C

Th 11.30a-12.30p CY BE 312C/D

F noon-1pm CY BE 312 C/D

but not today

AMS 131

1 Aug 16

doc ① Com. notes.

PLAN AHEA

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

This is like ELISA:

	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 $P(c1) = P(c3) = \frac{1}{3}$, (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 general form of product rule for and:

and

$$\frac{P(M2, Y1 \mid c1)}{P(M2, Y1 \mid c3)} = \frac{\cancel{P(Y1 \mid c1)} \cdot P(M2 \mid Y1, c1)}{\cancel{P(Y1 \mid c3)} \cdot P(M2 \mid Y1, c3)}$$

1

but Y_j 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 M2 (given y_i, c_j),

the posterior odds in favor of car behind door 3 are 2:1,

so $P(C_3 \mid M_2, y_i) = \frac{2}{3}$ & you should switch.

Case study: Cromwell's Rule

Dennis Lindley

for any D such that $P(D) > 0$ & any A,

(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
A = unknown	$P(A \mid D)$ posterior info about A	

Meaning | Anything you put prior probability $\boxed{1}$ on has to have posterior probability $\boxed{1}$ no matter how the dataset comes out; this destroys the possibility of learning from data.

(a) $$P(A \mid D) = \frac{P(A \text{ and } D)}{P(D)}$$

but if $P(A) = 0$ (ϕ)

then $P(A \cap D) = 0$ ✓

$$P(A \mid D) = \frac{P(A \text{ and } D)}{P(D)}$$

$P(A) = 1$

so $(A \cap D) = D$

and $P(D) / P(D) = 1$

A ↕ S

(9.50)

Case study: The Rasmussen Report

WASH-1400, "The Reactor Safety Study"

problem Estimate p (catastrophic accident at nuclear power plant) (*) at a moment in history (1975) when no such events had ever occurred (nuclear power began in about 1955, ...) per year

3 mile Island 1979

"Solution" Use expert judgment to break down (*) into a collection of simpler events connected together with and, or, ... ; for example

lots

(*) = (thing 1 breaks & alarm 1 fails to go off & thing 2 breaks & ...)

Result | Their Estimate of $P(\circledast)$ was extremely small: 10^{-12}, & yet only 4 years later: Three Mile Island

---

What went wrong? | Right calculation:

$$P(\circledast) = P\left(\begin{matrix}\text{thing 1}\\ \text{breaks}\end{matrix}\right) \cdot P\left(\begin{matrix}\text{alarm}\\ \text{fails}\end{matrix} \,\middle|\, \begin{matrix}\text{thing}\\ 1\\ \text{breaks}\end{matrix}\right)$$

$$\cdot\, P\left(\begin{matrix}\text{thing}\\ 2\ \text{breaks}\end{matrix} \,\middle|\, \begin{matrix}\text{thing } 1\\ \text{breaks}\end{matrix} \ \&\ \begin{matrix}\text{alarm}\\ 1\ \text{fails}\end{matrix}\ \&\right) + \cdots$$

What they did instead: they assumed independence!

$$P(\circledast) = \underbrace{P\left(\begin{matrix}\text{thing } 1\\ \text{breaks}\end{matrix}\right)}_{\text{small}} \cdot \underbrace{P\left(\begin{matrix}\text{alarm}\\ \text{fails}\end{matrix}\right)}_{\text{small}} \cdot \underbrace{P\left(\begin{matrix}\text{thing } 2\\ \text{breaks}\end{matrix}\right)}_{\text{small}} \cdots$$

= tiny just because many numbers close to 0 multiplied