April 28, 2017

This time: probability

Next time: probability models for sums + means

read: DD ch. 9

L 127-135

HW #2 due Wednesday May 3rd @ 11:59pm on canvas

* Take home midterm will be handed out next <u>wed</u>, and will be due one ~~wea~~ week later

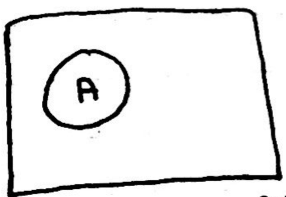
today: LN p. L 111

2nd draw depends on first draw w/ simple random sampling so we need a new version of probability ~~of~~ the form P(A given B), for ex., $P(y_2 = 9 \text{ given } y_1 = 9)$

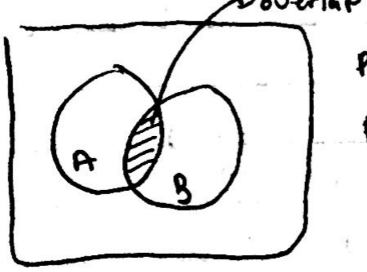
P(A given B) = P(A|B) conditional prob. of A given B

first people to think about this A. deMeivre (1720) + T. Bayes (1760)

def: P(A given B) = ?

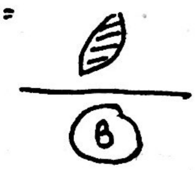


$$P(A) = \frac{A}{\square} \rightarrow P(\square) = 1 = 100\%$$



P(A given B) =

P(A/B) =



def: As long as ~~P(A~~ P(B) > 0,

$$P(A/B) = \frac{P(A \text{ and } B)}{P(B)}$$

multiply both sides of * by P(B): + reverse left + right:

$$P(A \text{ and } B) = P(B) \cdot P(A/B)$$

general form of product rule for working w/ and

ex. $\underset{SRS}{P(y_1 = 9}$ and $y_2 = 9) =$

$$\underbrace{P(y_1 = 9)}_{} \cdot \underset{SRS}{P(y_2 = 9 / y_1 = 9)} =$$

$$1/5 \cdot 0 = 0 \checkmark$$

True/False statements

def: A, B are independent if info about the truth of A doesn't change the chances for B, + vice versa

using at random w/ replacement sampling

The first and 2nd draws are independent (this is why this kind of sampling is called IID (independent identically distributed) sampling (all the IID draws have same prob. behavior)

~~pretty ct~~ if A, B independent, then

$P(A|B) = P(A)$ +

$P(B|A) = P(B)$

This observation reconsiles IID + SRS

IID $P(A \text{ and } B) = P(A) \cdot P(B|A)$
$= P(B \text{ and } A) = P(B) \cdot P(A|B)$ } chain rule for prob.

but under IID A, B indep. so under IID $P(B|A) = P(B)$ and $P(A|B) \underset{\text{if indep.}}{=} P(A)$ + $P(A \text{ and } B) = P(A) \cdot P(B|A)$

$\boxed{= P(A) \cdot P(B)}$

If independent

$= P(B) \cdot P(A|B)$

$\boxed{= P(B) \cdot P(A)}$

case study: UCLA 1992

MLP (Y/N)	gender (M/F)	
Y ⋮ Y	F ⋮ F	29
N ⋮ N	F ⋮ F	20
Y ⋮ Y	M ⋮ M	52
N ⋮ N	M ⋮ M	5

both categorical (qual.)

2x2 contigency table

gender \ MLP	Y	N	
F	29	20	49
M	52	5	57
	81	25	= 106

Q₁ = is there a relationship in this data set between gender + MLP?

Q₂ = Are gender and MLP associated?

Q₃ = Are gender, MLP, independent?

chance 1 person ~~to~~ at random from table → ELM

$P(Y) = \frac{81}{106} = 76\%$

$P(Y/M) = \frac{52}{57} = 91\%$

$P(Y|F) = \frac{29}{49} = 59\%$

This relationship is <u>strong</u> b/c 59% and 91% are quite diff. from 76%.

R-51 death penalty (DP)

outcome: death penalty or not? dich

"treatment variable": ~~dependent~~ defendant white (DW) or black (DB) dich

This is an observational study: enemy is bias ~~arising~~ arising from PCFs

PCF: race of ~~victum~~ victim VW VB

how control for PCF: hold it constant

$P(DP) = \frac{36}{326} = 11\%$

$P(DP/DW) = \frac{19}{160} = 12\%$

$P(DP/DB) = \frac{17}{166} = 10\%$

top table

white defendents got DP more than black defendants

$P(DP/\cancel{DW}\ VW) = \cancel{\frac{19}{151}}\quad \frac{30}{214} \approx 14\%$

$P(DP/VW + DW) \approx \frac{19}{151} = 12.6\%$

$P(DP/VW + DB) = \frac{11}{63} \approx 17.5\%$

2nd table

$P(DP/VB) = \frac{6}{112} = 5.4\%$

$P(DP/VB + DW) = \frac{0}{9} = 0\%$

$P(DP/VB + DB) = \frac{6}{103} = 5.8\%$

3rd table