Disc. Sec.
week of
6-10 Apr 20

gender vs. MLP

STAT 131
6 Apr 20

~~sample~~ pop.

y = MLP?	gender
yes Y	F
Y	M
N	M
Y	F
⋮	⋮

1 student at random

n = ~~sample~~ pop. size = 106

1 row for each individual (subject)

1 column for each variable

(sort)

Y	F	29
⋮		
Y	F	
N	F	20
⋮		
N	F	
Y	M	52
⋮		
Y	M	
N	M	5
⋮		
N	M	

MLP

Gender	Y	N	
F	29	20	49
M	52	5	57
	81	25	106

2x2 contingency table

Choose 1 student at random from 106:

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

conditional

ELM? yes

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

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

Q1: Are gender & MLP <variables> independent of each other in this data set?

A1: No, they're dependent; they are associated with each other; an association exists between them

Q2: Is assoc. strong (moderate ~~weak~~ nonexistent) because 91% is quite different from 59%

In reality ALL probabilities are conditional, at least on the assumptions made in calculating them

$P(Y) = P(Y \mid$ 1 student chosen at random from 186 $)$

Dr. Schram

P(inf. on any single act) $= p = \frac{1}{500}$

# acts (n)	P(inf.)	
100	20%	$= np$
500	100%	$= np$
501		
1000	200%	bullshit

surprising fact about HIV: it's possible to be infected with HIV more than once p(inf. in 100 acts) = ?

Math: All math can ever give is

If A then C

assumptions

conclusions

2 ways to measure quality:

① process: how we solve problems

② outcome: whether we succeeded in solving or not

useful process tip

When faced with a new problem, try to find an old problem with 2 properties:

① it's similar in all relevant ways to new prob.

② you already know how to solve

Q: Is Dr. Schramm similar to any problems we've already solved?

A: Our only success so far is T-S disease

T-S: A

P(1 or more T-S babies in family of 5, both p. carriers)

P(T-S on any one child) = $p = \frac{1}{4}$

Schramm:

P(inf. in 100 acts) = P(1 or more inf. in 100 unprotected acts)

P(inf on 1 act) $= \frac{1}{500} = p$

A: P(1 or more inf. in 100 acts)

$= 1 - P(\text{not } A)$

$= 1 - P(\text{no inf in 100 acts})$

$= 1 - P\left(\begin{matrix}\text{not}\\ \text{inf.}\\ \text{on 1st}\end{matrix} \text{ and } \begin{matrix}\text{not}\\ \text{inf.}\\ \text{on 2nd}\end{matrix} \text{ and} \cdots \text{and } \begin{matrix}\text{not}\\ \text{inf}\\ \text{on 100th}\end{matrix}\right)$

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

$\boxed{\text{Special case:}}$ if A, B indep.

then $P(A \text{ and } B) = P(A) \cdot P(B \cancel{\mid A})$
$= P(B) \cdot P(A \cancel{\mid B})$

$\overset{\text{indep}}{=} 1 - P\left(\begin{matrix}\text{not}\\ \text{inf.}\\ \text{on 1st}\end{matrix}\right) P\left(\begin{matrix}\text{not}\\ \text{inf.}\\ \text{on 2nd}\end{matrix}\right) \cdots P\left(\begin{matrix}\text{not}\\ \text{inf.}\\ \text{on 100th}\end{matrix}\right)$

constant risk

$$= 1 - \left(1 - \frac{1}{500}\right)\left(1 - \frac{1}{500}\right) \cdots \left(1 - \frac{1}{500}\right)$$

$$= 1 - \left(1 - \frac{1}{500}\right)^{100} = 1 - (1 - p)^n$$

n	P(1 or more inf.)	Schran
100	18%	20%
500	62% (?)	100%