this	probability
time:	models for
next	runs &
time:	graphs

read: DJ ch. 10

hwk 2 due Tue 20 oct in class

lab 2 due by 5pm Fri 23 oct in class or lab or disc. sec or at my office (BE 135)

MLP	gender (F, M)
Y	F
Y	M
N	F
Y	F
⋮	⋮

n = 106

MLP = marijuana legalization preference (Y, N)

Y F ⋮ Y F		29
Y M ⋮ Y M		52
N F ⋮ N F		20
N M ⋮ N M		5
		106

MLP

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

2x2 <u>contingency</u> <u>table</u> (example of categorical data analysis)

②

Q: Are gender, MLP independent in this data set, or are they dependent (associated)?

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

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

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

choose 1 person at random from 106 ELM ✓

so gender, MLP are (highly) dependent in this data set because

$76\% \neq 59\%$ or 91%

finish T-S

$\overbrace{P(\text{1 or more T-S in 5})}^{A} =$

$1 - P(\text{0 T-S in 5})$

$= 1 - P(\text{not T-S on 1st and not T-S on 2nd and} \dots \text{and not T-S on 5th})$

independent

$= 1 - P(\text{not T-J on 1st}) \cdot P(\text{not T-J on 2nd}) \cdots P(\text{not TJ on 5th})$ (3)

identical distribution

$= 1 - (1-\frac{1}{4}) \cdot (1-\frac{1}{4}) \cdots (1-\frac{1}{4})$

$= 1 - (1-\frac{1}{4})^5 \doteq 76\%$

~~each~~ penalty

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

defendant white

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

def. black

$P(DP \mid DB) = \frac{17}{166} \doteq 10\%$!

outcome: DP vs. not (Y)

treatment: (X) race of defendant (W vs. B)

basic design: obs. study

enemy: bias from PCFs

a PCF: race of victim (W vs. B) (Z)

Z, Y assoc? ✓

Z, X assoc? ✓

how defeat PCF at analysis time:
hold it constant

4

$P(DP \mid VW) = \frac{30}{214} \doteq 14\%$

$P(DP \mid DW \text{ and } VW) = \frac{19}{151} \doteq 13\%$

$P(DP \mid DB \text{ and } VW) \sim \frac{11}{63} \doteq 17\%$

$P(DP \mid VB) = \frac{6}{112} = 5\%$

$P(DP \mid DW \,\&\, VB) = \frac{0}{9} = 0\%$

$P(DP \mid DB \,\&\, VB) = \frac{6}{103} \doteq 6\%$

---

direction of relationship between X & Y switched when Z was controlled for: Simpson's paradox

P(coming out ahead | spin, single #) 5

$= \frac{1}{38}$ (ELM ✓) P(split) $= \frac{2}{38}$

$\doteq 2.5\%$ $\doteq 5\%$

$$\mu = \frac{\overbrace{(-\$1) + (-\$1) + \cdots + (-\$1)}^{37} + (+\$35)}{38}$$

$= \frac{-\$2}{38} = -\$0.0526$: each time I bet

\$1 on a single # I expect to

lose about 0.05, give or take

about $\sigma = \$5.76$ | pop SD σ: | math fact: if pop.

has only 2 values,

$$\sigma = \left[\begin{pmatrix} \text{larger} \\ \# \end{pmatrix} - \begin{pmatrix} \text{smaller} \\ \# \end{pmatrix} \right] \cdot \sqrt{ \begin{pmatrix} \text{proportion} \\ \text{of larger} \\ \#s \end{pmatrix} \begin{pmatrix} \text{proportion} \\ \text{of smaller} \\ \#s \end{pmatrix} }$$

6

pop

possible outcomes of a single spin

(single #)

your net gain

your net gain	single #
-$1	0
-1	00
⋮	1
⋮	⋮
-1	5
+35	6
-1	7
⋮	⋮
-1	36

N = 38

mean μ = -$0.05

SD σ = $5.76

pop. hist.

37/38

1/38

-1   +35

sample

observed spins

IID

your net gain

$\begin{bmatrix} -1 \\ -1 \\ +35 \\ \vdots \\ -1 \end{bmatrix}$ n = 1000

sum $ = ?

(ex. -$64)

IID

$\begin{bmatrix} \vdots \end{bmatrix}$ n = 1000

sum $ = ?

(ex. -$28)

⋮

imaginary ~~data set~~

all possible values of $

$\begin{bmatrix} -\$64 \\ -28 \\ \vdots \\ \vdots \end{bmatrix}$

a $ ↑

another $ ↓

M = ∞

long run mean = expected value of $ = -$53

long run SD

long run hist

-$53

here $\sigma = \left[(+\$35) - (-\$1)\right]\sqrt{\left(\frac{1}{38}\right)\left(\frac{37}{38}\right)}$

$= \$5.76$

$P(\text{coming out ahead}) = P(S > \$0)$

$= ?$

1000 spins, $\frac{1}{38}$ of $(+\$35)$, $\frac{37}{38}$ of $(-\$1)$

I expect around 26 $(+\$35)$ & 974 $(-\$1)$,

sum $= 26(+\$35) + 974(-\$1) = -\$64$

---

another possibility: $27(+\$35) + 973(-\$1)$

$= -\$28$

$64 - 28 = 36$

← 36 →

-1 $\quad$ $+35$

<u>math fact:</u>

expected value of S = EV = $\boxed{E_{IID}(S) = n\mu}$

$= \left(\#\ \text{draws}\right)\left(\text{pop mean}\right) = (1000)(-\$0.05) = -\$53$

after 1000 ~~21~~ plays of single #
I expect to be behind by
about $53 (EV) give or take