Julia Larson

AMS 7
Notes
10/15
1

10/15 exp. Probability model

Marijuana Legalized Preference

Gender	Yes	No	total
Female	29	20	49
male	52	5	57
total	81	25	106

MLP	Gender (M+F)
N	F
Y	M
Y	M
N	F
Y	M

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

MLP

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

2x2 Contingency table
(exp. of categorical data analysis)

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

A: $P(Y) = \frac{81}{106} \doteq 76\%$ (person chosen at random so ELM applies)

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

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

so: Gender and MLP are (Highly) dependent in this data set because
$76\% \neq 59\%$ or 91%

Back to T-S babies

P(1 or more T-S babies in 5) =

1 − P(0 T-S in 5) =

1 − P(Not T-S on 1st and ... and Not T-S on 5th) = cont...

Julia Larson

continued...

$= 1 - P\left(\begin{smallmatrix}\text{not T-S}\\ \text{on 1st}\end{smallmatrix}\right) \cdot P\left(\begin{smallmatrix}\text{not T-S}\\ \text{on 2nd}\end{smallmatrix}\right) \cdots\cdots P\left(\begin{smallmatrix}\text{Not T-S}\\ \text{on 5th}\end{smallmatrix}\right) =$

Babies with same parents is identically distribution

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

$= 1 - (1-\frac{1}{4})^5 \doteq \boxed{76\%}$ chance of having 1 or more T-S babies

Death Penalty exp.

$P(\text{Death Penalty}) = \frac{36}{326} \doteq 11\%$

$P(\text{DP} \mid \text{Defendent white}) = \frac{19}{160} \doteq 12\%$!

$P(\text{DP} \mid \text{Defendent Black}) = \frac{17}{166} \doteq 10\%$

Outcome: DP vs. Not (Y)

Treatment: race of defendent (X)

Basic Design: observational study

Enemy: bias from PCF's

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

Are Z, Y associated? Yes

are Z, X associated? Yes

How to defeat PCF at analysis time? HOLD IT CONSTANT

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

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

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

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

$P(\text{DP} \mid \text{DW and Victim B}) = \frac{0}{9} \doteq 0\%$

$P(\text{DP} \mid \text{DB and VW}) = \frac{6}{103} \doteq 6\%$

The direction of relationship between X and Y switched when Z was controlled for = Simpson's Paradox

Julia Larson

Roullete

P(coming out ahead | spin, singles) = $\frac{1}{38}$ (ELM) = 2.5%

P(" " " " " , split) = $\frac{2}{38}$ (ELM) $\doteq$ 5%

Pop.

All possible outcomes of a single spin

My Net Gain

\$-1	0
\$-1	00
⋮	⋮
\$ 35	6
⋮	
\$-1	36

Single # bet

IID

N=1000

Sample

observed spins

My Net game

\$-1
\$-1
\$+35
⋮
\$-1

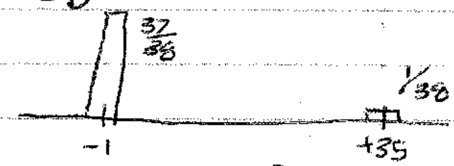
walk out here

Sum S = ?

Pop. Histogram

Mean

SD



μ: $\overbrace{(\$-1)+(\$-1)+(\$-1)+\ldots}^{37}+(\$35) = \frac{\$-2}{38} = \$-0.0526$:

each time I bet \$1 on a single # I expect to lose about μ 0.05, give or take

P(coming out ahead) = P(S > 0) = ?

IID

[ ] n=1000

Sum S = ?

⋮

Repeat Again and Again put the S in a table

SD σ: (Math fact: if pop. has only 2 values:

$\sigma = \left[\left(\text{larger \#}\right)-\left(\text{smaller \#}\right)\right]\sqrt{\left(\text{Proportion of larger \#s}\right)\left(\text{Proportion of smaller \#s}\right)}$

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

σ = \$ 5.76

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

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

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

Another possibility: 27(\$+35) + 973(\$-1) = \$-28

64 - 28 = 36

-1 ←36→ +35

\$-64
\$-28

A # Integer of S

m = ∞

Long run mean = E

expected value of S = E

long run SD

$E_{IID}(S) = \left(\text{\# spins}\right)\left(\text{pop mean}\right) = 1000(-.05) \doteq \$-53$

After a 1000 \$1 plays I expect to be behind by \$53 give or take...

\$-53

← long run hist.