

this time: conditional densities
 next time: multivariate distributions, transformations

please finish reading all of JS ch. 3 except section 3.10 (later) ①

STAT 131
 7 May 20
 (lecture)

UCLA case study:

gender and marijuana

X (G) legalization preference

$\begin{cases} 1 & \text{if female} \\ 0 & \text{not} \end{cases}$

Y (MLP) $\begin{cases} 1 & \text{if favor legalization} \\ 0 & \text{not} \end{cases}$

(MLP) Y | X (gender)

0	1
0	0
1	1
⋮	⋮

$n=106$

sort

joint PMF

	1	0	
X (G)	1	0	
	29	20	49
	52	5	57
	81	25	106

marginal for Y

marginal for X

2x2 contingency table:

bivariate discrete PMF

1	1	} 29
⋮	1	
1	0	} 52
⋮	0	
0	1	} 20
⋮	1	
0	0	} 5
⋮	0	
		106

$P(Y=y | X=x) = ?$

$$P(\underline{Y}=1) = P(\text{yes to MLP}) = \frac{81}{106} \stackrel{2}{=} \underline{79\%}$$

$$P(Y=1 | X=1) = P\left(\begin{matrix} \text{yes} \\ \text{to} \\ \text{MLP} \end{matrix} \mid \text{female}\right) = \frac{29}{49} = \underline{60\%}$$

$$P(Y=1 | X=0) = P\left(\begin{matrix} \text{yes} \\ \text{to} \\ \text{MLP} \end{matrix} \mid \text{male}\right) = \frac{52}{57} = \underline{90\%}$$

Are X and Y independent in this joint PMF?

no

$$P(Y=1 \text{ and } X=1) = \frac{29}{106}$$

they are strongly dependent

$$\neq \cancel{P} P(Y=1) \cdot P(X=1) \\ \left(\frac{81}{106}\right) \cdot \frac{49}{106}$$

dependent = strongly associated

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

population: all possible (n, b) pairs
 stationarity the observed (n, b) pair
 defective?

↑
 ↓
 mean $\theta = ?$
 (unknown)

1
 2
 0

like IID

D_1 $d_1 = 0$ 0%
 D_2 $d_2 = 0$ 0%
 D_3 $d_3 = 1$ 33%
 D_4 $d_4 = 0$ 35%

defective
 $= N = \sum_{i=1}^m D_i$

$\hat{\theta}$ = proportion of defective

$= \frac{N}{m} = \frac{1}{m} \sum_{i=1}^m D_i$

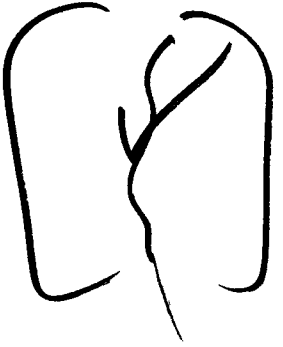
using $\hat{\theta}$ to estimate

θ is an example of statistical inference

case I

pop. known whole general

deduction sample unknown (don't have part particular yet)



like IID

[?]

this is probability

easier

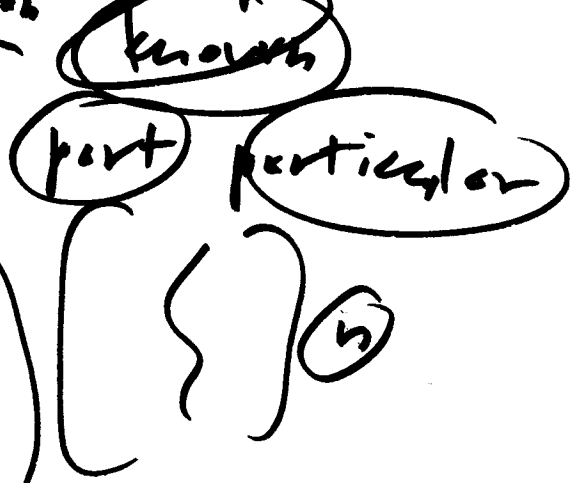
STAT 131 before

Case II prop.



induction

sample



this is statistics:
 statistical inference
 induction =

harder

Sheeplock Holmes: ~~A.C.J. said deduction~~
 induction

posterior information about nuclear accidents before 3 mile island in 1979

Bayesian = prior info about the possibility of such accidents

into content = 0
 ↓
 data into all previous accidents

external

internal