

JD  
Office  
1.5 hour  
Thu 16  
Apr 20

Quiz 3

$$P(U|D) = \frac{P(U)P(D|U)}{P(D)}$$

STAT 131  
16 Apr 20

$P(D) > 0$  assume

①

(a) suppose we put

$P(U) = 0$   
~~just: leg? no~~

$$\rightarrow \frac{P(U)P(D|U)}{P(D)} = 0$$

so if  $P(U) = 0$  then  $P(U|D) = 0$   
no matter what  $D$  is

(b) suppose we put  $P(U) = 1$

$$P(U|D) = \frac{\overset{1}{P(U)}P(D|U)}{P(D)} = \frac{P(D|U)}{P(D)} = 1$$

but if  $P(U) = 1$ ,  $P(D|U) = P(D)$

for any  
T/F

$$P(D | \underline{I=1}) = P(D)$$

prop. D,

$$I(\text{outcome}) = \begin{cases} 1 & \text{if dead in 1992} \\ 0 & \text{not} \end{cases}$$

$X =$  ("treatment")

(~~supposedly~~ supposedly

causal factor

$$= \begin{cases} 1 & \text{if current smoker in 1972} \\ 0 & \text{not} \end{cases}$$

$Z =$  (CF) =  
confounding factor

$$= \begin{cases} 1 & \text{if 15-64 in 1972} \\ 0 & \text{not} \end{cases}$$

tests to see if  $Z$  (age) is a CF

plausible that  $Z, X$  assoc. ? yes  
as  $Z$  goes from 0 (65+) to 1 (15-64),  $P(\text{dead}) \downarrow$

② plausible not  $Z_1, Z_2$  - soc.? (74) ③

as  $Z_1$  goes from 0 (65+) to 1  
 (18-64),  $P(\text{smoker} \uparrow)$   
 in 1972

~~$Z_1 = 0$~~   $Z_1 = 1$

$$P(S' | \text{64}) = \frac{533}{1072} = 49.7\%$$

$$P(S' | \text{65+}) = \frac{49}{242} = 20.2\%$$

$$O_A = \frac{P_A}{1 - P_A}$$

if  $P_A = 0$  then

$$O_A = P_A$$

$$P_A = \frac{1}{38.32} \rightarrow 0.026$$

$$O_A = \frac{1}{\frac{37.32}{25.32}} = \frac{1}{37.32} = 0.027$$

have to know 3 things:  $P(\odot) = 0.01$  (4)

① prevalence of COVID-19 in

U.S. pop. : case 1:  $\frac{64,186}{328,200,000} = 0.2\%$

case 2: 1.0%

②  $P(\oplus | \odot) = 93.8\%$   
 = sensitivity of cellex test

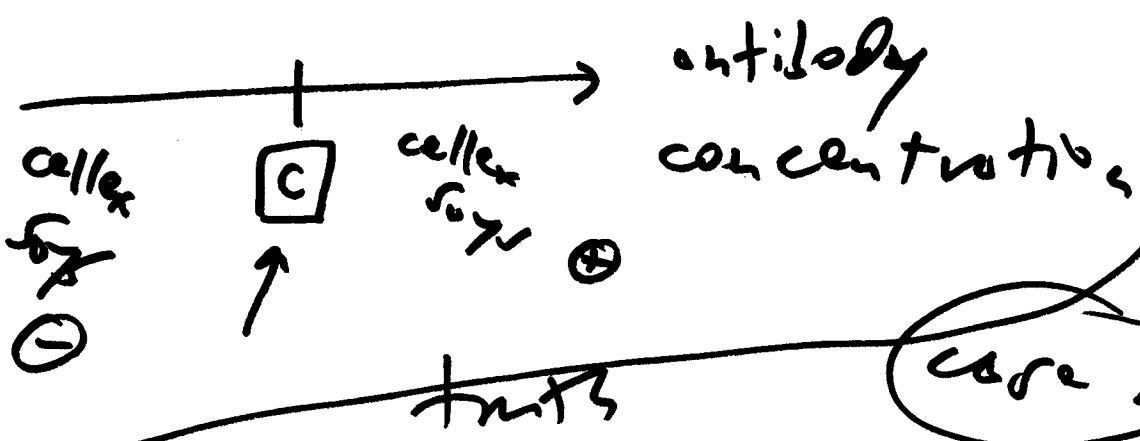
③  $P(\ominus | \text{not } \odot) = \text{specificity of } \uparrow$   
 = 95.6% culprit

method I

	$\odot$	$\text{not } \odot$	
cellex test says $\oplus$	9,380	43,560	52,940
cellex test says $\ominus$	620	946,440	947,060
	10,000	990,000	1,000,000

false positive: 43,560  
 false negative: 620  
 $P(\odot) = 1\% = \frac{10,000}{1,000,000}$   
 $P(\text{not } \odot) = \frac{990,000}{1,000,000} = 99\%$

$P(\odot | \oplus) = \frac{9,380}{52,940} = 0.177 = 18\% (!)$



	(C)	(not C)	
cellx test (+)	1,876	43,912	45,788
cellx test (-)	124	954,088	954,212
	2000	99,800	1,000,000

$P(C) = 0.002$

$P(C | +) = \frac{1,876}{45,788} = 0.04 (!)$

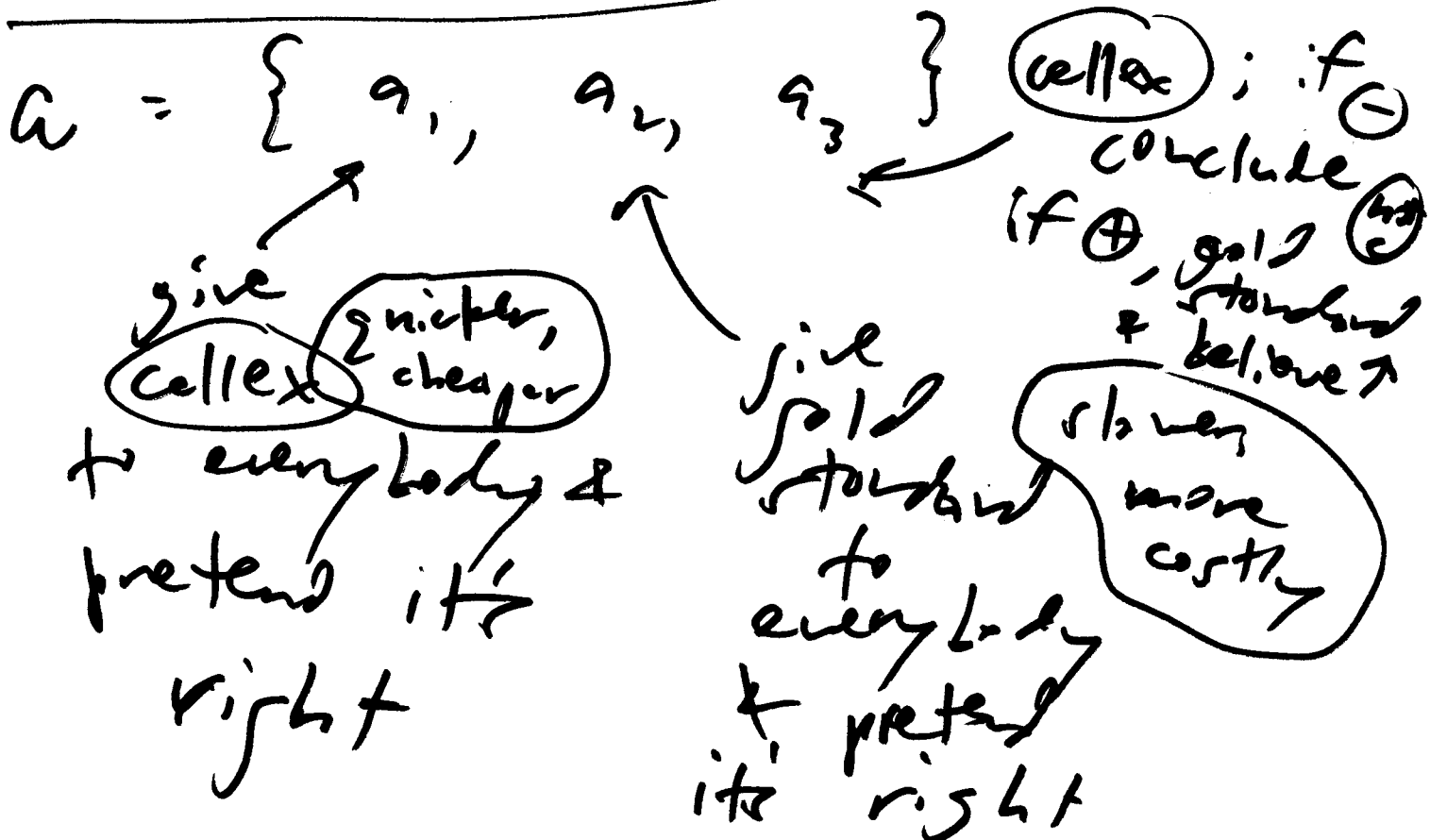
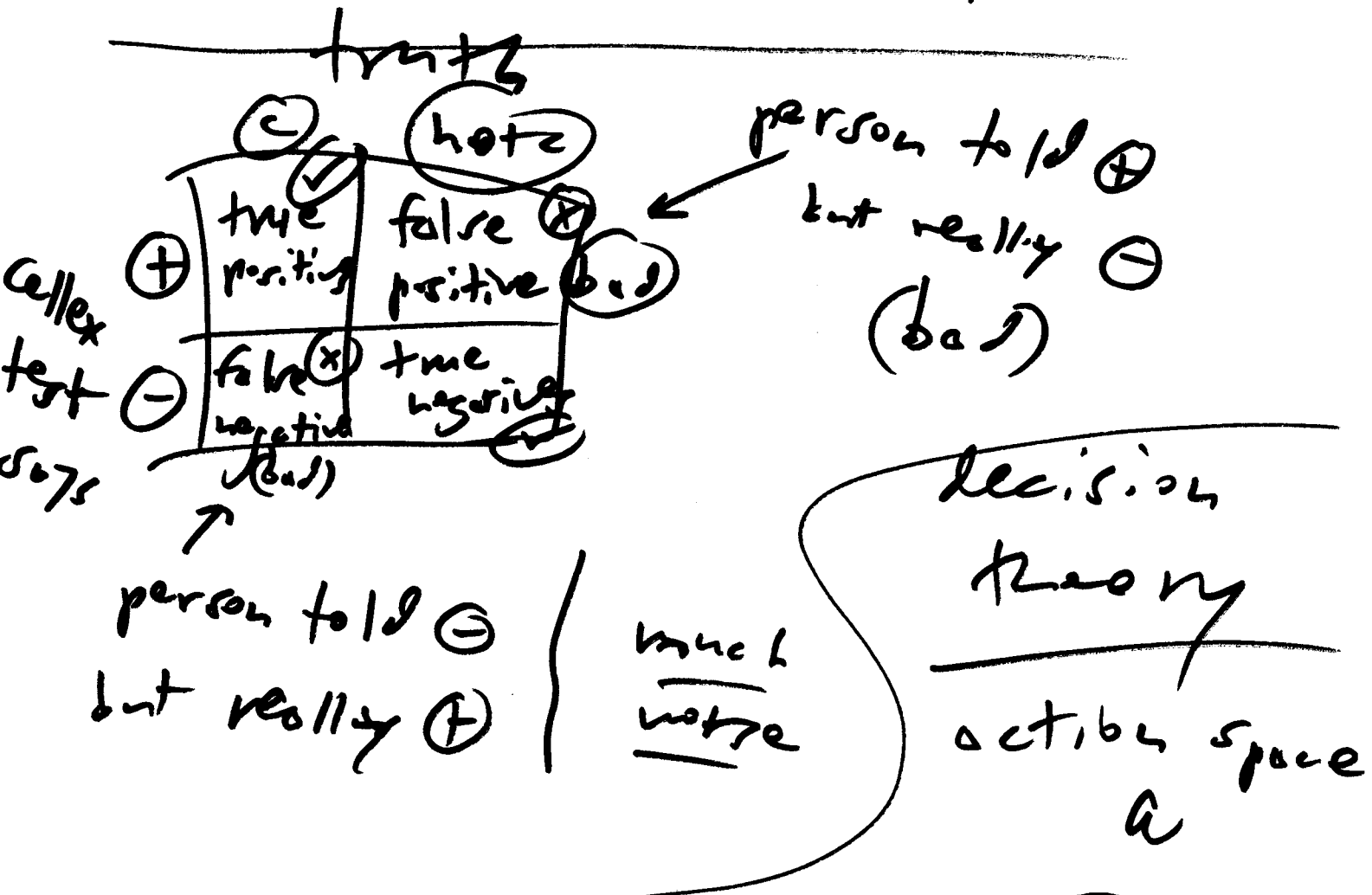
$P(\text{not } C | -) = \frac{954,088}{954,212} = .99987$

case	$P(C   +)$	$P(\text{not } C   -)$	prevalence
1	<del>0.04</del> 4%	99.99%	0.2%
2	18%	99.93%	1.0%

cellx test almost perfect when (-)

but it's dup when it says ⊕

⑥



know:

want ②

$$P(\ominus)$$

$$P(\ominus | \oplus)$$

$$P(\oplus | \ominus) = \text{sens.}$$

$$P(\text{hot } \ominus | \ominus)$$

$$P(\ominus | \text{hot } \ominus) = \text{spec.}$$

$$P(A | B) \stackrel{?}{=} P(B | A) \quad \text{ho}$$

$$P(\text{rain} | \text{clouds overhead}) = \text{low to in between}$$

$$P(\text{clouds overhead} | \text{rain}) = \text{high}$$

$$P(\ominus | \oplus) = \frac{P(\ominus) P(\oplus | \ominus)}{P(\oplus)} \quad (1)$$

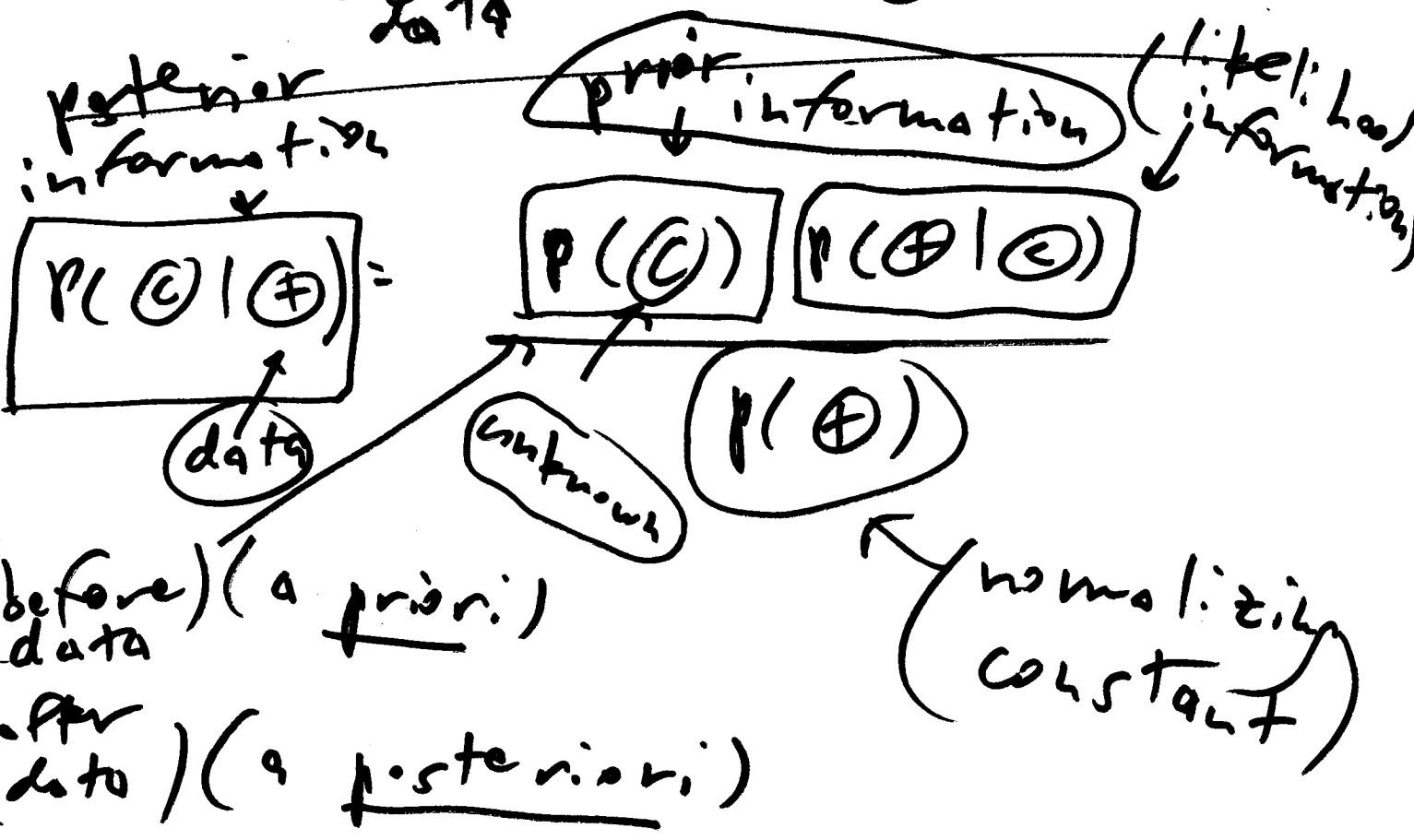
↓ prevalence
↓ sensitivity

(annoying) normalizing without

method II /  $P(\text{not } \ominus | \oplus) = \frac{P(\text{not } \ominus)P(\oplus | \text{not } \ominus)}{P(\oplus)}$  (2)

(1)  $\frac{P(\ominus | \oplus)}{P(\text{not } \ominus | \oplus)} = \frac{P(\ominus)}{P(\text{not } \ominus)} \left[ \frac{P(\oplus | \ominus)}{P(\oplus | \text{not } \ominus)} \right]$   
 (2)  $\frac{P(\ominus | \oplus)}{P(\text{not } \ominus | \oplus)} = \frac{P(\ominus)}{P(\text{not } \ominus)} \left[ \frac{P(\oplus | \ominus)}{P(\oplus | \text{not } \ominus)} \right]$   
 (Likelihood ratio)

(posterior odds w/ N in favor of  $\ominus$  given data) = (prior odds ratio in favor of  $\ominus$ ) (Bayes factor)





$$p(\text{posterior info}) \propto \frac{p(\text{prior info}) \cdot p(\text{likelihood info})}{\text{normalizing constant}}$$

(normalizing constant)

$$p(\text{unknown} | \text{data}) =$$

$$\frac{p(\text{unknown}) \cdot p(\text{data} | \text{unknown})}{\text{normalizing constant}}$$

(normalizing constant)

$$\frac{P(+|+)}{P(+|\text{not } +)}$$

prevalence (case 2)      sensitivity

$$= \begin{pmatrix} .01 \\ .99 \end{pmatrix} \begin{pmatrix} 0.938 \\ 1 - .956 \end{pmatrix} \quad (+)$$

1 - Specificity

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

$$\begin{pmatrix} 1 \\ 99 \end{pmatrix} \begin{pmatrix} 469 \\ 22 \end{pmatrix}$$