

this random  
time: usually;  
next CDF,  
time: PMF, PDF

read: JS ch. 3

STAT 131

21 Apr 20  
(lecture 1)

last time: (case 2)

	<div> <div>⊖</div> <div>⊕</div> </div>		
	9380	4356	52,940
Cells	⊕	⊖	
5.75	620	946,440	947,060
	10,000	990,000	1,000,000

Method I

(contingency table)

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

$$\frac{9380}{52,940} = \frac{469}{2,647} \approx 0.177$$

method II: Bayes'

Theorem in odds ratio form

$$\frac{P(\ominus | \oplus)}{P(\oplus | \ominus)} = \left[ \frac{P(\ominus)}{P(\oplus)} \right] \left[ \frac{P(\oplus | \ominus)}{P(\ominus | \oplus)} \right] = \text{odds ratio}$$

$$= \left( \frac{0.01}{0.99} \right) \left( \frac{0.938}{1 - 0.956} \right) = \frac{469}{2,178}$$

we already  
saw that for

any T/F statement A,  $o_A = \frac{P(A)}{P(\text{not } A)} = \frac{P_A}{1 - P_A}$

means

that  $P_A = \frac{o_A}{1 + o_A}$ , so we can easily convert  $o_{\ominus | \oplus}$  to  $P(\ominus | \oplus) =$

$$P(+|+) = \frac{0 \oplus 10}{1 + 0 \oplus 10} = \frac{469}{2,178} = \frac{469}{2,647} \approx 0.177 \quad (2)$$

method III

Bayes's theorem directly

$$P(+|+) = \frac{P(+|+)P(+)}{P(+)}$$

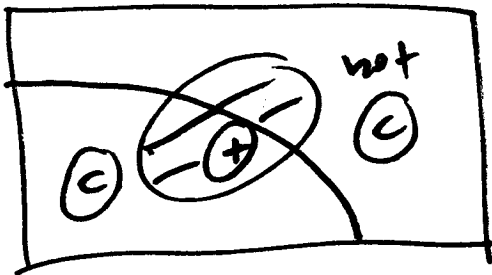
prevalence (case 2: 1%)

$P(+)$  annoying denominator

sensitivity of call test (0.938)

key idea:

$P(+)$  =  $P(\text{data})$  is hard to think about without knowing the unknown truth:  $\oplus$  or  $(\text{not } \oplus)$ , so extend the conversation (Jerris Lindley): ~~the~~ bringing  $\oplus$  and  $(\text{not } \oplus)$  into the calculation by partitioning over them:



$$P(+|+) = P(+ \text{ and } \oplus) + P(+ \text{ and } \text{not } \oplus)$$

$$P(+)=P(+ \text{ and } \odot)+P(+ \text{ and } \odot^{\text{not}}) \quad (3)$$

So far this still looks hard, but remember what we know: sensitivity =  $P(+|\odot)$  and specificity =  $P(-|\odot^{\text{not}})$ . Evidently We want to get

$\odot$  and  $\odot^{\text{not}}$  on the right-hand side of the conditioning bar, but this is easily done:

$$P(+ \text{ and } \odot) = P(\odot) P(+|\odot) \quad \left\{ \begin{array}{l} \text{we know} \\ \text{all of} \end{array} \right.$$

$$P(+ \text{ and } \odot^{\text{not}}) = P(\odot^{\text{not}}) P(+|\odot^{\text{not}}) \quad \left\{ \begin{array}{l} \text{these} \\ \text{probabilities} \end{array} \right.$$

$$= P(\odot^{\text{not}}) \cdot [1 - P(-|\odot^{\text{not}})] \quad \left\{ \begin{array}{l} \text{specificity} \end{array} \right.$$

so finally

$$P(+)=P(\odot) P(+|\odot)+P(\odot^{\text{not}})[1-P(-|\odot^{\text{not}})]$$

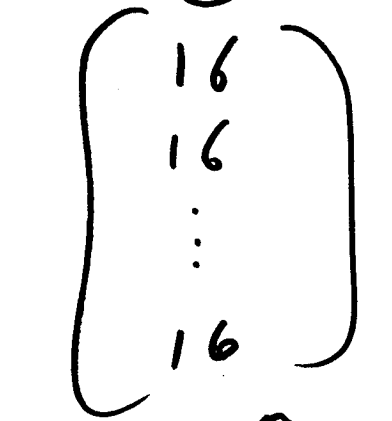
$$= (.01)(.938)+(.99)(1-.956) \doteq .053$$

$$= \frac{2,647}{50,000}$$

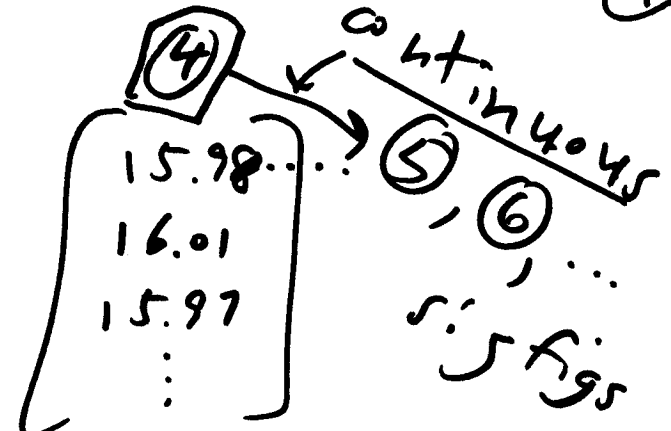
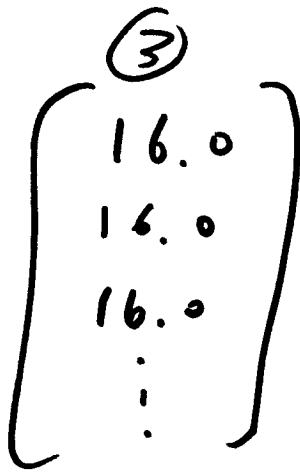
and

$$P(\odot|+)=\frac{(.01)(.938)}{.053}=\frac{469}{2,647} \doteq 0.177$$

weight (oz.)

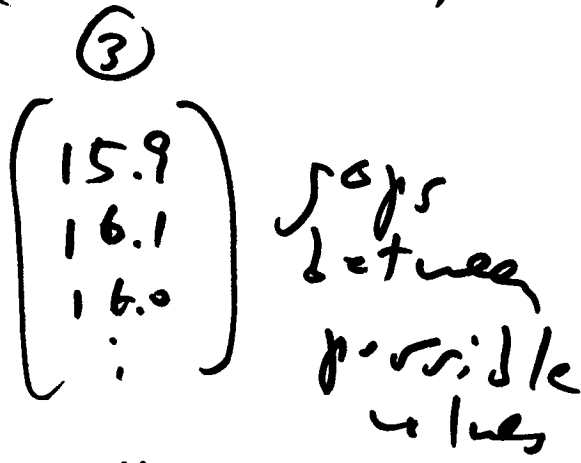


non-probabilistic  
(deterministic)

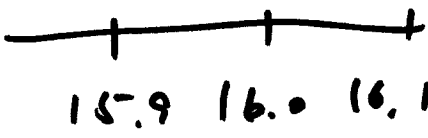


probabilistic

no conceptual  
gaps between  
possible values



discrete



we might choose to treat this  
as (approximately) continuous

Fact | All real-world measurement processes  
discretize things that may be

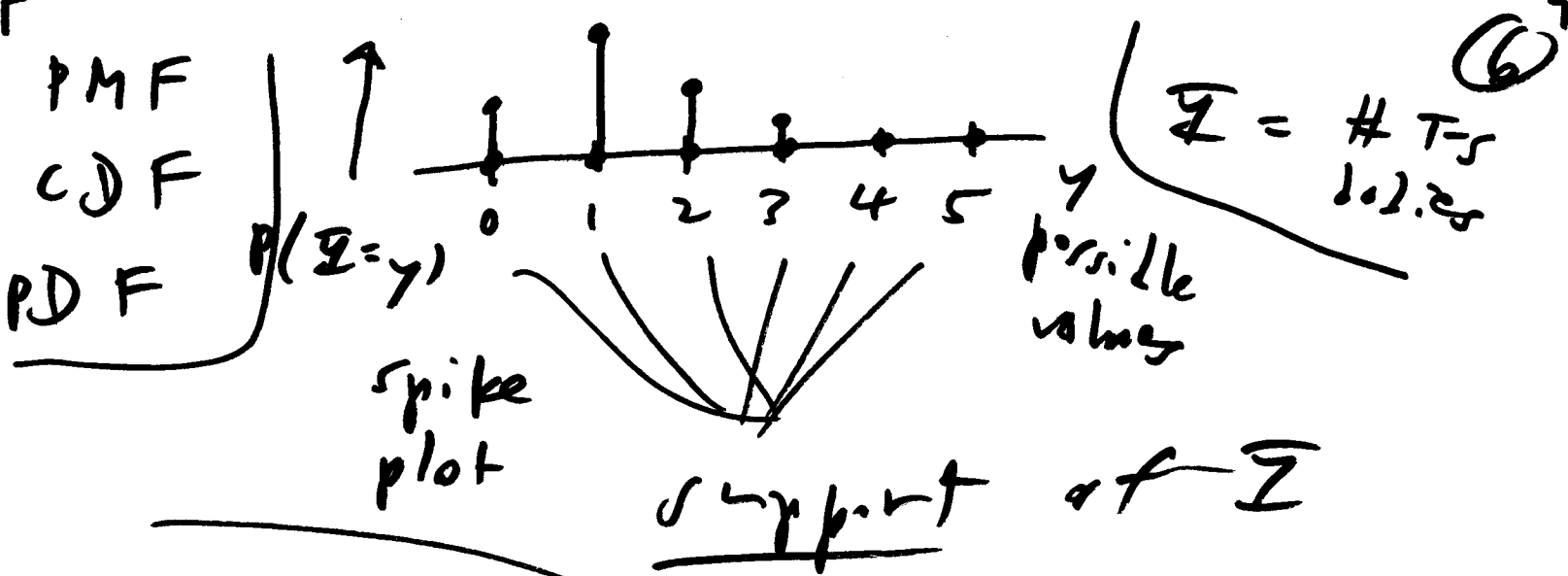
Conceptually continuous. ⑤

If <sup>gaps</sup> ~~gaps~~ between possible values are small in relation to magnitude we are about, we can choose to model a technically - discrete quantity as continuous.

ex. <sup>annual</sup> family income in U.S. ) we should treat this as \$82,148.93 cont.   
 ~~\_\_\_\_\_~~ \$ billions   
 \$0

ex. # kids with T-S disease

has to be discrete: structural gap between possible values ~~1/2~~



special bar graph

"is distributed as"

$X \sim \text{Bernoulli}(p)$

or

$(X|p) \sim \text{Bernoulli}(p)$

(freq)

$\updownarrow$

$0 \leq p \leq 1$

PMF of  $X$

$f_X(x) = \begin{cases} p & x=1 \\ 1-p & x=0 \\ 0 & \text{otherwise} \end{cases}$

else (math)

Computer science (if-then-else)

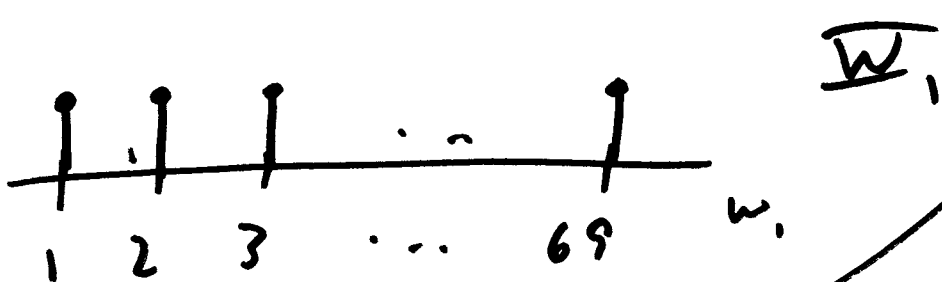
$$= p^x (1-p)^{1-x} \cdot I_{\{0,1\}}(x)$$

indicator  
function

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{else} \end{cases} \quad (7)$$

~~this is~~  
a uniform dist.

PMF  
of  $\Sigma$



$\Sigma = \underline{\text{\# animals at bank}}$   
in time interval  
 $\text{ran}(\vec{0}, T]$

Support of  $\Sigma = \{0, 1, 2, \dots\}$

$\Sigma$  is discrete but has

countably infinite support

(ex. Poisson)