this time: probability

read: DJ (B) Ch. 9–11, LN pp. 2–119 – 136

AMS 7
15 oct 15

next time: prob. models for sums & means

hwk 2 due next Thu in class

$$\begin{array}{c}\text{pop}\\ \begin{bmatrix}1\\2\\9\end{bmatrix}\end{array} \xrightarrow[\text{random}]{\text{at}} [\,y\,] \quad n=1$$

ELM ✓

$$P(y \text{ is odd}) = \frac{2}{3} \doteq 67\% \doteq 0.67$$

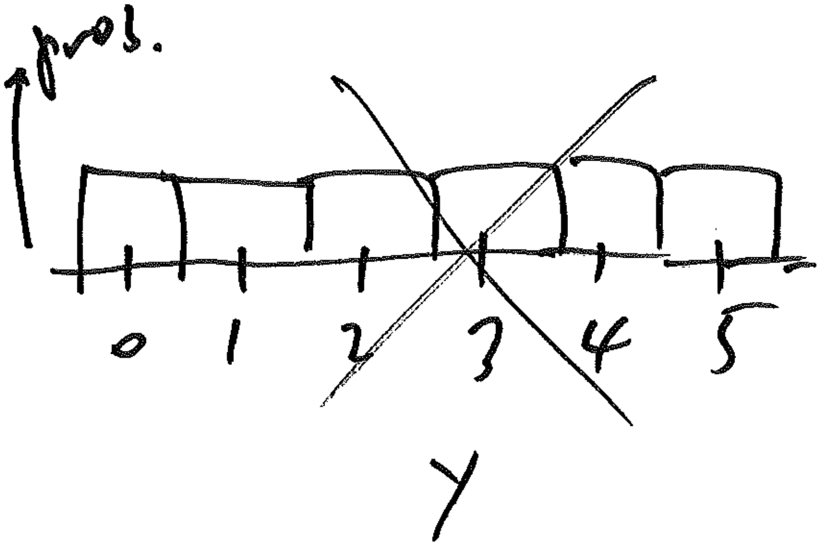
y = # of ~~their~~ children with T-S disease

$$\begin{bmatrix}0\\1\\2\\3\\4\\5\end{bmatrix}$$

if ELM applies

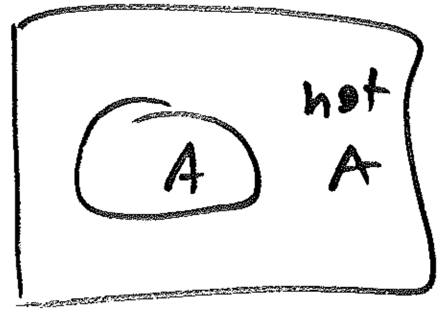
$$P(1 \text{ or more T-S}) = \frac{5}{6} \doteq 83\% \doteq 0.83$$

but ELM does <u>not</u> apply



ELM

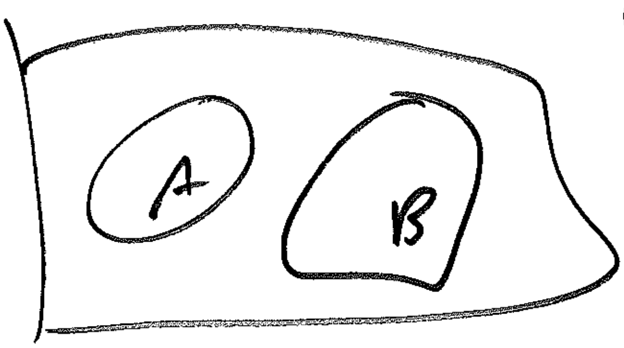
but not so



$$P(A) + P(\text{not } A) = 1$$
$$= 100\%$$

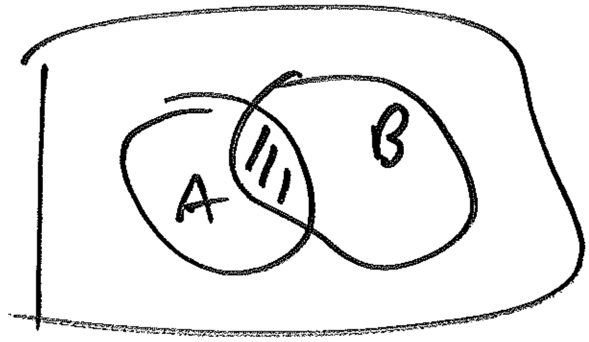
$$P(A) = 1 - P(\text{not } A)$$

addition rule, special case



$$P(A \text{ or } B) = P(A) + P(B)$$

no overlap $\Leftrightarrow$ A, B are mutually exclusive



overlay

$P(A \text{ or } B)$

$P(A) + P(B) - P(A \text{ and } B)$

general additional rule for* or

pop

$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$ ~~at~~ random $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ $n = 2$

$P(y_1 = 9 \text{ and } y_2 = 9) = ?$

at random → with replacement

→ without replacement

Independent identically distributed IID

Simple random sampling SRS

PLAN AHEA

case 1

$$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix} \xrightarrow{\text{IID}} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$P(Y_1 = 9 \text{ and } Y_2 = 9) = ?$$

IID

$Y_1 \backslash Y_2$	1	2	9
1	(1, 1)	(1, 2)	(1, 9)
2	(2, 1)	(2, 2)	(2, 9)
9	(9, 1)	(9, 2)	(9, 9)

ELM ? Yes

$$P(\overbrace{Y_1 = 9}^{A} \text{ and } \overbrace{Y_2 = 9}^{B}) = \frac{1}{9} \checkmark$$

$$P(Y_1 = 9) = P(A) = \frac{1}{3} = \frac{3}{9}$$

$$P(Y_2 = 9) = P(B) = \frac{1}{3} = \frac{3}{9}$$

our conjecture : $P(A \text{ and } B) = P(A) \cdot P(B)$

this works for IID $\frac{1}{9} = \frac{1}{3} \cdot \frac{1}{3}$

(8.53 5

$$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix} \xrightarrow{\text{SRS}} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

common sense

$$P(Y_1 = 9 \text{ and } Y_2 = 9) = 0$$

SRS

Y_1 \ Y_2	1	2	9
1	~~(1,1)~~	(1,2)	(1,9)
2	(2,1)	~~(2,2)~~	(2,9)
9	(9,1)	(9,2)	~~(9,9)~~

$$P(Y_1 = 9) = \frac{1}{3} = \frac{2}{6}$$

ELM? yes

$$P(Y_2 = 9) = \frac{1}{3} = \frac{2}{6}$$

by our conj

$$P_{SRS}(Y_1 = 9 \text{ and } Y_2 = 9) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$$

fails for SRS

but actually it's $0 = \frac{0}{6}$

conditional probability

$P(B \mid A) = ?$

given

Bayes
1760

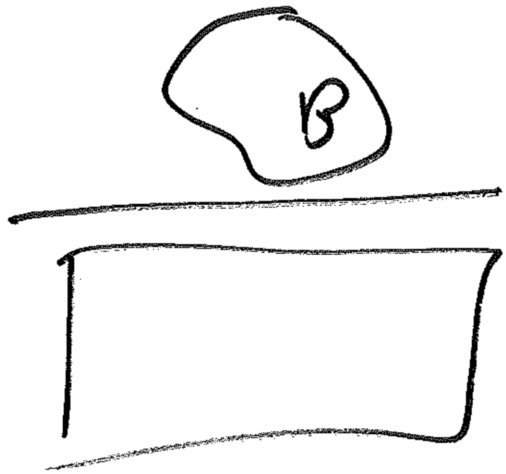
Bayesians

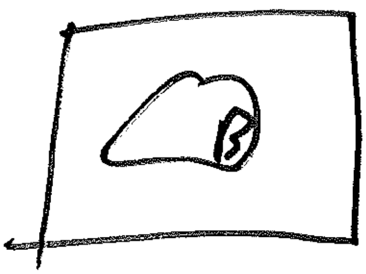
and
or
not

Pascal,
Fermat
1650

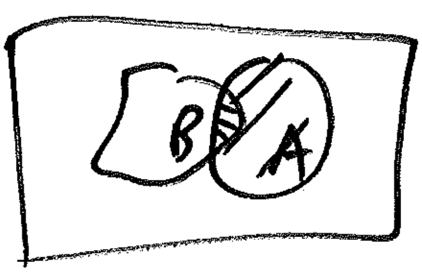
ex. $P(Y_2 = 9 \mid Y_1 = 2) = ?$

$P(B) = \dfrac{\text{area of } B}{\text{total area } (1)} =$





$P(B \mid A) = \dfrac{B \text{ and } A}{}$



def. $P(B \mid A) = \dfrac{P(A \text{ and } B)}{P(A)}$

(provided $P(A) > 0$; $P(B \mid \underline{A})$ is <u>undefined</u> if $P(A) = 0$

---

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

general product rule for and $= P(B) \cdot P(A \mid B)$

$P(A \mid B) = \dfrac{P(A \text{ and } B)}{P(B)}$

def. A, B independent if information about A does not change your predictions about B, & vice versa product rule, IID sampling:

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

$(y_1 = 9) \quad (y_2 = 9) \qquad \frac{1}{3} \cdot \frac{1}{3}$

---

if A, B indep then $P(A \text{ and } B) =$

$$P(A) \cdot P(B \mid A)$$

but $P(B \mid A) = P(B)$

so if A, B indep: $P(A \text{ and } B) = P(A) \cdot P(B)$

$P(\text{1 or more TJ in 5}) = ?$

P(1 or more T-S in family of 5)

$= 1 - P(0 \text{ T-S babies})$

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

indep

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

identical dist

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

$= 1 - \left(\frac{3}{4}\right)^5 = 0.76 = 76\%$

MLP	gender
Yes	M
No	F
⋮	⋮
Y	X
⋮	⋮

$n = 106$

sort

MLP	gender	
Y ⋮ Y	M ⋮ M	52
Y ⋮ Y	F ⋮ F	29
N ⋮ N	M ⋮ M	5
N ⋮ N	F ⋮ F	20

1 ~~person~~ row for each UCLA stat student in my class in 1992 who participated

Q: Are gender & MLP independent in this data set?

A: no, they're (strongly) dependent

choose a student at random

MLP

G

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

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

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

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