this time: basic probability; probability models

next time: probability models for sums & means; statistical models

read: DD (0) ch. 11

LN pp. L-137–160

AMS 7
1 Jul 2016

①

homework 2 target date: tue 5 july

take home mid term handed out tue 5 july, due mon 11 jul

makeup class Tue 5 jul 5–7.30 pm: will announce location in section Tue morning & on web site

P(1 or more T-s in family of 5) = ?

~~if ELM applied to this list of possible outcomes, then P(1 or more) = $\frac{5}{6} \doteq 83\%$~~

of T-s babies
0
1
2
3
4
5

but ELM doesn't apply to this list: 0 more likely than 5, ...

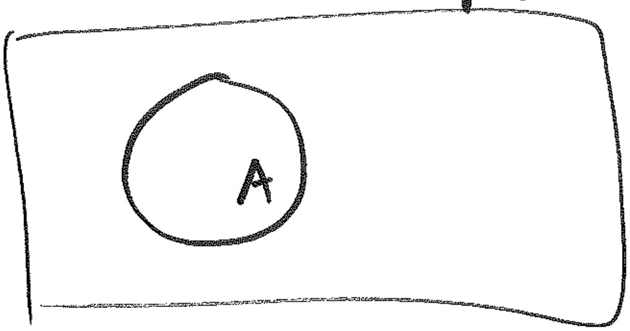
$P(A \boxed{\text{or}} B) \;?\; P(A) \quad P(B)$ ✓

$P(A) \;?\; P(\text{not } A)$ ✓

$P(A \boxed{\text{and}} B) \;?\; P(A) \quad P(B)$

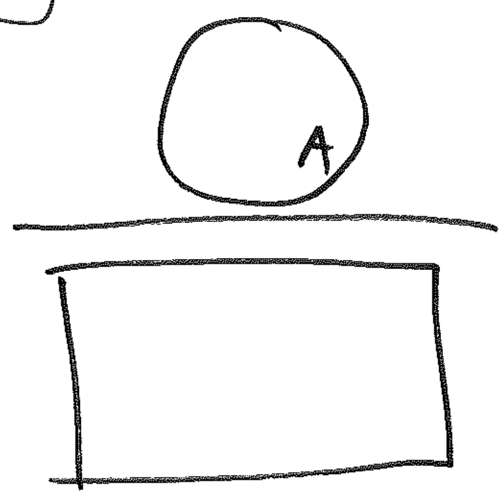
---

all possibilities



John Venn (1850)

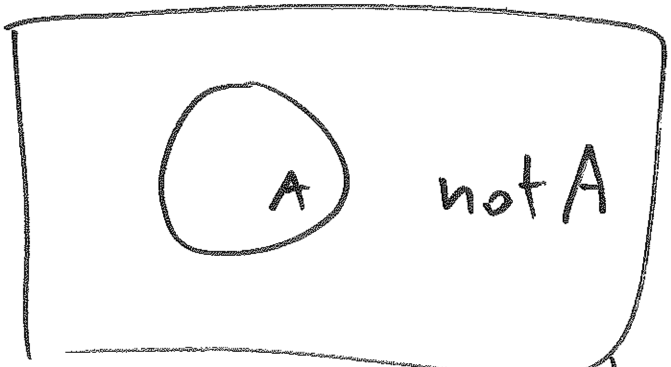
$P(A) =$



area 100% = 1

area of blob A

$$\frac{\text{area of blob A}}{1} = \text{area of blob A}$$



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

$= 100\%$

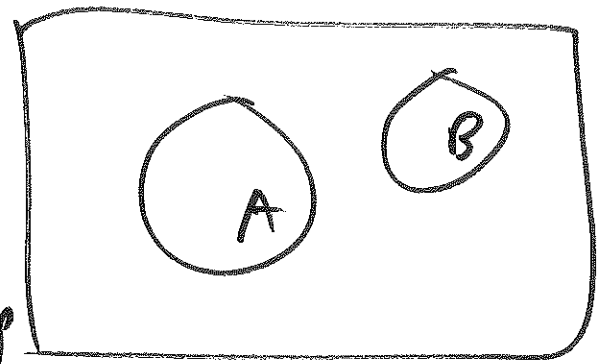
$P(A) = 1 - P(\text{not } A)$

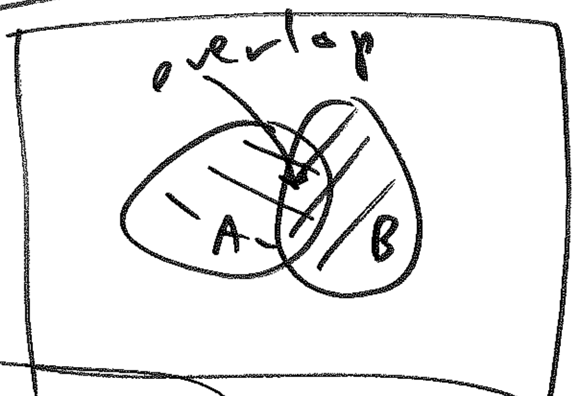
$0\% \le P(A) \le 100\%$

$0 \qquad\qquad 1$

special case when no overlap

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





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

general addition rule for $\boxed{\text{or}}$

if (A and B) can't happen, A, B mutually exclusive

(9.47)

and

pop.		sample
$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$	$\xleftrightarrow{\text{at random}}$	$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$

2 possibilities: at random with/without replacement

with replacement

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

		Y_2: 1	2	9
	1	(1, 1)	(1, 2)	(1, 9)
Y_1	2	(2, 1)	(2, 2)	(2, 9)
	9	(9, 1)	(9, 2)	(9, 9)

ELM with this list of 9 possibilities? yes

$$P(Y_1 = 9 \text{ and } Y_2 = 9) = \frac{1}{9}$$

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

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

$$P(Y_1 = 9 \text{ and } Y_2 = 9) = \frac{1}{9}$$

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

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

$$= P(Y_1 = 9) \cdot P(Y_2 = 9)$$

$$= \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9} \checkmark$$

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

without replacement

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)~~

ELM with this list of 6 possibilities?

yes

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

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

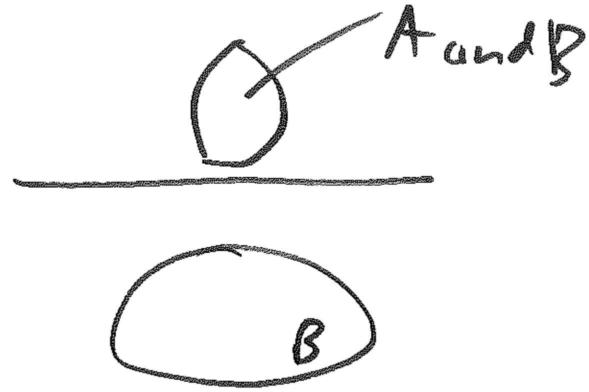
$$P(Y_1 = 9 \text{ and } Y_2 = 9) = 0$$

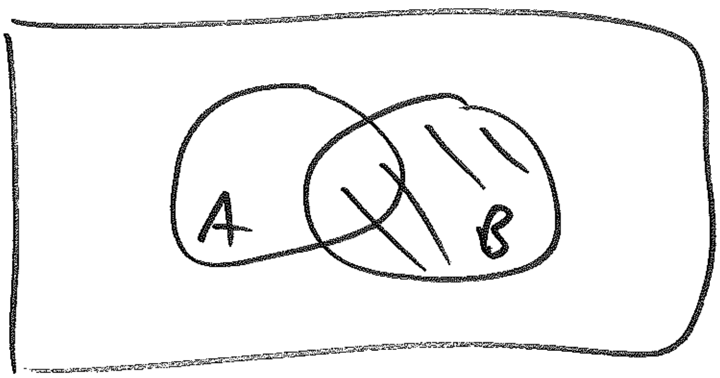
$$P(Y_1 = 9) \cdot P(Y_2 = 9) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9} \neq 0$$

conditional probability

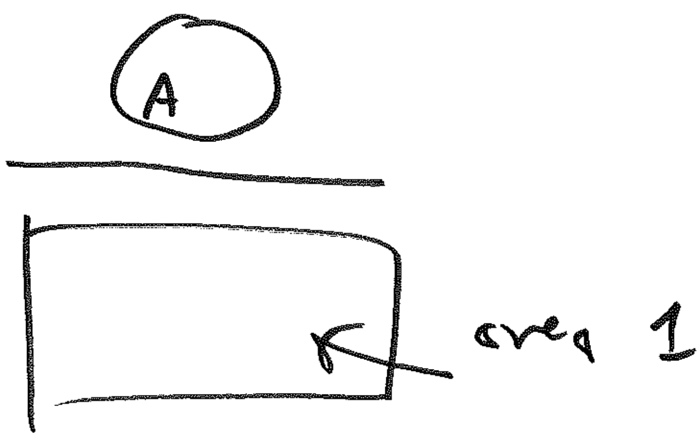
~~Bayes (1740)~~

$P(A \text{ given } B) =$





$P(A) =$



definition (Bayes)

$$P(A \text{ given } B) = P(A \mid B) = \begin{cases} \dfrac{P(A \text{ and } B)}{P(B)} & \text{if } P(B) > 0 \\ \text{undefined} & P(B) = 0 \end{cases}$$

multiply by $P(B)$:

product rule for and

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

def

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

so (mult. by $P(A)$ to get

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

without replacement

$$P(\underbrace{Y_1 = 9}_{A} \text{ and } \underbrace{Y_2 = 9}_{B}) = 0$$

$$P(Y_1 = 9) \cdot P(Y_2 = 9 \mid Y_1 = 9)$$

$$= \frac{1}{3} \cdot 0 = 0 \checkmark$$

---

with replacement

$$P(Y_1 = 9 \text{ and } Y_2 = 9) = \frac{1}{9}$$

$$= P(Y_1 = 9) \cdot P(Y_2 = 9 \mid Y_1 = 9)$$

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

$$= \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9} \checkmark$$

definition A, B are probabilistically independent if information about B doesn't change the chances for A, and vice versa

A, B independent

$\to$ $P(A \mid B) = P(A)$ and

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

product rule when A, B indep.

$$
\begin{aligned}
P(A \text{ and } B) &= P(A) \cdot P(B \mid A) \\
&= P(A) \cdot P(B) \\
&= P(B) \cdot P(A \mid B) \\
&= P(B) \cdot P(A)
\end{aligned}
$$

(at random with replacement) $\longleftrightarrow$ (independent identically distributed) (IID)

outcome

predictor

MLP	gender
Y	M
Y	F
N	F
⋮	⋮

1 row for each UCLA student

n = 106

qual. non. ord. dich.

qual. nominal dich

sort

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

MLP?

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

2x2 contingency table

10.49

$\boxed{q:}$ Are gender & MLP independent in this dataset?

(probability)

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

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

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

if all 3 #s are similar, gender & MLP are (approx.) indep.

choose 1 person at random from this dataset - chance of Y = ?

$76\% \neq 59\% \neq 91\%$ so gender & MLP are (strongly) <u>dependent</u> in this dataset

$P(A) = P(\text{1 or more T-s in 5})$

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

$= 1 - P(\text{exactly 0 T-s})$

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

$\overset{\text{indep}}{=} 1 - P\begin{pmatrix}\text{1st}\\\text{not}\\\text{T-s}\end{pmatrix} \cdot P\begin{pmatrix}\text{2nd}\\\text{not}\\\text{T-s}\end{pmatrix} \cdot \cdots \cdot P\begin{pmatrix}\text{5th}\\\text{not}\\\text{T-s}\end{pmatrix}$

$= 1 - \left[1 - P\begin{pmatrix}\text{1st}\\\text{T-s}\end{pmatrix}\right] \cdot \left[1 - P\begin{pmatrix}\text{2nd}\\\text{T-s}\end{pmatrix}\right] \cdot$

$\cdots \cdot \left[1 - P\begin{pmatrix}\text{5th}\\\text{T-s}\end{pmatrix}\right]$

$\overset{\text{iden. dist.}}{=} 1 - \left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{4}\right) \cdots \left(1 - \frac{1}{4}\right)$

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