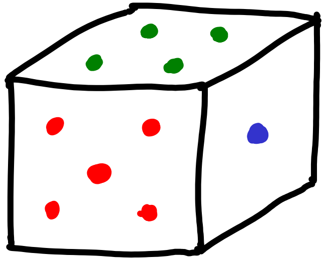
DISCRETE PROBABILITY

Roll a die ... Possible outcomes : $\overbrace{\{1,2,3,4,5,6\}}^{\text{sample space}}$



↳ each has a probability: $\underbrace{\{\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6}\}}_{\text{sum to } 1}$

Roll 2 dice ... sample space → $\{(1,1),(1,2),(1,3),(1,4),(1,5),\underline{(1,6)},$

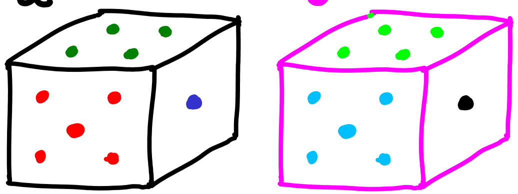
(a,b)

$(2,1),(2,2),\ldots$

$\vdots$

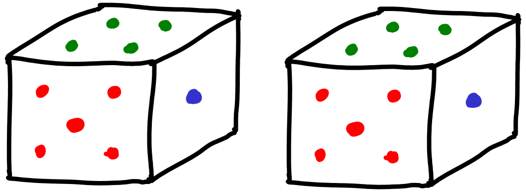
$\underline{(6,1)},(6,2),\ldots \qquad \ldots(6,6)\}$

a b



prob. → $\frac{1}{36}$

Roll 2 indistinguishable dice ...



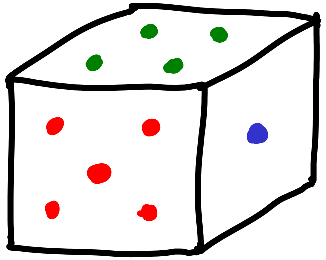
sample space $\rightarrow$ $\{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6),$
$(2,2), (2,3), (2,4), (2,5), (2,6),$
$(3,3), (3,4), (3,5), (3,6),$
$(4,4), (4,5), (4,6),$
$(5,5), (5,6),$
$(6,6)\}$

probability of observing some specific outcome (a,b) ?

- if $a=b$, $\frac{1}{36}$ $\rightarrow$ $\left(6 \cdot \frac{1}{36} = \frac{6}{36}\right)$
- if $a \neq b$, $\frac{2}{36}$ $\rightarrow$ $\left(15 \cdot \frac{2}{36} = \frac{30}{36}\right)$

$+ \;\; = 1$

Roll a die

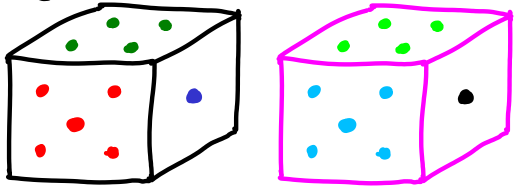


$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$$

$$P(\underbrace{\text{roll even}}_{\text{event}}) = P(\{2,4,6\}) = P(2) + P(4) + P(6) = \frac{1}{2}$$

Roll 2 dice...

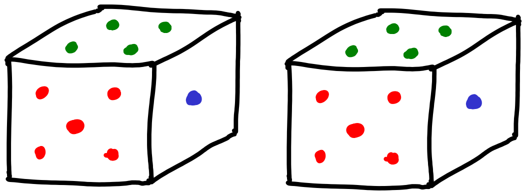
a b



$$P(\overbrace{\text{sum} = 7}^{\text{event}}) = P(\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\})$$

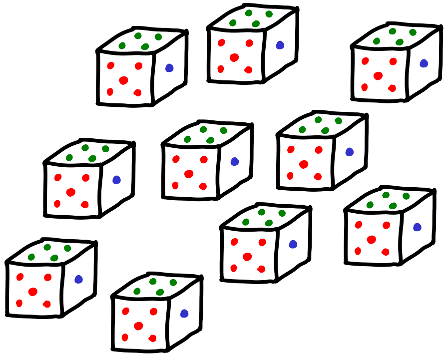
$$= P(1,6) + P(2,5) + P(3,4) + P(4,3) + P(5,2) + P(6,1)$$

$$= 6 \cdot \frac{1}{36} = \frac{1}{6}$$



$$P(\text{sum} = 7) = P(\{(1,6),(2,5),(3,4)\}) = 3 \cdot \frac{2}{36} = \frac{1}{6}$$

Roll 10 dice. (or 1 die 10 times) Probability we observe no 1's ?



- Sample space size: 6^{10} >60 million

including duplicates

- How many outcomes have no 1's ? $\rightarrow 5^{10}$

so we get $\dfrac{5^{10}}{6^{10}} = \left(\dfrac{5}{6}\right)^{10}$ ~16%

OR...

- for each roll, $P(\text{no } 1) = \dfrac{5}{6}$
- each roll/die is independent

$\left(\dfrac{5}{6}\right)^{10}$

(more later)

Poker: 52 cards (13 types; 4 of each) You get 5 cards

- Probability 4 of 5 are the same type? "4 of a kind"

  e.g., 3, 3, 3, 3, 7 OR 8, 8, 8, J, 8

- How many ways can you get 5 cards? $\binom{52}{5}$ sample space

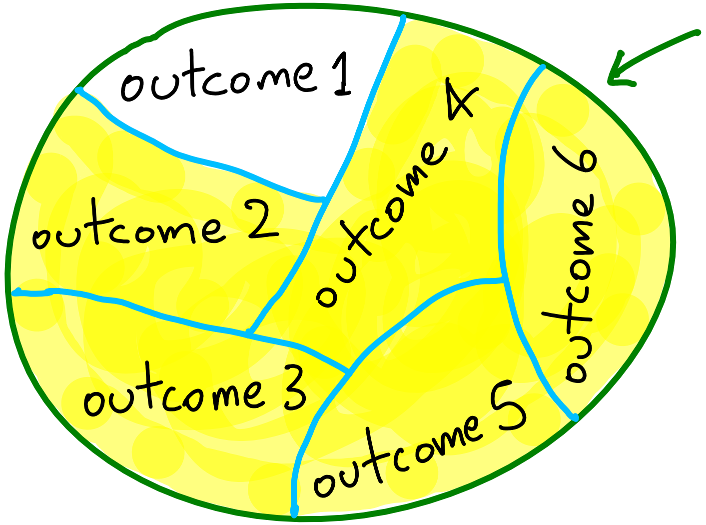
  ↳ order doesn't matter, no repetition

- How many "4 of a kind"? $13 \cdot 48$

  13 types. For each, 48 remaining options as 5th card

↳ $\dfrac{13 \cdot 48}{\binom{52}{5}} = \dfrac{1}{4165} \sim 0.00024 = 0.024\%$

PROBABILITY VISUALIZATION



sample space → area = 1

outcomes partition sample space

e.g., roll one die → all areas equal $\frac{1}{6}$

P(outcome i) = area(i)

P(event) = sum of appropriate areas

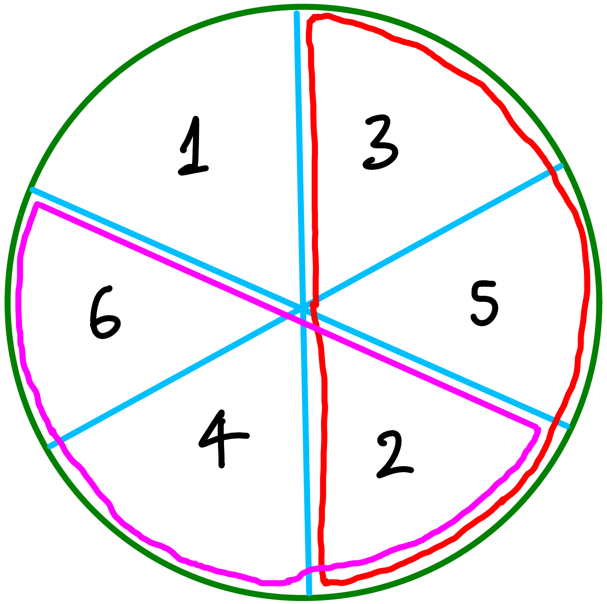
e.g. P(prime OR even) $= \frac{5}{6}$

$\underbrace{2, 3, 5}_{P(\text{prime}) = \frac{3}{6}}$ $\underbrace{2, 4, 6}_{P(\text{even}) = \frac{3}{6}}$

avoid doublecounting

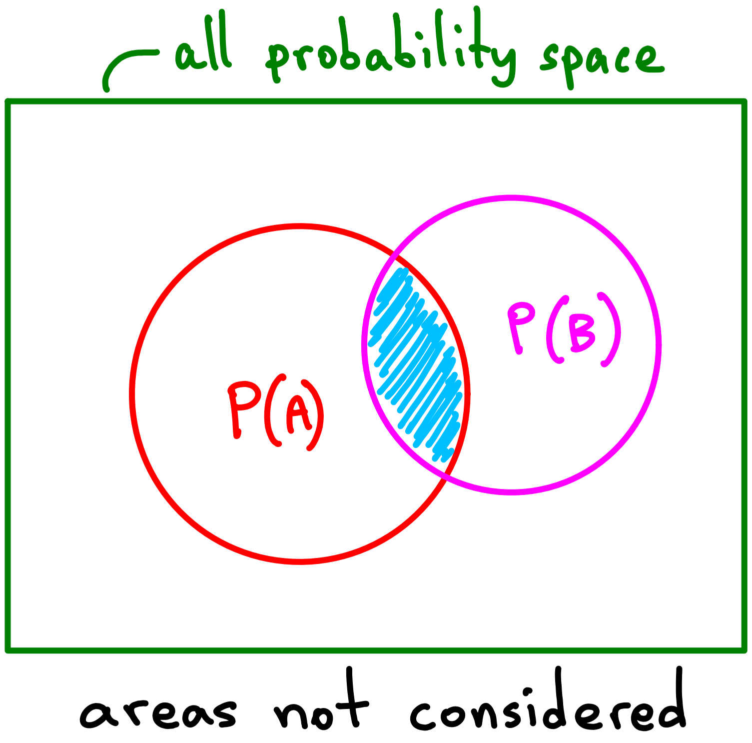
NOT $\frac{3}{6} + \frac{3}{6}$

areas of events can overlap

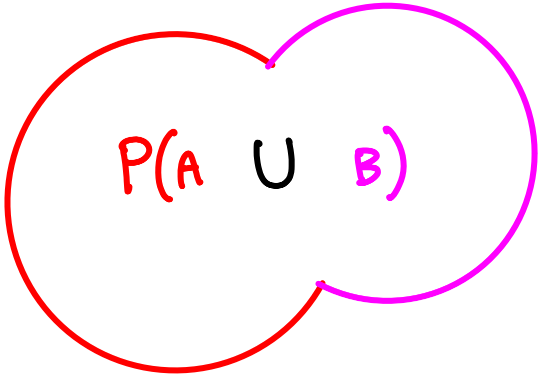


$P(\text{prime}) = P(A)$ $\quad$ $P(\text{even}) = P(B)$

$P(\text{prime OR even}) = P(A \cup B)$

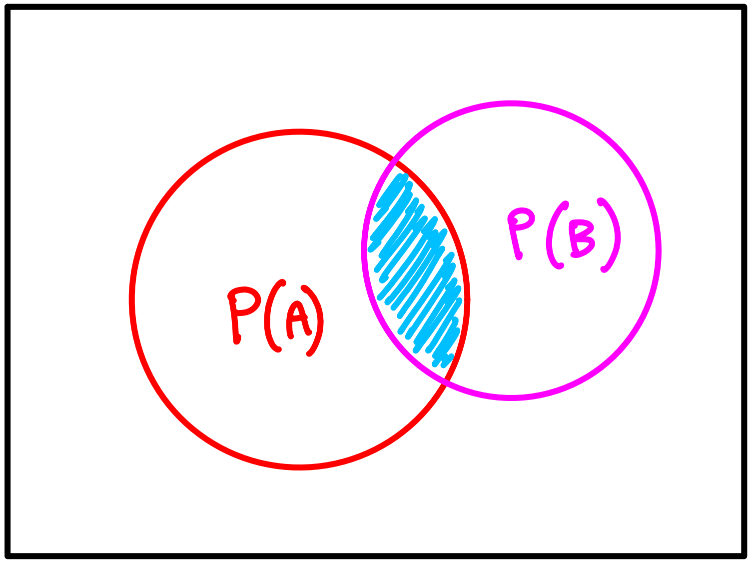


"OR" :



"AND" : $P(A \cap B)$

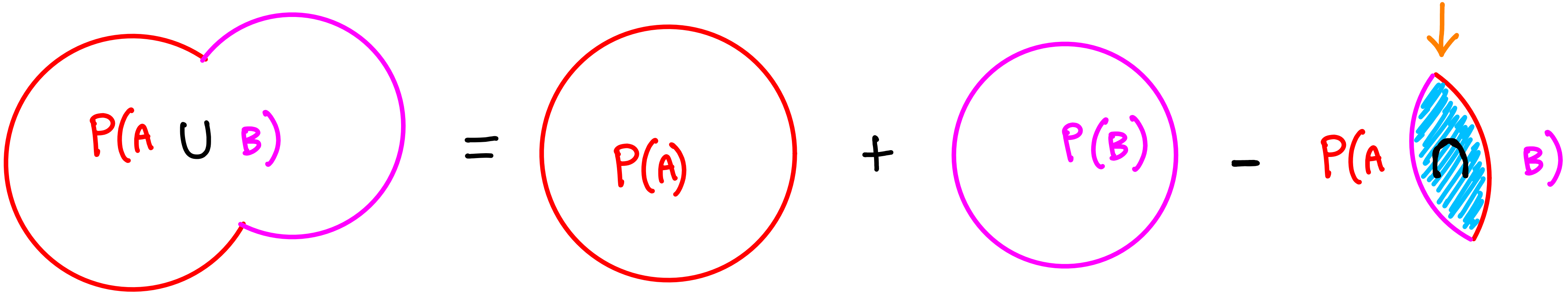
all probability space



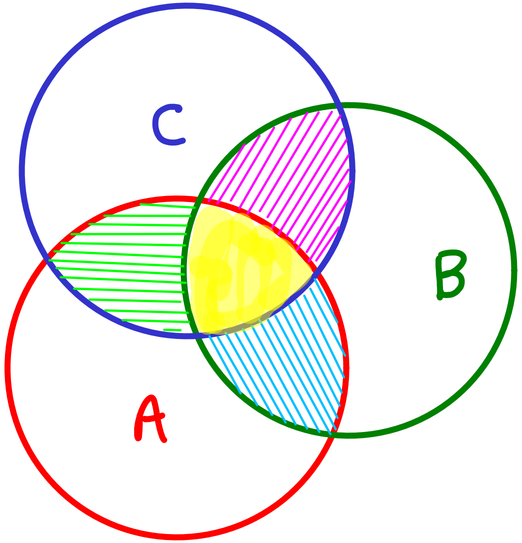
$$P(A) + P(B) = P(A \cup B) + P(A \cap B)$$

INCLUSION-EXCLUSION principle

avoid doublecounting



General INCLUSION-EXCLUSION principle for sets

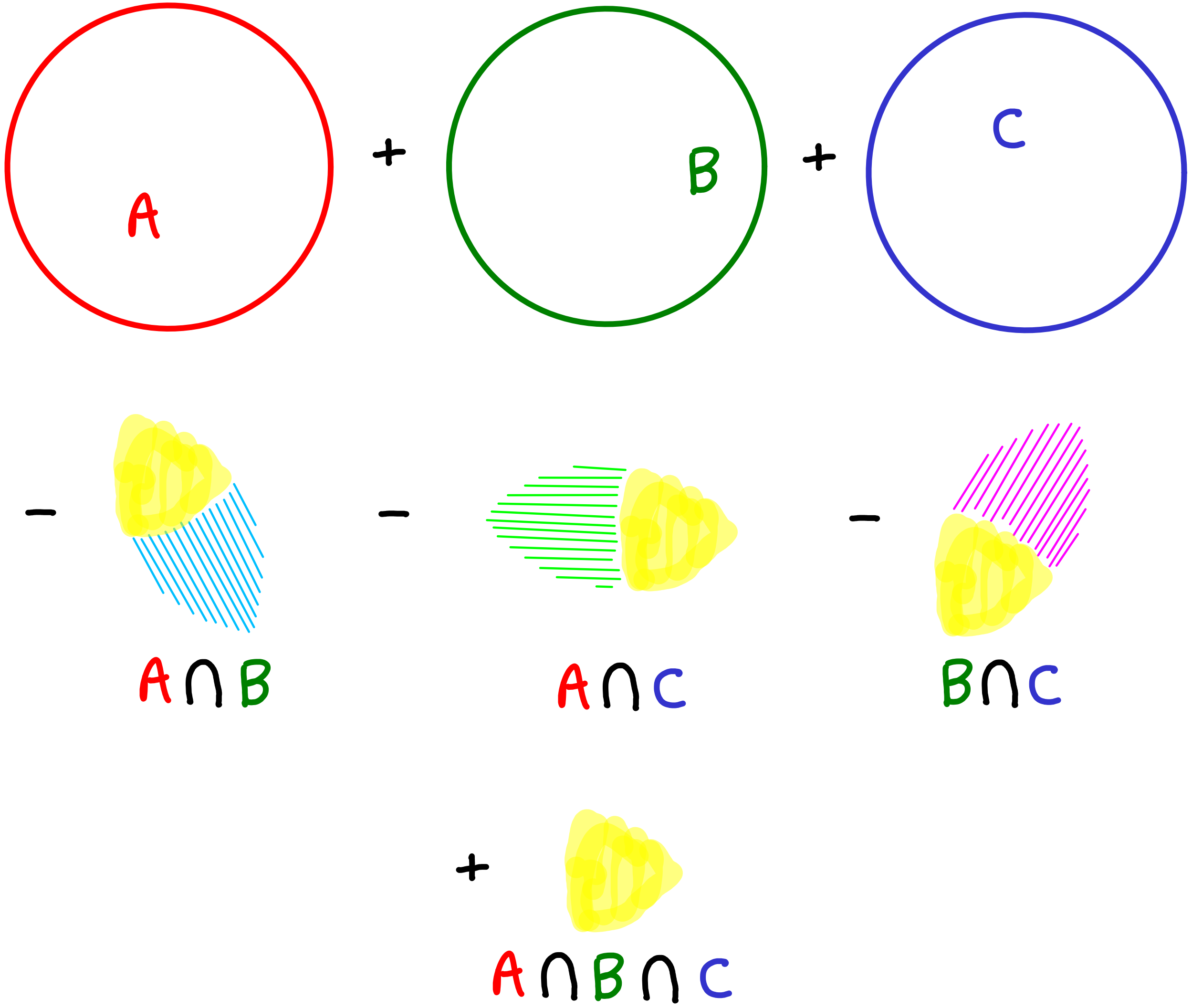


$$|A \cup B \cup C| =$$

$$= |A| + |B| + |C|$$

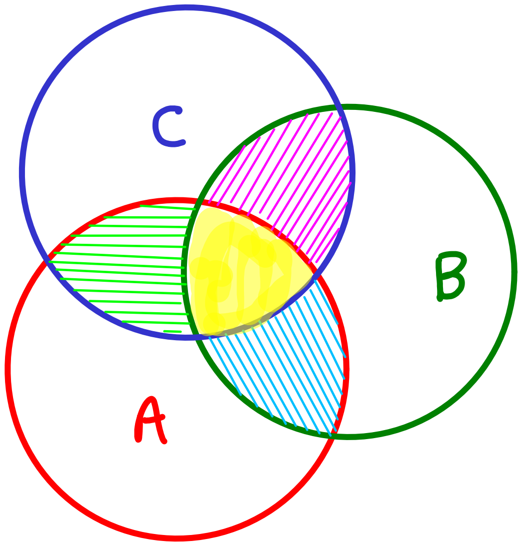
$$- |A \cap B| - |A \cap C| - |B \cap C|$$

$$+ |A \cap B \cap C|$$



General INCLUSION-EXCLUSION principle for sets

E_i : set i



can be proved using Binomial theorem

$|A \cup B \cup C| =$

$= |A| + |B| + |C|$

$- |A \cap B| - |A \cap C| - |B \cap C|$

$+ |A \cap B \cap C|$

$$\left|\bigcup_{i=1}^{n} E_i\right| = + \sum_{i=1}^{n} |E_i|$$

all pairs → $$- \sum_{1 \le i < j \le n} |E_i \cap E_j|$$

alternate sign

all triples → $$+ \sum_{1 \le i < j < k \le n} |E_i \cap E_j \cap E_k|$$

$\mp$

$\pm$

(+1) if n is odd

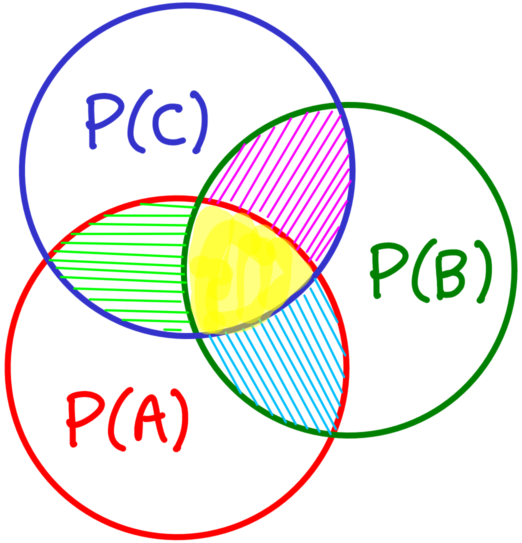
(−1) if n is even

"all" groups of size n

(there is only one group) → $$(-1)^{n-1} |E_i \cap \cdots \cap E_n|$$

General INCLUSION-EXCLUSION principle for probability

E_i : event i



$$P(A \cup B \cup C) =$$

$$= P(A) + P(B) + P(C)$$

$$- P(A \cap B) - P(A \cap C) - P(B \cap C)$$

$$+ P(A \cap B \cap C)$$

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$$

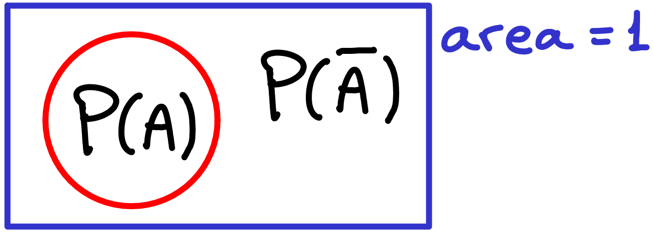
$$- \sum_{1 \le i < j \le n} P(E_i \cap E_j)$$

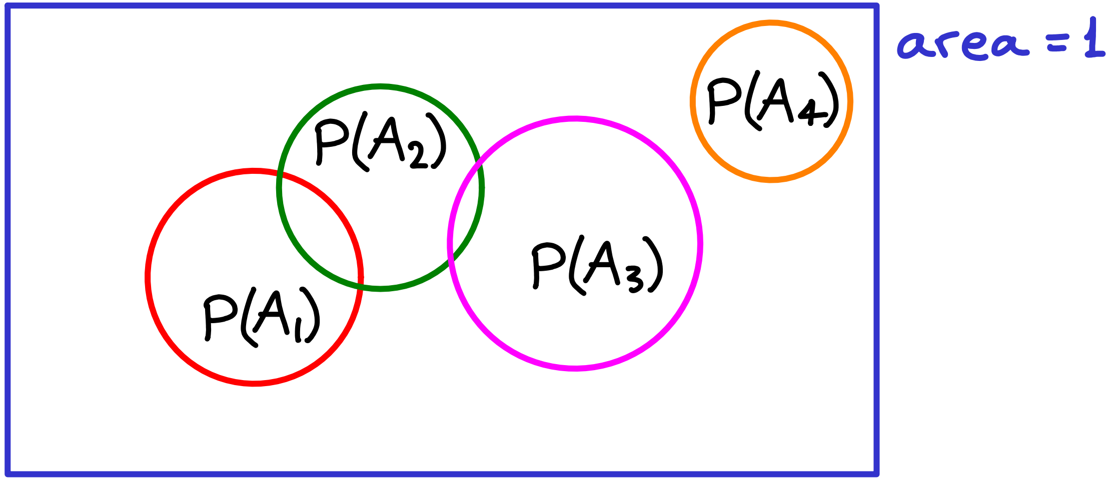
$$+ \sum_{1 \le i < j < k \le n} P(E_i \cap E_j \cap E_k)$$

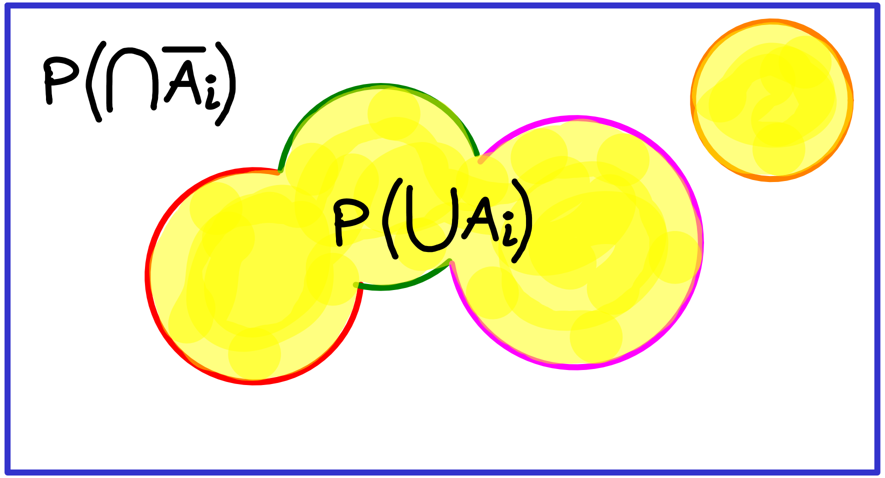
$$\vdots$$

$$(-1)^{n-1} P\left(\bigcap_{i=1}^{n} E_i\right)$$

$$P(A) + P(\bar{A}) = 1$$







$$P\underbrace{\left(\bigcup A_i\right)}_{\text{inside any circle}} = 1 - P\left(\underbrace{\bigcap \bar{A}_i}_{\text{outside every circle}}\right)$$

P(someone in a group of size k was born on April 1)

= P(person1 born on Apr.1 OR person2 born on Apr.1 ...

... OR person k-1 born on Apr.1 OR person k born on Apr.1)

Let e_i : person i born on Apr.1 $\rightarrow$ $P(e_1 \text{ OR } e_2 \text{ OR } e_3 \dots \text{ OR } e_k)$

$$= P\left(\bigcup_{i=1}^{k} e_i\right) = 1 - P\left(\bigcap_{i=1}^{k} \overline{e_i}\right)$$

= 1 − P(person1 NOT born on Apr.1 AND person2 NOT born on Apr.1 ...

...person k-1 NOT born on Apr.1 AND person k NOT born on Apr.1)

= 1 − P(nobody in a group of size k was born on April 1)

$P(\text{someone in a group of } k \text{ was born on April 1}) = P\left(\bigcup_{i=1}^{k} e_i\right)$

$= 1 - P(\text{nobody in a group of } k \text{ was born on April 1}) = 1 - P\left(\bigcap_{i=1}^{k} \overline{e_i}\right)$

minor approximation: assume no Feb. 29

$P(\text{person } i \text{ not born on Apr. 1}) = \frac{364}{365} = P(\overline{e_i})$

(in practice, not all days are equally likely)

$1 - P\left(\bigcap_{i=1}^{k} \overline{e_i}\right) = 1 - \left(\frac{364}{365}\right)^k$ → because of independence

$k = 80$: $\sim 19.7\%$ | $k = 365$: $\sim 63.3\%$

approximation (upper bound) $\Big\}$ $P\left(\bigcup_{i=1}^{k} e_i\right) \leq \sum_{i=1}^{k} P(e_i) = k \frac{1}{365}$

$\sim 21.9\%$ | bad 100%