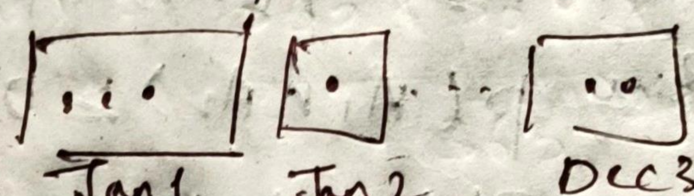
Lecture-3: Birthday problem, ... of probability.

Birthday problem

K people, find probability that 2 have same birthday.

Exclude Feb 29, assume other 365 days equally likely. assume independence of births.

(i) if $K > 365$, probability is 1



(pigeonhole principle)

let $K \leq 365$.

$$P(\text{no match}) = \frac{365 \cdot 364 \cdot 363 \cdots (365-K+1)}{365^K}$$

---

more clear

K people → in a group

feb 29 → with some probability p

Any of the other 365 days, each with probability $\frac{1-p}{365}$

Goal: find the probability that at least two people in the group share the same birthday (including leap day) have feb only check for 365 days.

Key Idea: instead of calculating the prob of shared bday we calculate the compliment.

→ probability that all k people have unique bdays, then subtract it from 1.

1) <u>Total possible combinations</u>

Each person can have a bday on any of 365 days

Total outcomes $= 365^k$

2) <u>Unique birthday</u>:

The first person has 365 choices, the second has 364, the third has 363. etc

3) <u>probability</u> for all <u>birthdays</u> are <u>unique</u>

$$P(\text{unique}) = \frac{365 \times 364 \times \cdots \times (365-k+1)}{365^k}$$

4) prob at least 2 share a bday

$$P(\text{shared}) = 1 - P(\text{unique})$$

Ex1) $K=2$

prob both have unique bdays

$$\frac{365 \times 364}{365^2} = \frac{364}{365} \approx \cancel{0.99}\ \underline{0.99726}$$

Shared birthday

$$1-0.99726 = 0.00274\ (\approx 0.27\%)$$

Ex2) $K=23$

$$P(\text{unique}) = \frac{365 \times 364 \times \cdots \times 343}{365^{23}}$$

$$\approx \underline{0.4927}$$

$$P(\text{Shared}) = 1-0.4927 = \underline{0.5073}$$

with 23 people the prob exceeds with 50%.

General:

$K \le 365$:

$$P(\text{Shared}) = 1 - \frac{365!}{(365-K)! \times 365^K}$$

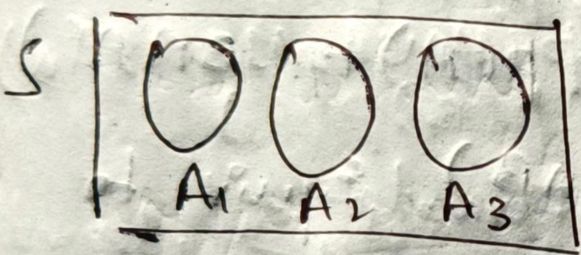
$$p(\text{match}) = \begin{cases} 50.7\% & \text{if } k=23 \\ 97\% & \text{if } k=50 \\ 99.99\% & \text{if } k=100 \end{cases}$$

$$\binom{k}{2} = \frac{k(k-1)}{2}, \quad \binom{23}{2} = \frac{23\cdot 22}{2} = 253$$

1) $P(\phi) = 0$, $P(S) = 1$

2) $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$

if $A_1, A_2 \ldots$ are disjoint events



Properties:

1) $P(A^c) = 1 - P(A)$

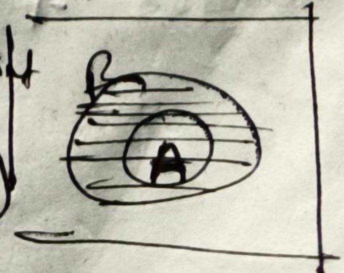
proof

$$1 = P(S) = P(A \cup A^c)$$
$$= P(A) + P(A^c) \text{ since } A \cap A^c = \phi$$

2) If $A \subseteq B$, then $P(A) \leq P(B)$

Proof: $B = A \cup (B \cap A^c)$ disjoint

$P(B) = P(A) + P(B \cap A^c) \geq P(A)$



1) $P(A^c) = 1 - P(A)$

Ex: $S = 1, 2, 3, 4, 5, 6$

Let

$A = 2, 4$

$B = 2, 4, 6$

Then $A^c = 1, 3, 5, 6$

$B \cap A^c = 6$

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

Proof, $P(A \cup B) = P(A \cup (B \cap A^c))$

$= P(A) + P(B \cap A^c)$

$= P(A) + P(B) - P(A \cap B)$, Equiv to

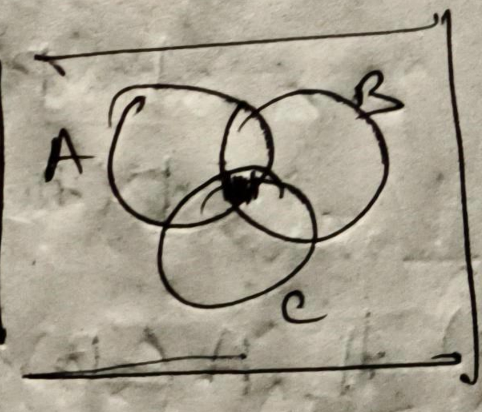
$P(A \cap B) + P(B \cap A^c) = P(B)$,

true since $A \cap B$, $A^c \cap B$ are disjoint union is B

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

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

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



$$p(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{j=1}^{n} P(A_j) - \sum_{i<j} P(A_i \cap A_j)$$

$$+ \sum_{i<j<k} P(A_i A_j A_k)$$

$$\dots + (-1)^{n+1} p(A_1 \cap \dots \cap A_n)$$

de Montmort's problem (1713) matching problem

n cards labelled 1, 2 ... n. Let A_j be the event "j^{th} card matches".

$$P(A_1 \cup A_2 \cup \dots \cup A_n)$$

$$P(A_j) = \frac{1}{n}$$ Since all positions are equally likely for card labelled j

$$P(A_1 \cap A_2) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$

$$P(A_1 \cap \dots \cap A_k) = \frac{(n-k)!}{n!}$$

Key Idea: Counting Permutations with at least one fixed point

→ A fixed point (or match) in a permutation when card is in its original position

we want the probability that at least one card is the right place.

1) Define Events:

A_k → card k is in position k after shuffling

$P(\text{atleast one match}) = P(A_1 \cup A_2 \cup \cdots \cup A_n)$

2) Inclusion - Exclusion Formula

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum P(A_k) - \sum P(A_k \cap A_l) + \sum P(A_k \cap A_l \cap A_m) - \cdots + (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots A_n)$$

Prob that card k is in position k

Shuffling is random, $P(A_k) = \frac{1}{n}$

Sum over all: $\sum P(A_k) = n \cdot \frac{1}{n} = 1$

~~$P(A_1 \cup$~~

General term for j intersection

$$P(A_{k_1} \cap \cdots \cap A_{k_j}) = \frac{(n-j)!}{n!}$$

Sum $\binom{n}{j} \cdot \frac{(n-j)!}{n!} = \frac{1}{j!}$

$$P(A_1 \cup A_2 \cdots \cup A_n) = \cancel{n} \cdot \frac{1}{\cancel{n}} - \frac{n\cancel{(n-1)}}{2!} \frac{1}{\cancel{n(n-1)}} + \frac{n\cancel{(n-1)(n-2)}}{3!} \frac{1}{\cancel{n(n-1)(n-2)}}$$

$$\simeq 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + (-1)^{n} \frac{1}{n!}$$

$$\simeq \boxed{1 - \frac{1}{e}}$$

$$P(\text{nomatch}) = P\left(\bigcap_{j=1}^{n} A_j^c\right) = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n} \frac{1}{n!}$$

$$\longrightarrow \boxed{\simeq \frac{1}{e}}$$