1. Probability 101

- Sample space Ω: set of all possible outcomes of a rand. experiment.
  - flip a coin: $\Omega = \{H, T\}$
  - roll a die: $\Omega = \{1, \cdots, 6\}$,
- Prob. distribution: $p: \Omega \to \mathbb{R}^+$
  - $\forall w \in \Omega$: $\Pr(w) := p(w) \geq 0$
  - $\sum_{w \in \Omega} p(w) = 1$
- Event: $E \subseteq \Omega$
  - $\Pr(E) := \sum_{w \in E} p(w)$
  - $\overline{E} := \Omega \setminus E$: complement event

Ex:
- die: $p(w) = \frac{1}{6}$ $w = 1, \cdots 6$. uniform.
- $E = \{1, 3, 5\}$, $\Pr(E) = p(1) + p(3) + p(5)$
  $F = \{1, 2\}$. $= \frac{1}{2}$

- Union bound:

$$\Pr(E \cup F) \leq \Pr(E) + \Pr(F)$$

- Conditional prob.
  - A, B events $Pr(A) > 0$.
  - $Pr(B|A) = \dfrac{Pr(A \cap B)}{Pr(A)}$

$$\boxed{\begin{array}{l}\text{Baye's Law} \\ A, B \text{ events}, \; Pr(A), Pr(B) > 0. \\ Pr(B|A) = Pr(A|B) \cdot \dfrac{Pr(B)}{Pr(A)}\end{array}}$$

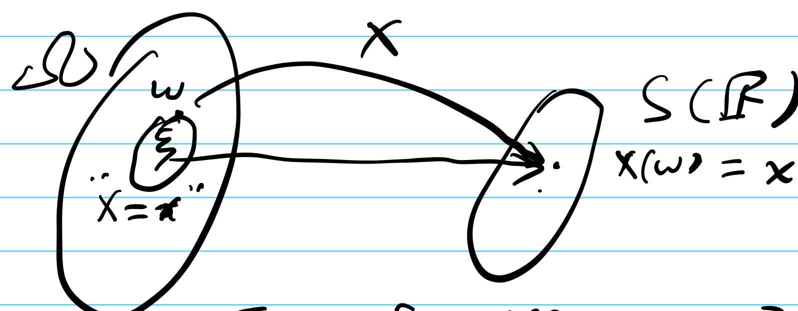
- Independence: A, B indep.

  iff $Pr(B|A) = Pr(B)$

  $\equiv Pr(A \cap B) = Pr(A) \cdot Pr(B)$

- Random variables (r.v.)
  - $X : \Omega \longrightarrow S \; (\mathbb{R})$

    $X(w)$ assign real value to an outcome.



- "$X = x$": event $E := \{w : X(w) = x\}$

- Indep. R.V. : X, Y RVs.

  say X & Y are indep. iff.
  for all possible x & y.
  events: $\underline{X=x}$ & $\underline{Y=y}$ indep.

  $$\forall x,y \quad \Pr(X=x \cap Y=y) = \Pr(X=x)\cdot\Pr(Y=y)$$

- Expectation: (期望) weighted average.

  $$\mathbb{E}[X] := \sum_{x \in S} \Pr(X=x)\cdot x$$

  - linearity of exp. (LoE)

  $$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] \quad *$$

- Ex.: - You choose a card from deck.
  - I pay you $¥ = \begin{cases} 5 & \heartsuit \\ 0 & \text{o.w.} \end{cases}$
  - X: your earning

  $$\Omega = \{\heartsuit, \overline{\heartsuit}\}$$

  $$X : \Omega \to \mathbb{R}$$
  $$\omega \mapsto X(\omega) \in \{0, 5\}$$

  $\mathbb{E}[X] = ?$

X	$\Pr(X=x)$
5	1/4
0	3/4

  $$\mathbb{E}[X] = \sum_x \Pr(X=x)\cdot x = \Pr(X=5)\cdot 5 + \Pr(X=0)\cdot 0 = 5/4 \ \#$$

Ex.: same setup

- play the game 100 rounds.
- Y: total earning.
- $\mathbb{E}[Y] = ?$

→ Soln: let X_i: earning in i^{th} round. $i = 1, \dots, 100$

$$Y = X_1 + X_2 + \cdots + X_{100}.$$

$$\mathbb{E}[X_i] = 5/4$$

$$\mathbb{E}[Y] = \mathbb{E}[X_1 + X_2 + \cdots + X_{100}]$$

$$= \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_{100}]$$

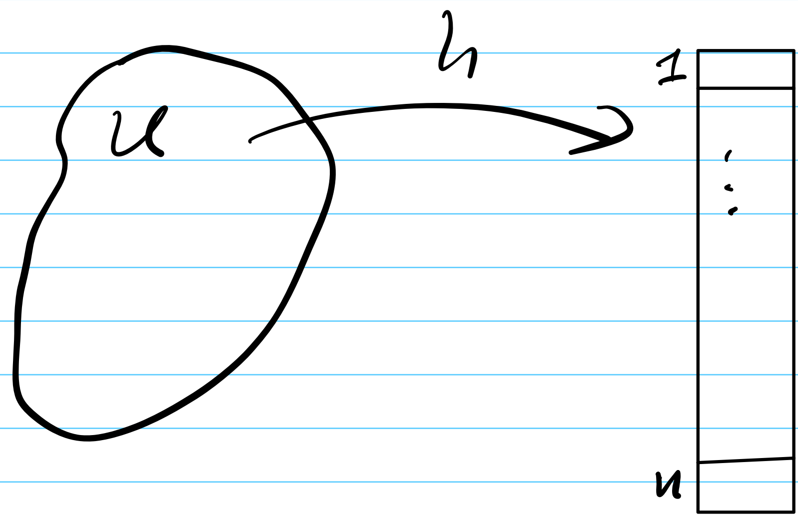
LoE

$$= 5/4 \times 100 = 125$$

#

2. Hash function/ Table analysis.

a.



$S \subseteq U$: obj's we need to manage

<u>Collision</u>: $i, j \in U$, $h(i) = h(j)$

h: $i \mapsto h(i) \leftarrow [n]$ unif. rand.

Fix 1^{st} bucket

X : # (collisions obj's) in 1^{st} bucket.

- $\forall i \in S$, $x_i := \begin{cases} 1 & \text{if } h(i) = 1 \\ 0 & \text{o.w.} \end{cases}$
- $X := \sum_{i=1}^{|S|} x_i$
- $\mathbb{E}[X] := \mathbb{E}[\sum x_i]$

LoE $\to$ $= \sum_i \mathbb{E}[x_i]$

$= \sum_i \frac{1}{n} = \frac{|S|}{n}$.

$\forall i$: $\mathbb{E}[x_i]$

$= Pr(x_i = 1) \cdot 1$

$+ Pr(x_i = 0) \cdot 0$

$= Pr(x_i = 1)$

$= Pr[h(i) = 1]$

$= 1/n$

6

$|S| = c \cdot n$, $\mathbb{E}[X] = \frac{c \cdot n}{n} = \underline{c}$

Same for every bucket.

b. <u>Birthday problem</u>

- q people w/ random BDays.

  each of 365 days of years equally likely

? How large q needs to be before there is at least 50% chance that 2 people have same birthday.

A 23   B 57   C 180   D 366

upper bound

<u>Claim</u>: Let N be size of a universe U.
(e.g. $N = 365$)

Pick $y_1, \dots, y_q \leftarrow U$, indep. unif random.

<u>Define</u>: $\underbrace{Col(q, N)}_{\text{event}} := \exists\, i \neq j$, s.t. $y_i = y_j$

<u>Then</u>: $\Pr[Col(q,N)] = \Theta\left(\frac{q^2}{N}\right)$

$$\frac{q(q-1)}{4N} \le \qquad \le \frac{q(q-1)}{2N}$$

<u>Cor</u>: Plug in $N = 365$.

$$\Pr[Col(q,N)] \ge \boxed{\frac{q(q-1)}{4N} \ge \frac{1}{2}}$$

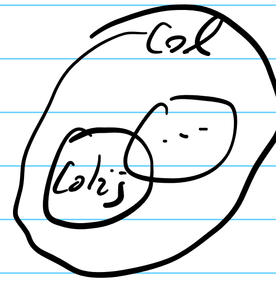
suffices to pick $q \ge 23$.

Pf: (U.B). (L.B. exercise)

Goal: $\Pr[\,Col(q,N)\,] \le \frac{q(q-1)}{2N}$.

$\forall i \ne j$, define Col_{ij} : $y_i = y_j$

Then $\underset{\exists i \ne j\; y_i = y_j}{Col} = \bigcup_{i \ne j} \underset{y_i = y_j}{Col_{ij}}$ (union)



$$\Pr[Col] = \Pr[\bigcup_{i \ne j} Col_{ij}]$$

union bound

$$\le \sum_{i \ne j} \Pr[Col_{ij}]$$

$$\underset{\text{fix } i,j}{\Pr[Col_{ij}]} = \frac{1}{N} \quad \frac{1}{N^2} \quad \forall i,j.$$

$$= \sum_{i \ne j} \frac{1}{N}$$

$$= \binom{q}{2} \cdot \frac{1}{N} = \frac{q(q-1)}{2 \cdot N}$$

$\frac{q^2}{N}$: $q = |S|$

$N = n$ (# buckets).

$q \ll \sqrt{n}$: unlikely to see collison.