- $Pr(z_i$ is compared to $z_j) = Pr($either z_i or z_j are the first pivots chosen in $Z_{ij})$
    1. If no element in Z_{ij} has been chosen, no 2 elements in Z_{ij} have been compared thus all of Z_{ij} are in the same subproblem.
    2. If some element in Z_{ij} is chosen other z_i + z_j, then $z_i + z_j$ will be split into two subarrays.

- $Pr($either z_i or z_j are the first pivots chosen in $Z_{ij}) = Pr(z_i$ is chosen as the first pivot$) + Pr(z_j$ is chosen as the first pivot$)$.

- Z_{ij} has $j-i+1$ elements ($\frac{1}{j-i+1}$ is the prob)

- $Pr(z_i$ is compared to $z_j) = \frac{1}{j-i+1} + \frac{1}{j-i+1}$

$$= \frac{2}{j-i+1}$$

- $\sum_{i=0}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1}$ $\quad k = j-i$

- $\sum_{i=0}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1}$ $\quad$ lower bound

$\sum_{i=0}^{n-1} \boxed{\sum_{k=1}^{n-i} \frac{2}{k}}$ $\quad$ Harmonic Series bound

$$\sum_{i=0}^{n-1} O(\log n)$$

$$\boxed{= O(n \log n)}$$

The Birthday Paradox

- What is the expected # of pairs of individuals that share the same birthday?

- $X_{ij} = \begin{cases} 1 & \text{if persons share the same birthday} \\ 0 & \text{otherwise.} \end{cases}$

Recall

$$E(X_{ij}) = Pr(\text{persons } i + j \text{ having the same birthday})$$

$$= 1/\ell$$

$\left(\frac{365}{365}\right)$ → person i having a unique birthday

$\left(\frac{364}{365}\right)$ → person j would have a unique birthday

$$E(X) = E\left(\sum_{ij} X_{ij}\right)$$

$$= \sum_{ij} E(X_{ij})$$

$$= \sum_{ij} 1/\ell$$

$$= \binom{n}{2} \frac{1}{\ell} \leftarrow \binom{n}{k} \quad \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!}$$

$$= \frac{n(n-1)}{2l}$$

We expect to see a pair of people with the same birthday when $n(n-1) \geq 2l$

$l = 365$

$n = 28$ 1.04 = Expected # of pairs.

$n = 200$ 54.5 = Expected # of pairs.

$n = 55$ 4.1 = Expected # of pairs.

Replace "birthday" w/ key + people with elements we'd see the same result would hold for hash tables in the sparse case.

As a function of $n + l$, what is the probability of finding at least one pair of individuals that share the same birthday?

Pr(at least 1 pair shares a birthday) = 1 − Pr(no pair share the same birthday)

Assume $k-1$ people have a distinct birthday.

$\therefore$ k^{th} person distinct birthday $= 1 - \frac{(k-1)}{l}$

$$\Pr(\text{At least 1 pair shares a birthday}) = \begin{cases} 1 - \prod_{i=1}^{n-1}\left(1 - \frac{i}{\ell}\right) & n \le \ell \\ 1 & n > \ell. \end{cases}$$

Taylor expansion $e^{-x} = 1 - x + O(n^2)$

implies $e^{-x} > 1 - x$ $\qquad x = i/\ell$

$$\Pr(\text{no pair shares a birthday}) = \prod_{i=1}^{n-1}\left(1 - \frac{i}{\ell}\right)$$

$$< \prod_{i=1}^{n-1} e^{-i/\ell}$$

$$= \left(e^{-1/\ell}\right)^{\frac{n(n-1)}{2}}$$

How large does n need to be before the probability of collisons falls below 1/2?

$$\frac{1}{2} > \left(e^{-1/\ell}\right)^{n(n-1)/2}$$

$\ell = 365$

$n = \underline{23}$.