Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación

Ayudante: Nicholas Mc-Donnell
Email: `namcdonnell@uc.cl`

Ayudantía 02

Paradoja de los Cumpleaños

IIC3253 – Criptografía

Fecha: 2021-04-06

Dado una función de `HASH` con output space O ($n = |O|$), queremos encontrar dos cosas:

1) Un $d \in \mathbb{N}$ "pequeño" tal que $\mathbb{P}[X = d] \geq \frac{1}{2}$, donde $\mathbb{P}[X = d]$ es la probabilidad de que haya al menos dos colisiones después de d intentos distintos.

2) Un $d \in \mathbb{N}$ "pequeño" tal que $\mathbb{E}[Y_d] \geq 1$, donde Y_d es la variable aleatoria que cuenta la cantidad de colisiones después de d intentos distintos.

$\mathbb{P}[X = d] \geq \frac{1}{2}$

Para el primero notemos que es más fácil calcular $1 - \mathbb{P}[X = i]$ (la probabilidad de que no hayan colisiones en i intentos):

$$1 - \mathbb{P}[X = i] = \prod_{j=0}^{i-1} \left(1 - \frac{j}{n}\right)$$

Usando esto, y el hecho de que $\mathrm{e}^x \geq 1 + x$, se tienen las siguientes desigualdades

$$\begin{aligned}
1 - \mathbb{P}[X = i] &= \prod_{j=0}^{i-1} \left(1 - \frac{j}{n}\right) \\
&\leq \prod_{j=0}^{i-1} \exp\left(-\frac{j}{n}\right) \\
&\leq \exp\left(\sum_{j=0}^{i-1} -\frac{j}{n}\right) \\
&\leq \exp\left(\frac{-d(d-1)}{2n}\right) \\
&\leq \exp\left(\frac{-(d-1)^2}{2n}\right)
\end{aligned}$$

Por lo que si $\exp\left(\frac{-(d-1)^2}{2n}\right) \leq \frac{1}{2}$ se tiene que $\mathbb{P}[X = d] \geq \frac{1}{2}$. Desarrollemos la primera expresión:

$$\begin{aligned}
\exp\left(\frac{-(d-1)^2}{2n}\right) &\leq \frac{1}{2} \\
\frac{-(d-1)^2}{2n} &\leq \log\frac{1}{2} \\
(d-1)^2 &\geq -2n\log\frac{1}{2} \\
(d-1)^2 &\geq 2n\log 2 \\
d - 1 &\geq \sqrt{2n\log 2} \\
d &\geq \sqrt{2n\log 2} + 1
\end{aligned}$$

Con lo que tomando $d = \left\lceil \sqrt{2n\log 2} + 1 \right\rceil$ se tiene la desigualdad.

$\mathbb{E}[Y_d] \geq 1$

Para calcular el segundo, notemos que $Y_d = \sum_{i=1}^{d}\sum_{j=i+1}^{d} X_{ij}$ donde $\mathbb{P}[X_{ij} = 1] = \frac{1}{n}$ y $\mathbb{P}[X_{ij} = k] = 0$ para $k \neq 1$. Por lo que $\mathbb{E}[Y_d] = \sum_{i=1}^{d}\sum_{j=i+1}^{d} \mathbb{E}[X_{ij}] = \frac{d(d-1)}{2n} \geq \frac{(d-1)^2}{2n}$, por lo que tomando $d = \left\lceil \sqrt{2n} + 1 \right\rceil$ es suficiente.