Suppose you throw $r$ balls into $n$ bins. If each ball is equally likely to land in each bin, how large does $r$ need to be such that a ball lands in every bin with probability at least $1 - 1/n$. We’ll show $r = 2n \ln n$ are sufficient.

- Let $A_i$ be the event that the $i$th bin is empty after $r$ balls are thrown. Then,

$$P[A_i] = (1 - 1/n)^r = (1 - 1/n)^{2n \ln n} \leq e^{-2 \ln n} = 1/n^2$$

- Then $A_1 \cup A_2 \cup \ldots \cup A_n$ is the event that there is an empty bin:

$$P[A_1 \cup A_2 \cup \ldots \cup A_n] \leq P[A_1] + P[A_2] + \ldots + P[A_n] = n \times 1/n^2 = 1/n$$
Problem
You observe a long stream of $T$ values in the range $\{0, 1, \ldots, m - 1\}$,

$$3, 5, 2, 9, 10, 101, 17, \ldots,$$

Because $T$ and $m$ are very large you can’t remember all the values. At the end of the stream, there’s a quiz with a single question:

How many times did the value “x” appear?

How well can you estimate the number of occurrences of $x$? The catch is that you don’t know $x$ in advance.
\textbf{Definition}

Let $U = \{0, 1, \ldots, m - 1\}$ and $V = \{0, 1, \ldots, n - 1\}$ where $m \geq n$. A family of hash functions $\mathcal{H}$ from $U$ to $V$ is said to be $k$-universal if, for any distinct elements $x_1, \ldots, x_k \in U$ and for a hash function $h$ chosen uniformly at random from $\mathcal{H}$, we have

\[
P[\text{ } h(x_1) = h(x_2) = \ldots = h(x_k) \text{ }] \leq \frac{1}{n^{k-1}}.
\]

\textbf{Example}

Fix some prime $p \geq m$ and define

\[ h_{a,b}(x) = ((ax + b) \mod p) \mod n. \]

The family $\mathcal{H} = \{h_{a,b}|1 \leq a \leq p - 1, 0 \leq b \leq p - 1\}$ is 2-Universal.
Count-Min Sketch

Let $H_1, \ldots, H_d : \{0, 1, \ldots, m - 1\} \to \{0, 1, \ldots, w - 1\}$ be $d$ functions chosen randomly from a set of 2-universal hash functions.

During the stream, we maintain a table of $d \times w$ counters where

$$c_{i,j} = \text{number of elements } e \text{ in the stream with } H_i(e) = j$$

Note that for any $x$, $c_{i,H_i(x)} \geq f_x = \text{"frequency of } x\text{"}$ and so,

$$f_x \leq \tilde{f}_x := \min(c_{1,H_1(x)}, \ldots, c_{d,H_d(x)})$$

We’ll show that if $w = 2/\epsilon$ and $d = \log_2 \delta^{-1}$ then,

$$\mathbb{P}\left[\tilde{f}_x \leq f_x + \epsilon T\right] \geq 1 - \delta.$$
Count-Min Sketch Analysis (1/2)

- Define random variables $Z_1, \ldots, Z_k$ such that $c_{i,H_i(x)} = f_x + Z_i$, i.e.,

$$Z_i = \sum_{y \neq x: H_i(y) = H_i(x)} f_y$$

- Define $X_{i,y} = 1$ if $H_i(y) = H_i(x)$ and 0 otherwise. Then,

$$Z_i = \sum_{y \neq x} f_y X_{i,y}$$

- By 2-universality,

$$\mathbb{E}[Z_i] = \sum_{y \neq x} f_y \mathbb{E}[X_{i,y}] = \sum_{y \neq x} f_y \mathbb{P}[H_i(y) = H_i(x)] \leq T/w$$

- By Markov inequality,

$$\mathbb{P}[Z_i \geq \epsilon T] \leq 1/(w \epsilon) = 1/2$$
Count-Min Sketch Analysis (2/2)

- Since each $Z_i$ is independent
  \[ \mathbb{P}[Z_i \geq \epsilon T \text{ for all } 1 \leq i \leq d] \leq (1/2)^d = \delta \]
- Therefore, with probability $1 - \delta$ there exists an $j$ such that
  \[ Z_j \leq \epsilon T \]
- Therefore,
  \[ \tilde{f}_x = \min(c_1, H_1(x), \ldots, c_j, H_j(x), \ldots, c_d, H_d(x)) \]
  \[ = \min(f_x + Z_1, \ldots, f_x + Z_j, \ldots, f_x + Z_d) \leq f_x + \epsilon T \]

**Theorem**

*We can find an estimate $\tilde{f}_x$ for $f_x$ that satisfies,

\[ f_x \leq \tilde{f}_x \leq f_x + \epsilon T \]

*with probability $1 - \delta$ while only using $O(\epsilon^{-1} \log \delta^{-1})$ memory.*
Example
Fix some prime $p \geq m$ and define

$$h_{a,b}(x) = ((ax + b) \mod p) \mod n.$$ 

The family $\mathcal{H} = \{h_{a,b}|1 \leq a \leq p - 1, 0 \leq b \leq p - 1\}$ is 2-Universal.
Proof of 2-Universality

- Fix \( x_1 \neq x_2 \in U \). Since \( |\mathcal{H}| = p(p - 1) \), if we can show that at most \( p(p - 1)/n \) pairs \((a, b)\) that satisfy \( h_{a,b}(x_1) = h_{a,b}(x_2) \) then,

\[
\mathbb{P}[h_{a,b}(x_1) = h_{a,b}(x_2)] \leq \frac{p(p - 1)/n}{p(p - 1)} = \frac{1}{n}.
\]

- Let \( u = (ax_1 + b \mod p) \) and \( v = (ax_2 + b \mod p) \). Then \((u, v)\) with \( u \neq v \) uniquely specifies \( a \) and \( b \):

\[
a = \left( \frac{v - u}{x_2 - x_1} \mod p \right) \quad \text{and} \quad b = (u - ax_1 \mod p).
\]

- Hence, we need to count how many pairs \( 0 \leq u, v \leq p - 1 \) satisfy \( u \neq v \) and \( u = v \mod n \):

\[
p((p - 1)/n + 1 - 1) = p(p - 1)/n.
\]