Estimating Frequency Moments $F_0$ and $F_2$

Anil Maheshwari

anil@scs.carleton.ca
School of Computer Science
Carleton University
Canada
Outline

Frequency Moments

Estimating $F_0$

Algorithm

Correctness

Further Improvements

Estimating $F_2$

Correctness

Improving Variance

Complexity
Frequency Moments
**Definition**

Let $A = (a_1, a_2, \ldots, a_n)$ be a stream, where elements are from universe $U = \{1, \ldots, u\}$. Let $m_i = \#$ of elements in $A$ that are equal to $i$. The $k$-th frequency moment $F_k = \sum_{i=1}^{u} m_i^k$, where $0^0 = 0$.

Example: $A = (3, 2, 4, 7, 2, 2, 3, 2, 1, 4, 2, 2, 2, 1, 1, 2, 3, 2)$ and $U = \{1, \ldots, 7\}$

$m_1 = m_3 = 3, m_2 = 10, m_4 = 2, m_7 = 1, m_5 = m_6 = 0$

\[
F_0 = \sum_{i=1}^{7} m_i^0 \\
= m_1^0 + m_2^0 + m_3^0 + m_4^0 + m_5^0 + m_6^0 + m_7^0 \\
= 3^0 + 10^0 + 3^0 + 2^0 + 0^0 + 0^0 + 1^0 \\
= 5 \text{ (# of distinct elements in } A)\\n\]
Example: \( F_k = \sum_{i=1}^{u} m_i^k \)

\[
A = (3, 2, 4, 7, 2, 2, 3, 2, 2, 1, 4, 2, 2, 1, 1, 2, 3, 2).
\]

\[
m_1 = m_3 = 3, m_2 = 10, m_4 = 2, m_7 = 1, m_5 = m_6 = 0
\]

\[
F_1 = \sum_{i=1}^{7} m_i^1
\]

\[
= 3^1 + 10^1 + 3^1 + 2^1 + 0^1 + 0^1 + 1^1
\]

\[
= 19 \text{ (# of elements in } A)\]

\[
F_2 = \sum_{i=1}^{7} m_i^2
\]

\[
= 3^2 + 10^2 + 3^2 + 2^2 + 0^2 + 0^2 + 1^2
\]

\[
= 123 \text{ (Surprise Number)}\]
## Frequency Moments Streaming Problem

<table>
<thead>
<tr>
<th>Find frequency moments in a stream</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> A stream $A$ consisting of $n$ elements from universe $U = {1, \ldots, u}$.</td>
</tr>
<tr>
<td><strong>Output:</strong> Estimate Frequency Moments $F_k$'s for different values of $k$.</td>
</tr>
</tbody>
</table>

Our Task: Estimate $F_0$ and $F_2$ using sublinear space

Estimating $F_0$
**Estimating $F_0$**

<table>
<thead>
<tr>
<th>Computation of $F_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> Stream $A = (a_1, a_2, \ldots, a_n)$, where each $a_i \in U = {1, \ldots, u}$.</td>
</tr>
<tr>
<td><strong>Output:</strong> An estimate $\hat{F}_0$ of number of distinct elements $F_0$ in $A$ such that $\Pr\left(\frac{1}{c} \leq \frac{\hat{F}_0}{F_0} \leq c\right) \geq 1 - \frac{2}{c}$ for some constant $c$ using sublinear space.</td>
</tr>
</tbody>
</table>

**Example:** $A = (3, 2, 4, 7, 2, 2, 3, 2, 2, 1, 4, 2, 2, 1, 1, 2, 3, 2)$

$m_1 = m_3 = 3, m_2 = 10, m_4 = 2, m_7 = 1, m_5 = m_6 = 0$

\[
F_0 = \sum_{i=1}^{7} m_i^0 = m_1^0 + m_2^0 + m_3^0 + m_4^0 + m_5^0 + m_6^0 + m_7^0 \\
= 3^0 + 10^0 + 3^0 + 2^0 + 0^0 + 0^0 + 1^0 \\
= 5 \text{ (# of distinct elements in } A) \]
Algorithm
Algorithm for Estimating $F_0$

**Input:** Stream $A$ and a hash function $h : U \rightarrow U$

**Output:** Estimate $\hat{F}_0$

**Step 1:** Initialize $R := 0$

**Step 2:** For each elements $a_i \in A$ do:
   1. Compute binary representation of $h(a_i)$
   2. Let $r$ be the location of the rightmost 1 in the binary representation
   3. if $r > R$, $R := r$

**Step 3:** Return $\hat{F}_0 = 2^R$

Space Requirements $= O(\log u)$ bits
An Example

**Step 1:** Initialize $R := 0$

**Step 2:** For each elements $a_i \in A$ do:

1. Compute binary representation of $h(a_i)$
2. Let $r$ be the location of the rightmost 1 in the binary representation
3. if $r > R$, $R := r$

**Step 3:** Return $\hat{F}_0 = 2^R$

Let $A = (3, 2, 4, 7, 2, 2, 3, 2, 2, 1, 4, 2, 2, 2, 1, 1, 2, 3, 2)$ and assume $h(a) = a$.

$h(1) = 0001$
$h(2) = 0010$
$h(3) = 0011$
$h(4) = 0100$
$h(7) = 0111$

Thus, $R = 3$ (given by $h(4)$) and $\hat{F}_0 = 2^R = 2^3 = 8$ is the estimate on the number of distinct items in the stream.
Correctness
Observation 1

Let $d$ to be smallest integer such that $2^d \geq u$ ($d$-bits are sufficient to represent numbers in $U$)

<table>
<thead>
<tr>
<th>Observation 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Pr$ (rightmost 1 in $h(a_i)$ is at location $\geq r + 1$) $= \frac{1}{2^r}$</td>
</tr>
</tbody>
</table>

**Proof:** For that to happen the last $r$ bits in $h(a_i)$ must be 0. Since $h$ is a hash function from universal family of hash functions, this happens with probability $(\frac{1}{2})^r$. 

□
Observation 2

For $a_i \neq a_j$, $Pr(\text{rightmost 1 in } h(a_i) \geq r + 1 \text{ and rightmost 1 in } h(a_j) \geq r + 1) = \frac{1}{2^{2r}}$

Proof: $h(a_i)$ and $h(a_j)$ are independent as $a_i \neq a_j$.

$$Pr(\text{rightmost 1 in } h(a_i) \geq r + 1 \text{ and rightmost 1 in } h(a_j) \geq r + 1) = Pr(\text{rightmost 1 in } h(a_i) \geq r + 1) \times Pr(\text{rightmost 1 in } h(a_j) \geq r + 1) = \frac{1}{2^r} \times \frac{1}{2^r} = \frac{1}{2^{2r}}$$
Observations 3

Fix \( r \in \{1, \ldots, d\} \). \( \forall x \in A \), define indicator r.v:

\[
I^r_x = \begin{cases} 
1, & \text{if the rightmost 1 is at location } \geq r + 1 \text{ in } h(x) \\
0, & \text{otherwise}
\end{cases}
\]

Let \( Z^r = \sum I^r_x \) (sum is over distinct elements of \( A \))

Observation 3
The following holds:

1. \( E[I^r_x] = \frac{1}{2^r} \)
2. \( Var[I^r_x] = \frac{1}{2^r} \left( 1 - \frac{1}{2^r} \right) \)
3. \( E[Z^r] = \frac{F_0}{2^r} \)
4. \( Var[Z^r] \leq E[Z^r] \)
Observation 3.1

\[ E[I_x^r] = \frac{1}{2^r} \]

**Proof:**  
\[ E[I_x^r] = 1 \times Pr(I_x^r = 1) + 0 \times Pr(I_x^r = 0) = \frac{1}{2^r} \]

Note that \( Pr(I_x^r = 1) \) corresponds to \( Pr(\text{rightmost 1 in } h(x) \text{ is at location } \geq r + 1) = \frac{1}{2^r} \) by Observation 1.
Observation 3.2

\[ V ar[I_x^r] = E[I_x^{r^2}] - E[I_x^r]^2 = \frac{1}{2^r} \left( 1 - \frac{1}{2^r} \right) \]

**Proof:** Note that the variance of a random variable \( X \) is given by

\[ V ar[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2. \]

\[ E[I_x^{r^2}] = 1^2 Pr(I_x^r = 1) = \frac{1}{2^r} \]

Then \( E[I_x^{r^2}] - E[I_x^r]^2 = \frac{1}{2^r} - \left( \frac{1}{2^r} \right)^2 = \frac{1}{2^r} \left( 1 - \frac{1}{2^r} \right) \)
Observation 3.3

\[ E[Z^r] = \frac{F_0}{2^r} \]

**Proof:** Let \( A' \subseteq A \) be the set of distinct elements of \( A \).

Note that \( F_0 = |A'| \).

By definition \( Z^r = \sum_{x \in A'} I_x^r \)

Then, \( E[Z^r] = E[\sum_{x \in A'} I_x^r] = \sum_{x \in A'} E[I_x^r] = \sum_{x \in A'} \frac{1}{2^r} = \frac{F_0}{2^r} \)

\( \square \)
Observation 3.4

\[
\text{Var}[Z^r] = \frac{F_0}{2r} \left(1 - \frac{1}{2^r}\right) \leq \frac{F_0}{2r} = E[Z^r]
\]

**Proof:** \( \text{Var}[Z^r] = \text{Var}\left[ \sum_{x \in A'} I_x^r \right] \)

For two independent random variables \( X \) and \( Y \),
\( \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \).

\[
\text{Var}[Z^r] = \text{Var}\left[ \sum_{x \in A'} I_x^r \right] \leq \sum_{x \in A'} \text{Var}[I_x^r] = F_0 \left(1 - \frac{1}{2^r}\right) \leq \frac{F_0}{2r} = E[Z^r]
\]

\( \square \)
When does the algorithm fail?

Step 1: Initialize $R := 0$

Step 2: For each elements $a_i \in A$ do:

1. Compute binary representation of $h(a_i)$
2. Let $r$ be the location of the rightmost 1 in the binary representation
3. if $r > R$, $R := r$

Step 3: Return $\hat{F}_0 = 2^R$

We want to show that $Pr\left(\frac{1}{c} \leq \frac{\hat{F}_0}{F_0} \leq c\right) \geq 1 - \frac{2}{c}$ for some constant $c$.

- Let $F_0 < \frac{2^r}{c}$ and $Z^r > 0 \implies \frac{1}{F_0} \geq \frac{c}{2^r}$ and $\hat{F}_0 > 2^r \implies \frac{\hat{F}_0}{F_0} > c$.
  Observation 4 will show that this can happen with probability at most $1/c$.

- Let $F_0 > c2^r$ and $Z^r = 0 \implies \frac{1}{F_0} \leq \frac{1}{c2^r}$ and $\hat{F}_0 < 2^r \implies \frac{\hat{F}_0}{F_0} < \frac{1}{c}$.
  Observation 5 shows that this can happen with probability at most $1/c$.

Hence, the probability that any of these two events happen is at most $\frac{2}{c}$ by the union bound $\implies Pr\left(\frac{1}{c} \leq \frac{\hat{F}_0}{F_0} \leq c\right) \geq 1 - \frac{2}{c}$. 
Observation 4

If \( F_0 < \frac{2^r}{c} \), \( Pr(Z^r > 0) < \frac{1}{c} \)

**Proof:** Recall Markov’s inequality for a random variable \( X \),
\( Pr(X \geq s) \leq \frac{E[X]}{s} \), where \( s > 0 \) and \( X \) takes positive values.

What is the number of distinct elements \( x \in A \), whose hash map \( h(x) \) has its rightmost 1 in position \( \geq r + 1 \)?

\[
= Z^r = \sum_{x \in A^t} I^r_x
\]

What is \( Pr(Z^r > 0) \)? \( \Leftrightarrow \) What is \( Pr(Z^r \geq 1) \).

By Markov’s inequality: \( Pr(Z^r \geq 1) \leq \frac{E[Z^r]}{1} = E[Z^r] = \frac{F_0}{2^r} < \frac{1}{c} \).

\( \square \)
Chebyshev’s Inequality

\[ Pr(|X - E[X]| \geq \alpha) \leq \frac{Var[X]}{\alpha^2} \]

**Proof:** Recall Markov’s inequality for a random variable \( X \),
\[ Pr(X \geq s) \leq \frac{E[X]}{s} \], where \( s > 0 \) and \( X \) takes positive values.

Now

\[
Pr(|X - E[X]| \geq \alpha) = Pr((X - E[X])^2 \geq \alpha^2) \\
\leq \frac{E[(X - E[X])^2]}{\alpha^2} \\
= \frac{Var[X]}{\alpha^2}
\]
Observation 5

If \( F_0 > c2^r \), \( \Pr(Z^r = 0) < \frac{1}{c} \)

**Proof:** Recall Chebyshev’s inequality \( \Pr(|X - E[X]| \geq \alpha) \leq \frac{Var[X]}{\alpha^2} \).

For a random variable \( X \), \( \Pr(X = 0) \leq \Pr(|X - E[X]| \geq E[X]) \), as the event \( |X - E[X]| \geq E[X] \) includes \( X \leq 0 \) and \( X \geq 2E[X] \).

Now, \( \Pr(Z^r = 0) \leq \Pr(|Z^r - E[Z^r]| \geq E[Z^r]) \).

\[
\begin{align*}
\Pr(Z^r = 0) & \leq \Pr(|Z^r - E[Z^r]| \geq E[Z^r]) \\
& \leq \frac{Var[Z^r]}{E[Z^r]^2} \\
& \leq \frac{E[Z^r]}{E[Z^r]^2} \\
& = \frac{1}{E[Z^r]} = \frac{2^r}{F_0} < \frac{1}{c}
\end{align*}
\]

\( \square \)
**Observation 6**

**Claim**

Set $\hat{F}_0 = 2^R$. We have $Pr \left( \frac{1}{c} \leq \frac{\hat{F}_0}{F_0} \leq c \right) \geq 1 - \frac{2}{c}$

**Proof**

We have that

Observation 4: if $F_0 < \frac{2^r}{c}$, $Pr(Z^r > 0) < \frac{1}{c}$

Observation 5, if $F_0 > c2^r$, $Pr(Z^r = 0) < \frac{1}{c}$

When do we produce the wrong answer?

Case 1: $\hat{F}_0 = 2^R > cF_0$, but this happens with $Pr(Z^R > 0) < \frac{1}{c}$

Case 2: $c2^R = c\hat{F}_0 < F_0$, but this happens with $Pr(Z^R = 0) < \frac{1}{c}$

Therefore, with probability $\leq \frac{2}{c}$, we produce a wrong answer.

$\implies$ with probability $\geq 1 - \frac{2}{c}$, we produce the right answer, i.e.,

$Pr \left( \frac{1}{c} \leq \frac{\hat{F}_0}{F_0} \leq c \right) \geq 1 - \frac{2}{c}$
Further Improvements
Improving success probability

Main Idea
Execute the algorithm $s$ times in parallel (with independent hash functions)
Return $\hat{F}_0 = 2^R$, where $R$ is the median value among these runs.

$i$-th Run of the Algorithm:

- **Step 1:** Initialize $R_i := 0$
- **Step 2:** For each elements $a_i \in A$ do:
  1. Compute binary representation of $h(a_i)$
  2. Let $r$ be the location of the rightmost 1 in the binary representation
  3. if $r > R_i$, $R_i := r$
- **Step 3:** Return $R_i$

Set $R = \text{Median}(R_1, R_2, \ldots, R_s)$
**Analysis**

**Space Bound**

The algorithm uses $O(s \log u)$ bits.

**Claim**

For $c > 4$, there exists $s = O(\log \frac{1}{\epsilon})$, $\epsilon > 0$, such that

$$\Pr\left(\frac{1}{c} \leq \frac{\hat{F}_0}{F_0} \leq c\right) \geq 1 - \epsilon.$$

**Technique:** Median + Chernoff Bounds
Setting Up Indicator Random Variables

\textbf{i}-th Run of the Algorithm:

- **Step 1:** Initialize $R_i := 0$
- **Step 2:** For each elements $a_i \in A$ do:
  1. Compute binary representation of $h(a_i)$
  2. Let $r$ be the location of the rightmost 1 in the binary representation
  3. if $r > R_i$, $R_i := r$
- **Step 3:** Return $R_i$

Define $X_1, \ldots, X_s$ be indicator random variables:

$$X_i = \begin{cases} 
0, & \text{if success, i.e. } \frac{1}{c} \leq \frac{2^{R_i}}{F_0} \leq c \\
1, & \text{otherwise}
\end{cases}$$

1. $E[X_i] = Pr(X_i = 1) \leq \frac{2}{c} = \beta < \frac{1}{2}$ (Since $c > 4$)
2. Let $X = \sum_{i=1}^{s} X_i$ = Number of failures in $s$ runs
3. $E[X] \leq s\beta < \frac{s}{2}$
4. If $X < \frac{s}{2}$, then $\frac{1}{c} \leq \frac{2^R}{F_0} \leq c$, where $R = \text{Median}(R_1, R_2, \ldots, R_s)$. 

23
Chernoff Bounds

If r.v. $X$ is sum of independent identical indicator r.v. and $0 < \delta < 1$,  

$$Pr(X \geq (1 + \delta)E[X]) \leq e^{-\frac{\delta^2 E[X]}{3}}$$

Proof: See my notes

An example: Toss a fair coin $n$-times. Let $X$ be the total number of heads obtained in these $n$-trials. Evaluate $Pr(X \geq \frac{3}{4}n)$

\[
Pr(X \geq \frac{3}{4}n) = Pr(X \geq (1 + \frac{1}{2})\frac{n}{2}) \\
= Pr(X \geq (1 + \frac{1}{2})E[X]) \\
\leq e^{-\frac{(\frac{1}{2})^2 E[X]}{3}} \\
= e^{-\frac{n}{12}}
\]
Main Result

Claim
For any $\epsilon > 0$, if $s = O(\log \frac{1}{\epsilon})$, $Pr(X < \frac{s}{2}) \geq 1 - \epsilon$

Proof: We will consider the complementary event and show that $Pr(X \geq \frac{s}{2}) < \epsilon$. We know that $E[X] = s\beta < \frac{s}{2}$.

$$Pr(X \geq \frac{s}{2}) = Pr(X - E[X] \geq \frac{s}{2} - E[X])$$
$$= Pr(X - E[X] \geq \frac{s}{2} - s\beta) \text{ as } E[X] < s\beta$$
$$= Pr(X - E[X] \geq \frac{1}{2} - \frac{\beta}{s\beta} s\beta)$$
$$= Pr(X - E[X] \geq \frac{1}{2} - \frac{\beta}{E[X]})$$
$$= Pr(X \geq \left(1 + \frac{1}{2} - \frac{\beta}{E[X]}\right) E[X])$$
Proof (contd.)

\[ Pr(X \geq \frac{s}{2}) = Pr(X \geq \left(1 + \frac{1}{\beta} \right) E[X]) \leq e^{-\frac{1}{3} \left( \frac{1-\beta}{\beta} \right)^2 E[X]} \]

We want \( e^{-\frac{1}{3} \left( \frac{1-\beta}{\beta} \right)^2 E[X]} \leq \epsilon \)

Substitute \( E[X] = s\beta \) and we have \(-\frac{1}{3} \left( \frac{1-\beta}{\beta} \right)^2 s\beta \leq \ln \epsilon\)

\[ \Leftrightarrow s \geq \frac{3}{\beta} \left( \frac{\beta}{\frac{1}{2}-\beta} \right)^2 \ln \frac{1}{\epsilon} \]

\[ \implies \text{if } s \in O(\ln \frac{1}{\epsilon}), \ Pr(X \geq \frac{s}{2}) < \epsilon. \]
Estimating $F_2$
Estimating \( F_2 \)

**Input:** Stream \( A \) with elements from the universe \( U \).

**Output:** Estimate \( \hat{F}_2 \) of \( F_2 = \sum_{i=1}^{u} m_i^2 \)

\[
A = (3, 2, 4, 7, 2, 2, 3, 2, 2, 1, 4, 2, 2, 2, 1, 1, 2, 3, 2).
\]
\[
m_1 = m_3 = 3, \quad m_2 = 10, \quad m_4 = 2, \quad m_7 = 1, \quad m_5 = m_6 = 0
\]

\[
F_2 = \sum_{i=1}^{7} m_i^2
\]
\[
= 3^2 + 10^2 + 3^2 + 2^2 + 0^2 + 0^2 + 1^2
\]
\[
= 123 \text{ (Surprise Number)}
\]
**Tug of War Algorithm**

**Input:** Stream $A$ and hash function $h : U \rightarrow \{-1, +1\}$

**Output:** Estimate $\hat{F}_2$ of $F_2 = \sum_{i=1}^{u} m_i^2$

**Algorithm (Tug of War)**

**Step 1:** Initialize $Y := 0$.

**Step 2:** For each element $x \in U$, evaluate $r_x = h(x)$.

**Step 3:** For each element $a_i \in A$, $Y := Y + r_{a_i}$

**Step 4:** Return $\hat{F}_2 = Y^2$
Step 1: Initialize $Y := 0$.

Step 2: For each element $x \in U$, evaluate $r_x = h(x)$.

Step 3: For each element $a_i \in A$, $Y := Y + r_{a_i}$

Step 4: Return $\hat{F}_2 = Y^2$

Let $A = (3, 2, 4, 7, 2, 2, 3, 2, 2, 1, 4, 2, 2, 2, 1, 1, 2, 3, 2)$,

$m_1 = m_3 = 3, m_2 = 10, m_4 = 2, m_7 = 1, m_5 = m_6 = 0$, and let

$r_1 = r_3 = r_5 = r_6 = r_7 = +1$ and $r_2 = r_4 = -1$.

\[
Y = \sum_{i=1}^{7} r_i m_i \\
= 3 - 10 + 3 - 2 + 0 + 0 + 1 \\
= -5
\]

Thus $Y^2 = 25$ will be the approximate value of $\hat{F}_2$ returned by the algorithm.
Correctness
Observation 1

$E[r_i] = 0$ and $E[r_i^2] = 1$.

Proof:

\[
E[r_i] = -1 \times P(r_i = -1) + 1 \times P(r_i = +1) \\
= -1 \times \frac{1}{2} + 1 \times \frac{1}{2} \\
= 0
\]

\[
E[r_i^2] = (-1)^2 \times P(r_i = -1) + (1)^2 \times P(r_i = +1) \\
= 1 \times \frac{1}{2} + 1 \times \frac{1}{2} \\
= 1
\]
Observation 2

\[ E[Y^2] = \sum_{i=1}^{u} m_i^2 = F_2, \text{ where } Y = \sum_{i=1}^{u} r_i m_i. \]

Proof:

\[
E[Y^2] = E \left[ \sum_{i=1}^{u} r_i m_i \sum_{j=1}^{u} r_j m_j \right] = E \left[ \sum_{i=1}^{u} r_i^2 m_i^2 + \sum_{i,j:i \neq j} r_i r_j m_i m_j \right]
\]

\[
= \sum_{i=1}^{u} E[r_i^2 m_i^2] + \sum_{i,j:i \neq j} E[r_i r_j m_i m_j]
\]

\[
= \sum_{i=1}^{u} m_i^2 E[r_i^2] + \sum_{i,j:i \neq j} m_i m_j E[r_i r_j]
\]

\[
= \sum_{i=1}^{u} m_i^2 E[r_i^2] + \sum_{i,j:i \neq j} m_i m_j E[r_i] E[r_j]
\]

\[
= \sum_{i=1}^{u} m_i^2 = F_2 \quad \square
\]
**Observation 3**

$$Pr \left( |Y^2 - E[Y^2]| \geq \sqrt{2}cE[Y^2] \right) \leq \frac{1}{c^2} \text{ for any positive constant } c. \text{ (i.e., } Y^2 \text{ approximates } F_2 = E[Y^2] \text{ within a constant factor with } Pr \geq 1 - \frac{1}{c^2} \text{)}$$

**Proof:** Recall Chebyshev’s inequality $$Pr(|X - E[X]| \geq \alpha) \leq \frac{Var[X]}{\alpha^2}.$$  

Now, $$Pr \left( |Y^2 - E[Y^2]| \geq \sqrt{2}cE[Y^2] \right) \leq \frac{Var[Y^2]}{(\sqrt{2}cE[Y^2])^2}.$$


\[
E[Y^4] = E \left[ \sum_{i=1}^{u} r_im_i \sum_{j=1}^{u} r_jm_j \sum_{k=1}^{u} r_km_k \sum_{l=1}^{u} r_lm_l \right] \\
= \sum_{i=1}^{u} E[r_i^4m_i^4] + 6 \sum_{1 \leq i < j \leq u} E[r_i^2r_j^2m_i^2m_j^2] \\
= \sum_{i=1}^{u} m_i^4 + 6 \sum_{1 \leq i < j \leq u} m_i^2m_j^2
\]
\[ = \sum_{i=1}^{u} m_i^4 + 6 \sum_{1 \leq i < j \leq u} m_i^2 m_j^2 - \left( \sum_{i=1}^{u} m_i^2 \right)^2 \]
\[ = 4 \sum_{1 \leq i < j \leq u} m_i^2 m_j^2 \]
\[ \leq 2F_2^2 \]

Now,
\[ \frac{Var[Y^2]}{(\sqrt{2c}E[Y^2])^2} = \frac{2F_2^2}{(\sqrt{2c}E[Y^2])^2} = \frac{2F_2^2}{2c^2F_2^2} = \frac{1}{c^2} \]

Thus,
\[ Pr \left( |Y^2 - E[Y^2]| \geq \sqrt{2c}E[Y^2] \right) \leq \frac{Var[Y^2]}{(\sqrt{2c}E[Y^2])^2} = \frac{1}{c^2} \]

\[ \square \]
Improving Variance
Improving the Variance

Execute the algorithm $k$ times (using independent hash functions) resulting in $Y_1^2, Y_2^2, \ldots, Y_k^2$.

Output $\bar{Y}^2 = \frac{1}{k} \sum_{i=1}^{k} Y_i^2$

Observations:

2. $Var[\bar{Y}^2] = \frac{1}{k} Var[Y^2]$
   (Note: $Var[cX] = c^2 Var[X]$)
3. $Pr \left( |\bar{Y}^2 - E[\bar{Y}^2]| \geq \sqrt{\frac{2}{k} cE[\bar{Y}^2]} \right) \leq \frac{1}{c^2}$
4. Set $k = O\left(\frac{1}{\epsilon^2}\right)$, we have
   $Pr \left( |\bar{Y}^2 - E[\bar{Y}^2]| \geq \epsilon cE[\bar{Y}^2] \right) \leq \frac{1}{c^2}$
Complexity
## Space Complexity

### Algorithm (Tug of War)

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>Initialize $Y := 0$.</td>
</tr>
<tr>
<td>Step 2</td>
<td>For each element $x \in U$, evaluate $r_x = h(x)$.</td>
</tr>
<tr>
<td>Step 3</td>
<td>For each element $a_i \in A$, $Y := Y + r_{a_i}$</td>
</tr>
<tr>
<td>Step 4</td>
<td>Return $\hat{F}_2 = Y^2$</td>
</tr>
</tbody>
</table>

- Need to store $Y$ and $(r_1, r_2, \ldots, r_u)$.
  - $Y$ requires $O(\log n)$ bits.
- We needed $r_i$’s to be 2-wise and 4-wise independent hash functions.
- 4-wise independent functions can be maintained using $O(\log u)$ bits.
- Total space required is $O(\log n + \log u)$. 

---

35


3. Notes on Algorithm Design by A.M

4. Several Lecture Notes (Tim Roughgarden, Ankush Moitra, Lap Chi Lau, Yufei Tao, John Augustine,...)