Probability & Computing

Concentration
Expectation Management

What does it mean?

“QuickSort has an *expected* running time of $O(n \log(n))$.”
Expectation Management

What does it mean?

- “QuickSort has an \textit{expected} running time of $O(n \log(n))$.”
- “The vertex has an \textit{expected} degree of $c$.\textquotedblright"
Expectation Management

What does it mean?

- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?

I “expect” the sniper to hit the target...
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?

![Graph showing probability distribution of hairs in soup with expectation at 1]

Every soup contains 1 hair
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?
Expectation Management

What does it mean?
- “QuickSort has an expected running time of \(O(n \log(n))\).”
- “The vertex has an expected degree of \(c\).”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?

Half of the soups 2 hairs, the rest none

Hairs in Soup

Probability

Expectation

0 1 2 3 4 5 6 7 8 9 10
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$."
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?

![Graph showing expected value](image)

Knowing that the expected value is 1 hair:

How likely is it that I get at least 10?
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?

"QuickSort has an expected running time of $O(n \log(n))$."
"The vertex has an expected degree of $c$."
"In expectation there is one hair in my soup."

<table>
<thead>
<tr>
<th>Hairs in Soup</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Knowing that the expected value is 1 hair:
- How likely is it that I get at least 10?
  - Not at all

Every soup contains 1 hair
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?

<table>
<thead>
<tr>
<th>Hairs in Soup</th>
<th>Probability</th>
<th>Expectation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Knowing that the expected value is 1 hair:
How likely is it that I get at least 10?

- Not at all
- Somewhat
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?

### Hairs in Soup

<table>
<thead>
<tr>
<th>Probability</th>
<th>Expectation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Some soups 10 hairs, most have none</td>
</tr>
</tbody>
</table>

Knowing that the expected value is 1 hair:

- How likely is it that I get at least 10?
  - Not at all
  - Somewhat
- How likely is it that I get less than 2?

```
$\text{Probability}$

### 0 1 2 3 4 5 6 7 8 9 10

$\text{Hairs in Soup}$

- $\text{How useful is that information in practice?}$
- $\text{In expectation there is one hair in my soup.}$
- $\text{QuickSort has an expected running time of } O(n \log(n)).$
- $\text{The vertex has an expected degree of } c.$
- $\text{In expectation there is one hair in my soup.}$
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?

Knowing that the expected value is 1 hair:
- How likely is it that I get at least 10? Somewhat
- How likely is it that I get less than 2? Extremely

Every soup contains 1 hair

Probability

Expectation

Hairs in Soup

0 1 2 3 4 5 6 7 8 9 10
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?

In expectation there is one hair in my soup.

How likely is it that I get at least 10?
Not at all  Somewhat

How likely is it that I get less than 2?
Extremely  Somewhat
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?
- Does not tell us much about the shape of the distribution

Knowing that the expected value is 1 hair:
- How likely is it that I get at least 10? Not at all
- How likely is it that I get less than 2? Extremely

Graph showing distribution of hairs in soup with expected value of 1 hair.
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?
- Does not tell us much about the shape of the distribution
- Does not come with a level of certainty

Knowing that the expected value is 1 hair:
- How likely is it that I get at least 10?
  - Not at all
  - Somewhat
- How likely is it that I get less than 2?
  - Extremely
  - Somewhat
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$."
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?
- Does not tell us much about the shape of the distribution
- Does not come with a level of certainty

Concentration
- In practice, expectation is often a good start
Expectation Management

What does it mean?
- “QuickSort has an expected running time of \( O(n \log(n)) \).”
- “The vertex has an expected degree of \( c \).”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?
- Does not tell us much about the shape of the distribution
- Does not come with a level of certainty

Concentration
- In practice, expectation is often a good start
- But for meaningful statements, we need to know how likely we are close to the expectation
Expectation Management

What does it mean?
- “QuickSort has an expected running time of $O(n \log(n))$.”
- “The vertex has an expected degree of $c$.”
- “In expectation there is one hair in my soup.”

Expectation
- The average of infinitely many trials
- How useful is that information in practice?
- Does not tell us much about the shape of the distribution
- Does not come with a level of certainty

Concentration
- In practice, expectation is often a good start
- But for meaningful statements, we need to know how likely we are close to the expectation

---

**Definition**: A concentration inequality bounds the probability of a random variable to deviate from a given value (typically its expectation) by a certain amount.
Markov’s Inequality

About Markov
- Andrei “The Furious” Andreyevich Markov (Russian mathematician)
- Unhappy with the state of living at the time (1921)
- Informed St. Petersburg Academy of Sciences that he could not attend meetings due to not having shoes
- After getting shoes from the Communist Party he replied:

  Finally, I received footwear. However, it is stupidly stitched together and does not accord with my measurements. Thus I cannot attend the meetings. I propose placing the footwear in a museum, as an example of the material culture of the current time.

“Shape, The hidden geometry of absolutely everything”, Jordan Ellenberg
Markov’s Inequality

Theorem (Markov’s inequality): Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$. 
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$. 

**Visual Proof**
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

**Visual Proof**

\[
\begin{align*}
\Pr[X = 1] &> 1/2 \\
\Pr[X = 2] &> 1/2 \\
\Pr[X = 3] &> 0 \\
\Pr[X = 4] &> 0
\end{align*}
\]
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**Visual Proof**
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**Visual Proof**

![Visual Proof Diagram]

$x \cdot \Pr[X = x]$
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**Visual Proof**

\[
x \cdot \Pr[X = x]
\]
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

**Visual Proof**

![Visual Proof Diagram]
Markov’s Inequality

Theorem (Markov’s inequality): Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

Visual Proof

$$\sum_x x \cdot \Pr[X = x]$$
Markov’s Inequality

Theorem (Markov’s inequality): Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$. 

Visual Proof

\[
\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x]
\]
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**Visual Proof**

\[
\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x] \quad \text{fits into} \quad 1 \cdot \Pr[X \geq 1]
\]
Markov’s Inequality

Theorem (Markov’s inequality): Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

Visual Proof:

\[
\mathbb{E}[X] = \sum_{x} x \cdot \Pr[X = x]
\]

fits into

$2 \cdot \Pr[X \geq 2]$
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{E[X]}{a}$.

**Visual Proof**

![Visual proof of Markov's inequality](image-url)
**Markov’s Inequality**

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**Visual Proof**

$$\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x] \quad \text{fits into} \quad 4 \cdot \Pr[X \geq 4]$$
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

**Visual Proof**

![Visual Proof Diagram]

$\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x] \geq a \cdot \Pr[X \geq a]$
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

**Proof**

\[
\mathbb{E}[X] = \mathbb{E}[X | X < a] \cdot \Pr[X < a] + \mathbb{E}[X | X \geq a] \cdot \Pr[X \geq a]
\]

**Law of Total Expectation**
Markov’s Inequality

Theorem (Markov’s inequality): Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

Visual Proof

Proof $\mathbb{E}[X] = \mathbb{E}[X \mid X < a] \cdot \Pr[X < a] + \mathbb{E}[X \mid X \geq a] \cdot \Pr[X \geq a]$
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

**Proof**

\[
\mathbb{E}[X] = \mathbb{E}[X | X < a] \cdot \Pr[X < a] + \mathbb{E}[X | X \geq a] \cdot \Pr[X \geq a] \geq 0
\]

Visual Proof

![Visual Proof Diagram](image)
**Markov’s Inequality**

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{E[X]}{a}$.

**Visual Proof**

**Proof**

$$E[X] = E[X \mid X < a] \cdot \Pr[X < a] + E[X \mid X \geq a] \cdot \Pr[X \geq a]$$

$\geq 0$  $\geq 0$  $\geq a$
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

**Visual Proof**

![Graph showing probability distribution]

**Proof**

$$
\mathbb{E}[X] = \mathbb{E}[X \mid X < a] \cdot \Pr[X < a] + \mathbb{E}[X \mid X \geq a] \cdot \Pr[X \geq a] \geq a \cdot \Pr[X \geq a]
$$

$\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x] \geq a \cdot \Pr[X \geq a]$ fits into $\geq a$.
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**Visual Proof**

**Proof** $\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x] \geq a \cdot \Pr[X \geq a]$

**Corollary:** Let $X$ be a non-negative random variable and $a > 0$. Then, $\Pr[X \geq a \cdot \mathbb{E}[X]] \leq 1/a$. 
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**Proof**

$$
\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x] \geq a \cdot \Pr[X \geq a] 
$$

**Visual Proof**

In expectation there is one hair in my soup.”

- How likely is it that I get at least 10?
- How likely is that I get less than 2?

**Corollary:** Let $X$ be a non-negative random variable and $a > 0$. Then, $\Pr[X \geq a \cdot \mathbb{E}[X]] \leq 1/a$. 
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

**Visual Proof**

- **Proof**
  \[
  \mathbb{E}[X] = \sum_x x \cdot \Pr[X = x] \geq a \cdot \Pr[X \geq a]
  \]

- **Corollary**: Let $X$ be a non-negative random variable and $a > 0$. Then, $\Pr[X \geq a \cdot \mathbb{E}[X]] \leq 1/a$.

- “In expectation there is one hair in my soup.”
  - How likely is it that I get at least 10? $\Pr[X \geq 10] \leq 1/10$
  - How likely is that I get less than 2?
Markov’s Inequality

**Theorem (Markov’s inequality):** Let $X$ be a non-negative random variable and let $a > 0$. Then, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.

**Visual Proof**

**Proof**

$$\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x] \geq a \cdot \Pr[X \geq a]$$

**Corollary:** Let $X$ be a non-negative random variable and $a > 0$. Then, $\Pr[X \geq a \cdot \mathbb{E}[X]] \leq \frac{1}{a}$.

- “In expectation there is one hair in my soup.”
  - How likely is it that I get at least 10? $\Pr[X \geq 10] \leq \frac{1}{10}$
  - How likely is it that I get less than 2? $\Pr[X < 2] = 1 - \Pr[X \geq 2] \geq 1 - \frac{1}{2} = 1/2$

Oh no...
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \( X \sim \text{Bin}(20, \frac{1}{5}) \)
Application: Unfair Coins

- The sum of 20 unfair \( \{0, 1\} \)-coin tosses: \( X \sim \text{Bin}(20, \frac{1}{5}) \)

\[
\begin{array}{cccccccccccc}
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1
\end{array}
\]

\[ X = 8 \]
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \(X \sim \text{Bin}(20, \frac{1}{5})\)
- What is the probability of getting at least 16 ones?

\[X = 8\]
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \( X \sim \text{Bin}(20, \frac{1}{5}) \)
- What is the probability of getting at least 16 ones?

\[
\begin{array}{cccccccc}
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\end{array}
\]

\( X = 8 \)

Markov: \( X \) non-negative, \( a > 0 \):
\[
\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.
\]
Application: Unfair Coins

- The sum of 20 unfair \( \{0, 1\} \)-coin tosses: \( X \sim \text{Bin}(20, \frac{1}{5}) \)
- What is the probability of getting at least 16 ones?
  \[ \Pr[X \geq 16] \leq \frac{E[X]}{16} \]
  
  \[ X = 8 \]
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \( X \sim \text{Bin}(20, \frac{1}{5}) \)
- What is the probability of getting at least 16 ones?

\[
\Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16}
\]

\[
20 \cdot \frac{1}{5} = 4
\]

Markov: \( X \) non-negative, \( a > 0 \):

\[
\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.
\]
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \( X \sim \text{Bin}(20, \frac{1}{5}) \)
- What is the probability of getting at least 16 ones?

\[
\Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
\]

\[
20 \cdot \frac{1}{5} = 4
\]

Markov: \( X \) non-negative, \( a > 0 \):
\[
\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.
\]
Application: Unfair Coins

- The sum of 20 unfair \(\{0, 1\}\)-coin tosses: \(X \sim \text{Bin}(20, \frac{1}{5})\)
- What is the probability of getting at least 16 ones?
  \[ \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25 \]
  \[ 20 \cdot \frac{1}{5} = 4 \]
- How tight is that bound?

Markov: \(X\) non-negative, \(a > 0\):

\[ \Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}. \]
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \( X \sim \text{Bin}(20, \frac{1}{5}) \)
- What is the probability of getting at least 16 ones?
  \[
  \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
  \]
  \[
  20 \cdot \frac{1}{5} = 4
  \]
- How tight is that bound?
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left( \frac{1}{5} \right)^k \cdot \left( 1 - \frac{1}{5} \right)^{20-k}
  \]
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \(X \sim \text{Bin}(20, \frac{1}{5})\)
- What is the probability of getting at least 16 ones?
  \[
  \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
  \]
  \[
  20 \cdot \frac{1}{5} = 4
  \]
- How tight is that bound?
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138
  \]
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \( X \sim \text{Bin}(20, \frac{1}{5}) \)
- What is the probability of getting at least 16 ones?

\[
\Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
\]

\[20 \cdot \frac{1}{5} = 4\]

- How tight is that bound? Not very?

\[
\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138
\]
Application: Unfair Coins

- The sum of 20 unfair \(\{0, 1\}\)-coin tosses: \(X \sim \text{Bin}(20, \frac{1}{5})\)
- What is the probability of getting at least 16 ones?
  \[
  \Pr[X \geq 16] \leq \frac{E[X]}{16} = 0.25
  \]
  \[
  20 \cdot \frac{1}{5} = 4
  \]
- How tight is that bound? Not very?
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138
  \]
  \[
  \text{Markov: } X \text{ non-negative, } a > 0: \Pr[X \geq a] \leq \frac{E[X]}{a}.
  \]
  \[
  X = 8
  \]
  \[
  \text{Maybe it is just a weak bound?}
  \]
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \(X \sim \text{Bin}(20, \frac{1}{5})\)
- What is the probability of getting at least 16 ones?
  \[
  \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
  \]
  \[
  20 \cdot \frac{1}{5} = 4
  \]
- How tight is that bound? Not very?
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138
  \]

Fair Coin
- A single \{0, 1\}-coin toss: \(Y \sim \text{Ber}(\frac{1}{2})\)
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \(X \sim \text{Bin}(20, \frac{1}{5})\)
- What is the probability of getting at least 16 ones?
  \[
  \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
  \]
  \[20 \cdot \frac{1}{5} = 4\]
- How tight is that bound? Not very?
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138
  \]

Fair Coin

- A single \{0, 1\}-coin toss: \(Y \sim \text{Ber}(\frac{1}{2})\)
- What is the probability of getting at least 1?
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \(X \sim \text{Bin}(20, \frac{1}{5})\)
- What is the probability of getting at least 16 ones?
  \[
  \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
  \]
  \[
  20 \cdot \frac{1}{5} = 4
  \]
- How tight is that bound? Not very?
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138
  \]

Fair Coin
- A single \{0, 1\}-coin toss: \(Y \sim \text{Ber}(\frac{1}{2})\)
- What is the probability of getting at least 1?
  - Clearly: \(\Pr[Y \geq 1] = \Pr[Y = 1] = \frac{1}{2}\)
Application: Unfair Coins

- The sum of 20 unfair \{0, 1\}-coin tosses: \(X \sim \text{Bin}(20, \frac{1}{5})\)
- What is the probability of getting at least 16 ones?
  \[
  \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
  \]
  \[
  20 \cdot \frac{1}{5} = 4
  \]
- How tight is that bound? Not very?
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.000000138
  \]
- Fair Coin
  - A single \{0, 1\}-coin toss: \(Y \sim \text{Ber}(\frac{1}{2})\)
  - What is the probability of getting at least 1?
    - Clearly: \(\Pr[Y \geq 1] = \Pr[Y = 1] = \frac{1}{2}\)
    - Markov: \(\Pr[Y \geq 1] \leq \frac{\mathbb{E}[Y]}{1} = \mathbb{E}[Y] = \frac{1}{2}\)
Application: Unfair Coins

- The sum of 20 unfair $\{0, 1\}$-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$
- What is the probability of getting at least 16 ones?
  \[
  \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
  \]
- How tight is that bound? Not very?
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.000000138
  \]

Fair Coin

- A single $\{0, 1\}$-coin toss: $Y \sim \text{Ber}(\frac{1}{2})$
- What is the probability of getting at least 1?
  - Clearly: $\Pr[Y \geq 1] = \Pr[Y = 1] = \frac{1}{2}$
  - Markov: $\Pr[Y \geq 1] \leq \mathbb{E}[Y]/1 = \mathbb{E}[Y] = \frac{1}{2}$
Application: Unfair Coins

- The sum of 20 unfair \(\{0, 1\}\)-coin tosses: \(X \sim \text{Bin}(20, \frac{1}{5})\)
- What is the probability of getting at least 16 ones?

\[
\Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
\]

\[20 \cdot \frac{1}{5} = 4\]

- How tight is that bound? Not very?

\[
\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138
\]

**Fair Coin**

- A single \(\{0, 1\}\)-coin toss: \(Y \sim \text{Ber}(\frac{1}{2})\)

- What is the probability of getting at least 1?

\[
\begin{align*}
\text{Clearly: } \Pr[Y \geq 1] &= \Pr[Y = 1] = \frac{1}{2} \\
\text{Markov: } \Pr[Y \geq 1] &\leq \frac{\mathbb{E}[Y]}{1} = \mathbb{E}[Y] = \frac{1}{2}
\end{align*}
\]

There exists a random variable and an \(a > 0\) such that Markov's inequality is exact.
Application: Unfair Coins

- The sum of 20 unfair \( \{0, 1\} \)-coin tosses: \( X \sim \text{Bin}(20, \frac{1}{5}) \)

- What is the probability of getting at least 16 ones?

\[
\Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25
\]

\[
20 \cdot \frac{1}{5} = 4
\]

- How tight is that bound? Not very?

\[
\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138
\]

- Markov: \( X \) non-negative, \( a > 0 \):

\[
\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.
\]

Maybe it is just a weak bound?

Fair Coin

- A single \( \{0, 1\} \)-coin toss: \( Y \sim \text{Ber}(\frac{1}{2}) \)

- What is the probability of getting at least 1?

\[
\Pr[Y \geq 1] = \Pr[Y = 1] = \frac{1}{2}
\]

\[
\Pr[Y \geq 1] \leq \frac{\mathbb{E}[Y]}{1} = \mathbb{E}[Y] = \frac{1}{2}
\]

⇒ There is no better bound (that relies only on the expected value)

There exists a random variable and an \( a > 0 \) such that Markov’s inequality is exact.

We need more information about the shape of the distribution!
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?
Characterizing the Shape of a Distribution

How much information do we need to characterize the shape of a distribution?

Example

- $X, Y$ independent fair die-rolls, $D = X - Y$
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

Example
- $X, Y$ independent fair die-rolls, $D = X - Y$
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

Example
- $X, Y$ independent fair die-rolls, $D = X - Y$
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

Example
- $X, Y$ independent fair die-rolls, $D = X - Y$
- $U$ uniform distribution over \{-5, -4, \ldots, 5\}
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

Example

- \(X, Y\) independent fair die-rolls, \(D = X - Y\)
- \(U\) uniform distribution over \([-5, -4, \ldots, 5]\)
Characterizing the Shape of a Distribution

How much information do we need to characterize the shape of a distribution?

Example

- $X, Y$ independent fair die-rolls, $D = X - Y$
- $U$ uniform distribution over $\{-5, -4, \ldots, 5\}$
Characterizing the Shape of a Distribution

How much information do we need to characterize the shape of a distribution?

- Example
  - $X, Y$ independent fair die-rolls, $D = X - Y$
  - $U$ uniform distribution over $\{-5, -4, \ldots, 5\}$
  - Consider all probabilities individually

![Probability distribution graph](image-url)
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

**Example**
- $X, Y$ independent fair die-rolls, $D = X - Y$
- $U$ uniform distribution over $\{-5, -4, \ldots, 5\}$
- Consider all probabilities individually

Tedious... We need to aggregate!
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

**Example**
- \(X, Y\) independent fair die-rolls, \(D = X - Y\)
- \(U\) uniform distribution over \([-5, -4, \ldots, 5]\)
- Consider all probabilities individually

**Expectation?**
\[
\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0
\]
Characterizing the Shape of a Distribution

How much information do we need to characterize the shape of a distribution?

Example
- \(X, Y\) independent fair die-rolls, \(D = X - Y\)
- \(U\) uniform distribution over \([-5, -4, \ldots, 5]\)
- Consider all probabilities individually

Expectation?
- \(\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0\)
- \(\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0\)
Characterizing the Shape of a Distribution

How much information do we need to characterize the shape of a distribution?

**Example**

- \( X, Y \) independent fair die-rolls, \( D = X - Y \)
- \( U \) uniform distribution over \( \{-5, -4, \ldots, 5\} \)
- Consider all probabilities individually

**Expectation?**

\[
\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0
\]

\[
\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0
\]

Same value, different shapes

(also just seen with Markov: \( \mathbb{E} \) not enough)

Tedious... We need to aggregate!
Characterizing the Shape of a Distribution

How much information do we need to characterize the shape of a distribution?

Example

- $X, Y$ independent fair die-rolls, $D = X - Y$
- $U$ uniform distribution over $\{-5, -4, \ldots, 5\}$
- Consider all probabilities individually. Tedious... We need to aggregate!

Expectation?

$$E[D] = \sum_k \Pr[D = k] \cdot k = 0$$
$$E[U] = \sum_k \Pr[U = k] \cdot k = 0$$

Same value, different shapes

(Also just seen with Markov: $E$ not enough)

Problem: $+$ & $-$ terms cancel

$$
\begin{align*}
\Pr[D = k] & \quad \text{more concentr.} \\
\Pr[U = k] & \quad \text{less concentr.}
\end{align*}
$$
Characterizing the Shape of a Distribution

How much information do we need to characterize the shape of a distribution?

Example

- $X, Y$ independent fair die-rolls, $D = X - Y$
- $U$ uniform distribution over $\{-5, -4, \ldots, 5\}$
- Consider all probabilities individually. Tedious... We need to aggregate!

Expectation?

$E[D] = \sum_k \Pr[D = k] \cdot k = 0$

Same value, different shapes

$E[U] = \sum_k \Pr[U = k] \cdot k = 0$

(also just seen with Markov: $E$ not enough)

Problem: $+$ & $-$ terms cancel

$\Rightarrow$ Fix: absolute value

$E[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945$

$E[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727$
Characterizing the Shape of a Distribution

How much information do we need to characterize the shape of a distribution?

Example

- $X, Y$ independent fair die-rolls, $D = X - Y$
- $U$ uniform distribution over $\{-5, -4, \ldots, 5\}$
- Consider all probabilities individually

Expectation?

$$E[D] = \sum_k \Pr[D = k] \cdot k = 0$$
$$E[U] = \sum_k \Pr[U = k] \cdot k = 0$$

(Also just seen with Markov: $E$ not enough)

Problem: $+$ & $-$ terms cancel

\[ E[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945 \]
\[ E[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727 \]

Distance to $E$
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

**Example**

- $X, Y$ independent fair die-rolls, $D = X - Y$
- $U$ uniform distribution over $\{-5, -4, \ldots, 5\}$
- Consider all probabilities individually

**Expectation?**

$$E[D] = \sum_k \Pr[D = k] \cdot k = 0$$

$$E[U] = \sum_k \Pr[U = k] \cdot k = 0$$

Same value, different shapes

(also just seen with Markov: $E$ not enough)

**Problem:** + & − terms cancel

$\Rightarrow$ Fix: absolute value

$$E[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945$$

$$E[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727$$

Smaller expected distance to $E$

Distance to $E$
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

**Example**

- \(X, Y\) independent fair die-rolls, \(D = X - Y\)
- \(U\) uniform distribution over \([-5, -4, \ldots, 5]\)
- Consider all probabilities individually (Tedious... We need to aggregate!)

**Expectation?**

\[
\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0
\]

\[
\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0
\]

(Also just seen with Markov: \(\mathbb{E}\) not enough)

**Problem:** + & − terms cancel

⇒ Fix: absolute value

\[
\mathbb{E}[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945
\]

\[
\mathbb{E}[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727
\]

More concentrated! Smaller expected distance to \(\mathbb{E}\)

Distance to Expected

Pr[\(D = k\)] more concentr.

Pr[\(U = k\)] less concentr.
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

Example
- \(X, Y\) independent fair die-rolls, \(D = X - Y\)
- \(U\) uniform distribution over \((-5, -4, \ldots, 5)\)
- Consider all probabilities individually

Expectation?
\[
\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0
\]
\[
\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0
\]
(Also just seen with Markov: \(\mathbb{E}\) not enough)

Problem: + & − terms cancel
⇒ Fix: absolute value
\[
\mathbb{E}[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945
\]
\[
\mathbb{E}[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727
\]

Problem: Nobody likes absolute value

Distance to \(\mathbb{E}\)
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

Example
- $X, Y$ independent fair die-rolls, $D = X - Y$
- $U$ uniform distribution over $\{-5, -4, \ldots, 5\}$
- Consider all probabilities individually

Expectation?

\[
\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0
\]

\[
\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0\]  
Same value, different shapes  
(also just seen with Markov: $\mathbb{E}$ not enough)

Problem: $+$ & $-$ terms cancel  
⇒ Fix: absolute value

\[
\mathbb{E}[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945
\]

\[
\mathbb{E}[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727
\]


Problem: Nobody likes absolute value  
⇒ Fix: square instead

Distance to $\mathbb{E}$  
More concentrated!  
Smaller expected distance to $\mathbb{E}$  

Tedious... We need to aggregate!
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

**Example**

- $X, Y$ independent fair die-rolls, $D = X - Y$
- $U$ uniform distribution over $\{-5, -4, \ldots, 5\}$
- Consider all probabilities individually

**Expectation?**

\[
\mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0
\]
\[
\mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0
\]

(Also just seen with Markov: $\mathbb{E}$ not enough)

**Problem:** $+ \& -$ terms cancel

⇒ Fix: absolute value

\[
\mathbb{E}[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945
\]
\[
\mathbb{E}[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727
\]

Distance to $\mathbb{E}$

**Problem:** Nobody likes absolute value

⇒ Fix: square instead

\[
\mathbb{E}[D^2] = \sum_k \Pr[D = k] \cdot k^2 \approx 5.833
\]
\[
\mathbb{E}[U^2] = \sum_k \Pr[U = k] \cdot k^2 = 10.0
\]

Squared distance to $\mathbb{E}$

- More concentrated!
- Smaller expected distance to $\mathbb{E}$
Characterizing the Shape of a Distribution

- How much information do we need to characterize the shape of a distribution?

**Example**

- \(X, Y\) independent fair die-rolls, \(D = X - Y\)
- \(U\) uniform distribution over \([-5, -4, ..., 5]\)
- Consider all probabilities individually
  
  **Expectation?**
  
  \[
  f(k) = k
  \]
  
  \[
  \mathbb{E}[D] = \sum_k \Pr[D = k] \cdot k = 0
  \]
  
  **Same value, different shapes**
  
  \[
  \mathbb{E}[U] = \sum_k \Pr[U = k] \cdot k = 0
  \]
  
  (also just seen with Markov: \(\mathbb{E}\) not enough)

- Problem: + & – terms cancel
  
  \[
  \Rightarrow \text{Fix: absolute value } f(k) = |k|
  \]
  
  \[
  \mathbb{E}[|D|] = \sum_k \Pr[D = k] \cdot |k| \approx 1.945
  \]
  
  **More concentrated!**
  
  \[
  \mathbb{E}[|U|] = \sum_k \Pr[U = k] \cdot |k| \approx 2.727
  \]
  
  **Distance to \(\mathbb{E}\)**

  \[
  \mathbb{E}[D^2] = \sum_k \Pr[D = k] \cdot k^2 \approx 5.833
  \]
  
  **Smaller expected distance to \(\mathbb{E}\)**
  
  \[
  \mathbb{E}[U^2] = \sum_k \Pr[U = k] \cdot k^2 = 10.0
  \]
  
  **Squared distance to \(\mathbb{E}\)**

  **Problem: Nobody likes absolute value**

  \[
  \Rightarrow \text{Fix: square instead } f(k) = k^2
  \]

  \[
  \mathbb{E}[D^2] = \sum_k \Pr[D = k] \cdot k^2 \approx 5.833
  \]
  
  **More concentr.**

  \[
  \mathbb{E}[U^2] = \sum_k \Pr[U = k] \cdot k^2 = 10.0
  \]
Do you have a Moment?

Expectation and Functions
- Random variable $X$ taking values in a set $S$
- A function $f$, e.g. $f(X) = X$, $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $\mathbb{E}[f(X)] = \sum_{x \in S} \Pr[X = x] \cdot f(x)$
Do you have a Moment?

Expectation and Functions

- Random variable $X$ taking values in a set $S$
- A function $f$, e.g. $f(X) = X^1$, $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $E[f(X)] = \sum_{x \in S} \Pr[X = x] \cdot f(x)$

These turn out to be particularly useful!
Do you have a Moment?

Expectation and Functions

- Random variable $X$ taking values in a set $S$
- A function $f$, e.g. $f(X) = X^1$, $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $E[f(X)] = \sum_{x \in S} \Pr[X = x] \cdot f(x)$

Moments

**Definition**: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th raw moment is $E[X^n]$. These turn out to be particularly useful!
Do you have a Moment?

Expectation and Functions
- Random variable $X$ taking values in a set $S$
- A function $f$, e.g. $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $\mathbb{E}[f(X)] = \sum_{x \in S} \Pr[X = x] \cdot f(x)$

Moments

Definition: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th raw moment is $\mathbb{E}[X^n]$.

Just seen: For $\mathbb{E}[X] = 0$, this captures distances to $\mathbb{E}[X]$
Do you have a Moment?

Expectation and Functions

- Random variable $X$ taking values in a set $S$
- A function $f$, e.g. $f(X) = X^1$, $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $\mathbb{E}[f(X)] = \sum_{x \in S} \Pr[X = x] \cdot f(x)$

These turn out to be particularly useful!

Moments

Definition: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th raw moment is $\mathbb{E}[X^n]$.

Just seen: For $\mathbb{E}[X] = 0$, this captures distances to $\mathbb{E}[X]$ What if $\mathbb{E}[X] \neq 0$?

Definition: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th central moment is $\mathbb{E}[(X - \mathbb{E}[X])^n]$. 


Do you have a Moment?

Expectation and Functions

- Random variable $X$ taking values in a set $S$
- A function $f$, e.g. $f(X) = X^1$, $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $E[f(X)] = \sum_{x \in S} \Pr[X = x] \cdot f(x)$

These turn out to be particularly useful!

Moments

Definition: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th raw moment is $E[X^n]$.

- Just seen: For $E[X] = 0$, this captures distances to $E[X]$

Definition: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th central moment is $E[(X - E[X])^n]$.

- Just seen: the 2nd central moment captures squared distances to the expected value
Do you have a Moment?

Expectation and Functions
- Random variable $X$ taking values in a set $S$
- A function $f$, e.g. $f(X) = \frac{X}{1}$, $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $E[f(X)] = \sum_{x \in S} Pr[X = x] \cdot f(x)$

Moments

Definition: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th raw moment is $E[X^n]$.

Just seen: For $E[X] = 0$, this captures distances to $E[X]$  What if $E[X] \neq 0$?

Definition: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th central moment is $E[(X - E[X])^n]$.

Just seen: the 2nd central moment captures squared distances to the expected value

$$E[(X - E[X])^2] = Var[X]$$
Do you have a Moment?

Expectation and Functions

- Random variable $X$ taking values in a set $S$
- A function $f$, e.g. $f(X) = X^1$, $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $E[f(X)] = \sum_{x \in S} P[X = x] \cdot f(x)$

Moments

**Definition**: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th raw moment is $E[X^n]$.

- Just seen: For $E[X] \neq 0$, this captures distances to $E[X]$

**Definition**: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th central moment is $E[(X - E[X])^n]$.

- Just seen: the 2nd central moment captures squared distances to the expected value

\[
E[(X - E[X])^2] = \text{Var}[X]
\]

- The smaller the variance, the more concentrated the random variable

These turn out to be particularly useful!
Do you have a Moment?

Expectation and Functions

- Random variable $X$ taking values in a set $S$
- A function $f$, e.g. $f(X) = X^1$, $f(X) = |X|$, $f(X) = X^2$, $f(X) = \sqrt{X}$, $f(X) = X^3$, $f(X) = e^X$
- $\mathbb{E}[f(X)] = \sum_{x \in S} \Pr[X = x] \cdot f(x)$

These turn out to be particularly useful!

Moments

Definition: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th raw moment is $\mathbb{E}[X^n]$.

Just seen: For $\mathbb{E}[X] = 0$, this captures distances to $\mathbb{E}[X]$. What if $\mathbb{E}[X] \neq 0$?

Definition: For random variable $X$ and $n \in \mathbb{N}$ the $n$-th central moment is $\mathbb{E}[(X - \mathbb{E}[X])^n]$.

Just seen: the 2nd central moment captures squared distances to the expected value

$$\mathbb{E}[(X - \mathbb{E}[X])^2] = \text{Var}[X]$$

The smaller the variance, the more concentrated the random variable

... and with Markov’s help, we can turn that insight into a concentration inequality!
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

Markov: $Y \geq 0, a > 0: \Pr[Y \geq a] \leq \mathbb{E}[Y]/a$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

Proof

$\Pr[|X - \mathbb{E}[X]| \geq b]$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

Proof

\[
\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2]
\]

Markov: $Y \geq 0, a > 0: \Pr[Y \geq a] \leq \mathbb{E}[Y]/a$
Chebychev’s Inequality

Markov's teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

**Theorem (Chebychev’s inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

**Proof**

$$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2]$$

Markov: $Y \geq 0, a > 0: \Pr[Y \geq a] \leq \mathbb{E}[Y]/a$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

**Theorem (Chebychev’s inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

**Proof**

$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2] / b^2$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

Proof

$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr\left[(X - \mathbb{E}[X])^2 \geq b^2\right] \leq \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]/b^2 = \text{Var}[X]/b^2 \checkmark$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

Proof

\[ \Pr[|X - \mathbb{E}[X]| \geq b] = \Pr \left[ (X - \mathbb{E}[X])^2 \geq b^2 \right] \leq \mathbb{E} \left[ (X - \mathbb{E}[X])^2 \right]/b^2 = \text{Var}[X]/b^2. \]

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16] \approx 0.0000000138$
- $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$
- $\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left( \frac{1}{5} \right)^k \left( 1 - \frac{1}{5} \right)^{20-k} \approx 0.0000000138$
- Markov: $\Rightarrow \Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

**Theorem (Chebychev’s inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

**Proof**

$$
\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2]/b^2 = \text{Var}[X]/b^2.
$$

**Application: Unfair Coins**

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16]$?
  - $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$
  - $\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.000000138$
  - Markov: $\Rightarrow \Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$
  - Chebychev:
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

**Theorem (Chebychev’s inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

**Proof**

$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr\left[(X - \mathbb{E}[X])^2 \geq b^2\right] \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{b^2} = \frac{\text{Var}[X]}{b^2}$

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16]$?
  $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$

  $\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.000000138$

- Markov: $\Rightarrow \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25$

- Chebychev:
  $\Pr[X \geq 16]$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

Proof

\[
\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{b^2} = \frac{\text{Var}[X]}{b^2} \quad \checkmark
\]

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16]$?
  \[
  \mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4
  \]
  \[
  \Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k}(\frac{1}{5})^k \cdot (1 - \frac{1}{5})^{20-k} \approx 0.000000138
  \]

Markov: $\implies \Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$

Chebychev:

$\Pr[X \geq 16]$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.  

Proof

$$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2]/b^2 = \text{Var}[X]/b^2 \checkmark$$

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16]?$  
  $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$  
  $\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot (1 - \frac{1}{5})^{20-k} \approx 0.000000138$
- Markov: $\Rightarrow \Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$
- Chebychev:  
  $\Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$

$X \geq 16 \iff X - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]$
Chebychev’s Inequality

Theorem (Chebychev’s inequality): Let \( X \) be a random variable with finite variance and let \( b > 0 \). Then, \( \Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2 \).

Proof

\[
\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{b^2} = \frac{\text{Var}[X]}{b^2}.
\]

Markov: \( Y \geq 0, a > 0: \Pr[Y \geq a] \leq \mathbb{E}[Y]/a \)

Application: Unfair Coins

\( X \sim \text{Bin}(20, \frac{1}{5}) \), \( \Pr[X \geq 16] \)?

\[
\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left( \frac{1}{5} \right)^k \left( 1 - \frac{1}{5} \right)^{20-k} \approx 0.000000138
\]

Markov: \( \Rightarrow \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25 \)

Chebychev:

\[
\Pr[X \geq 16] \quad \Leftrightarrow \quad X - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]
\]

\[
\Leftrightarrow X - \mathbb{E}[X] \geq 12
\]
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

**Theorem (Chebychev’s inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

**Proof**

\[
\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2]/b^2 = \text{Var}[X]/b^2 \tag{✓}
\]

**Application: Unfair Coins**

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16] = \mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$

\[
\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.000000138
\]

- **Markov:** $\Rightarrow \Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$

- **Chebychev:**

\[
\Pr[X \geq 16]
\]

$x \geq 16$

$\iff x - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]$

$\iff x - \mathbb{E}[X] \geq 12$

$|x - \mathbb{E}[X]| \geq 12 \Rightarrow x \geq 16 \text{ or } x \leq -8$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

**Theorem (Chebychev’s inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$.

**Proof**

Markov: $Y \geq 0$, $a > 0$: $Pr[Y \geq a] \leq \mathbb{E}[Y]/a$

$$Pr[|X - \mathbb{E}[X]| \geq b] = Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2]/b^2 = \frac{\text{Var}[X]}{b^2} \checkmark$$

**Application: Unfair Coins**

- $X \sim \text{Bin}(20, \frac{1}{5})$, $Pr[X \geq 16]$, $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$
  $$Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.000000138$$
- Markov: $\Rightarrow Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25$
- Chebychev:
  $$Pr[X \geq 16] \leq Pr[X \geq 16 \lor X \leq -8]$$

- $X \geq 16 
  \iff X - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]$
  \iff $X - \mathbb{E}[X] \geq 12$
  \iff $|X - \mathbb{E}[X]| \geq 12 \Rightarrow X \geq 16 \text{ or } X \leq -8$
Chebychev's Inequality

Markov's teacher! (Markov's inequality actually appeared earlier in Chebychev's works)

**Theorem (Chebychev's inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

**Proof**

$$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2]/b^2 = \text{Var}[X]/b^2 \quad \checkmark$$

**Application: Unfair Coins**

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16]$?
  - $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$
  - $\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.000000138$
- Markov: $\Rightarrow \Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$
- Chebychev:
  - $\Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8]$
  - $= \Pr[|X - \mathbb{E}[X]| \geq 12]$
  - $X \geq 16 \Leftrightarrow X - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]$
  - $X - \mathbb{E}[X] \geq 12$
  - $|X - \mathbb{E}[X]| \geq 12 \Rightarrow X \geq 16 \text{ or } X \leq -8$
Chebychev’s Inequality

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$.

Proof

Markov: $Y \geq 0, a > 0: \Pr[Y \geq a] \leq \mathbb{E}[Y]/a$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2] / b^2 = \text{Var}[X] / b^2 \]

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16] \approx 0.0000000138$
- Markov: $\Rightarrow \Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$
- Chebychev:

\[
\Pr[X \geq 16] \leq \Pr[X \geq 16 \vee X \leq -8] = \Pr[|X - \mathbb{E}[X]| \geq 12] \leq \frac{\text{Var}[X]}{12^2}
\]

$x \geq 16 \Leftrightarrow X - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]$

$x \geq 12 \Leftrightarrow X - \mathbb{E}[X] \geq 12$

$|X - \mathbb{E}[X]| \geq 12 \Rightarrow X \geq 16 \text{ or } X \leq -8$
Chebychev’s Inequality

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$.

Proof

Markov: $Y \geq 0, a > 0: \Pr[Y \geq a] \leq \mathbb{E}[Y]/a$

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16] = \frac{20!}{16!4!} \left(\frac{1}{5}\right)^4 \cdot \left(1 - \frac{1}{5}\right)^{20-4} \approx 0.000000138$
- Markov: $\implies \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25$
- Chebychev:

$$\Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8] = \Pr[|X - \mathbb{E}[X]| \geq 12] \leq \frac{\text{Var}[X]}{12^2}$$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$.

Proof

$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2] / b^2 = \text{Var}[X] / b^2$.

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16]$?
  $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$
  $\Pr[X \geq 16] = \sum_{k=16}^{20} (\begin{pmatrix} 20 \\ k \end{pmatrix}) \left( \frac{1}{5} \right)^k \cdot \left( 1 - \frac{1}{5} \right)^{20-k} \approx 0.000000138$

- Markov: $\Rightarrow \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25$

- Chebychev:
  $\Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8] = \Pr[|X - \mathbb{E}[X]| \geq 12] \leq \frac{\text{Var}[X]}{12^2}$

$X \sim \text{Bin}(n, p) : \text{Var}[X] = np(1 - p)$
**Chebychev’s Inequality**

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

**Theorem (Chebychev’s inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

**Proof**

$$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \mathbb{E}[(X - \mathbb{E}[X])^2]/b^2 = \text{Var}[X]/b^2$$

Markov: $Y \geq 0$, $a > 0$: $\Pr[Y \geq a] \leq \mathbb{E}[Y]/a$

**Application: Unfair Coins**

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16]$?
  
  $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$  
  $\text{Var}[X] = 20 \cdot \frac{1}{5} \cdot (1 - \frac{1}{5})$

  $\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138$

- Markov: $\Rightarrow \Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$

- Chebychev:
  
  $\Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8]$
  
  $= \Pr[|X - \mathbb{E}[X]| \geq 12]$
  
  $\leq \frac{\text{Var}[X]}{12^2}$

$X \geq 16$  

$\Leftrightarrow X - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]$  

$\Leftrightarrow X - \mathbb{E}[X] \geq 12$  

$|X - \mathbb{E}[X]| \geq 12 \Rightarrow X \geq 16 \text{ or } X \leq -8$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

**Theorem (Chebychev’s inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$.

**Proof**

Markov: $Y \geq 0, a > 0: \Pr[Y \geq a] \leq \frac{\mathbb{E}[Y]}{a}$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr\left[(X - \mathbb{E}[X])^2 \geq b^2\right] \leq \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] / b^2 = \frac{\text{Var}[X]}{b^2} \checkmark
\]

**Application: Unfair Coins**

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16]$?
  - $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$
  - $\text{Var}[X] = 20 \cdot \frac{1}{5} \cdot (1 - \frac{1}{5}) = \frac{16}{5}$
  - $\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \cdot \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138$
  - $X \sim \text{Bin}(n, p) : \text{Var}[X] = np(1-p)$

- Markov: $\Rightarrow \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25$

- Chebychev:
  
  \[
  \Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8] = \Pr[|X - \mathbb{E}[X]| \geq 12] \leq \frac{\text{Var}[X]}{12^2}
  \]

\[
\iff X \geq 16 \iff X - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]
\]

\[
\iff X - \mathbb{E}[X] \geq 12 \iff |X - \mathbb{E}[X]| \geq 12 \Rightarrow X \geq 16 \text{ or } X \leq -8
\]
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

**Theorem (Chebychev’s inequality):** Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$.

**Proof**

\[
\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr[(X - \mathbb{E}[X])^2 \geq b^2] \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{b^2} = \frac{\text{Var}[X]}{b^2}
\]

Markov: $Y \geq 0, a > 0: \Pr[Y \geq a] \leq \frac{\mathbb{E}[Y]}{a}$

**Application: Unfair Coins**

- $X \sim \text{Bin}(20, \frac{1}{5}), \Pr[X \geq 16]$? $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$ $\text{Var}[X] = 20 \cdot \frac{1}{5} \cdot (1 - \frac{1}{5}) = \frac{16}{5}$

\[
\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{1}{5}\right)^k \left(1 - \frac{1}{5}\right)^{20-k} \approx 0.0000000138
\]

- Markov: $\Rightarrow \Pr[X \geq 16] \leq \frac{\mathbb{E}[X]}{16} = 0.25$

- Chebychev:

\[
\Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8] = \Pr[|X - \mathbb{E}[X]| \geq 12] \leq \frac{\text{Var}[X]}{12^2} = \frac{16}{5 \cdot 144}
\]

$X \sim \text{Bin}(n, p) : \text{Var}[X] = np(1 - p)$
Chebychev’s Inequality

Markov’s teacher! (Markov’s inequality actually appeared earlier in Chebychev’s works)

Theorem (Chebychev’s inequality): Let $X$ be a random variable with finite variance and let $b > 0$. Then, $\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2$.

Proof

$$\Pr[|X - \mathbb{E}[X]| \geq b] = \Pr \left[ (X - \mathbb{E}[X])^2 \geq b^2 \right] \leq \mathbb{E} \left[ (X - \mathbb{E}[X])^2 \right]/b^2 = \text{Var}[X]/b^2 \checkmark$$

Application: Unfair Coins

- $X \sim \text{Bin}(20, \frac{1}{5})$, $\Pr[X \geq 16]$?
  - $\mathbb{E}[X] = 20 \cdot \frac{1}{5} = 4$  
  - $\text{Var}[X] = 20 \cdot \frac{1}{5} \cdot (1 - \frac{1}{5}) = \frac{16}{5}$

  $$\Pr[X \geq 16] = \sum_{k=16}^{20} \binom{20}{k} \left( \frac{1}{5} \right)^k \cdot \left( 1 - \frac{1}{5} \right)^{20-k} \approx 0.000000138$$

- Markov: $\Rightarrow \Pr[X \geq 16] \leq \mathbb{E}[X]/16 = 0.25$

- Chebychev:
  $$\Pr[X \geq 16] \leq \Pr[X \geq 16 \lor X \leq -8]$$
  $$= \Pr[|X - \mathbb{E}[X]| \geq 12]$$
  $$\leq \frac{\text{Var}[X]}{12^2} = \frac{16}{5 \cdot 144} \approx 0.022 \checkmark$$

Order of magnitude better than Markov!

$X \sim \text{Bin}(n, p)$: $\text{Var}[X] = np(1 - p)$

- $X \geq 16$  
  - $\Leftrightarrow X - \mathbb{E}[X] \geq 16 - \mathbb{E}[X]$  
  - $\Leftrightarrow X - \mathbb{E}[X] \geq 12$  
  - $|X - \mathbb{E}[X]| \geq 12 \Rightarrow X \geq 16$ or $X \leq -8$
Recap

- \( G(n, p) \): Start with \( n \) nodes, connect any two with fixed probability \( p \), independently
- Probability distribution of the degree of a \textit{single} node \( v \): \( \deg(v) \sim \text{Bin}(n - 1, p) \)
Application: ER – Degree Distribution

Recap

- $G(n, p)$: Start with $n$ nodes, connect any two with fixed probability $p$, independently
- Probability distribution of the degree of a single node $v$: $\deg(v) \sim \text{Bin}(n - 1, p)$
- For $p = c/n$ with $c \in \Theta(1)$ the degree of a vertex is approximately Poisson-distributed
Application: ER – Degree Distribution

Recap

- $G(n, p)$: Start with $n$ nodes, connect any two with fixed probability $p$, independently
- Probability distribution of the degree of a single node $v$: $\deg(v) \sim \text{Bin}(n - 1, p)$
- For $p = c/n$ with $c \in \Theta(1)$ the degree of a vertex is approximately Poisson-distributed
  - Total variation distance of $X, Y$ taking values in a set $S$: $d_{TV}(X, Y) = \frac{1}{2} \sum_{x \in S} |\Pr[X = x] - \Pr[Y = x]|$
Application: ER – Degree Distribution

Recap

- $G(n, p)$: Start with $n$ nodes, connect any two with fixed probability $p$, independently.
- Probability distribution of the degree of a single node $v$: $\deg(v) \sim \text{Bin}(n - 1, p)$.
- For $p = c/n$ with $c \in \Theta(1)$ the degree of a vertex is approximately Poisson-distributed.
  - Total variation distance of $X, Y$ taking values in a set $S$:
    $$d_{TV}(X, Y) = \frac{1}{2} \sum_{x \in S} |\Pr[X = x] - \Pr[Y = x]|$$
  - For $\lambda = -n \log(1 - p) = c + O(1/n)$ and $X \sim \text{Pois}(\lambda)$ we have $d_{TV}(\deg(v), X) = o(1)$.
Application: ER – Degree Distribution

Recap

- $G(n, p)$: Start with $n$ nodes, connect any two with fixed probability $p$, independently
- Probability distribution of the degree of a single node $v$: $\text{deg}(v) \sim \text{Bin}(n - 1, p)$
- For $p = c/n$ with $c \in \Theta(1)$ the degree of a vertex is approximately Poisson-distributed
- Total variation distance of $X, Y$ taking values in a set $S$:
  $$d_{TV}(X, Y) = \frac{1}{2} \sum_{x \in S} |\Pr[X = x] - \Pr[Y = x]|$$
- For $\lambda = -n \log(1 - p) = c + O(1/n)$ and $X \sim \text{Pois}(\lambda)$ we have $d_{TV}(\text{deg}(v), X) = o(1)$
- Empirical distribution of the degrees of all vertices in a graph $G = (V, E)$
  $$N_d = \sum_{v \in V} 1_{\{\text{deg}(v) = d\}} \quad (\text{normalized: } \frac{1}{n} N_d, \text{ for } n = |V|)$$
Application: ER – Degree Distribution

Recap

- $G(n, p)$: Start with $n$ nodes, connect any two with fixed probability $p$, independently
- Probability distribution of the degree of a single node $v$: $\deg(v) \sim \text{Bin}(n - 1, p)$
- For $p = c/n$ with $c \in \Theta(1)$ the degree of a vertex is approximately Poisson-distributed
- Total variation distance of $X, Y$ taking values in a set $S$:
  $d_{TV}(X, Y) = \frac{1}{2} \sum_{x \in S} |\Pr[X = x] - \Pr[Y = x]|$
- For $\lambda = -n \log(1 - p) = c + O(1/n)$ and $X \sim \text{Pois}(\lambda)$ we have $d_{TV}(\deg(v), X) = o(1)$
- Empirical distribution of the degrees of all vertices in a graph $G = (V, E)$
  $N_d = \sum_{v \in V} \mathbb{1}_{\{\deg(v) = d\}}$ (normalized: $\frac{1}{n} N_d$, for $n = |V|$)

Empirical distribution of the degrees of all vertices in a graph $G = (V, E)$ for $n = 100$, $n = 1000$, $n = 10000$.
Theorem: Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\epsilon > 0$ we have
\[
\lim_{n \to \infty} \Pr \left[ |\Pr[X = d] - \frac{1}{n} N_d| \geq \epsilon \right] = 0.
\]
Theorem: Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\epsilon > 0$ we have
\[
\lim_{n \to \infty} \Pr\left[\left|\Pr[X = d] - \frac{1}{n} N_d\right| \geq \epsilon\right] = 0.
\]
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\epsilon > 0$ we have
\[
\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0.
\]

**Proof**
- Step 1: $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$
\[
\lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0.
\]
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\epsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0.$$ 

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$
  \[
  \lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0
  \]

- **Step 2:** $\frac{1}{n} N_d$ is concentrated
  \[
  \lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0
  \]
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\varepsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0.$$

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

$$\lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0$$

- **Step 2:** $\frac{1}{n} N_d$ is concentrated

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\varepsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0.$$  

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$. 
  $$\lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0$$

  $$\left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = \frac{1}{n} \mathbb{E}[N_d]$$

- **Step 2:** $\frac{1}{n} N_d$ is concentrated

  $$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0.$$
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\varepsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0.$$  

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

\[
\left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right|
= \frac{1}{n} \mathbb{E}[N_d]
= \frac{1}{n} \mathbb{E}[\sum_{v \in V} 1\{\deg(v) = d\}]
\]

- **Step 2:** $\frac{1}{n} N_d$ is concentrated

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\varepsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0.$$

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

  $$\lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0$$

  $$\Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] = \frac{1}{n} \mathbb{E} [N_d]$$

  $$= \frac{1}{n} \mathbb{E} [\sum_{v \in V} 1_{\{\text{deg}(v) = d\}}]$$

  $$= \frac{1}{n} \sum_{v \in V} \mathbb{E} [1_{\{\text{deg}(v) = d\}}]$$

- **Step 2:** $\frac{1}{n} N_d$ is concentrated

  $$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

  $$\lambda = c + O(1/n) \to c \text{ for } n \to \infty$$
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\varepsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0.$$

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

  $$\lim_{n \to \infty} \left| \Pr[X = d] - \frac{1}{n} \mathbb{E}[N_d] \right| = 0$$

  $$\Pr[X = d] - \mathbb{E}\left[ \frac{1}{n} N_d \right] = \frac{1}{n} \mathbb{E}[N_d]$$

  $$= \frac{1}{n} \mathbb{E}\left[ \sum_{v \in V} 1_{\{\text{deg}(v) = d\}} \right]$$

  $$= \frac{1}{n} \sum_{v \in V} \mathbb{E}[1_{\{\text{deg}(v) = d\}}]$$

  $$= \frac{1}{n} \sum_{v \in V} \Pr[\text{deg}(v) = d]$$

- **Step 2:** $\frac{1}{n} N_d$ is concentrated

  $$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E}\left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\epsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0.$$ 

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

$$\left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = \frac{1}{n} \mathbb{E}[N_d] = \frac{1}{n} \mathbb{E} \left[ \sum_{v \in V} 1_{\{\deg(v) = d\}} \right] = \frac{1}{n} \sum_{v \in V} \mathbb{E}[1_{\{\deg(v) = d\}}] = \frac{1}{n} \sum_{v \in V} \Pr[\deg(v) = d] = \Pr[\deg(v) = d]$$

- **Step 2:** $\frac{1}{n} N_d$ is concentrated

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0$$
Theorem: Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\epsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0.$$ 

Proof

Step 1: $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

$$\lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0$$

$$\Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] = \Pr[X = d] - \Pr[\deg(v) = d]$$

$$= \frac{1}{n} \mathbb{E}[N_d]$$

$$= \frac{1}{n} \mathbb{E} \left[ \sum_{v \in V} 1_{\{\deg(v) = d\}} \right]$$

$$= \frac{1}{n} \sum_{v \in V} \mathbb{E}[1_{\{\deg(v) = d\}}]$$

$$= \frac{1}{n} \sum_{v \in V} \Pr[\deg(v) = d]$$

$$= \Pr[\deg(v) = d]$$

Step 2: $\frac{1}{n} N_d$ is concentrated

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0$$
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\epsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0.$$ 

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

  $$\lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0$$

  
  
  $$\left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = \left| \Pr[X = d] - \Pr[\deg(v) = d] \right| \leq \sum_{d \geq 0} \left| \Pr[X = d] - \Pr[\deg(v) = d] \right|$$

  
  
  $$= \frac{1}{n} \mathbb{E}[N_d]$$

  
  
  $$= \frac{1}{n} \mathbb{E}[\sum_{v \in V} 1_{\{\deg(v) = d\}}]$$

  
  
  $$= \frac{1}{n} \sum_{v \in V} \mathbb{E}[1_{\{\deg(v) = d\}}]$$

  
  
  $$= \frac{1}{n} \sum_{v \in V} \Pr[\deg(v) = d]$$

  
  
  $$= \Pr[\deg(v) = d]$$

- **Step 2:** $\frac{1}{n} N_d$ is concentrated

  $$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0$$
Theorem: Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\varepsilon > 0$ we have
\[
\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0.
\]

Proof

Step 1: $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$
\[
\lim_{n \to \infty} \left| \Pr[X = d] - E \left[ \frac{1}{n} N_d \right] \right| = 0
\]
\[
= \frac{1}{n} E[N_d]
= \frac{1}{n} E[\sum_{v \in V} 1\{\deg(v) = d\}]
= \frac{1}{n} \sum_{v \in V} E[1\{\deg(v) = d\}]
= \frac{1}{n} \sum_{v \in V} \Pr[\deg(v) = d]
= \Pr[\deg(v) = d]
\]

Step 2: $\frac{1}{n} N_d$ is concentrated
\[
\lim_{n \to \infty} \Pr \left[ \left| E \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\varepsilon > 0$ we have

$$
\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0.
$$

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

  $$
  \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = \left| \Pr[X = d] - \Pr[\text{deg}(v) = d] \right| \leq \sum_{d \geq 0} \left| \Pr[X = d] - \Pr[\text{deg}(v) = d] \right|
  $$

  $$
  \quad = 2 \cdot d_{TV}(X, \text{deg}(v))
  $$

  $$
  \quad = o(1)
  $$

  Already shown last time!

- **Step 2:** $\frac{1}{n} N_d$ is concentrated

  $$
  \lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
  $$
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1-p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\varepsilon > 0$ we have

$$\lim_{n \to \infty} \Pr \left[ \left| \Pr[X = d] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0.$$ 

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

  $$\lim_{n \to \infty} \left| \Pr[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0 \checkmark$$

  $$\mathbb{E} \left[ \frac{1}{n} N_d \right] = \frac{1}{n} \mathbb{E} \left[ \sum_{v \in V} 1_{\{\deg(v) = d\}} \right]$$

  $$= \frac{1}{n} \mathbb{E} \left[ \sum_{v \in V} \mathbb{E} [1_{\{\deg(v) = d\}}] \right]$$

  $$= \frac{1}{n} \sum_{v \in V} \Pr[\deg(v) = d]$$

  $$= \Pr[\deg(v) = d]$$

  $$= 2 \cdot d_{TV}(X, \deg(v))$$

  $d_{TV}(X, Y) = \frac{1}{2} \sum_{x \in S} |\Pr[X = x] - \Pr[Y = x]|$

  **Already shown last time!**

- **Step 2:** $\frac{1}{n} N_d$ is concentrated

  $$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right]$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right]$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n}N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right]$$

$N_d \in \{0, \ldots, n\}$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n}N_d$

$$
\Pr\left[|\mathbb{E}\left[\frac{1}{n}N_d\right] - \frac{1}{n}N_d| \geq \varepsilon\right] \leq \text{Var}\left[\frac{1}{n}N_d\right] / \varepsilon^2
$$

**Chebychev:** $X$ finite variance, $b > 0$

$$
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2
$$

$$
\lim_{n \to \infty} \Pr\left[|\mathbb{E}\left[\frac{1}{n}N_d\right] - \frac{1}{n}N_d| \geq \varepsilon\right] = 0
$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\text{Pr} \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \text{Var} \left[ \frac{1}{n} N_d \right] / \varepsilon^2$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ (\frac{1}{n} N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$\lim_{n \to \infty} \text{Pr} \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\text{Pr}[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X] / b^2$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 = \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr \left[ \left| \frac{1}{n} \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\epsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
\]

\[
= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \frac{1}{n} \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0
\]

Chebychev: $X$ finite variance, $b > 0$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]
Step 2: Concentration of $\frac{1}{n} N_d$

\[
Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
= \frac{1}{n^2} \left( \mathbb{E} \left[ \left( N_d \right)^2 \right] - \mathbb{E} [N_d]^2 \right)
\]

\[
N_d = \sum_{v \in V} 1_{\{\text{deg}(v) = d\}} = \mathbb{E} \left[ \left( \sum_{v \in V} 1_{\{\text{deg}(v) = d\}} \right)^2 \right]
\]

\[
\lim_{n \to \infty} Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] = 0
\]

**Chebychev:** $X$ finite variance, $b > 0$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$N_d = \sum_{v \in V} 1_{\{\deg(v)=d\}} = \mathbb{E} \left[ \left( \sum_{v \in V} 1_{\{\deg(v)=d\}} \right)^2 \right]$$

$$= \mathbb{E} \left[ \sum_{v \in V} (1_{\{\deg(v)=d\}})^2 \right] + \sum_{v \in V} \sum_{u \neq v} 1_{\{\deg(v)=d\}} \cdot 1_{\{\deg(u)=d\}}$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$$\lim_{n \to \infty} \Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n}N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ (\frac{1}{n} N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$N_d = \sum_{v \in V} 1_{\{\text{deg}(v) = d\}}$$

$\mathbb{E} \left[ \left( \sum_{v \in V} 1_{\{\text{deg}(v) = d\}} \right)^2 \right]$ = $\mathbb{E} \left[ \sum_{v \in V} (1_{\{\text{deg}(v) = d\}})^2 \right] + \sum_{v \in V} \sum_{u \neq v} 1_{\{\text{deg}(v) = d\}} \cdot 1_{\{\text{deg}(u) = d\}}$

<table>
<thead>
<tr>
<th>Indicator RV $X$: $X^2 = X$, Lin. of Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{E} \left[ \sum_{v \in V} 1_{{\text{deg}(v) = d}} \right] + \mathbb{E} \left[ \sum_{v \in V} \sum_{u \neq v} 1_{{\text{deg}(v) = d}} \cdot 1_{{\text{deg}(u) = d}} \right]$</td>
</tr>
</tbody>
</table>

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n}N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n}N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n}N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n}N_d \right)^2 \right] - \left( \mathbb{E} \left[ \frac{1}{n}N_d \right] \right)^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)$$

$$N_d = \sum_{v \in V} \mathbb{1}_{\{\text{deg}(v)=d\}}$$

$$= \mathbb{E} \left[ \left( \sum_{v \in V} \mathbb{1}_{\{\text{deg}(v)=d\}} \right)^2 \right]$$

$$= \mathbb{E} \left[ \sum_{v \in V} \left( \mathbb{1}_{\{\text{deg}(v)=d\}} \right)^2 + \sum_{v \in V} \sum_{u \neq v} \mathbb{1}_{\{\text{deg}(v)=d\}} \cdot \mathbb{1}_{\{\text{deg}(u)=d\}} \right]$$

Indicator RV $X$: $X^2 = X$, Lin. of Exp.

Ind. of Exp.

$$= \mathbb{E} \left[ \sum_{v \in V} \mathbb{1}_{\{\text{deg}(v)=d\}} \right] + \mathbb{E} \left[ \sum_{v \in V} \sum_{u \neq v} \mathbb{1}_{\{\text{deg}(v)=d\}} \cdot \mathbb{1}_{\{\text{deg}(u)=d\}} \right]$$

$$= \sum_{v \in V} \mathbb{E} \left[ \mathbb{1}_{\{\text{deg}(v)=d\}} \right] + \sum_{v \in V} \sum_{u \neq v} \mathbb{E} \left[ \mathbb{1}_{\{\text{deg}(v)=d\}} \cdot \mathbb{1}_{\{\text{deg}(u)=d\}} \right]$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| E \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = E \left[ \left( \frac{1}{n} N_d \right)^2 \right] - E \left[ \frac{1}{n} N_d \right]^2$$

$$N_d = \sum_{v \in V} I_{\{\text{deg}(v) = d\}} = E \left[ \left( \sum_{v \in V} I_{\{\text{deg}(v) = d\}} \right)^2 \right]$$

$$\text{Ind. RV } X: X^2 = X,$$

Lin. of Exp. = $E \left[ \sum_{v \in V} I_{\{\text{deg}(v) = d\}} \right] + E \left[ \sum_{v \in V} \sum_{u \neq v} I_{\{\text{deg}(v) = d\}} \cdot I_{\{\text{deg}(u) = d\}} \right]$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - E[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$$\lim_{n \to \infty} \Pr \left[ \left| E \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ |E \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \epsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\epsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = E \left[ \left( \frac{1}{n} N_d \right)^2 \right] - E \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( E \left[ (N_d)^2 \right] - E \left[ N_d \right]^2 \right)$$

$$N_d = \sum_{v \in V} 1 \{ \deg(v) = d \}$$

$$= E \left[ \left( \sum_{v \in V} 1 \{ \deg(v) = d \} \right)^2 \right]$$

$$= E \left[ \sum_{v \in V} 1 \{ \deg(v) = d \} \right]^2 + \sum_{v \in V} \sum_{u \neq v} 1 \{ \deg(v) = d \} \cdot 1 \{ \deg(u) = d \}$$

Indicator RV $X$: $X^2 = X$, Lin. of Exp.

$$= \sum_{v \in V} E \left[ 1 \{ \deg(v) = d \} \right] + \sum_{v \in V} \sum_{u \neq v} E \left[ 1 \{ \deg(v) = d \} \cdot 1 \{ \deg(u) = d \} \right]$$

$$= \Pr[\deg(v) = d]$$

$$= 1 \text{ iff } \deg(v) = d \land \deg(u) = d$$

$\lim_{n \to \infty} \Pr \left[ |E \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \epsilon \right] = 0$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - E[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$N_d = \sum_{v \in V} 1_{\{\text{deg}(v) = d\}}$$

$$= \mathbb{E} \left[ \left( \sum_{v \in V} 1_{\{\text{deg}(v) = d\}} \right)^2 \right]$$

$$= \mathbb{E} \left[ \sum_{v \in V} 1_{\{\text{deg}(v) = d\}}^2 \right] + \sum_{v \in V} \sum_{u \neq v} 1_{\{\text{deg}(v) = d\}} \cdot 1_{\{\text{deg}(u) = d\}}$$

Indicator RV $X$: $X^2 = X$, Lin. of Exp.

$$= \mathbb{E} \left[ \sum_{v \in V} 1_{\{\text{deg}(v) = d\}} \right] + \mathbb{E} \left[ \sum_{v \in V} \sum_{u \neq v} 1_{\{\text{deg}(v) = d\}} \cdot 1_{\{\text{deg}(u) = d\}} \right]$$

$$= \sum_{v \in V} \mathbb{E} \left[ 1_{\{\text{deg}(v) = d\}} \right] + \sum_{v \in V} \sum_{u \neq v} \mathbb{E} \left[ 1_{\{\text{deg}(v) = d\}} \cdot 1_{\{\text{deg}(u) = d\}} \right]$$

$$= \Pr[\text{deg}(v) = d]$$

Lim. of Exp.

$$= 1 \text{ iff } \text{deg}(v) = d \land \text{deg}(u) = d$$

$$= \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr \left[ |X - \mathbb{E}[X]| \geq b \right] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left( \frac{1}{n} N_d \right) - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left( \frac{1}{n} N_d \right)}{\varepsilon^2}$$

$$\text{Var} \left( \frac{1}{n} N_d \right) = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$N_d = \sum_{v \in V} 1\{\deg(v) = d\}$$

$$= \mathbb{E} \left[ \left( \sum_{v \in V} 1\{\deg(v) = d\} \right)^2 \right]$$

$$= \mathbb{E} \left[ \sum_{v \in V} \left( 1\{\deg(v) = d\} \right)^2 \right] + \sum_{v \in V} \sum_{u \neq v} 1\{\deg(v) = d\} \cdot 1\{\deg(u) = d\}$$

**Indicator RV $X$:** $X^2 = X$, Lin. of Exp.

$$= \sum_{v \in V} \mathbb{E} \left[ 1\{\deg(v) = d\} \right] + \sum_{v \in V} \sum_{u \neq v} \mathbb{E} \left[ 1\{\deg(v) = d\} \cdot 1\{\deg(u) = d\} \right]$$

$$= \Pr[\deg(v) = d]$$

$$= 1 \text{ iff } \deg(v) = d \land \deg(u) = d$$

$$= \Pr[\deg(v) = d \land \deg(u) = d]$$

$$= n \cdot \Pr[\deg(v) = d] + n(n - 1) \cdot \Pr[\deg(v) = d \land \deg(u) = d]$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left( \frac{1}{n} N_d \right) - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n}N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \left( \mathbb{E} \left[ \frac{1}{n} N_d \right] \right)^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)$$

$$= (n \Pr[\deg(v) = d])^2 \quad \text{(see Step 1)}$$

$$N_d = \sum_{v \in V} 1_{\{\deg(v) = d\}} = \mathbb{E} \left[ \left( \sum_{v \in V} 1_{\{\deg(v) = d\}} \right)^2 \right]$$

$$= \mathbb{E} \left[ \sum_{v \in V} 1_{\{\deg(v) = d\}} \right]^2 + \sum_{v \in V} \sum_{u \neq v} 1_{\{\deg(v) = d\}} \cdot 1_{\{\deg(u) = d\}}$$

Indicator RV $X$: $X^2 = X$,
Lin. of Exp.

$$= \sum_{v \in V} \mathbb{E} \left[ 1_{\{\deg(v) = d\}} \right] + \sum_{v \in V} \sum_{u \neq v} \mathbb{E} \left[ 1_{\{\deg(v) = d\}} \right] \cdot 1_{\{\deg(u) = d\}}$$

$$= 1 \text{ iff } \deg(v) = d \land \deg(u) = d$$

$$= \Pr[\deg(v) = \deg(u) = d]$$

$$= n \cdot \Pr[\deg(v) = d] + n(n-1) \cdot \Pr[\deg(v) = d \land \deg(u) = d]$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ \left( N_d \right)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)$$

$$= \frac{1}{n^2} \left( n \Pr[\text{deg}(v) = d] + n(n-1) \Pr[\text{deg}(v) = d \land \text{deg}(u) = d] - (n \Pr[\text{deg}(v) = d])^2 \right)$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$= \frac{1}{n^2} \left( n \Pr[\text{deg}(v) = d] \right.$$

$$+ n(n-1) \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \left( n \Pr[\text{deg}(v) = d] \right)^2 \bigg)$$

$$= \frac{1}{n} \Pr[\text{deg}(v) = d]$$

$$+ \frac{n-1}{n} \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d]^2$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n}N_d$

$$\text{Pr} \left[ |E \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d | \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n}N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n}N_d \right] = E \left[ \left( \frac{1}{n}N_d \right)^2 \right] - E \left[ \frac{1}{n}N_d \right]^2$$

$$= \frac{1}{n^2} \left( E \left[ (N_d)^2 \right] - E [N_d]^2 \right)$$

$$= \frac{1}{n^2} \left( n \text{Pr}[\deg(v) = d] + n(n - 1) \text{Pr}[\deg(v) = d \wedge \deg(u) = d] - (n \text{Pr}[\deg(v) = d])^2 \right)$$

$$= \frac{1}{n} \text{Pr}[\deg(v) = d] \leq 1$$

$$+ \frac{n - 1}{n} \text{Pr}[\deg(v) = d \wedge \deg(u) = d]$$

$$- \text{Pr}[\deg(v) = d]^2$$

$\textbf{Chebychev: } X$ finite variance, $b > 0$

$$\text{Pr}[|X - E[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$$\lim_{n \to \infty} \text{Pr} \left[ |E \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d | \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n}N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E}[(N_d)^2] - \mathbb{E}[N_d]^2 \right)$$

$$= \frac{1}{n^2} \left( n \Pr[\deg(v) = d] \right.$$  
$$+ n(n - 1) \Pr[\deg(v) = d \land \deg(u) = d]$$  
$$\left. - (n \Pr[\deg(v) = d])^2 \right)$$

$$= \frac{1}{n} \Pr[\deg(v) = d]$$  
$$\leq 1$$

$$+ \frac{n-1}{n} \Pr[\deg(v) = d \land \deg(u) = d] \leq 1$$

$$- \Pr[\deg(v) = d]^2$$

\[ \lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0 \]

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n} N_d$

Let $\epsilon > 0$. We want to show that

$$\Pr\left[ \left| \mathbb{E}\left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] \leq \frac{\text{Var}\left[ \frac{1}{n} N_d \right]}{\epsilon^2}$$

By the Chebyshev inequality, we know that

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

for any random variable $X$ with finite variance and $b > 0$. Thus, for our case, we have

$$\left| \mathbb{E}\left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \leq \epsilon \implies \frac{\text{Var}\left[ \frac{1}{n} N_d \right]}{\epsilon^2} \leq \lim_{n \to \infty} \Pr\left[ \left| \mathbb{E}\left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0$$

We have

$$\text{Var}\left[ \frac{1}{n} N_d \right] = \mathbb{E}\left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E}\left[ \frac{1}{n} N_d \right]^2$$

and

$$= \frac{1}{n^2} \left( \mathbb{E}\left[ (N_d)^2 \right] - \mathbb{E}[N_d]^2 \right)$$

$$= \frac{1}{n^2} (n \Pr[\deg(v) = d] + n(n - 1) \Pr[\deg(v) = d \land \deg(u) = d] - (n \Pr[\deg(v) = d])^2)$$

$$= \frac{1}{n} \Pr[\deg(v) = d]$$

$$\leq 1$$

$$+ \frac{n-1}{n} \Pr[\deg(v) = d \land \deg(u) = d] \leq 1$$

$$- \Pr[\deg(v) = d]^2$$

$$\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d]^2$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d]^2$$

$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
\]
\[
= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)
\]
\[
\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d]
- \Pr[\deg(v) = d] \Pr[\deg(v) = d]
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]

**Chebychev:** $X$ finite variance, $b > 0$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

\[
(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\]
Step 2: Concentration of $\frac{1}{n}N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n}N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n}N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n}N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n}N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \wedge \text{deg}(u) = d] - \Pr[\text{deg}(v) = d] \Pr[\text{deg}(v) = d]$$

$$\text{deg}(v) \overset{d}{=} \text{deg}(u)$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] = 0$$

**Chebychev**: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
\]

\[
= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)
\]

\[
\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d] - \Pr[\text{deg}(v) = d] \Pr[\text{deg}(v) = d]
\]

\[
= \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d] - \Pr[\text{deg}(v) = d] \Pr[\text{deg}(v) = d]
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]

\textbf{Chebychev:} $X$ finite variance, $b > 0$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

\[
(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\]

\[
\deg(v) \overset{d}{=} \deg(u)
\]
Step 2: Concentration of $\frac{1}{n} N_d$

Pr \left[ \left| E \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}

\text{Var} \left[ \frac{1}{n} N_d \right] = E \left[ \left( \frac{1}{n} N_d \right)^2 \right] - E \left[ \frac{1}{n} N_d \right]^2
= \frac{1}{n^2} \left( E \left[ (N_d)^2 \right] - E \left[ N_d \right]^2 \right)
\leq \frac{1}{n} + \text{Pr}[\text{deg}(v) = d \land \text{deg}(u) = d]
- \text{Pr}[\text{deg}(v) = d] \cdot \text{Pr}[\text{deg}(u) = d]
\leq \frac{1}{n} + \Pr[\text{deg}(v) = d] \cdot \Pr[\text{deg}(u) = d]

\lim_{n \to \infty} \text{Pr} \left[ \left| E \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0

\textbf{Chebyshev:} X \text{ finite variance, } b > 0
\text{Pr}[|X - E[X]| \geq b] \leq \text{Var}[X]/b^2

(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\text{deg}(v) \overset{d}{=} \text{deg}(u)
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr \left[ \text{deg}(v) = d \land \text{deg}(u) = d \right]$$

$$- \Pr \left[ \text{deg}(v) = d \right] \Pr \left[ \text{deg}(u) = d \right]$$

$$\text{deg}(v) \overset{d}{=} \text{deg}(u)$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$

---

Consider $\deg(u)$ and $\deg(v)$
Step 2: Concentration of $\frac{1}{n}N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d^2 \right) \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E}[N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d] - \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]$$

$$\text{deg}(v) \overset{d}{=} \text{deg}(u)$$

Couplings
- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$ independent
- $X_1, X_2 \sim \text{Ber}(p)$ dependent

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \variance(X)/b^2$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n}N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ (\frac{1}{n} N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]$$

$$\text{deg}(v) \overset{d}{=} \text{deg}(u)$$

\[ \lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0 \]

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

For $k$ independent random variables $a_1, a_2, \ldots, a_k$:

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$

\begin{itemize}
  \item independent
\end{itemize}
Step 2: Concentration of $\frac{1}{n}N_d$

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n}N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n}N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n}N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n}N_d \right]^2
\]

\[
= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)
\]

\[
\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d] - \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]
\]

\[\text{deg}(v) \overset{d}{=} \text{deg}(u)\]

\[\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] = 0\]

**Chebyshev:** $X$ finite variance, $b > 0$

\[\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}\]

\[\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j\]

**Couplings**
- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$ independent
- $X_1, X_2 \sim \text{Ber}(p)$
Step 2: Concentration of $1/n N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\deg(v) = d \wedge \deg(u) = d]$$

$$- \Pr[\deg(v) = d] \Pr[\deg(u) = d]$$

$$\text{deg}(v) \overset{d}{=} \text{deg}(u)$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

Couplings

- Consider $\deg(u)$ and $\deg(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$ (independent)
Step 2: Concentration of $\frac{1}{n}N_d$

\[
\Pr \left[ \left| \frac{1}{n} \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
\]

\[
= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)
\]

\[
\leq \frac{1}{n} + \Pr [\text{deg}(v) = d \land \text{deg}(u) = d]
\]

\[
- \Pr [\text{deg}(v) = d] \Pr [\text{deg}(u) = d]
\]

\[\text{deg}(v) \overset{d}{=} \text{deg}(u)\]

**Couplings**

- Consider \(\text{deg}(u)\) and \(\text{deg}(v)\)
- \(Y_1, Y_2 \sim \text{Bin}(n - 2, p)\)
- \(X_1, X_2 \sim \text{Ber}(p)\)

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]

**Chebychev:** \(X\) finite variance, \(b > 0\)

\[
\Pr \left[ |X - \mathbb{E}[X]| \geq b \right] \leq \frac{\text{Var}[X]}{b^2}
\]

\[
(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\]
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ (\frac{1}{n} N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$-\Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]$$

$\text{deg}(v) \overset{d}{=} \text{deg}(u)$

Couplings

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$

$$\text{deg}(v) \overset{d}{=} \text{deg}(u)$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

Consider $\text{deg}(u)$ and $\text{deg}(v)$
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
\]

\[
= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)
\]

\[
\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d] - \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]
\]

\[
\text{deg}(v) \overset{d}{=} \text{deg}(u)
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]

**Chebychev:** $X$ finite variance, $b > 0$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

\[
(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_ia_j
\]

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$ independent
- $X_1, X_2 \sim \text{Ber}(p)$ independent
- $X_1 + Y_1 \overset{\text{dep}}{\longrightarrow} X_1 + Y_2$
- $v \overset{\text{dep}}{\longrightarrow} u$

Step 2: Concentration of $\frac{1}{n}N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n}N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n}N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n}N_d \right)^2 \right] - \left( \mathbb{E} \left[ \frac{1}{n}N_d \right] \right)^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr \left[ \text{deg} (v) = d \land \text{deg} (u) = d \right]$$

$$- \Pr \left[ \text{deg} (v) = d \right] \Pr \left[ \text{deg} (u) = d \right]$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] = 0$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr \left[ |X - \mathbb{E}[X]| \geq b \right] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

### Couplings
- Consider $\text{deg} (u)$ and $\text{deg} (v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg} (v), \text{deg} (u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \frac{1}{n} \mathbb{E} \left[ N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]$$

$\text{deg}(v) \overset{d}{=} \text{deg}(u)$

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

$$\lim_{n \to \infty} \Pr \left[ \left| \frac{1}{n} \mathbb{E} \left[ N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n}N_d$

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n}N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n}N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n}N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n}N_d \right]^2
\]

\[
= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)
\]

\[
\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]
\]

\[
- \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]
\]

\[
\text{deg}(v) \overset{d}{=} \text{deg}(u)
\]

**Couplings**

- Consider deg($u$) and deg($v$)
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$ independent
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1+Y_1, X_1+Y_2)$

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n}N_d \right] - \frac{1}{n}N_d \right| \geq \varepsilon \right] = 0
\]

**Chebychev:** $X$ finite variance, $b > 0$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

\[
(\sum_i a_i)^2 = \sum_i a_i^2 + \sum \sum_{j \neq i} a_i a_j
\]
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ |\mathbb{E} \left( \frac{1}{n} N_d \right) - \frac{1}{n} N_d | \geq \varepsilon \right] \leq \frac{\text{Var} \left( \frac{1}{n} N_d \right)}{\varepsilon^2}$$

$$\text{Var} \left( \frac{1}{n} N_d \right) = \mathbb{E} \left( \left( \frac{1}{n} N_d \right)^2 \right) - \mathbb{E} \left( \frac{1}{n} N_d \right)^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left( (N_d)^2 \right) - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \wedge \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]$$

$$\text{deg}(v) \overset{d}{=} \text{deg}(u)$$

### Couplings

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

$$\lim_{n \to \infty} \Pr \left[ |\mathbb{E} \left( \frac{1}{n} N_d \right) - \frac{1}{n} N_d | \geq \varepsilon \right] = 0$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n} N_d$

\[ \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2} \]

\[ \text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \left( \mathbb{E} \left[ \frac{1}{n} N_d \right] \right)^2 \]

\[ = \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E}[N_d]^2 \right) \]

\[ \leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d] \Pr[\deg(u) = d] \]

\[ \text{deg}(v) \overset{d}{=} \text{deg}(u) \]

Couplings
- Consider \( \deg(u) \) and \( \deg(v) \)
- \( Y_1, Y_2 \sim \text{Bin}(n - 2, p) \)
- \( X_1, X_2 \sim \text{Ber}(p) \)
- \( (\deg(v), \deg(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2) \)

\[ \lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0 \]

Chebychev: \( X \) finite variance, \( b > 0 \)
\[ \Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2} \]

\[ (\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j \]
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \left( \mathbb{E} \left[ \frac{1}{n} N_d \right] \right)^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E}[N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d] \Pr[\deg(u) = d]$$

$\deg(v) \overset{d}{=} \deg(u)$

**Couplings**

- Consider $\deg(u)$ and $\deg(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\deg(v), \deg(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_{i \neq j} a_i a_j$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n} N_d$

Pr $\left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\epsilon^2}$

$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$

$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$

$\leq \frac{1}{n} + \text{Pr}[\text{deg}(v) = d \land \text{deg}(u) = d]$

$- \text{Pr}[\text{deg}(v) = d] \text{Pr}[\text{deg}(u) = d]$

$= \frac{1}{n} + \text{Pr} [X_1 + Y_1 = d \land X_1 + Y_2 = d]$

$\text{Chebychev: } X \text{ finite variance, } b > 0$

$\text{Pr}[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$

$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

$\text{deg}(v) \overset{d}{=} \text{deg}(u)$

$\lim_{n \to \infty} \text{Pr} \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0$
Step 2: Concentration of $\frac{1}{n}N_d$

$\Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$

$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$

$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$

$\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d] \Pr[\deg(u) = d]$

$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d] - \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d]$

$\text{deg}(v) \overset{d}{=} \text{deg}(u)$

**Couplings**

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

$\lim_{n \to \infty} \Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] = 0$

**Chebyshev:** $X$ finite variance, $b > 0$

$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$

$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ (\frac{1}{n} N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$- \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d]$$

$v \overset{\text{d}}{=} u$

Couplings

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{\text{d}}{=} (X_1 + Y_1, X_1 + Y_2)$

$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr\left[|\mathbb{E}\left[\frac{1}{n} N_d\right] - \frac{1}{n} N_d| \geq \varepsilon \right] \leq \frac{\text{Var}\left[\frac{1}{n} N_d\right]}{\varepsilon^2}$$

$$\text{Var}\left[\frac{1}{n} N_d\right] = \mathbb{E}\left[(\frac{1}{n} N_d)^2\right] - \mathbb{E}[\frac{1}{n} N_d]^2$$

$$= \frac{1}{n^2} \left(\mathbb{E}[N_d]^2 - \mathbb{E}[N_d]^2\right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$- \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$- \Pr[X_1 + Y_1 = d \land X_2 + Y_2 = d]$$


\[\lim_{n \to \infty} \Pr\left[|\mathbb{E}\left[\frac{1}{n} N_d\right] - \frac{1}{n} N_d| \geq \varepsilon \right] = 0\]

\[\text{Chebychev: } X \text{ finite variance, } b > 0 \]

\[\Pr[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2\]

\[(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j\]

\[\text{Consider } \text{deg}(u) \text{ and } \text{deg}(v)\]

\[Y_1, Y_2 \sim \text{Bin}(n - 2, p)\]

\[X_1, X_2 \sim \text{Ber}(p)\]

\[(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)\]
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| E \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \text{Var} \left[ \frac{1}{n} N_d \right] / \varepsilon^2$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = E \left[ \left( \frac{1}{n} N_d \right)^2 \right] - E \left[ \frac{1}{n} N_d \right]^2$$

$$\leq \frac{1}{n} \left( E \left[ (N_d)^2 \right] - E [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \text{Pr}[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \text{Pr}[\text{deg}(v) = d] \text{Pr}[\text{deg}(u) = d]$$

$$= \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d]$$

$$- \text{Pr}[X_1 + Y_1 = d] \text{Pr}[X_2 + Y_2 = d]$$

Couplings

- Consider $\text{deg}(u)$ and $\text{deg}(v)$
- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\text{deg}(v), \text{deg}(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \neg B]$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - E[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\deg(v) = d \wedge \deg(u) = d]$$

$$- \Pr[\deg(v) = d] \Pr[\deg(u) = d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \wedge X_1 + Y_2 = d]$$

$$- \Pr[X_1 + Y_1 = d] \Pr[X_2 + Y_2 = d]$$

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \wedge X_1 + Y_2 = d]$$

$$- \Pr[X_1 + Y_1 = d \wedge X_2 + Y_2 = d]$$

$$\wedge (X_1 + Y_1 \neq d \vee X_2 + Y_2 \neq d)$$

Couplings

- Consider $\deg(u)$ and $\deg(v)$
- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
- $(\deg(v), \deg(u)) \overset{d}{=} (X_1 + Y_1, X_1 + Y_2)$

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \wedge \overline{B}]$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

Fréchet: $\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \epsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\epsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
\]

\[
\leq \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right)
\]

Chebychev: $X$ finite variance, $b > 0$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \neg B]$

\[
\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0
\]
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
\]
\[
= \frac{1}{n^2} (\mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2)
\]
\[
\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d] - \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d]
\]
\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]

**Chebychev:** $X$ finite variance, $b > 0$
\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

\[
(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\]

**Fréchet:** $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \wedge \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d] \quad \text{deg}(v) \overset{d}{=} \text{deg}(u)$$

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \wedge X_1 + Y_2 = d \wedge X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \wedge X_1 + Y_2 = d \wedge X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 + Y_1 = d \wedge X_1 + Y_2 = d \wedge X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]$$

- $Y_1, Y_2 \sim \text{Bin}(n-2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$ independent

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \wedge \bar{B}]$

Law of total probability

$$\lim_{n \to \infty} \Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d]$$

$$- \Pr[\deg(v) = d] \Pr[\deg(u) = d] \quad \text{deg}(v) \overset{d}{=} \text{deg}(u)$$

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]$$

$$\leq 1 \quad \text{Law of total probability}$$

- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$ independent
- $X_1, X_2 \sim \text{Ber}(p)$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

Frechét: $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$
Step 2: Concentration of $\frac{1}{n} N_d$

\[ \Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] \leq \frac{\operatorname{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2} \]

\[ \operatorname{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 \]
\[ = \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right) \]
\[ \leq \frac{1}{n} + \Pr [\deg(v) = d \land \deg(u) = d] - \Pr [\deg(v) = d] \Pr [\deg(u) = d] \quad \text{deg}(v) \overset{d}{=} \text{deg}(u) \]
\[ \leq \frac{1}{n} + \Pr [X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d] \]
\[ = \frac{1}{n} + \Pr [X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr [X_1 = 0] \]
\[ + \Pr [X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr [X_1 = 1] \]
\[ \leq \frac{1}{n} + \Pr [X_1 = 1] \leq 1 \]

Fréchet: \( \Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}] \)

Law of total probability

\[ \lim_{n \to \infty} \Pr \left[ |\mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d | \geq \varepsilon \right] = 0 \]

Chebychev: \( X \) finite variance, \( b > 0 \)
\[ \Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\operatorname{Var}[X]}{b^2} \]

\( (\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j \)
Step 2: Concentration of $\frac{1}{n} N_d$

\[ \text{Pr}\left[\left|\frac{1}{n} \mathbb{E}\left(\frac{1}{n} N_d\right) - \frac{1}{n} N_d\right| \geq \varepsilon \right] \leq \frac{\text{Var}\left[\frac{1}{n} N_d\right]}{\varepsilon^2} \]

\[ \text{Var}\left[\frac{1}{n} N_d\right] = \mathbb{E}\left[\left(\frac{1}{n} N_d\right)^2\right] - \left(\mathbb{E}\left[\frac{1}{n} N_d\right]\right)^2 \]

\[ = \frac{1}{n^2} \left(\mathbb{E}\left[(N_d)^2\right] - \mathbb{E}[N_d]^2\right) \]

\[ \leq \frac{1}{n} + \text{Pr}[\text{deg}(v) = d \land \text{deg}(u) = d] \]

\[ - \text{Pr}[\text{deg}(v) = d] \text{Pr}[\text{deg}(u) = d] \quad \text{deg}(v) \overset{d}{=} \text{deg}(u) \]

\[ \leq \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d] \]

\[ = \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \text{Pr}[X_1 = 0] \]

\[ + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \text{Pr}[X_1 = 1] \]

\[ \leq \frac{1}{n} + \text{Pr}[X_1 = 1] \leq 1 \]

\[ \text{Chebychev: } X \text{ finite variance, } b > 0 \quad \text{Pr}[|X - \mathbb{E}[X]| \geq b] \leq \text{Var}[X]/b^2 \]

\[ (\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j \]

\[ \text{Fréchet: } \text{Pr}[A] - \text{Pr}[B] \leq \text{Pr}[A \land \bar{B}] \]

\[ \lim_{n \to \infty} \text{Pr}\left[\left|\frac{1}{n} \mathbb{E}\left(\frac{1}{n} N_d\right) - \frac{1}{n} N_d\right| \geq \varepsilon \right] = 0 \]

\[ \text{Law of total probability} \]

\[ Y_1, Y_2 \sim \text{Bin}(n - 2, p) \]

\[ X_1, X_2 \sim \text{Ber}(p) \]
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ (\frac{1}{n} N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d] \quad \text{deg}(v) \overset{d}{=} \text{deg}(u)$$

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0]$$

$$+ \Pr[X_1 = 1]$$

$$\leq 1$$

Chebychev: $X$ finite variance, $b > 0$
$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

Law of total probability

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

$X_1, X_2 \sim \text{Ber}(p)$

$Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d]$$

$$- \Pr[\deg(v) = d] \Pr[\deg(u) = d]$$

$$\deg(v) \overset{d}{=} \deg(u)$$

$$\leq \frac{1}{n} + \Pr[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[ X_1 = 0]$$

$$+ \Pr[ X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[ X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[ Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0]$$

$$+ \Pr[ X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[ Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0]$$

Law of total probability

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$\text{Var}[1/n N_d] = \mathbb{E}[(1/n N_d)^2] - \mathbb{E}[1/n N_d]^2$

$\text{Law of total probability}$

$\text{independent}$

$Y_1, Y_2 \sim \text{Bin}(n-2, p)$

$X_1, X_2 \sim \text{Ber}(p)$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \left( \mathbb{E} \left[ \frac{1}{n} N_d \right] \right)^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d]$$

$$- \Pr[\deg(v) = d] \Pr[\deg(u) = d] \quad \text{deg(v)} \overset{d}{=} \text{deg(u)}$$

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0]$$

$$+ \Pr[X_1 = 1] \quad \Rightarrow X_2 = 1$$

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_ia_j$$

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \overline{B}]$

Law of total probability

- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$ independent
- $X_1, X_2 \sim \text{Ber}(p)$ independent
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ (\frac{1}{n} N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
\]

\[
\leq \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)
\]

\[
\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d] \Pr[\deg(u) = d] \text{ deg}(v) = \text{ deg}(u)
\]

\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land \underline{X_1 + Y_2 = d} \land X_2 + Y_2 \neq d]
\]

\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]
\]

\[
+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]
\]

\[
\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Rightarrow X_2 = 1
\]

\[
\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 = 1 | X_1 = 0] + \Pr[X_1 = 1]
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]

**Chebychev**: $X$ finite variance, $b > 0$

\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

\[
(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\]

**Fréchet**: $\Pr[A] - \Pr[B] \leq \Pr[A \land \tilde{B}]$

\[
\text{Law of total probability}
\]

- $Y_1, Y_2 \sim \text{Bin}(n - 2, p)$
- $X_1, X_2 \sim \text{Ber}(p)$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr\left[|E\left[\frac{1}{n} N_d\right] - \frac{1}{n} N_d| \geq \varepsilon\right] \leq \frac{\text{Var}\left[\frac{1}{n} N_d\right]}{\varepsilon^2}$$

$$\text{Var}\left[\frac{1}{n} N_d\right] = E\left[(\frac{1}{n} N_d)^2\right] - E\left[\frac{1}{n} N_d\right]^2$$

$$= \frac{1}{n^2} \left(E\left((N_d)^2\right) - E[N_d]^2\right)$$

$$\leq \frac{1}{n} + \Pr[\text{deg}(v) = d \land \text{deg}(u) = d]$$

$$- \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d] \quad \text{deg}(v) \overset{d}\equiv \text{deg}(u)$$

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 \neq d | X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 = 1] \quad \Rightarrow X_2 = 1$$

$$= \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 = 1 | X_1 = 0] + \Pr[X_1 = 1]$$

$$\leq 1$$

$$\text{Law of total probability}$$

$$\text{Frechet: } \Pr[A] - \Pr[B] \leq \Pr[A \land \bar{B}]$$

$$\text{Chebychev: } X \text{ finite variance, } b > 0$$

$$\Pr[|X - E[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

$$Y_1, Y_2 \sim \text{Bin}(n - 2, p)$$

$$X_1, X_2 \sim \text{Ber}(p)$$

$\text{independent}$

$$\lim_{n \to \infty} \Pr\left[|E\left[\frac{1}{n} N_d\right] - \frac{1}{n} N_d| \geq \varepsilon\right] = 0$$
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} \left[ \Pr[\text{deg}(v) = d \land \text{deg}(u) = d] \right. - \Pr[\text{deg}(v) = d] \Pr[\text{deg}(u) = d] \; \text{deg}(v) = \text{deg}(u)$$

$$\leq \frac{1}{n} \left[ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d] \right.$$  

$$+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} \left[ \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \right.$$  

$$+ \Pr[X_1 = 1] \Rightarrow X_2 = 1$$

$$\leq 1 \begin{cases} 
Y_1, Y_2 \sim \text{Bin}(n-2, p) \end{cases}$$

$$X_1, X_2 \sim \text{Ber}(p)$$

Law of total probability

$$\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0$$

Chebychev: $X$ finite variance, $b > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \overline{B}]$
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr\left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ (\frac{1}{n} N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2
\]

\[
= \frac{1}{n^2} \left( \mathbb{E} [(N_d)^2] - \mathbb{E}[N_d]^2 \right)
\]

\[
\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d] \Pr[\deg(u) = d]
\]

\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]
\]

\[
= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]
\]

\[
+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]
\]

\[
\leq \frac{1}{n} + \text{Pr}[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] + \text{Pr}[X_1 = 1]
\]

\[
\leq \frac{1}{n} + \text{Pr}[X_2 = 1]
\]

\[
\Rightarrow X_2 = 1
\]

\[
\Rightarrow Y_1, Y_2 \sim \text{Bin}(n - 2, p)
\]

\[
X_1, X_2 \sim \text{Ber}(p)
\]

\[
\text{Law of total probability}
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]

\[
\text{Chebychev: } X \text{ finite variance, } b > 0 \Rightarrow \Pr[\left| X - \mathbb{E}[X] \right| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

\[
(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\]
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 \leq \frac{1}{n} + 2p$$

$$= \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} [N_d]^2 \right)$$

$$\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d] \Pr[\deg(u) = d]$$

$$\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]$$

$$= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0]$$

$$+ \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]$$

$$\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Rightarrow X_2 = 1$$

$$\leq 1$$

$$\Rightarrow X_2 = 1$$

$$= \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 = 1 | X_1 = 0] + \Pr[X_1 = 1] \leq \frac{1}{n} + \Pr[X_2 = 1] + \Pr[X_1 = 1]$$

$$\Rightarrow X_2 = 1$$
Step 2: Concentration of $\frac{1}{n}N_d$

\[ \text{Pr} \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2} \]

\[ \text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 \right] \leq \frac{1}{n} + 2p = \frac{1}{n} + 2\frac{\varepsilon}{n} \]

\[ = \frac{1}{n^2} \left( \mathbb{E} \left[ (N_d)^2 \right] - \mathbb{E} \left[ N_d \right]^2 \right) \]

\[ \leq \frac{1}{n} + \text{Pr}[\deg(v) = d \land \deg(u) = d] \]

\[ - \text{Pr}[\deg(v) = d] \cdot \text{Pr}[\deg(u) = d] \quad \text{deg}(v) \neq \text{deg}(u) \]

\[ \leq \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d] \]

\[ = \frac{1}{n} + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d \mid X_1 = 0] \cdot \text{Pr}[X_1 = 0] \]

\[ + \text{Pr}[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d \mid X_1 = 1] \cdot \text{Pr}[X_1 = 1] \]

\[ \leq \frac{1}{n} + \text{Pr}[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d \mid X_1 = 0] \]

\[ + \text{Pr}[X_1 = 1] \quad \Rightarrow X_2 = 1 \quad \text{independent} \]

\[ = \frac{1}{n} + \text{Pr}[Y_1 = d \land Y_2 = d \land X_2 = 1 \mid X_1 = 0] + \text{Pr}[X_1 = 1] \leq \frac{1}{n} + \text{Pr}[X_2 = 1] + \text{Pr}[X_1 = 1] \]

\[ \text{Chebychev: } X \text{ finite variance, } b > 0 \]

\[ \text{Pr}[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2} \]

\[ (\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_ia_j \]

\[ \text{Fréchet: } \text{Pr}[A] - \text{Pr}[B] \leq \text{Pr}[A \land \overline{B}] \]

\[ \text{Law of total probability} \]

\[ \text{Pr}[Y_1, Y_2 \sim \text{Bin}(n - 2, p)] \]

\[ \text{independent} \]

\[ \text{Pr}[X_1, X_2 \sim \text{Ber}(p)] \]
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2}
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ (\frac{1}{n} N_d)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 \leq \frac{1}{n} + 2p = \frac{1}{n} + 2c_n \xrightarrow{n \to \infty} 0
\]

\[
\leq \frac{1}{n} + \Pr[\deg(v) = d \land \deg(u) = d] - \Pr[\deg(v) = d] \Pr[\deg(u) = d] \quad \text{deg}(v) \overset{d}{=} \text{deg}(u)
\]

\[
\leq \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d]
\]

\[
= \frac{1}{n} + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Pr[X_1 = 0] + \Pr[X_1 + Y_1 = d \land X_1 + Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 1] \Pr[X_1 = 1]
\]

\[
\leq \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 + Y_2 \neq d | X_1 = 0] \Rightarrow X_2 = 1
\]

\[
= \frac{1}{n} + \Pr[Y_1 = d \land Y_2 = d \land X_2 = 1 | X_1 = 0] + \Pr[X_1 = 1] \leq \frac{1}{n} + \Pr[X_2 = 1] + \Pr[X_1 = 1]
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]

Chebychev: $X$ finite variance, $b > 0$
\[
\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}
\]

\[
\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\]

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \land \overline{B}]$

Law of total probability
\[
\Pr[Y_1 = d \land Y_2 = d \land X_2 = 1 | X_1 = 0] \leq 1
\]

$Y_1, Y_2 \sim \text{Bin}(n-2, p)$ independent

$X_1, X_2 \sim \text{Ber}(p)$ independent
Step 2: Concentration of $\frac{1}{n} N_d$

$$\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2} \xrightarrow{n \to \infty} 0$$

$$\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 \leq \frac{1}{n} + 2p = \frac{1}{n} + 2 \frac{\varepsilon}{n} \xrightarrow{n \to \infty} 0$$

**Chebychev:** $X$ finite variance, $b > 0$

$$\Pr \left[ |X - \mathbb{E}[X]| \geq b \right] \leq \frac{\text{Var}[X]}{b^2}$$

$$(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j$$

**Fréchet:** $\Pr[A] - \Pr[B] \leq \Pr[A \wedge \neg B]$
Step 2: Concentration of $\frac{1}{n} N_d$

\[
\Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] \leq \frac{\text{Var} \left[ \frac{1}{n} N_d \right]}{\varepsilon^2} \xrightarrow{n \to \infty} 0
\]

\[
\text{Var} \left[ \frac{1}{n} N_d \right] = \mathbb{E} \left[ \left( \frac{1}{n} N_d \right)^2 \right] - \mathbb{E} \left[ \frac{1}{n} N_d \right]^2 \leq \frac{1}{n} + 2p = \frac{1}{n} + 2\frac{\varepsilon}{n} \xrightarrow{n \to \infty} 0
\]

Chebychev: $X$ finite variance, $b > 0$

\[
\Pr \left[ \left| X - \mathbb{E}[X] \right| \geq b \right] \leq \frac{\text{Var}[X]}{b^2}
\]

Fréchet: $\Pr[A] - \Pr[B] \leq \Pr[A \wedge \neg B]$

\[
\left( \sum_i a_i \right)^2 = \sum_i a_i^2 + \sum_i \sum_{j \neq i} a_i a_j
\]

\[
\lim_{n \to \infty} \Pr \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \varepsilon \right] = 0
\]
Application: ER – Degree Distribution

**Theorem:** Consider a $G(n, p)$ with $p = c/n$ for constant $c > 0$. For $\lambda = -n \log(1 - p)$, let $X \sim \text{Pois}(\lambda)$. Then for all $d > 0$ and every $\epsilon > 0$ we have

$$
\lim_{n \to \infty} \text{Pr} \left[ \left| \text{Pr}[X = d] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0.
$$

**Proof**

- **Step 1:** $\Pr[X = d]$ is close to the expectation of $\frac{1}{n} N_d$

$$
\lim_{n \to \infty} \left| \text{Pr}[X = d] - \mathbb{E} \left[ \frac{1}{n} N_d \right] \right| = 0 \quad \checkmark
$$

- **Step 2:** $\frac{1}{n} N_d$ is concentrated (via Chebychev)

$$
\lim_{n \to \infty} \text{Pr} \left[ \left| \mathbb{E} \left[ \frac{1}{n} N_d \right] - \frac{1}{n} N_d \right| \geq \epsilon \right] = 0 \quad \checkmark
$$
Concentration Bounds So Far

**Definition:** A **concentration inequality** bounds the probability of a random variable to deviate from a given value (typically its expectation) by a certain amount.
**Definition**: A **concentration inequality** bounds the probability of a random variable to deviate from a given value (typically its expectation) by a certain amount.

**Markov**
- based on expectation (first moment)
- $X$ non-negative random variable and $a > 0$

$$\Pr[X \geq a] \leq \mathbb{E}[X]/a$$
Concentration Bounds So Far

**Definition:** A **concentration inequality** bounds the probability of a random variable to deviate from a given value (typically its expectation) by a certain amount.

**Markov**
- based on expectation (first moment)
- $X$ non-negative random variable and $a > 0$
  - $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$
- tight
Concentration Bounds So Far

**Definition:** A concentration inequality bounds the probability of a random variable to deviate from a given value (typically its expectation) by a certain amount.

**Markov**
- based on expectation (first moment)
- $X$ non-negative random variable and $a > 0$
  \[ \Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a} \]
- tight

**Chebychev**
- based on variance (second moment)
- $X$ random variable with finite variance and $b > 0$
  \[ \Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2} \]
Concentration Bounds So Far

**Definition:** A concentration inequality bounds the probability of a random variable to deviate from a given value (typically its expectation) by a certain amount.

**Markov**
- based on expectation (first moment)
- $X$ non-negative random variable and $a > 0$
  $$\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$$
- tight

**Chebychev**
- based on variance (second moment)
- $X$ random variable with finite variance and $b > 0$
  $$\Pr[|X - \mathbb{E}[X]| \geq b] \leq \frac{\text{Var}[X]}{b^2}$$
- tight (stated without proof)
Concentration Bounds So Far

**Definition:** A **concentration inequality** bounds the probability of a random variable to deviate from a given value (typically its expectation) by a certain amount.

**Markov**
- based on expectation (first moment)
- $X$ non-negative random variable and $a > 0$
  \[ \Pr[X \geq a] \leq \frac{E[X]}{a} \]
- tight

**Chebyshev**
- based on variance (second moment)
- $X$ random variable with finite variance and $b > 0$
  \[ \Pr[|X - E[X]| \geq b] \leq \frac{Var[X]}{b^2} \]
- tight (stated without proof)

*Can we utilize higher-order moments for even stronger bounds?*
Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$
Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$.
- We can capture all moments of $X$ using a single function.

**Definition:** For a random variable $X$ the **moment generating function** is $M_X(t) = \mathbb{E}[e^{tX}]$.
Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$
- We can capture all moments of $X$ using a single function

**Definition**: For a random variable $X$ the **moment generating function** is $M_X(t) = \mathbb{E}[e^{tX}]$

Where the name comes from: For the $n$-th derivative $M_X^{(n)}(t)$ we have $M_X^{(n)}(0) = \mathbb{E}[X^n]$ (assuming the function exists in a neighborhood around 0)

Looks scary, but is again just $\mathbb{E}[f(X)]$ for $f(X) = e^{tX}$
Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$
- We can capture all moments of $X$ using a single function

**Definition:** For a random variable $X$ the **moment generating function** is $M_X(t) = \mathbb{E}[e^{tX}]$

Where the name comes from: For the $n$-th derivative $M_X^{(n)}(t)$ we have $M_X^{(n)}(0) = \mathbb{E}[X^n]$ (assuming the function exists in a neighborhood around 0)

**Theorem:** For independent random variables $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.
Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$
- We can capture all moments of $X$ using a single function

**Definition:** For a random variable $X$ the **moment generating function** is $M_X(t) = \mathbb{E}[e^{tX}]$

- Where the name comes from: For the $n$-th derivative $M_X^{(n)}(t)$ we have $M_X^{(n)}(0) = \mathbb{E}[X^n]$ (assuming the function exists in a neighborhood around 0)

**Theorem:** For independent random variables $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

**Proof**

$M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}]$
Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$
- We can capture all moments of $X$ using a single function

**Definition:** For a random variable $X$ the moment generating function is $M_X(t) = \mathbb{E}[e^{tX}]$

Where the name comes from: For the $n$-th derivative $M_X^{(n)}(t)$ we have $M_X^{(n)}(0) = \mathbb{E}[X^n]$ (assuming the function exists in a neighborhood around 0)

**Theorem:** For independent random variables $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

**Proof**

$M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}]$
The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$

- We can capture \textit{all} moments of $X$ using a single function

\textbf{Definition:} For a random variable $X$ the \textbf{moment generating function} is

$$M_X(t) = \mathbb{E}[e^{tX}]$$

- Where the name comes from: For the $n$-th derivative $M^{(n)}_X(t)$ we have $M^{(n)}_X(0) = \mathbb{E}[X^n]$ (assuming the function exists in a neighborhood around 0)

\textbf{Theorem:} For independent random variables $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

\textbf{Proof} $M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}]$

Looks scary, but is again just $\mathbb{E}[f(X)]$ for $f(X) = e^{tX}$.
Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $E[X^n]$
- We can capture all moments of $X$ using a single function

**Definition:** For a random variable $X$, the **moment generating function** is $M_X(t) = E[e^{tX}]$

Where the name comes from: For the $n$-th derivative $M_X^{(n)}(t)$ we have $M_X^{(n)}(0) = E[X^n]$ (assuming the function exists in a neighborhood around 0)

**Theorem:** For independent random variables $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

**Proof**

\[ M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX} \cdot e^{tY}] = E[e^{tX}] \cdot E[e^{tY}] = M_X(t) \cdot M_Y(t) \]

Looks scary, but is again just $E[f(X)]$ for $f(X) = e^{tX}$
Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$.
- We can capture all moments of $X$ using a single function $M_X(t) = \mathbb{E}[e^{tX}]$.

**Definition:** For a random variable $X$ the **moment generating function** is $M_X(t) = \mathbb{E}[e^{tX}]$.

Where the name comes from: For the $n$-th derivative $M_X^{(n)}(t)$ we have $M_X^{(n)}(0) = \mathbb{E}[X^n]$.

**Theorem:** For independent random variables $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

**Proof**

$M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}] = M_X(t) \cdot M_Y(t)$.

**Concentration Inequality**

**Theorem (Chernoff Bounds):** Let $X$ be a random variable and $a > 0$.
Then, $\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta}$ and $\Pr[X \leq a] \leq \min_{t<0} \mathbb{E}[e^{tX}]/e^{ta}$.

Had his 100th birthday in 2023! Thought the bound (now named after him) to be so trivial that he didn’t mention that it actually came from Herman Rubin.

Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$.
- We can capture all moments of $X$ using a single function $\mathbb{E}[e^{tX}]$.

**Definition:** For a random variable $X$ the **moment generating function** is $M_X(t) = \mathbb{E}[e^{tX}]$.

Where the name comes from: For the $n$-th derivative $M_X^{(n)}(t)$ we have $M_X^{(n)}(0) = \mathbb{E}[X^n]$ (assuming the function exists in a neighborhood around 0).

**Theorem:** For independent random variables $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

**Proof**

$M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}] = M_X(t) \cdot M_Y(t)$ $\checkmark$

**Concentration Inequality**

**Theorem (Chernoff Bounds):** Let $X$ be a random variable and $a > 0$.

Then, $\Pr[X \geq a] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$ and $\Pr[X \leq a] \leq \min_{t<0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$.

**Proof**

For all $t > 0$: $\Pr[X \geq a] = \Pr[e^{tX} \geq e^{ta}]$.
The \( n \)-th raw moment of a random variable \( X \) is \( \mathbb{E}[X^n] \).

We can capture \textit{all} moments of \( X \) using a single function

\textbf{Definition}: For a random variable \( X \) the \textbf{moment generating function} is \( M_X(t) = \mathbb{E}[e^{tX}] \).

Where the name comes from: For the \( n \)-th derivative \( M_X^{(n)}(t) \) we have \( M_X^{(n)}(0) = \mathbb{E}[X^n] \)

\((\text{assuming the function exists in a neighborhood around } 0)\)

\textbf{Theorem}: For independent random variables \( X, Y \): \( M_{X+Y}(t) = M_X(t) \cdot M_Y(t) \).

\textbf{Proof} \quad M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}] = M_X(t) \cdot M_Y(t) \checkmark

\textbf{Concentration Inequality}

\textbf{Theorem (Chernoff Bounds)}: Let \( X \) be a random variable and \( a > 0 \). Then, \( \Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta} \) and \( \Pr[X \leq a] \leq \min_{t<0} \mathbb{E}[e^{tX}]/e^{ta} \).

\textbf{Proof} \quad \text{for all } t > 0:\Pr[X \geq a] = \Pr[e^{tX} \geq e^{ta}] \leq \mathbb{E}[e^{tX}]/e^{ta} \checkmark

Markov: \( X \) non-negative, \( b > 0 \):
\( \Pr[X \geq b] \leq \mathbb{E}[X]/b. \)
Another Moment Please

- The \( n \)-th raw moment of a random variable \( X \) is \( \mathbb{E}[X^n] \)
- We can capture all moments of \( X \) using a single function

**Definition:** For a random variable \( X \) the **moment generating function** is \( M_X(t) = \mathbb{E}[e^{tX}] \)

Where the name comes from: For the \( n \)-th derivative \( M_X^{(n)}(t) \) we have \( M_X^{(n)}(0) = \mathbb{E}[X^n] \)

**Theorem:** For independent random variables \( X, Y \): \( M_{X+Y}(t) = M_X(t) \cdot M_Y(t) \).

**Proof**

\[
M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}] = M_X(t) \cdot M_Y(t) \checkmark
\]

**Concentration Inequality**

**Theorem (Chernoff Bounds):** Let \( X \) be a random variable and \( a > 0 \).
Then, \( \Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta} \) and \( \Pr[X \leq a] \leq \min_{t<0} \mathbb{E}[e^{tX}]/e^{ta} \).

**Proof**

for all \( t > 0 \): \( \Pr[X \geq a] = \Pr[e^{tX} \geq e^{ta}] \leq \mathbb{E}[e^{tX}]/e^{ta} \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta} \checkmark
\]

Markov: \( X \) non-negative, \( b > 0 \):
\[
\Pr[X \geq b] \leq \mathbb{E}[X]/b.
\]

Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $\mathbb{E}[X^n]$
- We can capture all moments of $X$ using a single function

**Definition:** For a random variable $X$ the **moment generating function** is $M_X(t) = \mathbb{E}[e^{tX}]$

Where the name comes from: For the $n$-th derivative $M_X^{(n)}(t)$ we have $M_X^{(n)}(0) = \mathbb{E}[X^n]$

(assuming the function exists in a neighborhood around 0)

**Theorem:** For independent random variables $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

**Proof** $M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}] = M_X(t) \cdot M_Y(t)$ ✓

**Concentration Inequality**

**Theorem (Chernoff Bounds):** Let $X$ be a random variable and $a > 0$.
Then, $\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta}$ and $\Pr[X \leq a] \leq \min_{t<0} \mathbb{E}[e^{tX}]/e^{ta}$.

**Proof** for all $t > 0$: $\Pr[X \geq a] = \Pr[e^{tX} \geq e^{ta}] \leq \mathbb{E}[e^{tX}]/e^{ta}$
for all $t < 0$: analogous. ✓

Markov: $X$ non-negative, $b > 0$: $\Pr[X \geq b] \leq \mathbb{E}[X]/b$.  


Had his 100th birthday in 2023! Thought the bound (now named after him) to be so trivial that he didn’t mention that it actually came from Herman Rubin.
Another Moment Please

- The $n$-th raw moment of a random variable $X$ is $E[X^n]$
- We can capture all moments of $X$ using a single function

**Definition:** For a random variable $X$ the **moment generating function** is $M_X(t) = E[e^{tX}]$

- Where the name comes from: For the $n$-th derivative $M_X^{(n)}(0) = E[X^n]$

**Theorem:** For independent random variables $X, Y$: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.  

**Proof**

$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX} \cdot e^{tY}] = E[e^{tX}] \cdot E[e^{tY}] = M_X(t) \cdot M_Y(t) \checkmark$

**Concentration Inequality**

**Theorem (Chernoff Bounds):** Let $X$ be a random variable and $a > 0$. Then, $Pr[X \geq a] \leq \min_{t>0} \frac{E[e^{tX}]}{e^{ta}}$ and $Pr[X \leq a] \leq \min_{t<0} \frac{E[e^{tX}]}{e^{ta}}$.

**Proof** for all $t > 0$: $Pr[X \geq a] = Pr[e^{tX} \geq e^{ta}] \leq \frac{E[e^{tX}]}{e^{ta}} \leq \min_{t>0} \frac{E[e^{tX}]}{e^{ta}} \checkmark$

for all $t < 0$: analogous. $\checkmark$

Get bounds for specific random variables by finding a good $t$!

Looks scary, but is again just $E[f(X)]$ for $f(X) = e^{tX}$.
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{\mathbb{E}[X]}.$$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$
\text{Pr}[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left( \frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}} \right) \mathbb{E}[X].
$$

**Proof**

Chernoff: Random variable $X$ and $a > 0$:

$\text{Pr}[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta}$.

Mom. Gen. Function: $M_X(t) = \mathbb{E}[e^{tX}]$.
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^{\varepsilon}}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{\mathbb{E}[X]}.$$ 

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}] / e^{ta}.$$ 

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}}\right)\mathbb{E}[X].$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$M_{X_i}(t) = \mathbb{E}[e^{tX_i}]$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}.$$

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left( \frac{e^{\varepsilon}}{(1 + \varepsilon)^{(1+\varepsilon)}} \right)^{\mathbb{E}[X]}.$$ 

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}$$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t > 0} \mathbb{E}[e^{tX}] / e^{ta}.$$ 

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}}\right) \mathbb{E}[X].$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}$$

$$= (1 - p) + pe^t$$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t > 0} \mathbb{E}[e^{tX}] / e^{ta}.$$  

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^{\varepsilon}}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{\mathbb{E}[X]}.$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}$$

$$= (1 - p) + pe^t = 1 + (e^t - 1)p$$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}] / e^{ta}.$$ 

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}}\right)^{\mathbb{E}[X]}.$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot 0} + \Pr[X_i = 1] \cdot e^{t\cdot 1}$$

$$= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta}.$$

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{\mathbb{E}[X]}.$$

**Proof**
Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}$$

$$= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}$$

$$M_X(t) = M_{\sum X_i}(t)$$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta}.$$

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}}\right)^{\mathbb{E}[X]}.$$ 

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}$$

$$1 + x \leq e^x$$

$$M_X(t) = M\sum X_i(t) = \prod_{i=1}^n M_{X_i}(t)$$

**Chernoff:** Random variable $X$ and $a > 0$: 

$$\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta}.$$ 

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$

**Moment Addition:** Independent $X, Y$: 

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t).$$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n,p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^{\varepsilon}}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{\mathbb{E}[X]}.$$  

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p},$$

$$M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t-1)p}.$$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$
\[
\Pr[X \geq (1 + \varepsilon) \mathbb{E}[X]] \leq \left(\frac{e^{\varepsilon}}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{\mathbb{E}[X]}.
\]

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$
\[
M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot0} + \Pr[X_i = 1] \cdot e^{t\cdot1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}
\]
\[
M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t-1)p} = e^{(e^t-1)\cdot np}
\]

**Chernoff:** Random variable $X$ and $a > 0$:
\[
\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}] / e^{ta}.
\]

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$

**Moment Addition:** Independent $X, Y$:
\[
M_{X+Y}(t) = M_X(t) \cdot M_Y(t).
\]
Application: Binomial Distribution

**Theorem**  Let $X \sim \text{Bin}(n, p)$. Then for any $\epsilon > 0$

$$\Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \left(\frac{e^\epsilon}{(1 + \epsilon)^{(1+\epsilon)}}\right)^{\mathbb{E}[X]}.$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

$$1 + x \leq e^x$$

$$M_X(t) = M \sum X_i(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t-1)p} = e^{(e^t-1)\cdot np} = e^{(e^t-1)\mathbb{E}[X]}$$

**Chernoff**: Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}] / e^{ta}.$$ 

**Mom. Gen. Function**: $M_X(t) = \mathbb{E}[e^{tX}]$

**Moment Addition**: Independent $X, Y$:

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t).$$
### Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}}\right)^{\mathbb{E}[X]}.$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot0} + \Pr[X_i = 1] \cdot e^{t\cdot1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

$$M_X(t) = M\sum X_i(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t-1)p} = e^{(e^t-1)\cdot np} = e^{(e^t-1)\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]]$$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\epsilon > 0$

$$\Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \left( \frac{e^\epsilon}{(1 + \epsilon)^{(1+\epsilon)}} \right)^{\mathbb{E}[X]}.$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}$$

$$= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

$$M_X(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t-1)p} = e^{(e^t-1)np}$$

$$Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\epsilon)\mathbb{E}[X]}}$$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{\mathbb{E}[X]}.$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot 0} + \Pr[X_i = 1] \cdot e^{t\cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}$$

$$M_X(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t - 1)p} = e^{(e^t - 1)np} = e^{(e^t - 1)\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\varepsilon)\mathbb{E}[X]}} \leq \min_{t > 0} \frac{e^{(e^t - 1)\mathbb{E}[X]}}{e^{t(1+\varepsilon)\mathbb{E}[X]}}.$$
### Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\epsilon > 0$

$$\Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \left(\frac{e^\epsilon}{(1 + \epsilon)^{1+\epsilon}}\right)^{\mathbb{E}[X]}.$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot0} + \Pr[X_i = 1] \cdot e^{t\cdot1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

$$1 + x \leq e^x$$

$$M_X(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t-1)p} = e^{(e^t-1)np} = e^{(e^t-1)\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\epsilon)\mathbb{E}[X]}} \leq \min_{t>0} \frac{e^{(e^t-1)\mathbb{E}[X]}}{e^{t(1+\epsilon)\mathbb{E}[X]}} = \min_{t>0} \left(\frac{e^{(e^t-1)}}{e^{t(1+\epsilon)}}\right)^{\mathbb{E}[X]}$$
Application: Binomial Distribution

**Theorem:** Let \( X \sim \text{Bin}(n, p) \). Then for any \( \varepsilon > 0 \)

\[
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left( \frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}} \right)^\mathbb{E}[X].
\]

**Proof** Consider \( X \) as the sum of independent \( X_i \sim \text{Ber}(p) \)

\[
M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}
= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}
\]

\[
M_X(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t - 1)p} = e^{(e^t - 1)np}
= e^{(e^t - 1)\mathbb{E}[X]}.
\]

\[
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\varepsilon)\mathbb{E}[X]}} \leq \min_{t > 0} \frac{e^{(e^t - 1)\mathbb{E}[X]}}{e^{t(1+\varepsilon)\mathbb{E}[X]}} = \min_{t > 0} \left( \frac{e^\varepsilon}{e^{t(1+\varepsilon)}} \right) \mathbb{E}[X] \leq \left( \frac{e^\varepsilon}{(1+\varepsilon)^{(1+\varepsilon)}} \right)^\mathbb{E}[X] \text{ for } t = \log(1 + \varepsilon)
\]
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}}\right)^{\mathbb{E}[X]}.$$ 

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot 0} + \Pr[X_i = 1] \cdot e^{t\cdot 1}$$

$$= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

$$1 + x \leq e^x$$

$$M_X(t) = M_{\sum X_i}(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t-1)p} = e^{(e^t-1)np} = e^{(e^t-1)\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\varepsilon)\mathbb{E}[X]}} \leq \min_{t>0} \frac{e^{(e^t-1)\mathbb{E}[X]}}{e^{t(1+\varepsilon)\mathbb{E}[X]}} = \min_{t>0} \left(\frac{e^{(e^t-1)}}{e^{t(1+\varepsilon)\mathbb{E}[X]}}\right)^{\mathbb{E}[X]} \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}}\right)^{\mathbb{E}[X]}$$

**Example**

- Sum of 20 unfair \{0, 1\}-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$, $\mathbb{E}[X] = 4$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta}.$$ 

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$ 

**Moment Addition:** Independent $X, Y$:

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t).$$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left( \frac{e^{\varepsilon}}{(1 + \varepsilon)^{1+\varepsilon}} \right)^{\mathbb{E}[X]}.$$

**Proof**
Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}$$

$$= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

$$1 + x \leq e^x$$

$$M_X(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{(e^t-1)p} = e^{(e^t-1)np} = (1 - p + pe^t)^n = (1 + x)^n = \left(1 + \frac{x}{1 + \varepsilon}\right)^{\mathbb{E}[X]}$$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\varepsilon)\mathbb{E}[X]}} \leq \min_{t>0} \frac{e^{(e^t-1)p} \mathbb{E}[X]}{e^{t(1+\varepsilon)\mathbb{E}[X]}} = \min_{t>0} \left( \frac{e^{(e^t-1)} \mathbb{E}[X]}{e^{t(1+\varepsilon)} \mathbb{E}[X]} \right) \leq \left( \frac{e^{\varepsilon}}{(1 + \varepsilon)^{1+\varepsilon}} \right)^{\mathbb{E}[X]}$$

**Example**
- Sum of 20 unfair $\{0, 1\}$-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$, $\mathbb{E}[X] = 4$
- $\Pr[X \geq 16] = \Pr[X \geq (1 + 3)\mathbb{E}[X]]$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\epsilon > 0$

$$
\Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \left( \frac{e^\epsilon}{(1 + \epsilon)^{1+\epsilon}} \right)^{\mathbb{E}[X]}
$$

**Proof**
Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$M_X(t) = \mathbb{E}[e^{tX}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1}
= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}
$$

$$M_X(t) = \prod_{i=1}^n X_i(t) \leq \prod_{i=1}^n e^{(e^t - 1)p} = e^{(e^t - 1)np} = e^{(e^t - 1)\mathbb{E}[X]}
$$

$$\Pr[X \geq (1 + \epsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\epsilon)\mathbb{E}[X]}} \leq \min_{t > 0} \frac{e^{(e^t - 1)\mathbb{E}[X]}}{e^{t(1+\epsilon)\mathbb{E}[X]}} = \min_{t > 0} \left( \frac{e^{(e^t - 1)\mathbb{E}[X]}}{e^{t(1+\epsilon)\mathbb{E}[X]}} \right) \leq \left( \frac{e^\epsilon}{(1 + \epsilon)^{1+\epsilon}} \right)^{\mathbb{E}[X]}
$$

**Example**
- Sum of 20 unfair $\{0,1\}$-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$, $\mathbb{E}[X] = 4$
- $\Pr[X \geq 16] = \Pr[X \geq (1 + 3)\mathbb{E}[X]] \leq \left( \frac{e^3}{(1 + \epsilon)^{1+3}} \right)^4$

**Chernoff:** Random variable $X$ and $a > 0$:

$$
\Pr[X \geq a] \leq \min_{t > 0} \mathbb{E}[e^{tX}]/e^{ta}.
$$

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$

**Moment Addition:** Independent $X, Y$:

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t).$$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}}\right)^{\mathbb{E}[X]}.
$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$
M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t \cdot 0} + \Pr[X_i = 1] \cdot e^{t \cdot 1} = (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t - 1)p}
$$

$$
M_X(t) = M\sum x_i(t) = \prod_{i=1}^n M_{X_i}(t) \leq \prod_{i=1}^n e^{((e^t - 1)p) = e^{(e^t - 1)p}}
$$

$$
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]^{\mathbb{E}[X]}\cdot e^{t\mathbb{E}[X]}}{e^{(t(1+\varepsilon))\mathbb{E}[X]}} = \min_{t > 0} \left(\frac{e^{(e^t - 1)p}}{e^{(1+\varepsilon)(1+\varepsilon)}}\right)^{\mathbb{E}[X]} \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}}\right)^{\mathbb{E}[X]}
$$

**Example**

- Sum of 20 unfair \(\{0, 1\}\)-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$, \(\mathbb{E}[X] = 4\)
- $\Pr[X \geq 16] = \Pr[X \geq (1 + 3)\mathbb{E}[X]] \leq \left(\frac{e^3}{(1+3)^{1+3}}\right)^4 = \frac{e^{12}}{4^{14}} \approx 0.00003789$
Application: Binomial Distribution

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left( \frac{e^\varepsilon}{(1 + \varepsilon)^{(1+\varepsilon)}} \right)^{\mathbb{E}[X]}
$$

**Proof** Consider $X$ as the sum of independent $X_i \sim \text{Ber}(p)$

$$
M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = \Pr[X_i = 0] \cdot e^{t\cdot 0} + \Pr[X_i = 1] \cdot e^{t\cdot 1}
$$

$$
= (1 - p) + pe^t = 1 + (e^t - 1)p \leq e^{(e^t-1)p}
$$

$$
M_X(t) = \prod_{i=1}^{n} M_{X_i}(t) \leq \prod_{i=1}^{n} e^{(e^t-1)p} = e^{(e^t-1)\cdot np} = e^{(e^t-1)\mathbb{E}[X]}
$$

$$
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\varepsilon)\mathbb{E}[X]}} \leq \min_{t>0} \frac{e^{(e^t-1)\cdot np}}{e^{t(1+\varepsilon)\cdot np}} = \min_{t>0} \left( \frac{e^{(e^t-1)}}{e^{t(1+\varepsilon)}} \right)^{\mathbb{E}[X]} \leq \left( \frac{e^\varepsilon}{(1+\varepsilon)^{(1+\varepsilon)}} \right)^{\mathbb{E}[X]}
$$

**Example** 1 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 1 0 1 0 0 0 0 1 1 0 0 0 0 0 0 1 1

- Sum of 20 unfair $\{0, 1\}$-coin tosses: $X \sim \text{Bin}(20, \frac{1}{5})$, $\mathbb{E}[X] = 4$
- $\Pr[X \geq 16] = \Pr[X \geq (1 + 3)\mathbb{E}[X]] \leq \left( \frac{e^3}{(1+3)^{1+3}} \right)^4 = \frac{e^{12}}{4^{12}} \approx 0.00003789$

**Chernoff:** Random variable $X$ and $a > 0$:

$$
\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}]/e^{ta}.
$$

**Mom. Gen. Function:** $M_X(t) = \mathbb{E}[e^{tX}]$

**Moment Addition:** Independent $X, Y$:

$$
M_{X+Y}(t) = M_X(t) \cdot M_Y(t).
$$

**Markov:** $\leq 0.25$

**Chebychev:** $\approx 0.022$

**Actual:** $\approx 0.000000138$
Chernoff – Simpler Versions

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}}\right)^{\mathbb{E}[X]}.$$

**Chernoff:** Random variable $X$ and $a > 0$:

$$\Pr[X \geq a] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}.$$
### Chernoff – Simpler Versions

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left( \frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}} \right)^{\mathbb{E}[X]}.$$

**Corollary:** Let $X \sim \text{Bin}(n, p)$. Then for any $t \geq 6\mathbb{E}[X]$, $\Pr[X \geq t] \leq 2^{-t}$.
Chernoff – Simpler Versions

**Theorem:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$

$$\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}}\right)^\mathbb{E}[X].$$

**Corollary:** Let $X \sim \text{Bin}(n, p)$. Then for any $t \geq 6\mathbb{E}[X]$, $\Pr[X \geq t] \leq 2^{-t}$.

**Corollary:** Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon \in (0, 1]$, $\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2/3\mathbb{E}[X]}$. 

**Chernoff:** Random variable $X$ and $a > 0$:

$\Pr[X \geq a] \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$. 
Chernoff – Simpler Versions

**Theorem:** Let \( X \sim \text{Bin}(n, p) \). Then for any \( \varepsilon > 0 \)
\[
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left( \frac{e^{\varepsilon}}{(1 + \varepsilon)^{(1+\varepsilon)}} \right)^{\mathbb{E}[X]}.
\]

**Corollary:** Let \( X \sim \text{Bin}(n, p) \). Then for any \( t \geq 6\mathbb{E}[X], \) \( \Pr[X \geq t] \leq 2^{-t} \).

**Corollary:** Let \( X \sim \text{Bin}(n, p) \). Then for any \( \varepsilon \in (0, 1], \) \( \Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2/3\cdot\mathbb{E}[X]} \).

**Corollary:** Let \( X \sim \text{Bin}(n, p) \). Then for any \( \varepsilon \in (0, 1), \) \( \Pr[X \leq (1 - \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2/2\cdot\mathbb{E}[X]} \).

**Chernoff:** Random variable \( X \) and \( a > 0 \):
\[
\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}] / e^{ta}.
\]
Chernoff – Simpler Versions

Theorem: Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon > 0$
\[
\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq \left( \frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}} \right)^{\mathbb{E}[X]}.
\]

Corollary: Let $X \sim \text{Bin}(n, p)$. Then for any $t \geq 6\mathbb{E}[X]$, $\Pr[X \geq t] \leq 2^{-t}$.

Corollary: Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon \in (0, 1]$, $\Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2/3}\cdot\mathbb{E}[X]$.

Corollary: Let $X \sim \text{Bin}(n, p)$. Then for any $\varepsilon \in (0, 1)$, $\Pr[X \leq (1 - \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2/2}\cdot\mathbb{E}[X]$.

In fact, these also work when the $X_i$ are Bernoulli random variables with different success probabilities.

Chernoff: Random variable $X$ and $a > 0$: $\Pr[X \geq a] \leq \min_{t>0} \mathbb{E}[e^{tX}] / e^{ta}$.
Conclusion

Concentration

- Is a random variable likely to yield values close to the expectation?
- Concentration inequalities bound the probability for a random variable to deviate from its expectation.
Conclusion

Concentration
- Is a random variable likely to yield values close to the expectation?
- Concentration inequalities bound the probability for a random variable to deviate from its expectation

Moments
- Used to characterize the shape of a distribution
- First moment: expected value
- Second moment: variance
- Moment generating functions to determine higher-order moments
Conclusion

Concentration
- Is a random variable likely to yield values close to the expectation?
- Concentration inequalities bound the probability for a random variable to deviate from its expectation

Moments
- Used to characterize the shape of a distribution
- First moment: expected value
- Second moment: variance
- Moment generating functions to determine higher-order moments

Concentration Inequalities
- Markov: Based on first moment
- Chebychev: Squaring within Markov (utilizing second moment)
- Chernoff: Exponentiating within Markov (utilizing moment generating functions)
- Examples: Sum of coin flips, empirical degree distribution of ER graphs